

Assessment of the Quality of Selected NRC Research Projects by the Advisory Committee on Reactor Safeguards - FY 2018

February 2019

**U.S. Nuclear Regulatory Commission
Advisory Committee on Reactor Safeguards
Washington, DC 20555-0001**



ABOUT THE ACRS

The Advisory Committee on Reactor Safeguards (ACRS) was established as a statutory Committee of the Atomic Energy Commission (AEC) by a 1957 amendment to the *Atomic Energy Act* of 1954. The functions of the Committee are described in Sections 29 and 182b of the Act. The *Energy Reorganization Act* of 1974 transferred the AEC's licensing functions to the U.S. Nuclear Regulatory Commission (NRC), and the Committee has continued serving in the same advisory role to the NRC.

The ACRS provides independent reviews of, and advice on, the safety of proposed or existing NRC-licensed reactor facilities and the adequacy of proposed safety standards. The ACRS reviews power reactor and fuel cycle facility license applications for which the NRC is responsible, as well as the safety-significant NRC regulations and guidance related to these facilities. The ACRS also provides advice on radiation protection, radioactive waste management, and earth sciences in the agency's licensing reviews for fuel fabrication and enrichment facilities, and waste disposal facilities. On its own initiative, the ACRS may review certain generic matters or safety-significant nuclear facility items. The Committee also advises the Commission on safety-significant policy issues and performs other duties as the Commission may request. Upon request from the U.S. Department of Energy (DOE), the ACRS provides advice on U.S. Navy reactor designs and hazards associated with DOE's nuclear activities and facilities. In addition, upon request, the ACRS provides technical advice to the Defense Nuclear Facilities Safety Board.

ACRS operations are governed by the *Federal Advisory Committee Act*, which is implemented through NRC regulations at Title 10, Part 7, of the *Code of Federal Regulations*. ACRS operational practices encourage the public, industry, State and local governments, and other stakeholders to express their views on regulatory matters.

MEMBERS OF THE ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

Dr. Ronald G. Ballinger, Professor Emeritus of Nuclear Science and Engineering and Materials Science and Engineering, Massachusetts Institute of Technology.

Dr. Dennis C. Bley, President, Buttonwood Consulting, Inc.

Mr. Charles H. Brown, Senior Advisor for Electrical Systems, Syntek Technologies, Inc.

Dr. Margaret Sze-Tai Y. Chu, Former Director of the Department of Energy's Office of Civilian Radioactive Waste Management.

Dr. Michael L. Corradini, Professor Emeritus of Engineering Physics, University of Wisconsin.

Dr. Vesna B. Dimitrijevic, Retired Technical Consultant, AREVA, Inc.

Dr. Walter L. Kirchner, Retired Technical Staff Member, Argonne National Laboratory and Los Alamos National Laboratory.

Dr. Jose March-Leuba, Principal of MRU and Associate Professor of Nuclear Engineering Department, University of Tennessee.

Mr. Harold B. Ray, Retired Executive Vice President, Southern California Edison Company.

Dr. Joy L. Rempe, (Member-at-Large) Principal, Rempe and Associates, LLC.

Dr. Peter C. Riccardella (Chairman), Senior Associate, Structural Integrity Associates, Inc.

Mr. Gordon R. Skillman, Principal, Skillman Technical Resources, Inc.

Mr. Matthew Sunseri (Vice-Chairman), Retired President and Chief Executive Officer of Wolf Creek Nuclear Operating Corporation.

ABSTRACT

In this report, the ACRS presents the results of its assessment of the quality of selected research projects sponsored by the NRC Office of Nuclear Regulatory Research. An analytic/deliberative methodology was adopted by the Committee to guide its review of research projects. The methods of multi-attribute utility theory were used to structure the objectives of the review and develop numerical scales for rating each project with respect to each objective. The results of the evaluations of the quality of the selected research projects are summarized as follows:

- NUREG/CR-7237, “Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)”
 - This project was found to be satisfactory, a professional work that satisfies research objectives.
- NUREG-2218, “An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)”
 - This project was found to be satisfactory, a professional work that satisfies research objectives.

CONTENTS

Page

ABSTRACT	iv
FIGURES	vi
TABLES	vi
ABBREVIATIONS	vii
1. INTRODUCTION	1
2. METHODOLOGY FOR EVALUATING THE QUALITY OF RESEARCH PROJECTS.....	3
3. RESULTS OF QUALITY ASSESSMENT.....	5
3.1 Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)	5
3.2 An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs).....	10
4. REFERENCES	14
APPENDIX A: Comments on PIRT Elicitation Process and Facilitation.....	16

FIGURES

Page

1. The Value Tree used for Evaluating the Quality of Research Projects 3

TABLES

1. Constructed Scales for the Performance Measures 4
2. Summary Results of ACRS Assessment of the Quality of the Project NUREG/CR-7237, "Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)" 7
3. Summary Results of ACRS Assessment of the Quality of the Project NUREG-2218, "An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)" 11

ABBREVIATIONS

ACRS	Advisory Committee on Reactor Safeguards
AEC	Atomic Energy Commission
ASME	American Society of Mechanical Engineers
BWR	boiling-water reactor
CDF	core damage frequency
FY	fiscal year
HEAF	high energy arcing fault
LBNL	Lawrence Berkeley National Laboratory
LWR	light-water reactor
NPP	nuclear power plant
NRC	Nuclear Regulatory Commission
PIRT	phenomena identification and ranking table
PRA	probabilistic risk assessment
PWR	pressurized-water reactor
RES	Office of Nuclear Regulatory Research
SPRA	seismic probabilistic risk assessment
SSCs	structures, systems, and components
SSHAC	Senior Seismic Hazard Analysis Committee
U.S.	United States

1 INTRODUCTION

The Nuclear Regulatory Commission (NRC) maintains a safety research program to ensure that the agency's regulations have sound technical bases. The research effort is needed to support regulatory activities and agency initiatives while maintaining an infrastructure of expertise, facilities, analytical tools, and data to support regulatory decisions.

The Office of Nuclear Regulatory Research (RES) is required to have an independent evaluation of the effectiveness (quality) and utility of its research programs. This evaluation is required by the NRC Strategic Plan that was developed as mandated by the Government Performance and Results Act. Since fiscal year (FY) 2004, the Advisory Committee on Reactor Safeguards (ACRS) has been assisting RES by performing independent assessments of the quality of selected research projects [1-14]. The Committee established the following process for conducting the review of the quality of research projects:

- RES submits to the ACRS a list of candidate research projects for review because they have reached sufficient maturity that meaningful technical review can be conducted.
- The ACRS selects a maximum of four projects for detailed review during the fiscal year.
- A panel of three to four ACRS members is established to assess the quality of each research project.
- The panel follows the guidance developed by the ACRS Full Committee in conducting the technical review. This guidance is discussed further below.
- Each panel assesses the quality of the assigned research project and presents an oral and a written report to the ACRS Full Committee for review. This review is to ensure uniformity in the evaluations by the various panels.
- The ACRS submits an annual summary report to the RES Director.

Based on later discussions with RES, the ACRS made the following enhancements to its quality assessment process:

- After familiarizing itself with the research project selected for quality assessment, each panel holds an informal meeting with the RES project manager and representatives of the user office to obtain an overview of the project and the user office's insights on the expectations for the project with regard to their needs.
- In addition, if needed, an additional informal meeting is held with the project manager to obtain further clarification of information prior to completing the quality assessment.

The purposes of these enhancements were to ensure greater involvement of the RES project managers and their program office counterparts during the review process and to identify objectives, user office needs, and perspectives on the research projects.

An analytic/deliberative decision-making framework was adopted for evaluating the quality of NRC research projects. The definition of quality research adopted by the ACRS includes two major characteristics:

- Results meet the objectives
- The results and methods are adequately documented

Within the first characteristic, the ACRS considered the following general attributes in evaluating the NRC research projects:

- Soundness of technical approach and results
 - Has execution of the work used available expertise in appropriate disciplines?
- Justification of major assumptions
 - Have assumptions key to the technical approach and the results been tested or otherwise justified?
- Treatment of uncertainties/sensitivities
 - Have significant uncertainties been characterized?
 - Have important sensitivities been identified?

Within the general category of documentation, the projects were evaluated in terms of the following measures:

- Clarity of presentation
- Identification of major assumptions

In this report, the ACRS presents the results of its assessment of the quality of the research projects associated with:

- NUREG/CR-7237: Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)
- NUREG-2218: An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)

These projects were selected from a list of candidate projects suggested by RES.

The methodology for developing the quantitative metrics (numerical grades) for evaluating the quality of NRC research projects is presented in Section 2 of this report. The results of the assessment and ratings for the selected projects are discussed in Section 3.

2 METHODOLOGY FOR EVALUATING THE QUALITY OF RESEARCH PROJECTS

To guide its review of research projects, the ACRS has adopted an analytic/deliberative methodology [15-16]. The analytical part utilizes methods of multi-attribute utility theory [17-18] to structure the objectives of the review and develop numerical scales for rating the project with respect to each objective. The objectives were developed in a hierarchical manner (in the form of a "value tree"), and weights reflecting their relative importance were developed. The value tree and the relative weights developed by the Full Committee are shown in Figure 1.

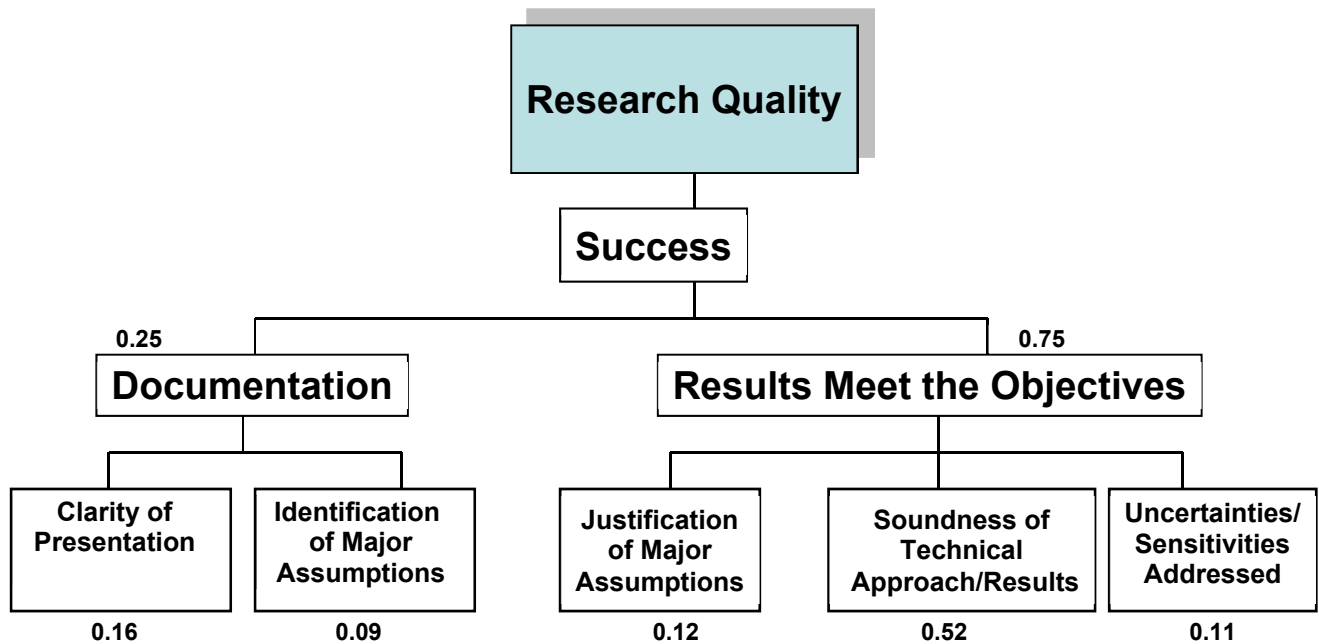


Figure 1. The Value Tree used for Evaluating the Quality of Research Projects

The quality of projects is evaluated in terms of the degree to which the results meet the objectives of the research and of the adequacy of the documentation of the research. It is the consensus of the ACRS that meeting the objectives of the research should have a weight of 0.75 in the overall evaluation of the research project. Adequacy of the documentation was assigned a weight of 0.25. Within these two broad categories, research projects were evaluated in terms of subsidiary "performance measures":

- Justification of major assumptions (weight: 0.12)
- Soundness of the technical approach and reliability of results (weight: 0.52)
- Treatment of uncertainties and characterization of sensitivities (weight: 0.11)

Documentation of the research was evaluated in terms of the following performance measures:

- Clarity of presentation (weight: 0.16)
- Identification of major assumptions (weight: 0.09)

To evaluate how well the research project was performed with respect to each performance measure, constructed scales were developed as shown in Table 1. The starting point is a rating of 5, Satisfactory (professional work that satisfies the research objectives). Often in evaluations of this nature, a grade that is less than excellent is interpreted as pejorative. In this ACRS evaluation, a grade of 5 should be interpreted literally as satisfactory. Although innovation and excellent work are to be encouraged, the ACRS realizes that time and cost place constraints on innovation. Furthermore, research projects are constrained by the work scope that has been agreed upon. The score was, then, increased or decreased according to the attributes shown in the table. The overall score of the project was produced by multiplying each score by the corresponding weight of the performance measure and adding all the weighted scores.

As discussed in Section 1, a panel of three to four ACRS members was formed to review each selected research project. Each member of the review panel independently evaluated the project in terms of the performance measures shown in the value tree. The panel deliberated the assigned scores and developed a consensus score, which was not necessarily the arithmetic average of individual scores. The panel's consensus score was discussed by the Full Committee and adjusted in response to ACRS members' comments. The final consensus scores were multiplied by the appropriate weights, the weighted scores of all the categories were summed, and an overall score for the project was produced. A set of comments justifying the ratings was also produced.

Table 1. Constructed Scales for the Performance Measures

SCORE	RANKING	INTERPRETATION
10	Outstanding	Creative and uniformly excellent
8	Excellent	Important elements of innovation or insight
5	Satisfactory	Professional work that satisfies research objectives
3	Marginal	Some deficiencies identified; marginally satisfies research objectives
0	Unacceptable	Results do not satisfy the objectives or are not reliable

3. RESULTS OF QUALITY ASSESSMENT

3.1 Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)

Introduction

The NRC regulations require that nuclear power plant structures, systems, and components (SSCs) important to safety be designed to withstand the effects of natural phenomena (such as earthquakes) without loss of capability to perform their safety functions. A common technique to increase the reliability of nuclear power plants is to increase redundancy, i.e., installing backup equipment to accomplish the safety function when the primary equipment fails. However, the redundancy may not be as effective in the case of “common cause initiators” such as earthquakes because a strong enough earthquake can simultaneously damage redundant pieces of equipment.

Probabilistic risk assessment (PRA) has been increasingly used by the NRC in all regulatory matters. This includes analyzing accident sequences initiated by earthquakes in the seismic PRA (SPRA).

Although seismic PRA is a mature analysis methodology, the treatment of dependencies or correlations in the seismic capacities of SSCs and in their responses to earthquakes continues to be a source of concern. Traditionally, the seismic impact on redundant equipment has been addressed in SPRA by a set of simplified “rule-based” assumptions: the total dependency, or the “if one fails, all fail”, rule is typically applied to all co-located similar redundant SSCs (e.g., all similar equipment located on the same building floor), the total independency, or the “zero dependency”, rule is typically applied to all non-co-located equipment, diverse or similar.

Experts have long questioned the validity of either the “if one fails, all fail” or “zero dependency” rule. These simplified approaches could lead to different degrees of conservatism or non-conservatism, in the risk results and could impact important risk insights. This is a reason that several technical approaches have been proposed and used to calculate the seismic correlation factors for redundant equipment.

RES sponsored a study at the Lawrence Berkeley National Laboratory (LBNL) to explore the correlation-dependency issue and, if feasible, to propose a more realistic approach. LBNL’s in-house expertise was supplemented by a team of highly experienced outside experts. Through workshops, that team, in turn, sought and received input, review, and advice from a larger group of SPRA experts. Four tasks were performed as directed by RES:

- Review SPRAs in the literature to understand the impact of correlation assumptions on risk estimates
- Review existing literature on seismic correlation analysis approaches
- Review existing data from earthquake experience and shake table tests for their usefulness in developing correlation factors
- With the help of invited experts in a series of workshops, recommend a methodology that better addresses correlation issues

The results of this study have been documented in NUREG/CR-7232 "Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components) [19].

General Observations

The following valuable tasks were accomplished as part of this project:

- The project team searched the existing literature on seismic correlation and dependency analysis, to understand the various methods that have been used in the industry. The most common current practice for treating correlations and dependencies in the seismic PRAs was reviewed.
- A review was done of about 10 SPRAs in the literature, to ascertain how correlations and dependencies were dealt with in each of them. Nearly all these SPRAs used the standard thumb rule (either 100% or 0% dependency) assumptions. The project team performed sensitivity studies to demonstrate the thumb rule approach may not be appropriate, especially for a few categories of SSCs.
- The project team also reviewed the existing earthquake-experience data base and the existing shake-table test data to see if they can be used to support a better approach for understanding and quantifying dependencies. The team concluded that these data are inadequate for defining dependency factors.
- The project team recommends adopting the "separation of independent and common variables" methodology (also referred to as the Reed-McCann methodology) for treating the dependency between component failures. This methodology requires the analyst to develop the fragility curves for the joint failure of components (cut sets) based on what are seen to be common and independent variabilities among these components. Once the methodologies were identified, the effort extended to obtain a consensus among field practitioners during a series of workshops.

The research team performed the required tasks competently. However, the value of this research project could benefit significantly if more effort was directed towards evaluating the impact of the correlation assumption on the SPRA results and insights. Without such evaluation and supporting data, it is difficult to conclude that the seismic risk evaluations would positively benefit from the new proposed methodology. The small percentage changes in the SPRA numerical results shown in the report do not seem sufficient to justify a time consuming and costly new methodology.

It is important to differentiate between the soundness of the methodology that ultimately was recommended by the expert panel and the soundness of the process used to obtain the consensus methodology. Although our review focuses on the process that followed RES guidance, we also considered the practicality/feasibility of implementing the recommended methodology.

Evaluation Scores for the Project

The consensus scores for this project are shown in Table 2. The score for the overall assessment of this work was evaluated to be 5.0 (satisfactory, a professional work that satisfies research objectives).

Table 2. Summary Results of ACRS Assessment of the Quality of the Project, "Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)"

Performance Measures	Consensus Scores	Weights	Weighted Scores
Clarity of presentation	6.0	0.16	0.96
Identification of major assumptions	5.0	0.09	0.45
Justification of major assumptions	4.0	0.12	0.48
Soundness of technical approach/results	5.0	0.52	2.60
Treatment of uncertainties/sensitivities	5.0	0.11	0.55
Overall Score:			5.0

Comments and conclusions within the evaluation categories are provided below.

Clarity of Presentation (Consensus Score: 6.0)

NUREG/CR-7237 is well written. The report clearly communicates the purpose of the project, the scope of the project, and the technical approach used, and the conclusions. The report documents the research project in sufficient detail to allow the reader to follow the work without having to refer to the original sources. The existing SPRA methodologies that address correlations are clearly presented and the rationale for selecting the recommended methodology is clearly articulated.

The report contains some minor deficiencies, such as its organization, which does not necessarily follow the order of the tasks, and it would benefit from a task flow diagram. Similarly, the tables and figures associated with each chapter are not presented in a consistent order. The captions in the tables and figures should define all terms and provide units, especially in cases where it is essential for understanding the presented results.

Despite these deficiencies, this report is a well-written, high quality professional work. Therefore, “Clarity of Presentation” is evaluated as above “satisfactory.”

Identification of Major Assumptions (Consensus Score: 5.0)

“Major assumptions” in this project are not clearly identified. Hence, they were subject to the reviewer’s interpretation. Several assumptions that we identified in this report include:

- The authors of NUREG/CR-7237 assumed that the “Correlation Assumption” is important to risk results and risk insights from SPRAs. Our review panel identified this as the most important assumption of the project
- In NUREG/CR-7237, the “Thumb Rule” assumption on dependency of similar components (100% or 0%) is deemed unsatisfactory and needs to be improved
- “Assumptions on Design, Qualification, and Installation of SSCs”, necessary for the Reed-McCann methodology (separation of independent and common variables), are evaluated in the report as important sources of uncertainty associated with the proposed methodology.

In summary, assumptions are identified as necessary through the report. Hence, the “Identification of Major Assumptions” category is evaluated as “satisfactory.”

Justification of Major Assumptions (Consensus Score: 4.0)

The project’s most important assumption, that the correlation assumption is important for SPRA results, was not well justified in the report.

It is stated multiple times in Section 4.3.1 of NUREG/CR-7237 that the correlation assumption **may not** significantly impact the SPRA results, but it could impact risk insights (without specifying or presenting examples of the risk insights that could be impacted). It stated that for some SPRAs, the difference in seismic CDF (based on how the dependency assumption was made) could be nearly a factor of two, but more typically a difference of 30% to 60% in overall seismic CDF. For some key accident sequences, the difference could be nearly a factor of two to four in the frequency, but the contribution to total CDF may be small. That may occur because in some SPRAs, the overall seismic CDF is dominated by an accident sequence consisting of a PRA “singleton” (a single failure), which would not be impacted by the correlation assumption. Given that the proposed method is likely to be time consuming and costly, the benefits of applying this new method are not clear.

Making a strong case is especially important, as the proposed approach would introduce new uncertainties that may limit the expert’s ability to identify sources of common variability for different components. The authors stated in Section 9.3: “There is the possibility that the more refined and insightful methodology will prove to be too difficult to use except in the hands of the most

experienced seismic PRA fragility analysts. There is no way to know now whether this will turn out to be true...., until the methodology has been applied several times by different analysts.”

Given that the authors did not make a strong case that the correlation assumption is important for SPRA results, “Justification of Major Assumptions” is evaluated as slightly above “Marginal.”

Soundness of Technical Approach/Results (Consensus Score: 5.0)

The technical approach of the project consisted of four major activities:

- 1) A review of several SPRAs to understand the impact that assumptions pertaining to correlations and dependencies have on risk estimates.
- 2) A review of existing literature on seismic correlation and dependency analysis, to understand the various methods used over the years, including the most common practice.
- 3) A review of existing data from earthquake experience and shake table tests to understand their usefulness to support the quantification of correlations and dependencies.
- 4) Recommendation by the project team, with the help of external experts, of a new analysis approach for treating dependencies in SPRA.

From reviews of existing SPRAs, seven categories of SSCs were found to dominate seismic risk contributors and were judged to have a high degree of potential correlation importance based on their numbers within the plant and their typical locations within the plant. A review of earthquake experience data and shake-table test data led to the conclusion that the data are inadequate for use in defining correlation factors for the selected categories of SSCs.

The project selected and examined four candidate methods for deriving the dependency between SSC fragilities: 1) the correlation coefficient method, 2) conditional probability of failure method, 3) split fraction method, and 4) separation of independent and common variables approach. The project team performed a thorough review of available methods by polling experts in the field in two workshops. The experts reached a consensus that the “Separation of Independent and Common Variables” method (Reed McCann method) is the most promising approach in modeling SSCs dependencies. Reed McCann method requires the analyst to develop the fragility curves for the joint failure of components based on what is judged to be common variabilities and independent variabilities among the components. The project team indicated that the fragility analyst should be well equipped to make this judgment assuming that he/she would have an intimate knowledge of how the components are designed, qualified, and installed.

From the above discussion it could be concluded that, although the recommended methodology appears to be promising, its implementation could be difficult because only the most experienced SPRA fragility analysts would be qualified to use it. Also, its reliance on expert judgment would introduce considerable analyst-to-analyst variability.

Thus, the project technical approach for evaluating seismic dependency is deemed adequate, and “Soundness of Technical Approach/Results” is evaluated as “satisfactory.”

Treatment of Uncertainties/Sensitivities (Consensus Score: 5.0)

In an application of the Reed and McCann method, the analyst deals directly with common variables and their epistemic uncertainty and aleatory variability. The report also describes how to obtain the uncertainty distribution associated with evaluated cut set frequencies by convolving the family of fragility curves with the family of seismic hazard curves.

However, because expert judgment is needed in the partitioning between the “independent” and the “dependent” variables, this method would introduce a new area of “model uncertainty”. In NUREG/CR-7237, each individual analyst is urged to try to identify how much uncertainty is associated with β -partitioning assignments and to do sensitivity studies on their effect.

Therefore, the project team has appropriately discussed uncertainty and sensitivity in this evaluation, and Treatment of Uncertainties/Sensitivities is evaluated as “satisfactory.”

3.2 An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)

The motivation for this report grew out of investigations into a series of events at operating nuclear power plants. These events generally announced themselves as explosions and fires, sometimes multiple fires in widely separated areas of the plant, and sometimes leaving the plant with very odd electrical alignments and unexpected operating configurations. Following investigation, it was found that the events involved two distinct phases: the first phase is a rapid energy release from high current arcs between electrical conductors – heat, light, and pressure – and the second phase involves ensuing fires in associated electrical equipment, oil from electrical transformers, nearby combustibles, and remote areas affected by momentarily high currents.

Review of U.S. and international operating experience revealed a significant number of high energy arc faults that have occurred in operating nuclear plants around the world. Approximately 10% of power plant fires were caused by high energy arc faults. Compared with other fires, high energy arc faults can create striking problems for electrical equipment – sudden very large currents that have overwhelmed protective features such as selective breaker coordination, simultaneous fires in multiple locations, and explosive damage, with burning oil spread over nearby areas.

The staff showed admirable creativity in organizing an international working group to continue the investigation of high energy arc faults, including a series of twenty-six full-scale experiments. The participating countries donated equipment used in the experiments. These exploratory experiments serendipitously exposed a much more energetic arc fault condition, not yet observed in actual power plant events: when aluminum is present in the vicinity of the high energy arc fault, the energy release can be multiplied dramatically – most monitoring equipment was destroyed during those experiments. Damage was much more substantial than existing models would predict. To support continuing experiments and analysis to develop a more thorough understanding of high energy arc faults and to allow appropriate modeling of these events, the working group suggested a second phase of experiments. A phenomena identification and ranking table (PIRT) was developed to provide a priority ordered list of phenomena to be investigated. That PIRT is the subject of the report [20] we reviewed.

General Observations

The idea of performing a PIRT to help set research priorities is sound. It appears that the research team applied it in a thorough manner. Some anomalies were noted and are described below. The prevalence of high energy arc faults, the severity of tests involving aluminum components, and the sometimes confusing plant conditions following these faults provide strong motivation for continuing research to develop reliable evaluation analysis tools.

We found the PIRT exercise was conducted in a satisfactory way to frame potential risk contribution in nuclear power plants from high energy arc fault events, as well as to characterize the experts' state of knowledge. The results of our evaluation are presented in Table 3.

Table 3. Summary Results of the ACRS Assessment of the Quality of the Project, “An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)” (NUREG-2218)

Performance Measures	Consensus Scores	Weights	Weighted Scores
Clarity of presentation	5	0.16	0.80
Identification of major assumptions	5	0.09	0.45
Justification of major assumptions	5	0.12	0.60
Soundness of technical approach/results	6	0.52	3.12
Treatment of uncertainties/sensitivities	3	0.11	0.33
Overall Score			5.3

Clarity of Presentation (Consensus Score = 5)

Parts of the report are very well written—Chapters 1, 2, and 4 are clear, and precise. Chapter 2 provides an excellent description of the PIRT process including direction to consider factors that identify and address uncertainty. Chapter 4 is a clear and ordered presentation of results at high level; however, no detailed ranking at the sub-phenomenon level was provided. Unfortunately, Chapter 3 is cryptic: it is no more than a collection of elicitation result tables taken from the appendices, with no text to explain what has been presented. The appendices too are poorly presented, with insufficient text to explain the data presented. Our score of 5 represents a compromise between the excellence of some chapters with the weak presentation in others. Also, a very interested reader can piece together an understanding of the appendices by cross-referencing other parts of the report making the objective, clarity of presentation, satisfactory.

We found that identification of “phenomena” as used in this report created some confusion between cause and effect for the reader and, perhaps, for the experts.

Identification of Major Assumptions (Consensus Score = 5)

The organizers of the PIRT process provided three generic scenarios for the experts to evaluate. The unstated assumption is that these three scenarios span the space of high energy arc fault phenomena. They do provide good descriptions of the bases of the phenomena and defend their importance.

The authors identify a number of other assumptions, although they are not actually called assumptions in the report. There are three unstated assumptions related to the rankings used in the study. The authors assume that the reader (and the experts) understand the meaning of each rank, although no statement of those meanings is provided. One ranking is "Unknown" and the hidden assumption is that phenomena ranked Unknown have no value; i.e., Unknown is valued less than a known "low value." Finally, they define a rank equation and assume it properly balances risk and state of knowledge over their ranges. It does have the nice property that rank goes up with increasing risk and decreases with decreasing state-of-knowledge.

Finally, the authors assume the experts can evaluate risk, with no specific guidance and no plant-specific PRA.

It is likely that some of these assumptions were discussed and clarified with the experts during the elicitation process. However, the report is silent about this possibility.

Justification of major assumptions (Consensus Score = 5)

Overall, the authors provided appropriate and useful justification of assumptions, leading to the satisfactory consensus score. However, there were gaps in the justification of assumptions. Treating the "Unknown" ranking as having no value is never justified and apparently not recognized. Likewise, the rank equation defined in Chapter 3 is not justified. Also, there is little explanation of why the three specific scenarios were selected and what issues could be left unaddressed, by limiting the scenarios.

Soundness of Technical Approach and Results (Consensus Score = 6)

If one mentally integrates the description of the methodology in Chapter 2, the coarse summary of results in Chapter 3, the detailed results in the Appendices, and the Conclusions and Recommendations of Chapter 4, it is possible to evaluate the overall soundness of the approach and results. The approach used in performing the elicitation from the six experts appears to be sound. It is well documented and produced a useful product for informing a roadmap moving forward on high energy arc fault research.

There are a number of issues associated with the proper role of the facilitator. We have chosen to evaluate all of these under the following section, although a number of them also affect several other criteria.

Uncertainties/Sensitivities Addressed (Consensus Score = 3)

The diverse background of the assembled panel provided a means for gaining different perspectives in addressing and ranking of important aspects of the three guiding scenarios. This should have enabled uncertainties and sensitivities to be identified and addressed objectively. The diversity of the experts, working essentially independently with the same data, precluded groupthink. This form of elicitation is effective because it enables objective assessment that account for uncertainties and sensitivities.

We question the “completeness” of the three scenarios derived from actual events: are there other possibilities for high energy arc fault events not covered here?

Most of the problems we see in the results tabulation is that there is a lack of evidence of good facilitation. Although Chapter 2 thoroughly addresses uncertainty, it appears that no one forced the experts to identify the uncertainty in their own evaluations. It is also not clear how the experts were advised to base their importance ranking on risk. Chapter 2 gives lip service to seeking consensus, but no discussion of consensus building is provided and some of the results imply that little effort was spent trying to reconcile divergent rankings. In cases where rankings span the full range from Low to High there is no text that indicates experts discussed their rankings. The facilitator should investigate such cases and seek resolution. Very often diverse ranking results from either some experts having access to information not available to the others or to experts ranking somewhat different situations. As an example, during our evaluation of the report, there was one case where one person scored a particular objective as an eight and another scored it a two. When we defended our independent scoring, we found that we included different issues in our evaluation. More carefully defining which issues were to be considered under each objective, our scores coalesced to a narrower range and it was possible to reach a consensus score all evaluators accepted.

This last point deserves a more thorough explanation. Often, guidance provided for an expert elicitation process, such as a PIRT, provides little help on the important issues of bias and effective elicitation. This is true for all the guidance we have seen for the PIRT process and for our own evaluation process as described in Chapter 2. The literature review in Appendix A is provided to further explain our evaluation process and to provide the basis for our comments on facilitation.

4. REFERENCES

1. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects," November 18, 2004 (ML043240107).
2. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2005," November 5, 2005 (ML053110211).
3. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2006," October 17, 2006 (ML062900517).
4. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2007," October 19, 2007 (ML072890365).
5. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2008," October 22, 2008 (ML082890373).
6. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2009," September 16, 2009 (ML091940352).
7. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2010," November 15, 2010 (ML103140150).
8. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2011," September 19, 2011 (ML11311A264).
9. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2012," October 22, 2012 (ML12293A451).
10. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2013," November 21, 2013 (ML13323B189).
11. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2014," November 25, 2014 (ML14322A844).
12. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2015," November 17, 2015 (ML15287A332).
13. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2016," January 11, 2017 (ML17011A227).
14. Advisory Committee on Reactor Safeguards, "ACRS Assessment of the Quality of Selected NRC Research Projects - FY 2017," January 23, 2018 (ML18022A054).
15. National Research Council, *Understanding Risk: Informing Decisions in a Democratic Society*. National Academy Press, Washington, DC, 1996.
16. Apostolakis, G.E., and S.E. Pickett, "Deliberation: Integrating Analytical Results into Environmental Decisions Involving Multiple Stakeholders," *Risk Analysis*, 18:621-634, 1998.

17. Clemen, R., *Making Hard Decisions*, 2nd Edition, Duxbury Press, Belmont, CA, 1995.
18. Keeney, R.L., and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Wiley, New York, 1976.
19. U.S. Nuclear Regulatory Commission, NUREG/CR-7237, "Correlation of Seismic Performance in Similar SSCs (Structures, Systems, and Components)," December 2017 (ML17348A155).
20. U.S. Nuclear Regulatory Commission, NUREG-2218, "An International Phenomena Identification and Ranking Table (PIRT) Expert Elicitation Exercise for High Energy Arcing Faults (HEAFs)," January 2018 (ML18032A318).

APPENDIX A

COMMENTS ON PIRT ELICITATION PROCESS AND FACILITATION

Controls for Unintentional Bias

One of the most important concerns associated with the use of a consensus expert judgment process is that of unintentional bias. In the subjective process of developing probability distributions, strong controls are needed to prevent bias from distorting the results (i.e., to prevent results that don't reflect the team's state of knowledge). Perhaps the best approach is to thoroughly understand how unintended bias can occur. With that knowledge, the facilitator and team can guard against its influence in their deliberations. A number of issues need to be considered, as discussed briefly below.

A number of studies present substantial evidence that people [both naive analysts and subject matter (domain) experts] are not naturally good at estimating probability (including uncertainty in the form of probability distributions or variance) [A1-A3]. For example, Hogarth [A3] notes that psychologists conclude that man has only limited information processing capacity. This in turn implies that his perception of information is selective, that he must apply heuristics and cognitive simplification mechanisms, and that he processes information in a sequential fashion. These characteristics, in turn, often lead to a number of problems in assessing subjective probability. Evaluators often:

- ignore uncertainty (this is a simplification mechanism); uncertainty is uncomfortable and complicating, and beyond most people's training
- lack an understanding of the impact of sample size on uncertainty; domain experts often give more credit to their experience than it deserves (e.g., if they have not seen it happen in 20 years, they may assume it cannot happen or that it is much more unlikely than once in 20 years)
- lack an understanding or fail to think hard enough about independence and dependence
- have a need to structure the situation, which leads people to imagine patterns, even when there are none
- are fairly accurate at judging central tendency, especially the mode, but tend to significantly underestimate the range of uncertainty (e.g., in half the cases, people's estimates of the 98% intervals fail to include the true values)
- are influenced by beliefs of colleagues and by preconceptions and emotions
- rely on a number of heuristics to simplify the process of assessing probability distributions; some of these introduce bias into the assessment process

Examples of this last area include:

- **Representativeness.** People assess probabilities by the degree to which they view a known proposition as representative of a new one. Thus, stereotypes and snap judgments can influence their assessment. In addition, representativeness also ignores the prior probability [A4]; i.e., what their initial judgment of the probability of the new proposition would be, before considering the new evidence – in this case their assumption of the representativeness of the known proposition. Clearly the prior should have an impact on the posterior probability but basing our judgment on similarity alone ignores that point. This also implies that representativeness is insensitive to sample size (since they jump to a final conclusion, based on an assumption of similarity alone).
- **Availability.** People assess the probability of an event by the ease with which instances can be recalled. This availability of the information is confused with its occurrence rate. Several associated biases have been observed:
 - biases from the retrievability of instances – recency, familiarity, and salience
 - biases from the effectiveness of a search set – the mode of search may affect the ability to recall
 - biases of imaginability – the ease of constructing inferences is not always connected with the probability.
- **Anchoring and Adjustment.** People start with an initial value and adjust it to account for other factors affecting the analysis. The problem is that it appears to be difficult to make appropriate adjustments. It is easy to imagine being locked to one's initial estimate, but anchoring is much more sinister than that alone. A number of experiments have shown that even when the initial estimates are arbitrary, and represented as such to the participants, the effect is strong. Two groups are each told that a starting point has been picked randomly “so that you have an anchor from which to make your adjustments.” The one given the higher arbitrary starting point generates higher probability. One technique found to be helpful is to develop estimates for the upper and lower bounds before addressing most likely values.

Lest we agree prematurely that people are irretrievably poor at generating subjective estimates of probability, it is significant to realize that many applications have been successful. Hogarth [A3] points out that studies of experienced meteorologists have shown excellent agreement with actual facts. Thus, an understanding is needed of what techniques can help make good assessments. In addition, in his comments published with the Hogarth paper, Edwards observes that humans use tools in all tasks, and tools can help us do a very good job in the elicitation process.

Winkler and Murphy [A5] make a useful distinction between two kinds of expertise or ‘goodness’. ‘Substantive’ expertise refers to knowledge of the subject matter of concern. ‘Normative’ expertise is the ability to express opinions in probabilistic form. Hogarth [A3] points out that the subjects in most of the studies were neither substantive nor normative experts. A number of studies have shown that normative experts (whose domain knowledge is critical) can generate appropriate probability distributions, but that substantive experts require significant training and experience, or assistance (such as provided with a facilitator), to do well.

The Facilitator

A facilitator is a normative expert with the interpersonal skills to control the elicitation process and ensure that it puts all available information on the table, that the experts are fairly heard and not allowed to hide behind others.

By understanding how the inadequacies in probability estimation and biases occur, the information can be used to combat their influence. The inadequacies of individuals can be dealt with by selecting analysts with a variety of expertise and by facilitating the process, challenging participants to explain the basis for their judgments. A facilitator can directly address biases. For example, representativeness bias involves ignoring available information and replacing a careful evaluation of that information with quick conclusions based on an over-focus on part of the information or allowing irrelevant information to affect conclusions. The facilitator must challenge analysts, asking them to explain their opinions. The facilitator must use his own judgment to sense when an individual is not using all available information.

Moreover, by understanding the heuristics that people often use to develop subjective probability distributions and the biases that attend those techniques, that awareness can help experts and analysts avoid the same traps. Through understanding which framings for eliciting distributions cause problems, we can use those that work better. Because the facilitator is familiar with the potential biases, it is possible to test the group's ideas and push them in the right direction. The strategies presented below should be used either explicitly or implicitly through the questioning of the facilitator, as described in the Senior Seismic Hazard Analysis Committee (SSHAC) report [A6]. In addition, Tversky and Kahneman [A7] give many detailed examples useful for helping facilitators develop awareness of such useful aids. Some of the simplest and best aids include:

- constructing simple models of the maximum and minimum points of the distribution, avoiding focus on the central tendency until the end points are studied to avoid anchoring; test these models to examine the evidence supporting them rather than relying on opinion alone
- seeking consensus on the evidence considered by the analysis team [A8]
- testing distributions by asking if the assessor agrees it is equally likely for the real answer to lie between the 25th to 75th percentiles or outside them; or between the 40th to 60th percentiles and outside the 10th and 90th percentiles. (Sometimes these questions must be phrased in ways to avoid suggesting the answer.)
- establishing a strong facilitator who ensures each participant must individually put his evidence on the table and justify it [A6]. (The facilitator must use judgment on when to push the participants, rather than going through a long and tedious checklist.)
- being careful when assessing parameters that are not directly observable. (The distribution is supposed to reflect the analyst's evidence concerning a particular parameter. If the analyst has little direct experience with the parameter, it can be difficult to justify an informative prior distribution.)

References:

- A1 Mosleh, A., V.M. Bier, and G. Apostolakis, "A Critique of Current Practice for the Use of Expert Opinions in Probabilistic Risk Assessment," *Reliability Engineering and System Safety*, vol. 20, pp. 63-85, 1988.
- A2 Cooke, R.M., *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press, 1991.
- A3 Hogarth, R.M., "Cognitive Process and the Assessment of Subjective Probability Distributions," *Journal of the American Statistical Association*, 1975;70(350): 271–94, 1975.
- A4 Siu, N.O., and D.L. Kelly, "Bayesian Parameter Estimation in Probabilistic Risk Assessment," *Reliability Engineering and System Safety* 1998;62: 89–116, 1998.
- A5 Winkler, R.L., and A.H. Murphy, "'Good' Probability Assessors," *Journal of Applied Meteorology*, 7 (1978) 751-8, 1978.
- A6 Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. "Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis," *Risk Analysis* 1998;18(4): 463–9,1998.
- A7 Tversky, A., and D. Kahneman. "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185 (1974) 1124-31, 1974.
- A8 Bley, D.C., S. Kaplan, and D.H. Johnson. "The Strengths and Limitations of PSA: Where We Stand," *Reliability Engineering and Systems Safety* 1992;38(1/2): 326, 1992.