INL/RPT-23-71699

Light Water Reactor Sustainability Program

Summary of Technical Peer Review on the Risk Assessment Framework proposed in Report INL/RPT-22-68656 for Digital Instrumentation and Control Systems

March 2023

U.S. Department of Energy

Office of Nuclear Energy

INL/RPT-23-71699

Light Water Reactor Sustainability Program

Summary of Technical Peer Review on the Risk Assessment Framework proposed in Report INL/RPT-22-68656 for Digital Instrumentation and Control Systems

Han Bao[1], Tate Shorthill[2], Edward Chen[3], Sai Zhang[1], Svetlana Lawrence[1]

March 2023

[1]Idaho National Laboratory
Idaho Falls, ID 83415

[2]University of Pittsburgh
Pittsburgh, PA 152601

[3]North Carolina State University
Raleigh, NC 27695

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

EXECUTIVE SUMMARY

This report summarizes the peer review activities initiated by Idaho National Laboratory during fiscal year (FY) 2023 for the evaluation and improvement of the methodology developed under the U.S. Department of Energy's Light Water Reactor Sustainability Program, Risk Informed Systems Analysis Pathway, digital instrumentation and control (DI&C) risk assessment project. In FY 2019, the Risk Informed Systems Analysis Pathway initiated a project to develop a risk assessment strategy for delivering a technical basis to support effective and secure DI&C technologies for digital upgrades/designs. A framework was proposed for this strategy, which aims to (1) provide a best-estimate, risk-informed capability to quantitatively and accurately estimate the risk impact of plant modernization, considering the introduction of high safety-significant safety-related DI&C systems, (2) support and supplement existing risk-informed DI&C design guides by providing quantitative risk information and evidence, (3) offer a capability of design architecture evaluation of various DI&C systems, (4) assure the long-term safety and reliability of high safety-significant safety-related DI&C systems, and (5) reduce uncertainty in costs and support integration of DI&C systems in the plant.

This project's research and development efforts from FY 2019 through FY 2022 were focused on methodology improvement and demonstration of the proposed framework for the risk assessment and design optimization of safety-critical DI&C systems. Collaboration with the nuclear industry have been initiated to support the reliability and risk assessment of their DI&C systems by using the proposed framework. In FY 2023, the framework reached a point for a technical peer review and stakeholder feedback. Peer review activities include coordinating the reviews performed by a group of industry stakeholders, documenting the peer review feedback, and providing resolutions and responses to the peer review comments.

The objective of this technical peer review is to obtain representative feedback on the proposed framework to improve the technical qualities of its methodology and readiness for deployment to the industry. Feedback may identify potential areas for improvement and further development. The subject-matter experts were invited to review the latest project report documenting the methodology developed in the project and provide evaluations of the technical qualities of the proposed framework and relevant methods. The reviewed project report is *An Integrated Framework for Risk Assessment of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants: Methodology and Demonstration* (i.e., INL/RPT-22-68656 in this report).

This peer review report documents the technical questions provided for supporting the technical peer review and introduces the technical peer reviewers representing industry stakeholders. Comments from technical peer reviewers and the resolutions and responses to these comments are outlined. Insights and lessons learned from the technical peer review are summarized in conclusions and future work.

The primary audience of this report are DI&C designers, engineers, and probabilistic risk assessment practitioners. This includes stakeholders, such as the nuclear utilities and regulators who consider the deployment and upgrade of DI&C systems, DI&C software developers and reviewers, and cybersecurity specialists.

ACKNOWLEDGEMENT

The peer review activities documented in this report are funded by the United States Department of Energy's Light Water Reactor Sustainability Program, Risk Informed Systems Analysis Pathway.

The authors would like to thank the technical peer reviewers from stakeholders including Nathan DeKett and Dennis Henneke at GE Hitachi Nuclear Energy, Christopher Hunter at the U.S. Nuclear Regulatory Commission, Hyun Gook Kang at Rensselaer Polytechnic Institute, Matt Gibson and John Weglian at the Electric Power Research Institute for their constructive comments.

The authors would like to acknowledge our collaborators from academia and the industry: Nam Dinh's research group at North Carolina State University, Heng Ban's research group at University of Pittsburgh, Carl Elks' research group at Virginia Commonwealth University, and Edward Quinn at Paragon Energy Solutions for their valuable support in methodology development and demonstration documented in the peer reviewed report INL/RPT-22-68656. The authors would also like to recognize the technical comments from Robert Youngblood and Zhegang Ma at Idaho National Laboratory to this peer review report.

CONTENTS

TABLES

ACRONYMS

APR-1400	Advanced Power Reactor 1400 MW
ATWS	anticipated transient without scram
BAHAMAS	Bayesian and HRA-Aided Method for the Reliability Analysis of Software
BBN	Bayesian Belief Network
BFM	beta-factor model
BP	bistable processor
BTP	branch technical position
CCCG	common cause component group
CCF	common cause failure
CDF	core damage frequency
DI&C	digital instrumentation and control
DOE	U.S. Department of Energy
EOC	error of commission
EOO	error of omission
EPRI	Electric Power Research Institute
ESFAS	Engineered Safety Features Actuation System
ET	event tree
FMEA	failure mode effect analysis
FT	fault tree
FTA	fault tree analysis
FY	fiscal year
GEH	GE Hitachi Nuclear Energy
HAZCADS	Hazards and Consequence Analysis for Digital Systems
HEP	Human-Error Probability
HRA	human reliability analysis
HSSSR	High Safety-significant Safety-related
IE	initiating event
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IFP	information/feedback pathway
INL	Idaho National Laboratory
KAERI	Korea Atomic Energy Research Institute
KAIST	Korea Advanced Institute of Science and Technology

LERF	large early release frequency
LOCA	loss-of-coolant accident
LOSC	loss-of-seal cooling
LP	logic processor
LPCI	low-pressure core injection
LWRS	Light Water Reactor Sustainability
MBLOCA	medium-break LOCA
ML	machine learning
NEI	Nuclear Energy Institute
NPP	nuclear power plant
NRC	U.S. Nuclear Regulatory Commission
ODC	orthogonal-defect classification
ORCAS	Orthogonal-defect Classification for Assessing Software reliability
PLC	programmable logic controller
PWR	pressurized-water reactor
PRA	probabilistic risk assessment
QIAS-P	qualified indication and alarm system – safety
R&D	research and development
RESHA	Redundancy-guided Systems-theoretic Hazard Analysis
RG	Regulatory Guide
RISA	Risk-Informed Systems Analysis
RPI	Rensselaer Polytechnic Institute
RPS	reactor protection system
SBLOCA	small-break LOCA
SDLC	software development life cycle
SAPHIRE	Systems Analysis Programs for Hands-on Integrated Reliability Evaluations
SIL	safety integrity level
SRGM	software reliability growth method
SSC	system, structure, and component
STPA	systems-theoretic process analysis
THERP	Technique for Human-Error Rate Prediction
UCA	unsafe control action
UIF	unsafe information flow

SUMMARY OF TECHNICAL PEER REVIEW ON THE RISK ASSESSMENT FRAMEWORK PROPOSED IN REPORT INL/RPT-22-68656 FOR DIGITAL INSTRUMENTATION AND CONTROL SYSTEMS

1. INTRODUCTION

This report summarizes the peer review activities initiated by Idaho National Laboratory (INL) during fiscal year (FY) 2023 for evaluating and improving the methodology developed under the U.S. Department of Energy's (DOE's) Light Water Reactor Sustainability (LWRS) Program, Risk Informed Systems Analysis (RISA) Pathway, digital instrumentation and control (DI&C) risk assessment project [1] [2] [3] [4] [5].

1.1 Scope and Objective

The LWRS program, sponsored by the U.S. DOE and coordinated through a variety of mechanisms and interactions with industry, vendors, suppliers, regulatory agencies, and other industry research and development (R&D) organizations, conducts research to develop technologies and other solutions to improve economics and reliability, sustain safety, and extend the operation of nation's fleet of nuclear power plants (NPPs). The LWRS program has two objectives to maintain the long-term operations of the existing fleet: (1) to provide science- and technology-based solutions to industry to implement technology to exceed the performance of the current business model and (2) to manage the aging of systems, structures, and components (SSCs) so NPP lifetimes can be extended, and the plants can continue to operate safely, efficiently, and economically. As one of the LWRS program's R&D pathways, RISA Pathway aims to support decision-making related to economics, reliability, and safety by providing integrated plant and systems analysis solutions through collaborative demonstrations to enhance economic competitiveness of the operating fleet. The RISA Pathway R&D's purpose is to support plant owner-operator decisions with the aim to improve economics and reliability and maintain the high levels of current NPPs' safety over periods of extended plant operations. To achieve this purpose, RISA Pathway conducts R&D to optimize safety margins and minimize uncertainties to achieve economic efficiencies while maintaining high levels of safety. This is accomplished in two ways: (1) by providing scientific basis to better represent safety margins and factors that contribute to cost and safety; and (2) by developing new technologies that reduce operating costs.

One research effort under the RISA Pathway is the DI&C Risk Assessment project, which was initiated in FY 2019 to develop a risk assessment strategy for delivering a strong technical basis to support effective and secure DI&C technologies for digital upgrades/designs [1]. An integrated risk assessment framework for the DI&C systems (i.e., "the LWRS-developed framework", or "the proposed framework" or "the framework" in this report) was proposed for this strategy which aims to:

- Provide a best-estimate, risk-informed capability to quantitatively and accurately estimate the safety margin obtained from plant modernization, especially for the high safety-significant safety-related (HSSSR) DI&C systems
- Support and supplement existing advanced risk-informed DI&C design guides by providing quantitative risk information and evidence
- Offer a capability of design architecture evaluation of various DI&C systems to support system design decisions and diversity and redundancy applications
- Assure the long-term safety and reliability of HSSSR DI&C systems

- Reduce uncertainty in costs and support integration of DI&C systems at NPPs.

The R&D efforts of this project from FY 2019 through FY 2022 were focused on methodology improvement and demonstration of the proposed framework for the risk assessment and design optimization of safety-critical DI&C systems. Collaboration with the nuclear industry has been initiated to support the reliability and risk assessment of their DI&C systems by using the proposed framework. In FY 2023, the framework has reached the point where a technical peer review and stakeholder feedback is beneficial. This peer review activity includes coordinating the reviews performed by a group of industry stakeholders, documenting peer review feedback, providing resolutions and responses to the peer review comments.

The objective of this technical peer review is to obtain representative feedback on the proposed framework to improve the technical qualities of its methodology and readiness for deployment to the industry. Feedback may identify potential areas for improvement and further development.

The subject matter experts were invited to review the latest project report documenting the methodology developed and provide a technical evaluation of the proposed framework and relevant methods. The reviewed project report is "An Integrated Framework for Risk Assessment of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants: Methodology and Demonstration" INL/RPT-22-68656 [4] (i.e., "INL/RPT-22-68656" in this report).

1.2 Target Audience

The primary audience of this report are DI&C designers, engineers, and probabilistic risk assessment (PRA) practitioners. This includes stakeholders, such as the nuclear utilities and regulators who consider the deployment and upgrade of DI&C systems, DI&C software developers and reviewers, and cybersecurity specialists.

1.3 Organization of the Report

The remaining sections of the report are organized as follows: Section 2 briefly reviews the sections and technical topics of INL/RPT-22-68656. Section 3 documents the technical questions provided for technical peer review. Section 4 introduces the technical peer reviewers. Section 5 lists the comments from technical peer reviewers and the resolutions and responses to these comments. Section 6 summarizes the insights and lessons learned from the technical peer review and outlines potential future work. The peer review narrative is included in Appendix A.

2. OVERVIEW OF THE REVIEWED PROJECT REPORT

This section briefly describes the sections and technical topics of the peer reviewed report (i.e., INL/RPT-22-68656) [4]. INL/RPT-22-68656 documents the activities performed by INL during FY 2022 for the LWRS-RISA project, DI&C Risk Assessment based on collaboration with the University of Pittsburgh, North Carolina State University, Virginia Commonwealth University and Technology Resources.

The R&D efforts of this project in FY 2022 were focused on methodology improvement and demonstration of the LWRS-developed framework for the risk assessment and design optimization of safety-critical DI&C systems. In FY 2022, this framework was further developed with a capability to trace software failures in digital feedback pathways in highly redundant safety-critical DI&C systems; potential failures to a DI&C system are organized in a fault tree (FT) for clear visual and linear traceability. Case studies demonstrated the identification of digital failure mechanisms in key instrumentation, the construction of software FT for highly complex DI&C systems, and the identification of potential single points of failure and common cause failures (CCFs). In FY 2022, this framework was further developed with a capability to trace software failures in digital feedback pathways in highly redundant safety-critical DI&C systems. All these capabilities offer a common and modularized platform to DI&C designers, software developers, cybersecurity analysts, and plant engineers for evaluating various design architectures of DI&C systems to support system design decisions in diversity and redundancy applications.

INL/RPT-22-68656 has eight sections:

- **Section 1: Introduction** briefly introduces the scope, objective, and layout.
- **Section 2: Technical Background** provides the technical background to identify and quantify risks associated with HSSSR DI&C systems. Subsections include:
    - Section 2.1 provides background details for relevant efforts performed in the last few years
    - Section 2.2 reviews regulatory positions and guidance, especially U.S. Nuclear Regulatory Commission (NRC)'s current DI&C CCF policy and future extension plan
    - Section 2.3 briefly introduces the proposed framework for the risk assessment of the HSSSR DI&C systems
    - Section 2.4 describes the value propositions of the proposed framework.
- **Section 3: Redundancy-Guided System-Theoretic Hazard Analysis (RESHA)** documents the methodology development and demonstration of a hazard analysis method incorporated in the framework, called redundancy-guided systems-theoretic method for hazard analysis (RESHA). Subsections include:
    - Section 3.1 provides an overview of RESHA introducing some basic concepts and terms
    - Section 3.2 discusses RESHA's capability to trace software failures in digital feedback pathways in highly redundant DI&C systems
    - Section 3.3 describes RESHA's methodology in detail
    - Section 3.4 discusses the results of case studies in the hazard analysis of a representative digital reactor trip system (RTS), engineered safety features actuation system (ESFAS), and safety-related human system interface (HSI).
- **Section 4: Multiscale Quantitative Reliability Analysis** documents the methodology to develop and demonstrate a multiscale, quantitative reliability analysis approach of the proposed

framework for DI&C risk assessment. The goal of the multiscale quantitative reliability analysis is to estimate the DI&C system reliability by calculating the integrated FT of DI&C systems obtained in the hazard analysis, then provide inputs for following consequence analysis. For the reliability analysis of large-scale DI&C systems, the quantitative small-scale reliability analysis of software and Type II interactions in DI&C systems are also included in the reliability analysis workflow. Subsections include:

  - Section 4.1 overviews the approach of the multiscale reliability analysis of DI&C system
  - Section 4.2 describes the technical backgrounds for the development of ORCAS (Orthogonal-defect Classification for Assessing Software reliability) and BAHAMAS (Bayesian and human reliability analysis [HRA]-Aided Method for the Reliability Analysis of Software)
  - Sections 4.3 and 4.4 respectively provide the technical details of ORCAS, a method for software reliability analysis when testing data is available and sufficient, and BAHAMAS, which was developed for software reliability analysis in data-limited conditions
  - Section 4.5 introduces the CCF modeling method applied for both hardware and software failures in safety-critical DI&C systems.

- **Section 5: Consequence Analysis of a Generic PWR With Advanced HSSSR DI&C Systems** documents the consequence analysis of a generic pressurized-water reactor (PWR) PRA model with improved RTS, ESFAS, and HSI fault trees (FTs), which quantitatively evaluates how the previously identified and quantified CCFs affect the overall plant safety. Subsections include:
  - Section 5.1 describes the generic PWR PRA model developed using the SAPHIRE (Systems Analysis Programs for Hands-on Integrated Reliability Evaluations) tool
  - Section 5.2 introduces the scenario to be analyzed as well as the original event tree (ET) models for these scenarios including a FT for the failure of an analog RTS and one CCF basic event for the ESFAS failure
  - Section 5.3 compares the original FTs for analog RTS and ESFAS and the new FTs for digital RTS and ESFAS
  - Section 5.4 discusses the results for consequence analysis of these selected ET models.
- **Section 6: Future Applications on AI-Aided Control Systems** discusses the motivation, applicability, limitations, and potential benefits of the proposed framework in the risk assessment and design optimization of potential artificial intelligence (AI)-aided control systems in future reactor designs and upgrades.
- **Section 7: Conclusions and Future Works** outlines the conclusions and future work of this project.
- **Section 8: References** lists the report's references.
- **Appendix A: Human Reliability Analysis in BAHAMAS** provides a demonstration of an HRA method, technique for human-error rate prediction, for the evaluation root nodes within the BAHAMAS Bayesian belief network (BBN). The root nodes of BAHAMAS represent general stages of the software development life cycle (SDLC). Quantification of the root nodes for the BBN requires that details of the SDLC for a given software must be provided. This appendix details the application of HRA as part of the case study for the reliability analysis of bistable processors found within the four-division digital RTS.

- **Appendix B: Selected Accident Scenarios for Consequence Analysis** presents the ET models of the five selected accident scenarios for consequence analysis, including INT-TRANS (initiating event - general plant transient) with ATWS (anticipated transient without scram), LOSC (loss-of-seal cooling), SBLOCA (small-break loss-of-coolant accident [LOCA]), and MBLOCA (medium-break LOCA).

3. TECHNICAL QUESTIONS FOR PEER REVIEW

This section documents the technical questions provided for technical peer review. The 12 technical questions were designed to focus on specific technical topics/sections (Section 2 - 7) of the report to assist with peer reviews. These questions are documented below.

For Section 2: Technical Background

1. The proposed framework presented in Figure 2 includes three steps (e.g., hazard, reliability, and consequence analysis) for risk analysis and defines three acceptance criteria for risk evaluation. Is its workflow clear and complete for DI&C system risk analysis and evaluation? Are the steps and acceptance criteria well defined and sufficient to provide insights to reduce risks and optimize designs?

2. Do you have any suggestions for overall framework improvements, if anything, on the identification, quantification, and evaluation of potential DI&C system failures?

3. The framework is expected to support existing industry guidance and methods (e.g., HAZCADS and DRAM) by providing quantitative risk information. In your opinion, which aspects of the framework can be utilized to support existing industry design and evaluation guidance? What adjustments and/or improvements are needed, if any, to provide better support to existing industry guidance and methods?

For Section 3: Redundancy-Guided System-Theoretic Hazard Analysis (RESHA)

4. The framework's hazard analysis method is called RESHA and its workflow is presented in Figure 11. With respect to the complexity of High Safety-significant Safety-related (HSSSR) DI&C systems considered, do you think Section 3 has clearly described and demonstrated the RESHA capability in identifying potential CCFs, especially software CCFs, for different levels of redundancy and failure mode types?

5. Based on the integration of STPA and HAZCADS, RESHA can identify software failure mechanisms in the control/actuation pathway using Unsafe Control Actions (UCAs). A novel concept called Unsafe Information Flow (UIF) has been developed and introduced in RESHA to complete the identification and tracing of software failure mechanisms in the information/feedback pathway (IFP), as discussed in Section 3.2. Please provide your feedback for the development and application of UIFs in identifying software failure mechanisms in the IFP.

For Section 4: Multiscale Quantitative Reliability Analysis

6. The framework's multiscale quantitative reliability analysis workflow presented in Figure 25 includes two methods for software reliability analysis, ORCAS and BAHAMAS.

   a. In Section 4.3, ORCAS is developed and demonstrated to estimate the probability of UCAs/UIFs for rich-data conditions. Please provide your feedback on: (1) are the assumptions of ORCAS reasonable? (2) are input requirements and steps of ORCAS clearly described, logical and practical for deployment with rich testing data available? (3) is ORCAS method well demonstrated in the case studies with all expected outcomes obtained?

   b. In Section 4.4, BAHAMAS is developed and demonstrated to estimate the probability of UCAs/UIFs for limited-data conditions. Please provide your feedback on: (1) are the assumptions of BAHAMAS reasonable? (2) are input requirements and steps of BAHAMAS clearly described, logical and practical for deployment with no testing data and very limited design information? (3) is BAHAMAS method well demonstrated in the case studies with all expected outcomes obtained?

7. In Section 4.5, a modified beta factor method was developed and modified to support CCF modeling and parameter estimation as part of this framework. Please provide your feedback on: (1) are the method assumptions reasonable? (2) are the input requirements and steps clearly described, logical

and practical for deployment to the industry? (3) is the method well demonstrated in the case studies with all expected outcomes obtained?

For Section 5: Consequence Analysis of a Generic PRW with Advanced HSSSR DI&C Systems

8. Section 5 documents the consequence analysis of a generic PWR SAPHIRE model with an improved digital reactor trip system (RTS) and Engineered Safety Features Actuation System (ESFAS) FTs. Please provide your feedback on: (1) Are the demonstrated consequence analyses on HSSSR DI&C systems sufficiently detailed to support industry needs in DI&C modeling and risk analysis? (2) Is it beneficial to perform similar consequence analyses for other DI&C systems that are safety-related but not safety-significant?

9. In addition to providing the changes of core damage frequency (CDF) and large early release frequency (LERF) due to the digital upgrades for HSSSR I&C systems, what other risk-informed insights from the quantitative consequence analysis would be beneficial for the evaluation and reduction of the plant-level risks?

For Section 6: Future Applications on AI-Aided Control Systems

10. Some technical gaps in applying the framework for the risk analysis of AI/ML-aided control systems have been reviewed and investigated in Section 6. With respect to the technical gaps considered, are they complete and significant in terms of identifying, quantifying, and evaluating potential failure modes of AI-aided control systems? Are there any other approaches to addressing these types of technical gaps that you recommend us to consider?

For Section 7: Conclusions and Future Works

11. Section 7 summarizes recent work and proposes future R&D. Do you agree with the identified needs for these activities? Are there other relevant short- and long-term industry needs in the area of DI&C risk assessment not addressed by this framework?

12. Please provide any additional suggestions for the framework improvements.

4. TECHNICAL PEER REVIEWERS

This section introduces the technical peer reviewers. A total of six technical peer reviewers from the GE Hitachi Nuclear Energy (GEH), the NRC, Rensselaer Polytechnic Institute (RPI), and the Electric Power Research Institute (EPRI) were invited and provided their comments to the INL/RPT-22-68656. Peer reviewers' information is listed in Table 1. Table 2 shows a matrix of the technical questions and feedback from the peer reviewers.

Table 1. The list of technical peer reviewers from stakeholders.

NO.	Technical Peer Reviewers			
	Name	**Affiliation**	**Title**	**Brief Bio**
1	Nathan DeKett	GEH	Senior PRA Engineer	Nathan DeKett is a Senior PSA Engineer at GEH Nuclear Energy and has 12 years of experience in nuclear power plant PSAs. He has experience developing PSAs and performing risk-informed analysis at an electric utility, as a consultant supporting the operating fleet, and in support of designing GEH's new BWRX-300. Nathan is the project lead responsible for developing the BWRX-300 PSA models and integrating analyses into the risk-informed design of the plant. Nathan received his BSE in Nuclear Engineering and Radiological Sciences from the University of Michigan.
2	Dennis Henneke	GEH	Consulting Engineer	Dennis Henneke is a Consulting Engineer at GEH Nuclear Energy, with 41 years of PSA experience. He is performing technical oversight for the PSAs supporting the BWRX-300, VTR, and Natrium Reactors, was the Principal Investigator for the DOE Funded project for the PRISM reactor on "Development/Modernization of an Advanced Non-LWR Probabilistic Risk Assessment," and was the Technical Lead for the UK ABWR PRA development completed in 2018. Dennis has authored or co-authored more than 100 publications and is the ANS-Chairman of the ANS/ASME Joint Committee on Nuclear Risk Management. Dennis received his MS and BS in Nuclear Engineering from the University of Florida.
3	Christopher Hunter	NRC	Senior Reliability and Risk Engineer	Chris Hunter joined the NRC in 2002 as a member of the Accident Sequence Precursor (ASP) Program. He is the current program manager and senior analyst. He has performed dozens of precursor analyses, hundreds of screening analyses, and has authored several annual ASP reports. Mr. Hunter has B.S. in Chemical Engineering from State University of New York at Buffalo and was an enlisted nuclear operator in the U.S. Navy for 6 years.
4	Hyun Gook Kang	RPI	Professor	Dr. Kang is a Professor of the Department of Mechanical, Aerospace, and Nuclear Engineering at Rensselaer Polytechnic Institute (RPI). Before joining RPI, he was an Associate Professor at Korea Advanced Institute of Science and Technology (KAIST) and a senior research staff of the Korea Atomic Energy Research Institute (KAERI). After his PhD from Nuclear Engineering Department of KAIST in 1999, Dr. Kang's research focus has been on innovations of dynamic risk assessment of safety-critical applications including digitalized I&C systems. The topics include digital I&C risk, passive safety features, human errors, intelligent control and protection, and advanced emergency procedures. His long-term research goal is to develop a risk-free autonomous operation scheme for nuclear power plants. He is the vice-chair of Human Factors and I&C division of American Nuclear Society and the chair of Safety review broad of RPI reactor. He authored more than 300 journal and conference articles.

5	Matt Gibson	EPRI	Technical Executive - EPRI Nuclear I&C Program	Matt Gibson is a Technical Executive at the Electric Power Research Institute (EPRI), with over 40 years of experience in Operational Technology as well as Digital I&C systems design, implementation, and management. Matt is a licensed Control Systems Engineer and certified cyber security professional who joined EPRI in 2013 after a long career with Progress Energy, now Duke Energy, where he had various roles in the design, implementation, and support of Digital I&C systems throughout the Duke Fleet, including an 11-year stint as the Fleet Digital Architect. Matt's research focus areas at EPRI are Digital Systems Engineering, digital I&C, hazards and reliability analysis, human factors engineering, and cyber security. Matt is currently leading the Risk -Informed Digital Systems Engineering portfolio at EPRI which seeks to integrate and modernize the digital engineering processes and methods used in the nuclear power industry.
6	John Weglian	EPRI	Principal Technical Leader	John Weglian is a Principal Technical Leader at EPRI. He is the project manager for the development of Probabilistic Risk Assessment (PRA) risk software including CAFTA, Phoenix Risk Monitor, Phoenix Architect, and the HRA Calculator. Prior to coming to EPRI, he was a PRA engineer for FirstEnergy where he developed risk models and performed risk assessments for the Perry Nuclear Power Plant. He spent a number of years working in the aerospace field for Lockheed Martin, as a NASA contractor through the Ohio Aerospace Institute, and at the U.S. Federal Aviation Administration. He also spent seven years in the U.S. Navy as a submarine officer. He holds a B.S. in aerospace engineering from the University of Cincinnati and an M.S. in aerospace engineering from the Georgia Institute of Technology.

Table 2. A matrix of technical questions and feedback from the technical peer reviewers.

NO.	**Technical Peer Reviewers**			
	EPRI	**GEH**	**NRC**	**RPI**
1	Y	Y	Y	Y
2	Y	Y	Y	Y
3	Y	Y	Y	Y
4	Y	Y	Y	Y
5	Y	Y	Y	Y
6	Y	Y	Y	Y
7	Y	Y	Y	Y
8	Y	Y	Y	Y
9	Y	Y	Y	Y
10	Y	Y	Y	Y
11	Y	Y	Y	Y
12	N	Y	Y	Y
Additional Comments	N	Y	N	Y
* Y = answered, N = not answered				

5. TECHNICAL COMMENTS AND RESPONSES

This section lists the comments from technical peer reviewers and the resolutions and responses from the LWRS technical team to these comments.

5.1 Technical Question #1

Technical Question #1 is "The proposed framework presented in Figure 2 includes three steps (e.g., hazard, reliability, and consequence analysis) for risk analysis and defines three acceptance criteria for risk evaluation. Is its workflow clear and complete for DI&C system risk analysis and evaluation? Are the steps and acceptance criteria well defined and sufficient to provide insights to reduce risks and optimize designs?" Technical comments and responses to Question #1 are summarized in Table 3.

Table 3. Summary of technical comments and responses to Question #1.

Reviewers	Comments	Responses/Resolutions
GEH	Yes, the figure is mostly clear. The first two questions of the acceptance criteria don't seem to fit the Yes/No output of that block and the third question is the only one that seems to really warrant an iteration of design optimization. However, I think the reader can understand the process reasonably well to figure out what the authors' intent is.	It is a good point to keep three criteria consistent. Acceptance criterion #1 and #2 will be revised as: • Is the function of digital system still available even with the identified individual failures? • Is the digital system still reliable even with the identified individual failures?
NRC	Yes, the workflow appears to be clear and complete. I am assuming that Criterion 3 includes the risk evaluation of both the loss of function for a digital system and the potential for failures to result in plant initiating events/transients.	Yes, Acceptance criterion #3 includes the risk evaluation of both the loss of function for a digital system and the potential for failures to result in plant initiating events/transients. ET quantification is conducted in consequence analysis to estimate CDF.
RPI	I think the report is addressing those elements in a reasonable detail. On the other hand, this single milestone report doesn't have to completely cover all related topics. My understanding on Figure 2 is conceptual logical diagram illustrating the scope of the study. Among those elements, software-induced CCF issues in Hazard analysis and Reliability analysis seem to be the focus of this report.	Yes, technical details of some approaches and methods that have been well defined and demonstrated in previous milestone reports are not included in this report.
EPRI	(1). i. You are using the wrong reference and concept for HAZCADS. HAZCADS is an EPRI product and should always refer to the current version which is currently 3002016698 (https://www.epri.com/research/products/000000003002016698). I have pointed this out before. While Sandia was a key collaborator for HAZCADS and did the analysis of the best methods to combine based on previous EPRI research (STPA and FTA) and	(1). i. Thanks for pointing it out. The HAZCADS reference will be updated in the revised version of INL/RPT-22-68656. ii. It is a good point to not limit hazard analysis only to component level. Hazard analysis may provide different outcomes at different levels, but under the context of this framework, hazard analysis is conducted for risk analysis mainly at component level. The authors

Reviewers	Comments	Responses/Resolutions
	discovered a novel way to integrate STPA and FTA via UCA insertion in a fault tree, the fully usable HAZCADS R0 and R1 and soon to be released Rev 2 are the products of extensive research and proof testing by EPRI and other collaborators. ii. For figure 2 directly, Hazard Analyses are effective at the System and plant Level and reliability flows up from the component level. Software is a component and so the allocation of system level risk should be identified by hazards analysis so that step of the process should be reversed to be consistent with international standards.	recognize that Figure 2 should be modified to avoid confusion. The key outcomes from each stage of risk assessment will be changed from "component-level risk evaluation", "system-level risk evaluation" and "plant-level risk evaluation" to the new outcomes "Potential Single Points of Failure", "DI&C Reliability Metrics", and "Plant Safety Metrics". These changes better highlight the key outcomes that are used to support the evaluation of the Acceptance Criteria which have also been changed. The Acceptance criteria are now given as: 1) "Is the function of digital system still available even with the identified individual failures?" 2) "Is the digital system still reliable even with the identified individual failures?" and 3) "Are the consequences of digital failures acceptable at the plant level (e.g., ΔCDF)?" Figure 2 will be updated in next version of the reviewed report.
	(2). Other than the reversal of the role of hazards and reliability analysis, the figure two outline is good. The overall process is redundant to the current HAZCADS/DRAM, which is 3 years old at this point, in production, and getting feedback.	(2). The proposed framework provides a diverse approach to EPRI's methods for DI&C risk assessment, and they have different scopes and focuses. The authors anticipate constructive benefits from both approaches that can support specific needs and interests of the industry. The two frameworks, i.e., EPRI's Digital Systems Engineering Framework (cited as "EPRI's framework" in the rest of the report) and LWRS framework, have different but relevant goals. The LWRS framework aims to develop an advanced PRA-based framework for DI&C risk assessment. It provides a best-estimate, risk-informed capability to address digital issues quantitatively, focusing on software CCFs in safety-critical DI&C systems. It also provides risk-informed metrics and insights to assist users to address and evaluate CCFs in the HSSSR DI&C systems of NPPs. It should be noted that PRA is used as a tool to provide evidence, quantitative risk analysis results can be used to support the diversity and redundancy applications in DI&C system design. The two frameworks also have different technical approaches. The LWRS framework follows the basic PRA logic from qualitative hazard analysis to quantitative reliability and consequence analysis. It answers the questions: what can go wrong? (Hazard analysis), how likely is it? (Reliability analysis) and what are the consequence? (Consequence analysis). Here PRA is used as an approach to provide various metrics and evidence to be used in reasoning, not as a

Reviewers	Comments	Responses/Resolutions
		presumptively authoritative result. EPRI's framework focuses on whether the I&C design meets stakeholder needs? (Design Engineering Guide), how risky is the I&C design? (HAZCADS) and is the I&C design good enough? (DRAM). From this point, the metrics and results of the LWRS framework can be used to support EPRI's framework. It can be used as a tool that offers various evidence and metrics for evaluation of various DI&C system design architectures to support system design decisions in diversity and redundancy applications. For example, for HAZCADS, the LWRS framework can provide detailed CCFs at different redundancy levels, quantifiable metrics to support risk importance analysis and ranking of risk reduction targets, and a quantitative consequence analysis to trace the impacts of individual failures. For DRAM, the LWRS framework can provide quantifiable software reliability metrics to check if and how much the control methods can mitigate consequences and reduce risks. More communication and collaboration between LWRS and EPRI would be beneficial to offer the best support to the industry.

5.2 Technical Question #2

Technical Question #2 is "Do you have any suggestions for overall framework improvements, if anything, on the identification, quantification, and evaluation of potential DI&C system failures?" Technical comments and responses to Question #2 are summarized in Table 4.

Table 4. Summary of technical comments and responses to Question #2.

Reviewers	Comments	Responses/Resolutions
GEH	I think it would be reasonable to introduce the potential for plant improvements that are not only focused on the DI&C platform being examined as a result of higher-than-desired risk from DI&C failures. For example, a design-phase plant could introduce another DI&C platform, introduce a new mechanical system, or change the way an existing mechanical system functions to address DI&C failures.	Agreed. The goal of this work is to provide risk-informed insights to support plant improvements. These insights could be at component level (e.g., eliminating single point failures such as some CCFs, or enhancing the reliability of some safety-critical components/software), and system level (e.g., changing to another DI&C platform that has a diverse design architecture). This risk-informed improvement should be conducted with the plant / system design teams and I&C teams from the vendors or utilities.
NRC	I do not have any suggestions on overall framework improvements.	No resolution is required/needed.

Reviewers	Comments	Responses/Resolutions
RPI	(1) INL's work will be referred to as an important basis of this field where the well-agreed method is not available yet, so providing very crisp basis of all data handled in the report is crucial. (2) This report suggests two possible software reliability quantification methods considering the availability of data (ORCAS and BAHAMAS). They would be practically useful, but we cannot claim they cover all possible domains of software reliability, so a clearer definition of the targets of those two methods should be provided. E.g., we may specify them as "SW reliability quantification method based on defect mode analysis and statistical prediction models". Based on this definition, we can add more theoretical basis to build a stronger foundation.	No resolution is required/needed for (1). For (2), the authors agree these two methods presented cannot be used to claim complete coverage for all software reliability domains. Complete coverage may not be possible considering the fault space is, in essence, infinite. The premise of these proposed methods is to develop evidence towards reliability by examining qualitative and quantitative aspects of good programming practices. Consideration of good programming practices can be those recommended in documents such as presented in 7-4.3.2-2016 IEEE. Indeed, more work is necessary to establish more evidence towards software reliability.
EPRI	i. DRAM is an EPRI product and is not correctly cited and does not contain its product ID. The proper way to site EPRI document is in the front matter of the document. (https://www.epri.com/research/products/000000003002018387) ii. You are seeking to quantify systemic issues like software or cyber based on historical statistical values because that is the tool you know but statistical quantification of systematic issues is not technically sound. The NASA paper, The Infeasibility of Quantifying the Reliability of Life-Critical Real-Time Software, Lay out those issues very well. Also, risk is an exercise in predicting the future. Statistical quantification using prior data can indeed provide a failure frequency for systematic error in a specific data set with specific bounding systematic criteria. But it cannot predict the future reliability because the organized complexity of the system is not completely defined (see "Introduction of General Systems Thinking, Weinberg, ISBN 978-0-932633-49-1). International Standards also demonstrate the random vs. systematic dichotomy by rating systems by statistical hardware capability and systematic capability separately. Software reliability must then be expressed with a rating based on the strength and completeness of the systematic	i. Thanks for pointing it out. The DRAM reference will be updated in the revised version of INL/RPT-22-68656. ii. This is a very valuable comment. The subject of software failure probability is highly contentious among the industries, academia, and regulatory bodies. To date, there does not exist a widely used industry standard or policy that accepts any type of software quantification methodology as a means for risk assessment. The reasoning is that software failures are viewed as systematic failures rather than random failures. One opinion is software failures are defined as systematic as they can be replicated with perfect accuracy if the exact conditions of failure are known. In many complex DI&C systems, it is rare to know in advance the exact input conditions to trigger a failure; therefore, software will still fail regardless of the magnitude of the development or testing effort. Many failures are discovered during operation under a unique combination of conditions. These conditions can arise spontaneously and unpredictably due to a range of complex factors such as cosmic rays, weather, human action, sequence/combination, or time. These are known as triggering events and are typically treated as random occurrences. Similar statements can be found in [6] [7]. For these reasons, in this work, software failures are treated as probabilistic due to the uncertainty of the events that can trigger them. It should be

Reviewers	Comments	Responses/Resolutions
	(qualitative) reliability characteristics and control measures applied. IEC-61508/61511.	noted that failure probability is just used as a metric to estimate software reliability for supporting the diversity and redundancy applications in DI&C system design. As suggested by the reviewers in later comments, BAHAMAS and ORCAS could be evolved to assist with software and other systematic issues, providing appropriate reliability metrics to represent the results would be very beneficial. BAHAMAS and ORCAS will be refined to provide more realistic software failure estimates, and other metrics to evaluate software reliability. Verification, validation, and uncertainty quantification (VVUQ) of BAHAMAS/ORCAS and CCF modeling approach will be conducted as one key future research in FY 2024. One of the possible solutions to address the systematic failure concern is implementation of a safety case approach where quantified reliability of DI&C system provides evidence (one of several) that supports the basis that the performance objectives of the technology are met by its design during postulated accidents or upset conditions. It will be one future work of the LWRS project.

5.3 Technical Question #3

Technical Question #3 is "The framework is expected to support existing industry guidance and methods (e.g., HAZCADS and DRAM) by providing quantitative risk information. In your opinion, which aspects of the framework can be utilized to support existing industry design and evaluation guidance? What adjustments and/or improvements are needed, if any, to provide better support to existing industry guidance and methods?" Technical comments and responses to Question #3 are summarized in Table 5.

Table 5. Summary of technical comments and responses to Question #3.

Reviewers	Comments	Responses/Resolutions
GEH	I do not foresee any limitations on which aspects have the potential to be useful. It is my judgment that the aspects of the guidance work with each other to make a complete framework for evaluation and design. I believe that it would be beneficial if the framework more explicitly considered DI&C platforms that perform several mitigation functions since failures in these systems have a greater potential to result in undesired consequences due to simultaneous failure of functions.	It is a very valuable comment. The reviewer mentions in later comments that coupled failure modes of an identical or multiple functions of digital processors may lead to complex and undesired accident scenarios. Methodology development and demonstration of a function-based risk assessment of DI&C system is an important topic and will be conducted under this project for FY 2024 R&D. The current framework and methods are mainly focused on the hazard/reliability/consequence analysis at the component,

Reviewers	Comments	Responses/Resolutions
		system, and plant levels, which will be coupled with the function-based risk assessment to provide a more comprehensive study. A collaboration with the nuclear industry will facilitate methodology demonstration and validation.
NRC	I am not familiar with the guidance and methods and, therefore, cannot comment on this question.	No resolution is required/needed.
RPI	Quantified results expressed in numerical values always matter in industry, especially for licensing. So traceable and transparent basis of data (as discussed in the previous question) should be crucial for applicability. Current industry guidance mainly covers deterministic factors like report items, team members, revision process, etc. Risk aspects of software applications are rarely covered by currently available industry standards. In that sense, this report will provide a unique guidance.	No resolution is required/needed.
EPRI	(1). i. This framework is substantially redundant to the existing Digital Systems Engineering Framework (DEG/HAZCADS/DRAM) but differs in its perspective/approach on software reliability and so would be a presumptive competitor. ii. RESHA also exhibits a misunderstanding of how STPA should be applied and has potentially made it into an FMEA. I&C systems inherit the risk of the Equipment/System under control and systems thinking is required to root out the emergent behaviors that can cause issues. iii. That said, BAHAMAS and ORCAS could be evolved to assist with software and other systematic issues. A transition to a reliability metric other than probability to represent the results would be very beneficial. (2). I think future work should concentrate on refining BAHAMAS and ORCAS to align better with International Standards and achieve compatibility and influence with the DEG/HAZCADS/DRAM ecosystem. Opportunity exists to integrate BAHAMAS/ORCAS more directly into the front end of the DEG in functional analysis and testing section as well as in the DRAM.	(1). i. The proposed framework provides a diverse approach to EPRI's methods for DI&C risk assessment, and they have different scopes and focuses. Consequently, rather than being competitive with EPRI's framework, the authors anticipate a collaborative engagement which would allow benefits from integration of both approaches that can support specific needs and interests of the industry. ii. Regarding whether RESHA reduces STPA to FMEA: It is true that the presented examples found within the report have the appearance of FMEA. But this is arguably due to the particular characteristics of the case study and chosen scope which emphasize the component-specific software failures. That said, the authors look forward to further dialogues with EPRI on this point to ensure that the authors properly understand it. For iii. and (2): The reviewers indicated that BAHAMAS and ORCAS could be evolved to assist with software and other systematic issues, refined to better align with international standards, and potentially support EPRI's framework, particularly on the front end. The authors look forward to constructive dialogue in this area that may lead to new and

Reviewers	Comments	Responses/Resolutions
		useful reliability metrics for BAHAMAS and ORCAS that benefit the nuclear industry.

5.4 Technical Question #4

Technical Question #4 is "The framework's hazard analysis method is called RESHA and its workflow is presented in Figure 11. With respect to the complexity of High Safety-significant Safety-related (HSSSR) DI&C systems considered, do you think Section 3 has clearly described and demonstrated the RESHA capability in identifying potential CCFs, especially software CCFs, for different levels of redundancy and failure mode types?" Technical comments and responses to Question #4 are summarized in Table 6.

Table 6. Summary of technical comments and responses to Question #4.

Reviewers	Comments	Responses/Resolutions
GEH	Yes, I think the authors' description and demonstration of RESHA is well done. The figures provided are helpful.	No resolution is required/needed.
NRC	I believe so; however, (1) Section 3 feels a little general in this regard. The first case example does not have any potential CCF so it's not a great case to show the identification of CCFs (although I understand it shows the process). (2) Regarding the second case example, I am not sure why it wasn't broken down into steps like the first case example, which I think hurts the description of the process a bit.	(1). Yes, case study 1 was performed to illustrate the process, no software CCFs were involved. The target systems in case study 2 and 3 have CCFs identified. (2). A full RESHA analysis of the HSI QIAS-P system was previously conducted in the FY 2021 milestone report, *Quantitative Risk Analysis of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants using IRADIC Technology* (i.e., INL/EXT-21-64039) [3]. To cut down on unnecessary repetition, only the key results drawn from that analysis were summarized in this report. INL/EXT-21-64039 will be added as a reference for this case study.
RPI	In many safety-critical digital applications, even though their system design is very complex due to interconnections and test/maintenance capabilities, their system function is very straightforward. I have not had any issues at identification of hazard states of digital safety-critical systems. Figure 11 is for illustrating the application process and easy to understand. It should however be noted how general this diagram is. That is, it matters if this flowchart covers all possible cases. If this is just for the suggested techniques, the limitation and applicability must be clarified.	Figure 11 shows a general workflow for any situation, especially for safety-critical DI&C systems that have highly redundant and diverse designs at the system and component levels. For this specific situation, various CCFs at different levels with different coupling factors can be well identified using the proposed RESHA method. RESHA results will also be used for the following reliability and consequence analysis. Of course, when the system architecture is quite simple and straightforward, some steps of RESHA may be not necessary.

Reviewers	Comments	Responses/Resolutions
EPRI	Figure is a comprehensive description of RESHA.	No resolution is required/needed.

5.5 Technical Question #5

Technical Question #5 is "Based on the integration of STPA and HAZCADS, RESHA can identify software failure mechanisms in the control/actuation pathway using Unsafe Control Actions (UCAs). A novel concept called Unsafe Information Flow (UIF) has been developed and introduced in RESHA to complete the identification and tracing of software failure mechanisms in the information/feedback pathway (IFP), as discussed in Section 3.2. Please provide your feedback for the development and application of UIFs in identifying software failure mechanisms in the IFP." Technical comments and responses to Question #5 are summarized in Table 7.

Table 7. Summary of technical comments and responses to Question #5.

Reviewers	Comments	Responses/Resolutions
GEH	While I typically don't think of feedback being critical aspect of the most safety-significant DI&C, the need for the introduction of the UIF concept is clear and well-supported. Perhaps the authors could have expanded the UCA concept for IFP, but I think introducing UIF makes the framework less confusing.	No resolution is required/needed.
NRC	UIF seems intuitive way to ensure the analysis is complete.	No resolution is required/needed.
RPI	In other previous studies, it was considered as 'function failure of safety digital systems'. In this study, it was named UIF. I believe the conventional definition of function failure for safety-critical systems well addresses UIF.	The "function failures of safety digital system" are assumed to be where the information device violates defined functional requirements. Technically, this definition does describe UIF in a general sense, but it does not identify in what ways the system has failed. For instance, digital output/input modules can fail functionally, and these failures can be detected by digital defensive measures (i.e., concurrency checking) and mitigated. However, if the digital output/input modules provide errant information, these are not necessarily detected by defensive measures and can lead to additional unscreened failures. UIFs are proposed to describe how incorrect design, assumptions, requirements, and implementation of the information software leads to feasible losses by dependent systems (e.g., controllers, operators, etc.). Specified requirements are not necessarily violated thus a UIF may not necessarily be a defined functional failure.
EPRI	i. I think RESHA does NOT integrate HAZCADS/DRAM as developed by EPRI. The UIF idea has potential, but it is essentially a	i. Indeed RESHA does not use HAZCADS/DRAM, rather RESHA incorporates the concept of using UCAs within a FT,

Reviewers	Comments	Responses/Resolutions
	loss scenario development aid. I would recommend describing it in that context and testing its efficacy in that context.	an idea that came from an early HAZCADS publication. RESHA has since expanded beyond the control action pathway (UCAs) to consider aspects of the information feedback pathway via UIFs. The UIFs/UCAs along with a complete failure pathway within a FT can represent STPA loss scenarios due to external incorrect information and internal software defects. The relationship of UIFs, UCAs and loss scenarios are important, it is recognized that they can inform each other. This will be clarified in the new version of the report. According to the communication with our collaborators, the interests in and needs for UIFs have been recognized as a critical issue in both conventional control and cybersecurity where information plays a continuously growing role in system behavior. The concept of UIF is envisioned to be used in two ways: (1) ensuring that the necessary safeguards (or self-checking mechanisms) are in place to guard against errant assumptions between the interface of components, and (2) if safeguards are not in place, the potential impact it can have on dependent systems. One may argue that UIFs are described with conventional failure modes (i.e., functional failure of digital system), however, UIFs also cover failure instances where perfect and continuous functional verification of information transmission cannot be guaranteed.
	ii. The STPA process identifies Unsafe Control Actions based on the effect the DI&C system has on the equipment under control. The concept of UIFs breaks the connection to the equipment under control, so it is more difficult to identify when an Information Flow can be unsafe. The report does not provide any steps that would be required to determine if the Information Flow would be unsafe in a particular context.	(ii). It is a very good point. In this work, the connection to components is explicitly defined in two ways. First, the RESHA-developed integrated FT is a closer representation to the true DI&C control loop than conventional FT. RESHA FTs track component upstream dependencies by explicitly defining how FTs should be organized into three categories: failures due to hardware, failures due to internal software defects, and failures due to external dependencies, such as software defects. For a particular controller, failures of dependencies, such as information systems (which generate UIFs), can be directly traced by examining the chain of failure events along the dependency branch. Second, the semantic construction of UIFs requires a designation of the user of the information. Therefore, the linkage of UIF events to equipment under control can be traced from the top down (via RESHA) but also bottom up (via construction of UIFs).

Reviewers	Comments	Responses/Resolutions
	iii. The processes described do not elaborate on the various composition and decomposition layers of the total plant functional stack. That will be useful going froward to establish context in which the methods are used.	(iii). UIFs are still under investigation but based on preliminary discussion with industry experts, can exist in two instances, (1) a control loop where a human makes decisions, and (2) a control loop where a software controller makes decisions. In instance (1), spurious alarms are a form of UIF and determine their danger and, would require a HRA analysis of operator response to errant information within the MCR. In instance (2), self-checking mechanisms within the software controller (i.e., consistency checking) are examined to see if they can successfully rule out if the UIF is credible. The authors anticipate further discussing it with the EPRI colleagues and working to ensure this point is resolved.

5.6 Technical Question #6 (a)

Technical Question #6 (a) is "In Section 4.3, ORCAS is developed and demonstrated to estimate the probability of UCAs/UIFs for rich-data conditions. Please provide your feedback on: (1) are the assumptions of ORCAS reasonable? (2) are input requirements and steps of ORCAS clearly described, logical and practical for deployment with rich testing data available? (3) is ORCAS method well demonstrated in the case studies with all expected outcomes obtained?" Technical comments and responses to Question #6 (a) are summarized in Table 8.

Table 8. Summary of technical comments and responses to Question #6 (a).

Reviewers	Comments	Responses/Resolutions
GEH	The method seems reasonable, but the reviewer doesn't have sufficient experience in software development to make a judgment as to whether the data requirements can be met in practice. It would seem as if the SDLC would need to be designed with ORCAS in mind. Further, it seems as if non-reporting of defect detection could make the resultant analysis non-conservative. If it would be "easy" to fail to report defect detection during the SDLC and under-reporting of defect detection could lead to underestimated defect escape probabilities, then I would be concerned with adoption of this method.	It is a very good point. While ORCAS does not need to be explicitly considered during those phases, a process to log defects will greatly reduce uncertainty in the ORCAS results. In fact, to ensure the reliability of the software deployment process in a safety-critical DI&C systems, the suggestion is software developers begin defect logging during the early development phases of the SDLC. It is recognized that non-compliance of defect logging is an issue, however, failure to log also leads to a large uncertainty in the final reliability of the software. This, however, is captured in the qualitative analysis of the completeness of the SDLC verification processes. In a separate industry study with ORCAS, the objective was to develop reliability metrics to justify failure probabilities below IEC 65108 safety integrity level (SIL) 4. However, it was found that the software development team did not log early phase

Reviewers	Comments	Responses/Resolutions
		defects due to technology infancy, only final product testing defects and resolutions were logged. Ultimately, due to the (1) lack of early phase data, and (2) an incomplete testing process (from qualitative analysis) it was determined that the reliability of the software was between SIL 3 and SIL 4 failure probability (due to uncertainty). A lower reliability could not be justified with ORCAS, and the target software was assigned a conservative failure probability based on the available evidence. Lastly, regardless of the reliability analysis approach used, unless sufficient evidence is available, the determined reliability value must be justified with sufficient evidence. Therefore, it is suggested the SDLC development evidence is collected as soon as possible to provide a realistic understanding of the true reliability, rather than a conservative estimation using SIL levels.
NRC	(1) As someone who is not a systems or software engineer, it was difficult to review this section. However, the assumptions appeared to be reasonable. (2) The ORCAS steps appear to be clearly described. (3) I think it would be beneficial to break down the test case using the ORCAS steps.	No resolution is required/needed for (1) and (2). It is a very good point. For (3), the authors have just completed an industry study where the exact steps and processes of ORCAS were broken down for a safety actuation system. It is anticipated that some details and lessons learned from this study will be made public in the upcoming months.
RPI	(1) It seems difficult to make a simple answer since the assumptions might be reasonable in some cases but not in other cases. If we consider the case suggested in the report as a pilot study, it seems to make sense. For (2) and (3), please refer my comments emailed earlier.	The authors agree with reviewer's comment (1). Assumptions should be reviewed and confirmed for different applications. For (2) and (3), reviewer's comments in email and respective responses can be found in Section 5.14.
EPRI	(1). i. The ORCA Assumption that future software reliability can be statistically quantified for representative test data results is incorrect in that this process cannot predict the future statistically as the organized complexity of the system is not being held constant.	(1). i. One of the fundamental assumptions of quantitative software reliability is the probability of defects resulting in a failure is driven by triggering events. While true, given the same errant input, the same software failure can be activated 100% of the time, the probability that the errant input will exist is a function of the test space covered vs the operational space

Reviewers	Comments	Responses/Resolutions
	ii. Recommend that the failure probability of a UCA/UIF based on test data be replaced with a non-probabilistic likelihood metric that tracks with the effectiveness of software reliability methods. iii. The current work misleads in drawing a conclusion that the methods would work using observations from commercial software. This reveals a dependency on specific development methods and metrics which would not be the same for safety critical safety software. iv. ORCAS is based on finding software coding errors such as typos or incorrect implementation of requirements. It will not be able to find systematic errors that are based on an incomplete set of software requirements. v. The results of ORCAS will need to integrate into a loss scenario to establish traceability to a UCA.	anticipated. As the triggering event probability is non-deterministic, this results also in a non-deterministic failure probability of software. Similar statements can be found in [6] [7]. It is recognized that SRGMs are useful methods to anticipate future reliability, however, are severely limited when failure data is unavailable. This problem is especially prevalent for safety-critical systems, hence the need for qualitative assessment of the programming practices and test coverage space to ensure comprehensive consideration of defect triggers. One future work is to improve the method to reflect this reality where software failures are driven by triggering event. ii. Within ORCAS, we qualitatively assess the coverage of the test data, but it is not currently used in the quantification of the reliability. Improving ORCAS to better incorporate the data and provide various metrics for software reliability estimation is one future work. iii. The ORCAS method is demonstrated in this report with very limited available software testing data from commercial software, which may not be the same as the data of safety-critical software. Commercial open-source software does not have the same professional development process and policy compliance as safety critical software. More experimental data is required to calibrate and validate the UCA/UIF correlations developed. VVUQ of these developed methods will be part of the research work in FY 2024. iv. Directly identifying errors during the development of the requirements development process is not the goal of ORCAS, and ORCAS is not designed to assess the human-related aspects of the SDLC. Completeness of the requirement specifications is primarily a human activity as developers must decide which assumptions and designs are applicable. In this respect, ORCAS can find defects in the code that are missing from the requirements specifications; however, it cannot find defects that are not specified in the requirements specification. As ORCAS

Reviewers	Comments	Responses/Resolutions
		is intended to assess the reportable and measurable verification and validation activities (i.e., testing) during the entire SDLC, missing requirements as systematic errors are not reportable or measurable. Consistency and completeness of the human development processes, such as those conducted in requirements specifications and design detailing is measured via the BAHAMAS process, which is suggested to be applied when testing data is quite limited and unavailable. v. The author assumes that the loss scenario is referencing the STPA definition; a set of causal factors that can lead to a loss. In this work, the causal factors are analogous to software failure mechanisms. Part of the ORCAS analysis is to look at the defect triggers that can activate root cause defects. In the UCA, if a process module causal factor is postulated, the ORCAS analysis can examine the possible defect trigger classes that can lead to a failure of the process module. What this achieves is (1) if no defects are detected within the trigger classes of the process module, then qualitative evidence can be established for reliability, or (2) if a detect is detected, then the defect can be resolved and provide reliability quantification metrics for completeness of testing. Either way, ORCAS establishes traceability by assessing the possible trigger classes (as failure mechanisms) that can activate a defect for a particular loss scenario and providing evidence for the reliability of the module through the completeness of the testing procedure.
	(2). ORCAS is not deployment ready. It will need additional proof testing by the target user demographic and the integration of that feedback/challenge to be deployment ready. Real world results will need to be evaluated or tested.	(2). More refinement and demonstration are needed for ORCAS to be ready for deployment. Right now, the authors are collaborating with the industry for method demonstration and improvement.
	(3). It is well demonstrated but with the caveats mentioned above.	No resolution is required/needed for (3).

5.7 Technical Question #6 (b)

Technical Question #6 (b) is "b. In Section 4.4, BAHAMAS is developed and demonstrated to estimate the probability of UCAs/UIFs for limited-data conditions. Please provide your feedback on: (1) are the assumptions of BAHAMAS reasonable? (2) are input requirements and steps of BAHAMAS clearly described, logical and practical for deployment with no testing data and very limited design information? (3) is BAHAMAS

method well demonstrated in the case studies with all expected outcomes obtained?" Technical comments and responses to Question #6 (b) are summarized in Table 9.

Table 9. Summary of technical comments and responses to Question #6 (b).

Reviewers	Comments	Responses/Resolutions
GEH	The method seems reasonable and well-demonstrated. The reviewer would like for some more discussion, potentially, on software CCF since perhaps it could be argued that escaping software defects could affect multiple aspects of the system and the coupling could be complete (i.e., $\beta = 1$). Perhaps this is addressed by the $\beta SW=0.429$ in Table 53? It was a surprising result to discover that software CCF was not a dominant contributor in the RTS example. Perhaps the function cannot be defeated by software CCFs in the same CCCG?	It is exactly right that defects, which have escaped removal during software development, could affect multiple aspects of the system. The authors focus primarily on the critical functions of interest and the redundancy employed to ensure those critical functions. For example, the BPs of the RTS case study are responsible for sending a trip signal and the analysis is focused on the failure mode associated with the BPs not sending that trip signal when needed. As a final point, the selection of coupling mechanisms and modeling parameters (e.g., beta) is still being developed and improved for the CCF model. More details of the development and improvements made to the CCF model have been published in a secondary report (INL/RPT-22-70056: *Risk Analysis of Various Design Architectures for High Safety-Significant Safety-Related Digital Instrumentation and Control Systems of Nuclear Power Plants During Accident Scenarios* [5]). Within this report, the authors provide details of the scoring and justification for why beta is not =1. The RTS has multiple diverse means of ensuring a trip signal can be sent when needed. Thus, the failure of automatic function does not prevent actuation of reactor trip because the diverse protection system and operators provide an alternate means of ensuring reactor trip.
NRC	(1) The assumptions appear to be reasonable. (2) Is there a reason that the steps were not documented in a similar manner to ORCAS section. Although, in this section the case study does have the steps. (3) I am not sure what the expected outcomes were, but it is demonstrated to fulfil its objective.	No resolution is required/needed for (1). For (2), the reason the sections do not follow the same presentation format was largely because the BAHAMAS procedure was published in a previous report published in FY 2021 (INL/EXT-21-64039: *Quantitative Risk Analysis of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants using IRADIC Technology* [3]) then BAHAMAS was further developed and modified this fiscal year. Therefore, only the modified parts were emphasized in this report. This is a good observation, and

Reviewers	Comments	Responses/Resolutions
		in future publications, the authors will make greater effort to follow a consistent format or presentation style. For (3), the expected outcomes are the probabilities of UCAs and UIFs, which will be clarified in the new version of this report.
RPI	Please refer to my comments emailed earlier. Test data for the target SW changes the game. The method suggested in this report is the case that we do not have clear evidence of SW reliability. If we have, we don't need this difficult approach.	Reviewer's comments in email and respective responses can be found in Section 5.14. For the last statement concerning data, the reviewer's comments are correct. With sufficient testing data, BAHAMAS is largely not needed. This follows the proposed framework the point is that when data is available, users will transfer from BAHAMAS to ORCAS.
EPRI	(1). BAHAMAS needs some refinement in the general composition and decomposition particularly in how STPA is used. (2). They are well presented. (3). It was difficult to tell. The case studies were not validations.	For (1) and (3) STPA is not used in BAHAMAS. Refinement and VVUQ of BAHAMAS will be one future work in FY 2024. No resolution is required/needed for (2).

5.8 Technical Question #7

Technical Question #7 is "In Section 4.5, a modified beta factor method was developed and modified to support CCF modeling and parameter estimation as part of this framework. Please provide your feedback on: (1) are the method assumptions reasonable? (2) are the input requirements and steps clearly described, logical and practical for deployment to the industry? (3) is the method well demonstrated in the case studies with all expected outcomes obtained?" Technical comments and responses to Question #7 are summarized in Table 10.

Table 10. Summary of technical comments and responses to Question #7.

Reviewers	Comments	Responses/Resolutions
GEH	The method seems mostly reasonable and clear to follow. See the relevant portions of the answer to 6.b for comments relative to CCF.	Responses are provided in Section 5.7.
NRC	(1) Yes, they appear reasonable. (2) I believe that the reader needs to have a firm background on the reference materials to fully understand the steps.	No resolution is required/needed for (1) and (2). (3). Agreed, the uncertainty of the approach has been mentioned in the report and will be part of the future work in FY-2024.

Reviewers	Comments	Responses/Resolutions
	(3) Yes, but I think it needs to be acknowledged that results will have some significant uncertainties that need to be evaluated. Also, the discussion was focused on software CCF, but the case examples integrate hardware CCF, which was unexpected.	Although of the focus of this work is software CCF identification and estimation, hardware CCFs are also considered to complete the case studies on the integrated risk analysis of safety-critical DI&C systems.
RPI	It is a bit unclear what are the assumptions used for this method. The report discussed the difficulties of asymmetric CCF analysis, but it did not share clear cases or theoretical examination regarding if the modified beta factor method makes reasonable and acceptable approximations. For additional comments, please refer my earlier email.	The modified beta factor (MBF) method assumes a CCCG (n) will fail with $\beta_n Q_T$ (as discussed in Section 4.5.2 of the reviewed report). Inherent to this model is the assumption that all components of the CCCG will fail. This assumption, based on the traditional beta factor model, is disputed by IEC 61508; where it is argued that for large groups of components, it is unrealistic to model all failing at one time. However, as a basis for the modeling choice, it is quite simple for redundant software elements to fail together so long as they receive identical operational inputs that activate a common defect. This mechanism for software failure forms a theoretical basis for keeping and working with the MBF model.
EPRI	This section seems to be an implementation of the Beta factor model and those assumptions hold but errors in using the baseless statistical probabilities generated earlier. There are also several generalized statements with no basis, for example "Our work assumes that the potential for combinations of failures with the CCCG is largely dependent on the existence of subtle differences in the coupling mechanisms". Clarification of these statements and assumptions will clarify the methods.	The comments to software reliability estimation have been responded to in previous sections. The following sentences are given in Section 4.5.1 of the reviewed report, which is sufficient to clarify the indicated statement: "*A software CCF will occur when a coupling mechanism creates a scenario for operational conditions to activate a common software defect. Given a group of redundant software components, variations in their operating conditions may lead to some, but not all, components failing together. Variation of maintenance activities, input variable sources, component locations, and installation teams influence the operational environment; ultimately, subtle differences in coupling mechanisms may influence which components fail together. Capturing asymmetry between components may be necessary for software CCF modeling, but it can be challenging with conventional methods.*" To further clarify the differences in parameter estimation and identification of coupling factors of hardware and software CCCGs: In the determination of hardware CCCGs via the beta factor method, qualitative coupling mechanisms that are known to contribute to higher beta factors between redundant trains

Reviewers	Comments	Responses/Resolutions
		were identified. However, these coupling mechanisms were developed for analog systems. Software systems fundamentally exhibit different coupling mechanisms that are not translatable. For instance, n-train redundancy implementation can protect against single stochastic failure. However, n-train software programming has not been shown to offer the same degree of protection. Therefore, in this work, the authors seek to specify and identify the subtle differences that may exist in software coupling mechanisms to obtain a more reasonable estimate on CCF potential.

5.9 Technical Question #8

Technical Question #8 is "Section 5 documents the consequence analysis of a generic PWR SAPHIRE model with an improved digital reactor trip system (RTS) and Engineered Safety Features Actuation System (ESFAS) fault trees. Please provide your feedback on: (1) Are the demonstrated consequence analyses on HSSSR DI&C systems sufficiently detailed to support industry needs in DI&C modeling and risk analysis? (2) Is it beneficial to perform similar consequence analyses for other DI&C systems that are safety-related but not safety-significant?" Technical comments and responses to Question #8 are summarized in Table 11.

Table 11. Summary of technical comments and responses to Question #8.

Reviewers	Comments	Responses/Resolutions
GEH	1) Yes, the examples were sufficiently detailed for this document. It would be helpful to have an example with a DI&C platform that could affect multiple functions, but it is understood that that may be just an extension from the foundational work being presented here. 2) I do not foresee that being materially helpful since it seems as if it would just be redundant or simpler to the examples provided.	For (1), yes, relevant methodology development and demonstration of a function-based risk assessment of DI&C system will be conducted in FY 2024, target systems will still be focused on HSSSR DI&C systems. No resolution is required/needed for (2).
NRC	(1) I believe you need to show the sub fault trees in the RTS and HSI fault trees. And the improved FT for ESFAS is not shown at all. This results in the reader to not see the basic events in the minimal cut sets at all for some of the examples. So while the examples are good, they don't provide enough detail to the reader. I understand that showing these details can be "clunky" in the report. They could possibly be put in the appendix.	For (1), yes, the sub-FTs are not shown in this report because there are many pages of them, it will make this report too large to share in a convenient way. The FTs will be released as an appendix of the revised report and available to the public.

Reviewers	Comments	Responses/Resolutions
	(2) Yes, I would think that these analyses would be what is needed to show that SR systems are or are not safety significant. I understand the use term "consequence" analysis probably has some considerable inertia; however, this term has a specific meaning to PRA practitioners. These analyses are just Level 1 risk analyses of the inclusion DI&C systems. I didn't see any discussion on how these systems could increase or decrease the likelihood of initiating events that would increase or decrease the CDF from DI&C systems.	(2). It is a very valuable comment. Yes, the authors understand that the term "consequence" is usually used in Level 3 PRA to describe the consequences to workers/public given the release of radioactive materials out of containment. In this project, consequence analysis focuses on evaluating how CDF can be affected by the identified digital-based failures, basically Level 1 risk analysis. The authors will investigate how to perform an extended-scope PRA for the DI&C risk assessment in this project. The authors also agree that, for the scope of a Level 1 PRA scope, the impacts of DI&C systems can be incorporated through both initiating events (IEs) and non-IE top events (i.e., mitigation systems) in the ETs. The current work scope is confined to incorporating DI&C impacts through mitigation systems, and the authors will start investigating how to evaluate the impacts on IEs.
RPI	It can be much more detailed and precise. The international benchmark project of safety-critical digital system risk assessment (by OECD) was recently published. Please refer to it to see how many details could be captured. On the other hand, I think the focus of this report is not on the system-level risk assessment, so I believe the current version is fine. If higher resolution FT figures and clearer explanation are provided, it would be great. If this kind of FT-based 'function failure' analysis must be done for non- safety-significant systems, the target function must be multiple, which makes analysis way more complex with limited gain. In the current failure-oriented PRA models, more detailed non-safety system failure models would not help to enhance the analysis accuracy.	Thanks for recommending this report. The OECD report will be referred to in the revised version of INL/RPT-22-68656. Higher resolution FT figures will be provided in the revised version of INL/RPT-22-68656.
EPRI	(1). Yes, this generally illustrates how improving functional reliability can reduce the likelihood of an accident and improve safety and outcomes. However, since it uses the baseless quantification (no actual reliability data or qualitative bounding assumptions), it is misleading when combined in a traditional PRA with statistically valid reliability for associated components. A method to allow systematic reliability insights to inform the random hardware reliability metrics used is needed.	(1). Comments on software reliability estimation have been responded to in previous sections. For (2) and the second comment of (1), using systematic reliability insights to develop metrics for hardware reliability analysis is not included in current project scope, relevant benefits and technical soundness will be investigated in the future. The authors look forward to collaborating with the EPRI colleagues on this topic.

Reviewers	Comments	Responses/Resolutions
	(2). This exercise does illustrate the challenge of demonstrating the improved reliability of digital systems within traditional PRA without forcing an arbitrary reliability value for systematic errors including software. This will require further R&D to be used for design or operational risk insights.	

5.10 Technical Question #9

Technical Question #9 is "In addition to providing the changes of core damage frequency (CDF) and large early release frequency (LERF) due to the digital upgrades for HSSSR I&C systems, what other risk-informed insights from the quantitative consequence analysis would be beneficial for the evaluation and reduction of the plant-level risks?" Technical comments and responses to Question #1 are summarized in Table 12.

Table 12. Summary of technical comments and responses to Question #9.

Reviewers	Comments	Responses/Resolutions
GEH	I think that importance measures and sequence contributions are also important. They may not solely be able to evaluate the risk benefit from DI&C implementation, but those insights might only be attained using the frameworks suggested by the authors and may have material impacts on risk communication, plant operation, and design activities.	Agreed. Importance measure and sensitivity analyses have been conducted in another report published in November 2023: INL/RPT-22-70056: *Risk Analysis of Various Design Architectures for High Safety-Significant Safety-Related Digital Instrumentation and Control Systems of Nuclear Power Plants During Accident Scenarios* [5]. As part of the sensitivity analysis, risk-significant sequences were identified, and their frequencies before and after the DI&C implementations were calculated and compared. The authors agree it is a good point to generate more insights on risk communication, plant operation, and design activities based on the results of risk analysis, relevant importance measure and sensitivity analysis. Current work relevant to these topics will be included in a report this coming August.
NRC	I think the evaluations need to consider the uncertainties. For example, the digital RTS shows a decrease overall CDF; however, the uncertainties associated with the digital system are much higher. Also, LERF isn't even mentioned in the report.	Agreed. The uncertainty from the methodology and input data should be analyzed to obtain a better estimation of the key outcomes, like CDF. Uncertainty quantification is part of the work in FY 2023 and relevant work will be included in a report this coming August. And yes, in current case studies, only CDF is used as the key metric for plant safety. LERF will be considered as another one in future case studies, depending on the selected accident scenarios and availability of PRA Level 2 models.

Reviewers	Comments	Responses/Resolutions
RPI	If it is about the EOOs (error of omission), the current metrics (CDF and LERF) would be good enough since they actually serve as the initial events of next stage of analysis. For EOC like in SW failures or for general control problems rather than safety decision making problems, there should be more insightful metrics than CDF and LERF. This issue is also correlated with how to define the goals of the software. Through this process, the conventional deterministic requirements are linked with risk evaluation.	It is a very good point. Errors of commission (EOCs) represent doing something that should not have been done, while errors of omission (EOOs) represent not doing something that should have been done. Currently, the safety-based metrics (like failure probability and CDF) are used for safety decision-making problems. To better support the analysis of systems whose metrics are not clearly tied to common safety metrics (i.e., LERF or CDF), the authors will investigate metrics to better capture or provide insight to EOC issues relevant to the performance, resilience, and robustness of software systems, and support the design optimization of both non-safety and safety-critical DI&C systems.
EPRI	Let me answer this question in the context of the earlier comments. While I consider a direct statistical quantification of software reliability technically unsound and unjustifiable, if a reliable systematic insight can be added to the hardware reliability metric, then this process would be quite powerful and could be used for any Systematic error, software or otherwise.	The comments to software reliability estimation have been responded in previous sections, mainly in Section 5.6. The authors agree that using systematic reliability insights to develop metrics for hardware reliability analysis could be beneficial, relevant benefits and feasibility will be investigated in the future. The authors look forward to collaborating with EPRI colleagues on this topic.

5.11 Technical Question #10

Technical Question #10 is "Some technical gaps in applying the framework for the risk analysis of AI/ML-aided control systems have been reviewed and investigated in Section 6. With respect to the technical gaps considered, are they complete and significant in terms of identifying, quantifying, and evaluating potential failure modes of AI-aided control systems? Are there any other approaches to addressing these types of technical gaps that you recommend us to consider?" Technical comments and responses to Question #1 are summarized in Table 13.

Table 13. Summary of technical comments and responses to Question #10.

Reviewers	Comments	Responses/Resolutions
GEH	It is the current opinion of the reviewer that an AI/ML-aided system would suffer from a lack of quality training data, both in terms of the kinds of parameters available and also in exploring the relevant parameter space where safety-critical I&C systems are needed. The discussion provided by the authors was good, but I'm not sure it addresses this problem. One wonders if ML techniques that trained on actual plant data, simulator data, and/or thermal-hydraulic code training sets simultaneously could mitigate such issues.	It is a very good point. Lack of qualified training data is one of the major challenges during the development lifecycle of AI/ML models. Meanwhile, it can be one of the major root causes leading to the failure of AI/ML-aided control systems. One ongoing work is to define and quantify metrics to measure the impacts of the lack of qualified training data to the performance/reliability of AI/ML-aided control systems. Results

Reviewers	Comments	Responses/Resolutions
		will be included in the milestone report to be published in August 2023.
NRC	I don't have the background regarding AI-aided control systems to answer whether the reported technical gaps are complete.	No resolution is required/needed.
RPI	We need to discuss first if there are any differences in the failure effects of digital or conventional control systems. Not only the deterioration of control performance, if it causes any safety actions activated, whole plant operation would be limited by poorly designed control systems. It should be noted if there is any unique difference due to AI-aided control system was used once it was well-verified and validated by fully implemented tested cases.	Yes, whether new or unanalyzed failure modes exist or how to identify them and evaluate their impacts to system reliability and plant safety is the motivation of this review activity. This is part of the ongoing research to assess if the failure modes of conventional software systems will be different than those of AI-aided systems. Importantly, the contextual conditions that the AI-aided system operates in will define in what ways the system will fail. Relevant work has been initiated and will be included in the coming August report.
EPRI	(1). This is a good narrative and the issues presented are indeed of concern. I would not agree that the narrative is complete, or the significance is understood, since the AI/ML environments change daily. The ChatGPT revolution has occurred just since this product was finalized.	(1). This narrative can be considered as a preliminary review and study on the potential technical gaps in the risk assessment of AI-aided control system. As the reviewer mentioned, the AI/ML techniques and applications are evolving every day; it is not easy to cover all potential issues. However, this effort is a good start for following relevant research activities.
	(2). It is too early for AI/ML to invade critical/safety space just yet, and I would not waste any time with it in this context. The problems will change many times between now and when AI/ML is credible for serious stuff. In the meantime, we should let it cut its teeth on more mundane things.	(2). Research on applying AI/ML to the operation and control of NPPs has been started for a long time in many DOE-led industry-involved programs and projects. ML-aided digital twin technology is a popular topic in nuclear engineering. Although current AI/ML techniques are not mature enough for real deployment, it is necessary to push relevant supporting work such as a reliability/risk assessment and trustworthiness study of AI/ML-aided reactor operation and control. The NRC also published several reports on this topic, one of which is *NUREG-2261: Artificial Intelligence Strategic Plan* (June 2022). Anticipating the industry's potential application of AI to NRC-regulated activities, the NRC has developed a strategic plan to ensure the agency's readiness to review such uses. In this context, it is meaningful and necessary to push forward relevant research work. In this project, it is about how to extend and adjust the proposed framework and relevant methods to support the risk assessment and design optimization of AI-aided control systems.

5.12 Technical Question #11

Technical Question #11 is "Section 7 summarizes recent work and proposes future R&D. Do you agree with the identified needs for these activities? Are there other relevant short- and long-term industry needs in the area of DI&C risk assessment not addressed by this framework?" Technical comments and responses to Question #11 are summarized in Table 14.

Table 14. Summary of technical comments and responses to Question #11.

Reviewers	Comments	Responses/Resolutions
GEH	Yes, this section identifies reasonable needs. Not new suggestions are provided here by the reviewer at this time.	No resolution is required/needed.
NRC	Yes, I agree with these activities. The risk evaluations need to be expanded to account for potential IEs that digital systems may introduce and to account for the uncertainties associated with the methods described in the report.	The impacts of DI&C failures incorporated through IEs to plant safety will be considered in FY 2024. Uncertainty, importance, and sensitivity analyses have been performed in FY 2023 and will be continued in future to better support DI&C risk management and design optimization.
RPI	Yes, without knowing them, we would keep checking if the newly developed SW satisfies the requirements. Quantification of its reliability is almost the only way to address this issue. I think this report will provide a very meaningful milestone.	No resolution is required/needed.
EPRI	(1). I don't agree with all the needs, but we do need a clear actionable process to evaluate digital I&C risk. (2). It's not clear that the LWRS Framework successfully addresses short- or long-term needs of the industry.	For (1) and (2), the authors look forward to further dialogue with EPRI and other stakeholders on this point to ensure that recent expectations for short- and long-term industry needs in DI&C risk assessment are clearly defined for the path forward. Currently, this project focuses on some of the essential elements of the relevant needs of industry: (1). Identifying and evaluating software CCFs in the safety-critical DI&C systems of NPPs. (2). Optimizing diversity and redundancy applications for the safety-critical DI&C systems. Adding diversity within a system or components is the primary means to eliminate and mitigate CCFs, but diversity also increases system complexity and may not address all sources of systematic failures. for instance, is it necessary to eliminate all identified potential CCFs, or just the ones that can significantly affect system reliability and plant safety based on a prioritized assessment? How to determine which CCF is more significant, especially when considering there are different types and scales of CCFs in the safety-critical highly redundant DI&C systems?

Reviewers	Comments	Responses/Resolutions
		Therefore, this LWRS-developed framework was proposed to address these digital issues quantitatively by focusing on software CCFs in safety-critical DI&C systems. It is expected to be used as a risk-informed tool that offers the capability of design architecture evaluation of various DI&C systems to support system design decisions in diversity and defense-in-depth applications. For instance, one collaborative study with the nuclear industry has demonstrated that, while the system-level CCF was intolerable and the major contributor for risk increase within the model, some CCFs at smaller levels are tolerable as the increase in risk is manageable. It is recognized that the methodology development for such a framework has many technical challenges. More communication and collaboration with peers are expected.
	(3). The LWRS risk framework has great potential to contribute to the overall analysis of Digital I&C Risk and Reliability. My primary concern is the Framework's single-minded goal of statistical quantification contrary to any input or feedback suggesting a different path. I do not think the positions of the framework can be defended if challenged to any technical depth. I recommend you reflow this otherwise good work to address compatibility with known technical positions in other safety industries. Be less concerned about what the regulatory guidance is and the need for a "number" but rather concentrate on a defendable method that is understandable, defendable, usable.	Regarding the reviewer's primary concern: the authors appreciate the reviewer's commitment to this point. In previous comments the reviewers indicated a potential to evolve the proposed reliability methods (e.g., BAHAMAS and ORCAS) to provide alternative metrics to the traditional probability estimation. Such changes may, as indicated earlier, enhance the capabilities for the LWRS-developed framework to address systematic issues, better align with international standards, and potentially support EPRI's framework. In response, the authors anticipate defining various quantifiable metrics to evaluate the interactions among software defects, failure mechanisms, triggering events, and failure modes, for which BAHAMAS and ORCAS are well suited. The authors look forward to continued work and dialogue in this area that may lead to new and useful capabilities to benefit risk management and design optimization of DI&C systems.

5.13 Technical Question #12

Technical Question #12 is "Please provide any additional suggestions for the framework improvements." Technical comments and responses to Question #12 are summarized in Table 15.

Table 15. Summary of technical comments and responses to Question #12.

Reviewers	Comments	Responses/Resolutions
GEH	No additional suggestions are identified by the reviewer that have not been discussed in earlier comments.	No resolution is required/needed.
NRC	Section 3.1 seems to limit UCAs definition to software failures alone. However, later in Section 3 both hardware and software issues are covered. Should this be changed?	In the STPA manual, UCAs can be used for both hardware and software. And in the framework, UCAs represent software failures, not hardware failures. Hardware failures, including hardware CCFs, are not defined or covered as UCAs or UIFs.
	The first sentence in Section 3.3, Step 5 should be rewritten.	Change will be made accordingly in the new version of the report.
	It would be beneficial in my opinion to continue the example through the steps in Section 3.3.	Changes will be made accordingly in the new version of the report.
	It might be beneficial to link the hardware pieces where they fit in Figure 16.	The authors appreciate the comment. The complete details of the hardware of the system can be found in reference [31] of the reviewed report. It will be clarified in the revised version.
	Section 3.4.1, 3rd sentence describes Figure 16 from left to right, but is actually doing it right to left. The same thing occurs for Figures 35 and 36.	Changes will be made accordingly in the new version of the report.
	Figure 17 has a typo in the second box from the top.	Changes will be made accordingly in the new version of the report.
	It may be beneficial for the Section 3.4.1, Step 7 to show the revised FT and/or explain the reduction in the minimal cut sets.	Changes will be made accordingly in the new version of the report.
	On page 31, the text discusses the top event "reactor fails to trip when needed causing core damage," but the subsequent text changes and then a different top event and fault tree are provided, which was confusing.	Changes will be made accordingly in the new version of the report.
	It may be hard for the reader to determine what the minimal cut sets are for the example in 3.4.2 without the full fault tree.	The full FT will be released as an appendix.
	Should the word "semantically" on page 42 be "systematically"?	Changes will be made accordingly in the new version of the report.
	I suggest changing the title of Figure 44...it doesn't represent a traditional event tree.	Title will be changed to: *Example of an HRA Event Tree with Recovery as Discussed in THERP*. This figure is based on HRA event trees as they are used within THERP.
	On page 73, it states that the BAHAMAS can provided similar values to HRA (THERP). Based on Figure 45 this is true at the higher dependence but can differ by over two orders of magnitude at lower dependence.	There is uncertainty in the human errors and correction of human errors by reviewers. THERP is a starting place. The authors acknowledge the difference shown in Figure 45 and will provide additional discussion as part of the revision to this

Reviewers	Comments	Responses/Resolutions
		report. As it stands, given the uncertainty associated with the model, the authors find it more acceptable that the BAHAMAS curves in Figure 45 are conservative compared to THERP's.
	Did the authors review the evaluation of THERP against the NRC HRA good practices (NUREG-1842). There are some significant limitations.	An in-depth review of these limitations has not been addressed, THERP is a starting point for the developed method, and there is, as the reviewer aptly pointed out, a need to address the limiting points of HRA. One that is identified in NUREG-1842 is THERP does not provide significant consideration of the cognitive complexity for scoring human error. These are areas that can be investigated as part of future work.
	In the second to last sentence in Section 4.5.1...I think this sentence should be modified. While it may be the best, it could not be. Perhaps it should focus that the modified MBF was selected due to have the fewest parameters.	This sentence will be removed. The MBF method was selected because it worked well with the assumption that redundant software elements will fail together. Specifically, it is unlikely that a subgroup of software elements from a CCCG will fail when the full group has identical coupling mechanisms. A model based on the beta-factor method fits well with this assumption.
	Page 82 often stated that a CCF requires a shared root cause. This is not necessarily true; it depends on the causal level chosen by the modelers. Note that the NPP CCF data is based on the proximate cause and not the root cause.	The in-depth discussion was not provided on proximate and root causes, though the authors are familiar with the concept that a proximate cause is considered the most readily identifiable cause of failure, but itself will have a specific root cause. The authors will use "shared cause" to keep the sentence in a general format. All other uses of root cause are non-restrictive and qualify as true statements.
	Second sentence in Section 5.2 states the IEs are limited to critical operation. That is true for this model but not in general.	Changes will be made from "while a plant is in critical operation" to "when a reactor is critical and at or above the point of adding heat" in the new version of the report.
RPI	Please refer my comments emailed earlier.	No resolution is required/needed.
EPRI	No comments provided.	No resolution is required/needed.

5.14 Additional Comments from Technical Peer Reviewers

This section records the additional comments from GEH and RPI in emails and relevant responses from the report authors.

5.14.1 Additional Comments from GEH

Additional comments from GEH and responses from the report authors are summarized in Table 16.

Table 16. Summary of additional comments from GEH and responses.

#	Comment	Responses/Resolutions
1.	In the executive summary, it mentions that the method is developed for HSSSR systems. Any reason to limit the approach to those that are safety-related? The methods seem generalizable enough to apply to non-SR control systems and future plants may have diverse I&C platforms from the SR platform where this approach may be of interest in application.	This project was initiated in 2019 when software CCF was considered a technical issue for HSSSR DI&C systems, so this framework and relevant methods were developed to fully address this issue by identifying potential software CCFs, quantifying their probabilities, and evaluating their impacts to system reliability and plant safety. Typically, redundancy and diversity are found in safety-related systems. But, of course, this framework and relevant methods can be used for all types of control systems.
2.	One additional technical goal you may want to mention in the executive summary is related to risk-informed design. For example, introducing a best-estimate approach to DI&C reliability reduces potential masking by conservative approaches and facilitates risk-informed design. This can lead to real safety improvements to plants that would otherwise be unable to be "notice" the benefit of making changes to functions that are masked by conservative treatment of the control systems. I think item (4) of paragraph 2 comes close to discussing this, but there is real benefit for design and improvement of systems that are not the I&C systems as well.	Yes, Section 2.4 of the reviewed report describes the value proposition of the proposed framework and relevant methods, especially Point 1: "[a] best-estimate, risk-informed capability to address digital issues quantitatively, focusing on software CCFs in HSSSR DI&C systems of NPPs" and Point 4 "[a] risk-informed tool that offers a capability of design architecture evaluation of various DI&C systems to support system design decisions in diversity and redundancy applications." It is a good idea to clarify the Executive Summary; changes will be made in the new version of this report.
3.	I think it would be important to stress to the audience that best-estimate quantitative reliability calculations for DI&C systems are absolutely critical to any risk-informed applications that a plant may pursue (e.g., risk-informed design, risk-informed license applications). A traditional example would be trying to calculate risk importance measures. With conservative approaches to DI&C reliability, importance measures for many components may be superficially small because the dominant failure modes are DI&C failures, leading the plant to make changes that are not in the real best interest for public safety. Alternatively, in plant design stages, conservative approaches may mask potential design improvements that would enhance public safety but are shielded from the view of the designer because conservative DI&C failures dominate the risk profile.	Yes, importance, sensitivity, and uncertainty analyses are a necessary portion of this framework to support risk-informed decision making for design optimization. Work, relevant to these topics, has been initiated in FY 2023, part of it was published in INL/RPT-22-70056: Risk Analysis of Various Design Architectures for High Safety-Significant Safety-Related Digital Instrumentation and Control Systems of Nuclear Power Plants During Accident Scenarios [5]. Additional details and results will be included in a report this coming August 2023.
4.	i. Does RESHA (or the authors) take a stance on whether categories of UCAs are disjoint?	Very insightful questions. UCAs represent failure modes and are not of themselves disjoint. A controller may perform multiple unique and separate control actions, each of which may fail at the same time with different failure modes. However, for a specific control action, failure

#	Comment	Responses/Resolutions
	ii. For example, in an integrated DI&C platform, do we think it's possible for software to fail to send control actions to some components while spuriously sending unsafe control actions to other components? iii. Does it make sense to discuss this when introducing RESHA UCA categories?	modes are disjoint. As an example, if a controller has a specific action, that action may either fail to occur when needed (UCA-A) or it may occur prior to when it is needed (UCA-B). For this single action it is impossible for both failure modes to occur at the same time. Therefore, in this work, both UCA and UIF categories are not considered disjoint, but the failure modes for a single control action or function are disjoint. Depending on its design, a controller may have multiple functions that may fail at the same time but with different modes (i.e., UCA/UIFs). Ultimately, multiple specific events that represent UCA or UIF can co-exist at the same time and can be modeled within the same FTs (if applicable). It is reasonable to identify scenarios where specific functions within a complex controller or information processor fail under different modes at the same time. These complex failure conditions represent an unanalyzed and interesting challenge for risk assessment. The authors agree that these points will help clarify how UCAs/UIFs should be assessed and incorporate such a discussion in the future.
5.	The discussion seems to center on I&C systems that perform control actions, receive feedback from the controlled process, and then modify control actions (e.g., through UCA and IFP in Figure 7). It's hard for me to map this onto real-life safety systems in nuclear plants, where there typically isn't an ongoing feedback loop. For example, if a scram system detects a high reactor pressure signal, it initiates the control action (inserting control rods) and that is the end of it. Is this the right conception? Perhaps an illustrative example in the document would help the reader map the concepts onto a familiar system? Maybe another way to state this is that it seems the process can be simplified when the controlled process is divorced from the sensed process? I think that's typical for safety-related I&C systems.	Indeed, in the APR-1400 HSI QIAS-P system, the sensing apparatus is divorced from the actuation mechanisms. In such instances, one can analyze only UIFs or only UCAs. The QIAS-P is one of those examples where only failures along the information feedback pathway are assessed, and the actuation mechanisms are not assessed. In one-off actuation systems (i.e., scram), there still exists a sensing apparatus that monitors reactor state. In such cases, the UCAs and UIFs describe failure events defined for when the system has not yet actuated. Continuous control is not necessary.
6.	In RESHA Step 5, it's noted that there may be software CCFs at different levels in the FT or for different functions simply because the PLCs, for instance, may be manufactured by the same company and use the same low-level operating system. Software is very much unlike hardware in these cases because low-level operating systems may be precisely cloned from PLC to PLC. So, common failures here aren't really CCFs; they're single failures that may have been propagated. This seems like a very different scenario than, say, two pumps in a system that may be very similar, but are not precise clones (e.g., through stochastic effects that	Duplication of software defects across PLCs are indeed seeded failures. Given the same IE for duplicated defects, redundant PLCs will fail simultaneously and result in a common failure event. In this work, the authors describe software CCF events by the likelihood that software will share common triggering events to these defects. For instance, suppose the communication software of duplicated PLCs share the same defect where if an acknowledgment (ACK) signal is received twice, the system will fail. If an IE occurs where the ACK signal is sent twice to all connected PLCs, then yes, it is a single failure event of the PLC, but

#	Comment	Responses/Resolutions
	impact failure likelihood). If there is an underlying software defect that causes one PLC to fail, then it's not really a CCF for another software system to fail under the same input circumstances. How would the authors respond to this concern?	the failure has occurred over all PLCs that share the same defect. If these PLCs are redundancies to each other, then their simultaneous failure is equivalent to a CCF.
7.	Should the reader assume that the BBN developed in Fig. 41 is generalizable for all software development relevant to NPPs?	The BBN shown in Figure 41 is of a general form. Many software development life cycles will follow contain the same five stages of development. The authors also performed an application of BAHAMAS for a case study with industry for which there were six stages that were assessed.
8.	I'm having a bit of trouble understanding triggers, for example in Table 36. It appeared to be used in a context early in the document to refer to errors that would be carried out by the system. But, in Table 36, it seems like it's an input to help the user judge the review quality. If triggers are errors, then it's not making sense to me how the analyst could know this *a priori* to judge the review quality.	The issues of trigger usage in the document will be corrected. Triggers may be considered as scenarios that activate faults (defects) and bring about failure. Triggers, when used in the context of BAHAMAS are based on orthogonal-defect classification (ODC), in which a trigger is the activity or process that brings about a failure and therefore serves to identify a software defect. In BAHAMAS, trigger coverage represents the percentage of the trigger activities that have been investigated during the software development. For example, code inspection is a type of review activity considered as a trigger and contributes to trigger coverage. Thus, if code inspection does not occur, then one can judge that the review quality is less than it could be. Ultimately trigger coverage provides an indication of how comprehensive a review activity was.
9.	The authors state that the "…BAHAMAS output is considered as a total value rather than an independent failure value." Could you explain how this is possible before performing the CCF analysis? For example, how can the BBN solver account for the dependency information between activities in the SDLC before the CCF analysis has been performed? Perhaps it is just my reading and interpretation, but the placement of that statement seems to suggest that the total failure probability is able to be attained before the CCF Modeling and Estimation (Section 4.5).	BAHAMAS provides an indication of the total remaining defects within a software (e.g., a version of application software), and from that probability of defects remaining provides an indication of software failure probability. When that software is implemented on redundant components, each component will, by default, have the same set of hidden defects. From the viewpoint of defects available to cause a failure, BAHAMAS has captured them all and can provide an indication of the total failure probability in terms of those defects. It is from this set of defects that failure can occur, either concurrently (via an activation event shared by the CCCG) or independently (via an activation event that affects only a single component). Finally, BAHAMAS assumes that defect activation given the redundant configuration, each soft experiences nearly identical operational conditions; BAHAMAS provides a direct indication of the common defects that exist between redundant software components.

#	Comment	Responses/Resolutions
10.	The insignificance of software CCF in the example in Section 5.3 seems like a surprising result, especially given the preceding section where a software modified β factor of 0.429 was calculated. In the example that is presented, is there diversity in the RTS such that no single software CCCG could cause a failure of the function (i.e., there is complete diversity)?	As mentioned for question 6b, the RTS has multiple diverse means of ensuring a trip signal can be sent when needed. Thus, the failure of automatic function does not prevent actuation of reactor trip because the diverse protection system and operators provide an alternate means of ensuring reactor trip. For the model shown in this report, besides a mechanical failure of control rods, there are no first order cut sets that will lead to failure of the reactor trip function.
11.	Similar to an earlier comment, the distributions of accident sequences in the examples in 5.4 are quite different when the modified ETs are employed from the original ETs. Reducing estimated CDF is typically good, but when the plant is faced with limited resources to address safety improvements, this can have a large impact. For example, with the original ET, INT-TRANS:20 and *02-02-09 were significant but were very much overshadowed by other sequences. In the revised tree, these deserve more attention. If the methods are a consensus best-estimate, then how the plant chooses how to improve safety, what to train operators on, what components get enhanced inspections (e.g., CDBI) may all be changed. I think it might help make the case for adoption of the authors methods to highlight these examples.	It is a very good point to develop a use case to demonstrate how insights can be generated based on the proposed framework to support risk management and design optimization. Various types of data, information and knowledge at various levels and scales can be obtained from this framework. How to acquire, couple, integrate these various types of metrics in an efficient way to provide sufficient, accurate and tidy information for supporting decision-making processes is one ongoing work in this project.
12.	In the discussion of AI/ML, it's not clear that the authors addressed the issue of active AI systems *never* being in a position to train itself for safety-related control actions. It seems that any ML algorithm operating during plant operation would always be assimilating data about a plant state for which it's not designed (e.g., LLOCAs). It also seems unlikely that ML would be a benefit without augmented "knowledge" about the plant. For example, setpoints of an RTS based on reactor level are based on "god-like" knowledge about everything else in the plant based on simulation software and would include things like DNBR, linear heat generation rate, reactor pressure, etc. Could an AI/ML system ever competently judge when it's right to scram the reactor on low level without a similar training parameter space dimension size?	The authors agree that there is an issue of "active AI systems *never* being in a position to train themselves for safety-related control actions" due to lack of data and knowledge. The motivation of this discussion in Section 6 of the reviewed report is to identify and estimate how this issue can affect the availability of system control functions and plant safety. It is infeasible to ensure the absolute safety and reliability of a complex control system. However, the question is raised, "how safe is safe enough?" To support efforts to answer this question, this work plans to establish metrics that represent and measure the plant safety, system availability and reliability. Some relevant work has been initiated in FY 2023, including defining and estimating some metrics that can measure the impacts of the lack of qualified training data to the plant safety, system availability and reliability. Results will be included in the milestone report to be published in August 2023.

5.14.2 Additional Comments from RPI

Additional comments from GEH and responses from the report authors are summarized in Table 17Table 16.

Table 17. Summary of additional comments from RPI and responses.

#	Comments	Responses/Resolutions
1	Can we make the report title more concise? For example, High Safety-significant Safety-related Digital Instrumentation and Control Systems --> Safety-critical digital systems.	The term HSSSR was leveraged from the NRC/NEI's definition in Report NEI 20-07: *Guidance for Addressing Software Common Cause Failure in High Safety-Significant Safety-related Digital I&C Systems* to represent the safety-related systems, structures, components (SSCs) that perform safety-significant functions (e.g., RTS and ESFAS). In NRC BTP-17, it is also suggested to use safety significance to determine whether a diversity and defense in-depth assessment is necessary. That is why both "safety-related" and "safety-significant" are emphasized in this work.
2	In section 4.3.1, (1). Table 15 needs to appear after explaining what UCA/UIF-A/B/C/D are. (2). Tables 15 and 16 can be shown in a consistent format for easier insights. (3). Table 16 can include probabilities in addition to event numbers. Can the title of Table 16 be revised to 'Number of defects reported …' rather than 'sample size …'? (4). The numbers in Equation 4 must be explained before they appear. For example, where does F_alg(1hr|s) =8.73e-4 come from? Moreover, without specified s, it cannot be quantified, so the readers cannot follow this calculation process. 0.0417 should be replaced by '1 hour'. Since the point of software reliability growth model is the progress of debugging provides a less-faulty software. The debugging time dominates the quantified result. (5). It would help the readers if is explained that Table 15 is the basis of the first vector.	Changes will be made accordingly in the new version of the report: (1). Table 15 will be moved for clearer logic. (2). Table 16 will be revised to include the number of UCA/UIF per software. (3). Table 15 is the culmination of Table 16. Sample size is used to specify (a) the number of defects used to develop the numbers from table 15 and (b) the data is not a population and may be biased. (4). F_alg(X|s); 'X=0.0417' comes from 1/24 hrs which is used to represent failures per hour as the data collected was measured in days and not hours. 0.0417 is the conversion between days to hours. S is also in 1-hour units and is the end state of the SRGM. 's' is known within the SRGM but has very little meaning. Additional description about formula 4 will be added. (5). A comment will be added to note that Table 15 is the basis for the first vector.
3	In section 4.3.4.3, "Using the correlations shown in Table 24, the individual UIF failure probabilities are determined, observed in Table 32" should be explained more since Table 24 gives different numbers from those in Table 32. "The total software failure probability regardless of failure mode was determined to be 8.77E-4 per hour (bottom right sum in Table 32)" doesn't seem correct since the total in Table 32 is 7.31E-04.	Table 24 is the global UCA/UIF correlation rate while Table 32 is Table 24 applied to the failure data from the VCU case study. The correct value is updated for Table 32. An explanation will also be provided to highlight the difference between these tables.
4	In section 4.3.4.4,	Changes will be made accordingly in the new version of the report.

#	Comments	Responses/Resolutions
	How the test coverage 75% was estimated? In the previous section, it was mentioned that 10951 tests were done and each was assumed to correspond to 1 hour of efforts. How can this be connected to 75%? And there are many possible ways to define a test coverage, so this must be clearly discussed.	In section 4.3.4.4, it is mentioned that 76% of defect triggering conditions were tested. This value is determined from Table 31. Qualitative test assessment does not consider the number of tests or quality of the tests, only if a test (test set) exists that addresses a specific defecting triggering scenario.
5	In section 4.4.1, "Table 24 shows the relationships for relating defect classifications to UCAs/UIFs" can be rephrased since Table 24 is about the quantification of portions of specific failures and does not address any relationships. Current HRA methods including THERP are for the quantification of 'human errors under specific procedures', i.e., the target is error of omission rather than error of commission. This work utilizes them for EOC-style errors (such as defect insertions). So the proper assumptions should be discussed somewhere in BAHAMAS section.	Table 24 provides conditional probabilities for different UCAs/UIFs for different defect types. In other words, this defines a relationship between defect types and UCAs/UIFs. Regarding THERP and EOC-style errors, the authors will add additional clarification for the assumptions associated with the usage in the subsequent revision of the report.
6	In section 4.4.2, In the recovery part, "y = exp(-R*x) where R=0-3" seems that it doesn't need to be defined separately from y = exp(-x) since if there were multiple reviews it would increase x (review efforts) even without R. Assuming 0-3 reviews only doesn't seem correct. Since Figure 45 was suggested as the comparison between BAHAMAS and THERP, the reason of difference at higher review quality. Caption of Figure 45 has an error. X-axis label (Reviewer dependence) is not explained in the main text at all. It is unclear where equation 13 was used in the following sections of the report. It is difficult to digest "Additionally, a conditional relationship between nodes representing defect type remaining and the software failure mode is assumed: the occurrence of a software failure results from an activated software defect." Please elaborate. Please pay attention at probability notations. E.g., in Pr (software failure mode), 'software failure mode' is in the place of a variable. 'software failure mode' and 'defect type remains' cannot be a variable either. It mentions "Table 24 provides the conditional probabilities used for these nodes", but Table 24 is for UCA/UIF conditional probabilities, so this probability quantification process needs to be explained in a crisper manner.	This area of the report will be revised to provide clarification for Figure 45 and the usage of Equation 13. It will be clarified as "Additionally, a conditional relationship between nodes representing defect type remaining and the software failure mode is assumed: the occurrence of a software failure results from an activated software defect." Specifically, software failure consists of the existence of a defect and the activation of that defect which results in a software failure. BAHAMAS essentially tracks the probability of defects remaining in the software. There needs to be a conditional probability that relates defects to failure; activation probability is needed. In this work activation is assumed to be unity. For convenience, Pr(defect type remaining) is assumed equivalent to Pr(defect type remaining and is active). For example, the probability of an algorithm defect remaining within the software, Pr(Algorithm defect), is assumed to represent the probability that the algorithm defect exists and is active. This has a conservative impact on the results from BAHAMAS. Future work will investigate the activation probability. More details concerning the relationships of each node of the BBN will be added in a future revision of the report. These details will clarify how table 24 is used.
7	In section 4.4.3,	A future version of the report will include more relevant information about the HRA analysis and specific tables used. Though, the reader is

#	Comments	Responses/Resolutions
	Since some THERP tables (like 20-3, 12-5) are mentioned and referred, the readers would appreciate it if they were included in the report. We cannot assume all readers have access to THERP tables. Are software development activities diagnostic actions? Please elaborate the assumptions regarding how THERP guidance can be applied to this task with clear details. The numerical values of tables should be traceable. Some of them might come from references, so it would be great if as much as possible details are clearly demonstrated in the report.	advised to see the details of the HRA application found in the report's Appendix A where the justification of each table has been indicated.
8	In section 4.5.1, In Equations 15 and others, based on conventional notation norm, [Q_1]3 means [Q_1]^3 (Q_1 to the power of 3), so better to revise like [Q_1](3) as in the Dr. Mosleh's original notation in NUREG/CR-5485. In Equations 15-19, the last term should be removed since Q1-Q3 cannot be determined without m. If the challenges of SW-based CCF quantification are explained systematically before the description of asymmetric CCF models, it would help the reader to digest the topics discussed in this section. I didn't clearly understand why the trials of asymmetric causes of CCF modeling should be explained here before I read example cases in section 4.5.4.	This will be corrected in the next revision of this report.
9	In section 4.5.2, Is Table 40 universal to any kind of hardware? If there are more specifications, they need to be clarified here. The quantitative basis of Table 41 needs to be clearly specified including assumptions and limitations.	It will be clarified that Table 40 was designed originally to assess electronic systems such as the components of a system, not for software. However, given its age, the table can be re-calibrated to modern digital hardware failure data. That is outside the current work plan for the current report. But future work may merit such investigations. The current research plan is currently focused on refining Table 41. The results of that effort will be documented in a future report in which the assumptions the limitations will be addressed more thoroughly.
10	In section 4.5.3, In Table 45, is it 'average probability'? In Table 49, can 'Division A&B racks are physically isolated' be the reason of poor score of 'input similarity'? It may correspond to 'isolation', but if it is the case, the evaluation should be 'good' rather than 'poor' because they are well isolated.	In Table 45, the last column refers to the averaged probability of occurrence modeled over different software reliability growth method (SRGM) algorithms presented in Table 17. The column was changed to "Average probability of occurrence". First question about Table 49, it is a typo. The score is poor because Division A&B have the same input source.

#	Comments	Responses/Resolutions
	In Table 49, is 'understanding' for the direct experience of using the target software? Otherwise, can 'Less than 10 operating years of software experience' be the reason of poor score? For Table 50, it would be helpful to the readers if the calculation basis (formula and etc.) is clarified in the text again. It seems like P(F) was calculated first and divided into beta*P(F) and (1-beta)*P(F), but clear explanation would help.	The second question about Table 49, was based on an old version of for scoring. The authors have since updated the scoring in a subsequent report. One aspect of that scoring was limited experience with the system. That old scoring assumed for understanding that if software was involved for the components, then the score would be given "A" for understanding. Table 50, the response is that the calculation to determine the values in the table will be reiterated.
11	In section 4.5.4, For Table 51, there is a reference for hardware failure probabilities, but no basis explained for software failure probabilities.	The software failure probabilities are a result of the case study performed by BAHAMAS. The authors will add a note to the table.
12	In section 5.3.3, The details of fault tree models and data are not available. But, regarding Table 56, I was wondering how the 1E-4 scale event (like LP SW CCF for all in Table 54) does not appear. The function of LP can be backed up by human operator's manual actuation, but assuming 1-10% of HEP, the cutset probability should be somewhere between 1E-5 to 1E-6, which is well above those of the dominant cutsets in Table 56. Thus, some more clear discussion would help the readers.	Changes will be added in the new version of the report.
13	In section 5.3.5, In table 58, 'top event' should be revised. Every FT model has its top event. Here it seems like HSI, but it is not clear.	Yes, it should be "HSI Failure". Changes will be made accordingly in the new version of the report.

6. CONCLUSIONS

This report summarizes the INL-initiated peer review activities during FY 2023 for evaluating and improving the methodology developed under the U.S. DOE LWRS-RISA's DI&C Risk Assessment project. This peer review activity includes coordinating the reviews from industry stakeholders, documenting the peer review feedback, and providing resolutions and responses to the peer review comments. This technical peer review's objective is to obtain representative feedback on the proposed framework to improve the technical qualities of its methodology and readiness for deployment to the industry. Feedback may identify potential areas for improvement and further development.

Six technical peer reviewers were invited to review one of the latest project reports, INL/RPT-22-68656, documenting the methodology developed in the project, and provide technical evaluations of the proposed framework and relevant methods. Comments from technical peer reviewers and resolutions and responses to these comments from the LWRS-RISA team are outlined and discussed. Insights obtained from the technical peer review and relevant future research activities are summarized in this section.

6.1 Insights from Peer Review

This section provides the insights obtained from the technical peer review:

1. There is a significant need and benefits to developing a best-estimate, risk-informed capability for addressing potential CCF issues of safety-critical digital systems and supporting decision-making processes of relevant risk management and design optimization.

As discussed in the peer reviewed report, technical challenges, and questions about potential software CCFs in the HSSSR DI&C systems of NPPs still remain. According to feedbacks from the stakeholders, there are significant benefits if a best-estimate risk assessment process can be developed to support a risk-informed and cost-efficient diversity and defense-in-depth applications for assuring the long-term safety of safety-critical digital systems and reducing uncertainties in costs, time, and supporting integration of digital systems during the design stage of the plant.

To better address these technical challenges, this LWRS-developed framework was proposed to provide a best-estimate risk-informed capability to address digital issues quantitatively, focusing on software CCFs in safety-critical DI&C systems of NPPs. It is worth developing such a capability to provide evidence to support risk and cost reduction during the DI&C design and upgrade. For instance, is it necessary to eliminate all identified potential CCFs, or just the ones that can significantly affect system reliability and plant safety based on a prioritized assessment? How to determine which CCF is more safety-significant, especially when considering there are different types and scales of CCFs in the safety-critical highly redundant DI&C systems? For instance, one current collaborative study with the nuclear industry has demonstrated that, while the system-level CCF was intolerable and the major contributor for risk increase within the model, some CCFs at smaller levels are tolerable as the increase in risk is manageable. Of course, answering these questions needs close collaboration among system designers, risk analysts and regulators. The proposed framework is expected to be used as a risk-informed tool that offers various quantifiable metrics to support relevant decision-making processes in diversity and defense-in-depth applications.

It is recognized that the development for such a comprehensive framework for DI&C risk assessment has many technical challenges. These metrics proposed in this work need sufficient validation which is difficult for now due to the lack of data, so results are only suggested to be used for comparison of different design architectures not as "truth" before validation is performed. Here PRA is used as an approach to provide various metrics and evidence to support reasoning, not as a presumptively authoritative result.

2. The metrics and results of the proposed framework can be used to support and supplement the implementation of other existing advanced risk-informed DI&C design guides such as EPRI's framework by providing quantifiable risk information and metrics.

While EPRI's framework provides a very detailed and reasonable risk-informed DI&C design guideline, the LWRS-developed framework proposes an advanced PRA-based framework for DI&C risk assessment and provides various quantifiable metrics that can be used to support and supplement EPRI's framework. The LWRS framework offers the capability of design architecture evaluation of various DI&C systems to support system design decisions in diversity and redundancy applications. For example, for HAZCADS, the LWRS-developed framework can provide detailed CCFs in different redundancy levels, quantifiable metrics to support the risk importance analysis and the ranking of risk reduction targets, and a quantitative consequence analysis to trace the impacts of individual failures. For DRAM, the LWRS-developed framework can provide quantifiable software reliability metrics to evaluate if and how much the control methods can mitigate consequences and reduce risks.

The reviewers also pointed out the opportunities to integrate BAHAMAS/ORCAS into the front end of the DEG in functional analysis and testing section as well as in the DRAM. More communication and collaboration between LWRS and EPRI are expected to realize these benefits.

3. Function-based risk assessment of DI&C system is an unanalyzed and realistic challenge; relevant methodology development and demonstration should be conducted in future research.

According to the reviewer's comments, it would be beneficial if the proposed framework can more explicitly consider DI&C platforms that perform several mitigation functions, since failures in these systems have a greater potential to result in undesired consequences due to simultaneous failure of functions. Coupled failure modes of an identical or multiple functions of digital processors may lead to complex and undesired accident scenarios.

Depending on its design, a digital controller may have multiple functions that may fail at the same time but with different modes (i.e., UCA/UIFs). Ultimately, multiple specific events that represent UCA or UIF can co-exist at the same time and can be modeled within the same FTs (if applicable). Therefore, it is reasonable to identify scenarios where specific functions within a complex controller or information processor fail under different modes at the same time. These complex failure conditions represent an unanalyzed and interesting challenge for risk assessment.

While current framework and relevant methods developed in this project focus on the risk analysis at component and system levels, the proposed framework and methods will be further developed and extended to provide a DI&C risk assessment at function level. This will also help build up a more comprehensive risk assessment framework that covers small scales (e.g., software elements and functions), middle scales (e.g., digital modules and components), and large scales (e.g., systems and plants). Risk data, information and metrics at different scales can be collected, coupled, and integrated to provide a clearer picture of DI&C risk status to support risk management and design optimization.

4. Two proposed reliability analysis methods, BAHAMAS and ORCAS should be further refined to (1) align better with international standards and achieve compatibility and influence with the DEG/HAZCADS/DRAM ecosystem and (2) provide various, appropriate and, quantifiable metrics to measure software reliability.

Future work will concentrate on refining BAHAMAS and ORCAS to align better with international standards and achieve compatibility and influence with the DEG/HAZCADS/DRAM ecosystem. Opportunity exists to integrate BAHAMAS/ORCAS more directly into the front end of the DEG in functional analysis and testing section as well as in the DRAM. More communication and collaboration in coupling the LWRS framework with EPRI's framework are expected.

The reviewers indicated a potential to evolve the proposed reliability methods (e.g., BAHAMAS and ORCAS) to provide alternative metrics to the traditional probability quantification. Such changes may enhance the capabilities for the LWRS framework to address systematic issues, better align with international standards, and potentially support EPRI's framework. The authors anticipate defining various quantifiable metrics and their associated uncertainty to evaluate the interactions among software defects, failure mechanisms, triggering events, and failure modes, for which BAHAMAS and ORCAS are well suited. The authors look forward to continued work and dialogue in this area that may lead to new and useful capabilities to benefit decision making for the risk management and design optimization of DI&C systems.

5. Consequence evaluation of digital failures should be extended to consider the impacts through both the IEs and non-IE top events in the ET models.

In this project, consequence analysis focuses on evaluating how CDF can be affected by the identified digital-based failures, basically Level 1 risk analysis. The current study is focused on evaluating the impacts of DI&C systems on mitigation systems (i.e., non-IE top events in the ET models). The reviewers suggested to extend the consequence analysis to consider the impacts through the IEs as well. This will be investigated and considered as one future work. The work can be started with a qualitative evaluation of the impacts from digital-based failures on the IEs. Potential impacts on the IEs include changing frequencies of existing IEs, introducing new IEs, and leading to concurrent or cascading IEs.

6. Suggestions are provided by the reviewers to extend the demonstration and application of the proposed framework to all safety-related and even some non-safety-related DI&C systems.

According to the reviewer's comments, some I&C systems of future plants may also have diverse and redundant designs where this framework could be of interest in application. The current framework is focused on the risk assessment of HSSSR DI&C systems.

Meanwhile, to better deal with these EOC issues relevant to the performance, resilience, and robustness of software systems, and support the design optimization of non-safety and safety-critical DI&C systems, some other metrics should also be developed or leveraged.

7. Considering the differences between AI-aided and traditional DI&C systems, novel methods and metrics should be developed. Specifically, technical issues in AI-aided control system risk assessment should be clarified and clearly addressed.

The motivation of the review activity documented in Section 6 of INL/RPT-22-68656 is to define (1) whether new or unanalyzed failure modes exist in the newly designed AI-aided control system, and (2) how to identify them and evaluate their impacts to system reliability and plant safety. Part of the ongoing research is to assess if the failure modes of conventional software systems will be different than those of AI-aided systems. Importantly, the contextual conditions that the AI-aided system operates in will define how the system will fail. Addressing technical issues in AI-aided control system risk assessment should focus on (1) identifying differences in the failure cause/modes/effects of AI-aided and traditional DI&C systems, (2) developing suitable concepts and metrics to describe the new potential failure cause/modes/effects, and (3) developing suitable methods to quantify and estimate these metrics.

Research on applying AI/ML to the operation and control of NPPs has been started for a long time in many DOE-led programs and projects. ML-aided digital twin technology is a popular topic in nuclear engineering. Although current AI/ML techniques are not mature enough for real deployment, it is necessary to push relevant supporting work such as reliability/risk assessment and trustworthiness study of AI/ML-aided reactor operation and control. The NRC also published several reports on this topic, one of which is *NUREG-2261: Artificial Intelligence Strategic Plan* from June 2022. Anticipating the industry's potential application of AI to NRC-regulated activities, the NRC has developed a strategic plan to ensure the agency's readiness to review such uses. In this context, it is meaningful and necessary to push forward relevant research work, in this project, it is about how to extend and adjust the proposed

framework and relevant methods to support the risk assessment and design optimization of AI-aided control systems.

6.2 Future Work

By considering reviewer's comments, Report INL/RPT-22-68656 will be revised and released in a new reversion. Based on reviewer's comments and insights obtained from this review, some potential future research activities are listed below and will be considered as an important reference to FY 2024 work planning.

1. Refine the current framework to align better with international standards and achieve compatibility and influence with existing advanced risk-informed approaches, perform a VVUQ of developed methods (i.e., BAHAMAS / ORCAS/ CCF modeling approach).
    a. Refine the RESHA process by clarifying (1) the connections between UIFs and UCAs, and (2) relationship among root cause, failure mechanism, triggering events and failure modes (UCA/UIF) for software failures.
    b. Refine BAHAMAS by (1) providing a detailed process, (2) improving the HRA applications, (3) developing relevant UQ process, and (4) providing various metrics for software reliability evaluation.
    c. Refine ORCAS by (1) developing relevant UQ process, specifically acquiring more data to validate and justify the UCA/UIF correlation table, which is currently based on non-safety critical software, (2) integrating qualitative test coverage metrics into the evaluation of software reliability, and (3) integrating deployment or operational triggering event probability into reliability quantification for a more realistic software reliability estimate.
    d. Define various quantifiable metrics using BAHAMAS/ORCAS to represent the interactions among software defects, failure mechanisms, triggering events, and failure modes.
    e. Refine the CCF modeling approach by (1) improving the scoring table, and (2) developing relevant UQ process.
2. Further develop the proposed framework and methods for function-based risk assessment of DI&C system.
3. Collaborate with the industry to couple the LWRS-developed framework with existing risk-informed approaches to better support the optimization of safety-critical DI&C designs and upgrades.
4. Extend consequence evaluation of digital failures to consider the impacts through both the IEs and non-IE top events (i.e., mitigation systems) in the ET models.
5. An integrated importance, sensitivity and prevention analysis will be investigated and conducted to better support the best-estimate decision-making processes of risk assessment and design optimization.
6. Propose metrics to evaluate the performance, resilience, and robustness of software systems, and support the design optimization of DI&C systems.
7. Improve and further develop current LWRS-developed framework to support risk evaluation and design optimization of machine-learning-based digital-twin-enabled control systems.

7. REFERENCES

[1] H. Bao, H. Zhang and K. Thomas, "An Integrated Risk Assessment Process for Digital Instrumentation and Control Upgrades of Nuclear Power Plants," Idaho National Laboratory, Idaho Falls, ID, August 2019.

[2] H. Bao, H. Zhang and T. Shorthhill, "Redundancy-guided System-theoretic Hazard and Reliability Analysis of Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants," Idaho National Laboratory, Idaho Falls, ID, August 2020.

[3] H. Bao, T. Shorthill, E. Chen and H. Zhang, "Quantitative Risk Analysis of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants using IRADIC Technology," Idaho National Laboratory, Idaho Falls, ID, August 2021.

[4] H. Bao, T. Shorthill, E. Chen, J. Park, S. Zhang, A. V. Jayakumar, C. Elks, N. Dinh, H. Ban, H. Zhang, E. Quinn and S. Lawrence, "An Integrated Framework for Risk Assessment of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants: Methodology and Demonstration," Idaho Nationla Laboratory, Idaho Falls, ID, August 2022.

[5] H. Bao, S. Zhang, R. Youngblood, T. Shorthill, P. Pandit, E. Chen, J. Park, H. Ban, M. Diaconeasa, N. Dinh and S. Lawrence, "Risk Analysis of Various Design Architectures for High Safety-Significant Safety-Related Digital Instrumentation and Control Systems of Nuclear Power Plants During Accident Scenarios," Idaho National Laboratory, Idaho Falls, ID, November 2022.

[6] B. Littlewood, I. Bainbridge and R. Bloomfield, "The use of Computers in Safety-critical Applications," The Health and Safety Commission, London, UK, 1988.

[7] R. Maguire, Safety Cases and Safety Reports: Meaning, Motivation and Management, CRC Press, 2017.

Page intentionally left blank

APPENDIX A: PEER REVIEW NARRATIVE

Light Water Reactor Sustainability Program – Risk-Informed Systems Analysis

Technical Peer Review

Risk Assessment of Digital Instrumentation and Control Systems

Dear Reviewer,

Research and development (R&D) is being conducted by the Light Water Reactor Sustainability (LWRS) Program to facilitate digitalization of existing nuclear power plants including the use of risk assessment techniques. The project "Risk Assessment of Digital Instrumentation and Control (DI&C) Systems" aims to develop a risk assessment framework for providing a technical basis to support effective, licensable, and secure DI&C technologies for digital upgrade and design. The framework has matured to a point that we would like to conduct a technical peer review and obtain stakeholder feedback.

We would invite you to participate in the technical peer review of the methodology development and demonstration of the "LWRS Program developed framework for DI&C risk assessment" (short as "proposed framework" or "the framework" in following paragraphs) and appreciate your assistance. More information on the peer review is provided below.

Goal:

The goal of this technical peer review is to obtain representative feedback on the proposed framework to improve the technical qualities of its methodology and readiness for deployment to the industry. Feedback may identify potential areas for improvement and further development.

Scope:

The subject-matter experts will review the latest project report documenting the methodology developed in the project and provide evaluations of the technical qualities of the proposed framework and relevant methods. The project report to be reviewed is "An Integrated Framework for Risk Assessment of High Safety-significant Safety-related Digital Instrumentation and Control Systems in Nuclear Power Plants: Methodology and Demonstration" INL/RPT-22-68656.

The results of this peer review will be documented in a report that will be sent to the peer review team for final review and comments before finalizing. The following questions focus on specific topics to assist with the peer review. However, feedback and comments outside of the offered questions are appreciated and welcomed.

Section 2. Technical Background

1. The proposed framework presented in Figure 2 includes three steps (e.g., hazard, reliability, and consequence analysis) for risk analysis and defines three acceptance criteria for risk evaluation. Is its workflow clear and complete for DI&C system risk analysis and evaluation? Are the steps and acceptance criteria well defined and sufficient to provide insights to reduce risks and optimize designs?

2. Do you have any suggestions for overall framework improvements, if anything, on the identification, quantification, and evaluation of potential DI&C system failures?

3. The framework is expected to support existing industry guidance and methods (e.g., HAZCADS and DRAM) by providing quantitative risk information. In your opinion, which aspects of the framework can be utilized to support existing industry design and evaluation guidance? What adjustments and/or improvements are needed, if any, to provide better support to existing industry guidance and methods?

Section 3. Redundancy-Guided System-Theoretic Hazard Analysis (RESHA)

4. The framework's hazard analysis method is called RESHA and its workflow is presented in Figure 11. With respect to the complexity of High Safety-significant Safety-related (HSSSR) DI&C systems considered, do you think Section 3 has clearly described and demonstrated the RESHA capability in identifying potential CCFs, especially software CCFs, for different levels of redundancy and failure mode types?

5. Based on the integration of STPA and HAZCADS, RESHA can identify software failure mechanisms in the control/actuation pathway using Unsafe Control Actions (UCAs). A novel concept called Unsafe Information Flow (UIF) has been developed and introduced in RESHA to complete the identification and tracing of software failure mechanisms in the information/feedback pathway (IFP), as discussed in Section 3.2. Please provide your feedback for the development and application of UIFs in identifying software failure mechanisms in the IFP.

Section 4. Multiscale Quantitative Reliability Analysis

6. The framework's multiscale quantitative reliability analysis workflow presented in Figure 25 includes two methods for software reliability analysis, ORCAS and BAHAMAS.

    a. In Section 4.3, ORCAS is developed and demonstrated to estimate the probability of UCAs/UIFs for rich-data conditions. Please provide your feedback on: (1) are the assumptions of ORCAS reasonable? (2) are input requirements and steps of ORCAS clearly described, logical and practical for deployment with rich testing data available? (3) is ORCAS method well demonstrated in the case studies with all expected outcomes obtained?

    b. In Section 4.4, BAHAMAS is developed and demonstrated to estimate the probability of UCAs/UIFs for limited-data conditions. Please provide your feedback on: (1) are the assumptions of BAHAMAS reasonable? (2) are input requirements and steps of BAHAMAS clearly described, logical and practical for deployment with no testing data and very limited design information? (3) is BAHAMAS method well demonstrated in the case studies with all expected outcomes obtained?

7. In Section 4.5, a modified beta factor method was developed and modified to support CCF modeling and parameter estimation as part of this framework. Please provide your feedback on: (1) are the method assumptions reasonable? (2) are the input requirements and steps clearly described, logical and practical for deployment to the industry? (3) is the method well demonstrated in the case studies with all expected outcomes obtained?

Section 5. Consequence Analysis of a Generic PRW with Advanced HSSSR DI&C Systems

8. Section 5 documents the consequence analysis of a generic PWR SAPHIRE model with an improved digital reactor trip system (RTS) and Engineered Safety Features Actuation System (ESFAS) fault trees. Please provide your feedback on: (1) Are the demonstrated consequence analyses on HSSSR DI&C systems sufficiently detailed to support industry needs in DI&C modeling and risk analysis? (2) Is it beneficial to perform similar consequence analyses for other DI&C systems that are safety-related but not safety-significant?

9. In addition to providing the changes of core damage frequency (CDF) and large early release frequency (LERF) due to the digital upgrades for HSSSR I&C systems, what other risk-informed insights from the quantitative consequence analysis would be beneficial for the evaluation and reduction of the plant-level risks?

Section 6. Future Applications on AI-Aided Control Systems

10. Some technical gaps in applying the framework for the risk analysis of AI/ML-aided control systems have been reviewed and investigated in Section 6. With respect to the technical gaps considered, are they complete and significant in terms of identifying, quantifying, and evaluating potential failure modes of AI-aided control systems? Are there any other approaches to addressing these types of technical gaps that you recommend us to consider?

Section 7. Conclusions and Future Works

11. Section 7 summarizes recent work and proposes future R&D. Do you agree with the identified needs for these activities? Are there other relevant short- and long-term industry needs in the area of DI&C risk assessment not addressed by this framework?

12. Please provide any additional suggestions for the framework improvements.