Independent External Peer Review of the Risk Policies and Methodologies for the U.S. Army Corps of Engineers, Bureau of Reclamation, and the Federal Energy Regulatory Commission

FINAL REPORT

15 September 2020

Authors
John W. France, PE, D.GE, D.WRE – Panel Chair
Gregory B. Baecher, PhD
Rudolf B. Jongejan, PhD
Shane G. McGrath, FIE Aust, CPEng, NER
Ali Mosleh, PhD

Report Prepared Under Contract To:
United States Army Corps of Engineers
Louisville District
600 Dr. Martin Luther King, Jr. Place
Louisville, Ky 40202-2239

Draft Report Submitted 27 May 2020
Final Report Submitted 15 September 2020

Prepared Jointly by Subconsultants for:

RJH Consultants, Inc.
Project Manager: A. Tom MacDougall, P.E.
9800 Mt. Pyramid Ct.
Englewood, CO 80112
United States
T: 303-225-4611

Contract No. W912QR-16-D-0018
Delivery Order No.: W912QR-19-F0309
Requisition No.: W22W9K91331824
RJH Project No. 19123

Wood Environment & Infrastructure Solutions, Inc.
Project Manager: Mary Knopf, P.E.
11003 Bluegrass Parkway
Suite 690
Louisville, Kentucky 40299
United States
T: 502-267-0700

Contract No. W912QR-16-D-0005
Delivery Order No.: W912QR-19-F0296
Requisition No.: W22W9K91331819
Wood Project No. 7382-19-3348

Cover Photos:
(Left) Folsom Main Dam, Folsom California. Photo by Tom MacDougall, RJH Consultants, Inc.
(Right) Kingsley Dam, Ogallala Nebraska. Photo by Central Nebraska Public Power and Irrigation District.

EXECUTIVE SUMMARY

A panel was assembled to provide an independent external peer review (IEPR) of dam safety risk policies and methodologies for three Federal agencies: the United States Army Corps of Engineers (USACE); US Department of the Interior, Bureau of Reclamation (Reclamation); and Federal Energy Regulatory Commission (FERC) (the Agencies). The Panel was selected to include professionals with a broad range of experience, perspectives, and expertise. It included academics and private sector consultants from the United States, the Netherlands, and Australia. Expertise included civil engineering, dam and levee engineering, dam and levee risk analysis, risk analysis practice in other industries, dam ownership, and dam safety regulation.

The Panel's objective was to evaluate the Agencies use of risk in their management and regulation of dams. The Panel was also to provide recommendations to improve the risk policies and methodologies of each agency. The Panel focused its evaluation on the following "Charge" questions:

1. Given each agency's missions, are there any governance, critical policies, or methodologies missing from the agencies?
2. Is the overall direction of each of the three agencies appropriate?
3. Is there anything the Panel would like the agencies to consider?

The charge instructions also encouraged the Panel to include observations beyond the scope of the charge questions to bring important issues to the attention of decision makers, and to identify any lessons learned in both the risk informed decision making (RIDM) process and / or design and construction.

The IEPR was undertaken in response to legislative direction from the United States Congress, at least in part stimulated by the 2017 Oroville Dam, California spillway incident. The Panel did not solely focus on the Agencies' practices, but also considered how those practices influence or could influence broader dam safety practices in the United States.

The review included the following tasks:

1. An initial kickoff meeting with the panel and agency representatives.
2. An initial review of key documents provided by each agency.
3. A briefing from the Agencies to the Panel.
4. Additional review and research into each agency's risk policies, methods, and practices.
5. Panel evaluations and discussions.
6. Preparation of a draft report.
7. A briefing of the Panel's draft findings and recommendations to the Agencies.
8. Responding to agency comments and updating the draft report.
9. Publishing this final report.

In 2018 according to the National Inventory of Dams (NID), there were more than 91,000 dams in the United States, about 17 percent of which are classified as high hazard potential dams. A dam classification of "high hazard" means that if the dam were to fail, unintentionally releasing its impounded water, the loss of at least one life is probable. Only about 4 percent of the dams in the NID are owned by the US federal government, with the remainder owned by state or local governments, utilities, or private entities. The largest ownership category by far is private entities at 70 percent. Although the percentage of federally-owned dams in the NID is small, federally-owned and federally regulated dams constitute a large portion of dams over 100 feet tall, with about 31 percent of those dams being federally-owned and about 19 percent being federally regulated: a total of 50 percent being either federally-owned or federally regulated.

Federal dam owners, such as USACE, Reclamation, and others are self-regulated; that is, the agencies are not subject to external regulation, but rather each agency establishes policies and procedures to manage their dams' safety in accordance with Congressional authorization and overall federal guidance. FERC regulates non-federally-owned dams that are part of projects that generate hydroelectric power. All other non-federally-owned dams are regulated by State dam safety programs[1].

Prior to the 1970s, dam safety was managed almost exclusively using defined standards. This led to binary conclusions regarding the safety of a dam as either safe or unsafe. The conclusion was based on a comparison of engineering calculations relative to a defined standard. If the calculation showed that the dam met the requirements, it was deemed "safe." Additionally, a dam whose failure might result in one fatality was typically held to the same safety requirements as a dam whose failure could result in thousands of fatalities. Following notable failure of a number of dams in the 1970s, the Federal Emergency Management Agency (FEMA) published guidelines for dam safety that encouraged Federal dam owners to use risk to manage dam safety.

Numerous entities, governments, and industries (i.e., chemical, petroleum, and nuclear) throughout the world have used risk analysis since the 1960s to inform decisions. Those that have used risk to inform safety decisions have found that it provides a rational and thorough process to manage uncertainty, communicate to decision makers, and support safety decisions. Risk is the combination of the probability of an event occurring and its consequence(s).

Applied to dam safety, risk analysis involves estimating:

1. The probability that a load is applied to a dam (e.g., water impounded by the dam, or a seismic event)
2. The probability that the dam does not perform adequately under the applied load, resulting in an unintended release of the impounded reservoir.
3. The magnitude of consequences, particularly the number of lives lost as a result of a dam failure.

Beginning in the mid-1980s, Reclamation was the first Federal agency to move toward considering risk in dam safety evaluations with their introduction of the performance parameters program. This later evolved into what is now called potential failure modes analysis (PFMA). Circa 2000, application of the PFMA methodology became much more widespread in dam safety practice after FERC added a requirement for completing PFMAs, as part of its 5-year independent consultant inspection program for dams under its jurisdiction.

In the 1990s, Reclamation evolved its risk methodology from PFMA to quantitative risk analysis (QRA), a numerical and more detailed method of risk estimation. Reclamation used the quantitative estimates of risks to support decisions concerning whether dam safety modifications are needed, and, if so, the severity, urgency and relative priority of addressing potential dam safety issues across their portfolio of different dams. USACE began to migrate to a risk-informed approach for their dams and levees in the early 2000s, an effort that accelerated following Hurricane Katrina in 2005 and its impacts on the New Orleans levee system. In 2015, FEMA published *Federal Guidelines for Dam Safety Risk Management (FEMA P-1025)*. In 2016, FERC published *Interim Guidance, Risk-Informed Decision Making (RIDM), Risk Guidelines for Dam Safety* to provide guidance for its licensees to apply risk analysis methods to dam evaluations. A number of risk analysis pilot projects are currently underway or planned by hydropower dam owners in response to the 2016 FERC guidance.

---

[1] Within the United States, all states, except for Alabama, have programs with authority to regulate the safety of dams.

As QRA became more common in Federal dam safety, guidelines were developed to help evaluate the results of risk analyses. Reclamation published *Interim Guidelines for Achieving Public Protection in Dam Safety Decision Making* in 1997 and subsequently updated those guidelines in 2003 and 2011. USACE published tolerable risk guidelines in *Policy and Procedures, Regulation No. 1110-2-1156*, a document originally published in 2011 and updated in 2014. Based on these Federal guidelines, the Agencies primarily evaluate risk based on two key metrics from the QRA method: the annualized probability of dam failure and annualized expected life loss. Additionally, USACE and FERC consider individual life safety flood risk, the probability distribution of potential life loss, economic risks, and environmental and other non-monetary risks.

Dam failures are low-probability high-consequence events. Public preferences concerning low-probability, high-consequence risks are difficult to characterize, and are influenced by factors such as dread, control, and equity. While decisions about the acceptability of dam safety risks are questions of public policy, the Agencies adopt the philosophy that life safety is paramount and dams need to be managed to have risks “as low as reasonably practicable” (ALARP).

Based on the Panel’s review, the directions of the Agencies’ risk management programs for dam safety are appropriate and sound, and the Agencies’ implementation of RIDM are consistent with federal guidance. However, the levels of maturity in the application of RIDM differ among the three agencies and moving forward with the risk programs there are issues that warrant attention in all three agencies.

Many of the following recommendations address important RIDM process details. Completing risk analyses and risk assessments as correctly as possible, communicating risks to decision makers clearly and consistently, and continually improving risk process are important to the credibility of the Agencies’ dam safety RIDM programs. The recommendations presented below are offered in the spirit of continuous improvement.

The findings and recommendations provided in the remainder of this section are grouped under the following three headings: Agency Policy, Agency Governance, and Agency Methodology. Only the recommendations are listed below. The bases for the findings and recommendations are presented in the body of the report and summarized in Section 8 of the report.

Agency Policy

The Panel recommends the following regarding Agency policies:

- USACE and Reclamation continue their use and development of RIDM in managing the safety of the dams they own and operate. The RIDM programs of these two agencies have improved prioritization of risk reduction efforts.
- FERC continue its movement toward the application of RIDM in regulating the dams under its purview. FERC will need to further develop RIDM staff capability and staff resources within its organization and develop or support development of increased RIDM capacity in the private sector.
- The Agencies clarify and broaden their use of the ALARP (risks reduced “As Low as Reasonably Practicable”) principle in dam safety decision making, taking into consideration whether the recent developments in Australian dam safety practice for applying the factors used in an ALARP judgment could be of benefit. While the federal and agency guidelines on dam safety risk management stress the importance of ALARP, the Panel did not see evidence of consistent and detailed ALARP application in dam safety decisions by the Agencies.
- The Agencies evaluate more formalized safety case approaches using ‘prevention, mitigation, and control’ methods employed in other hazardous industries, to help ensure all reasonably practicable measures have been taken to reduce risk for all potential failure mechanisms, including surveillance, maintenance and other activities that control dam safety. The Agencies

should consider building on the current practice of "making the dam safety case" (as documented in *Best Practices in Dam and Levee Safety Risk Analysis*) by considering the broader safety case approach as it is employed in several other major hazards industries.

- The Agencies recognize that the quantitative risk guidelines for life safety (USACE and FERC refer to the quantitative risk guidelines as *tolerable risk guidelines*) are best described as agency policy rather than public policy. The agencies have been thoughtful and diligent in developing their guidelines, by benchmarking to dam safety practices in other countries and to other daily risks experienced by the United States population, but the results cannot be considered public policy without a public or political process. In the Panel's opinion this underscores the importance of ALARP for agency decision making in dam safety.
- FERC explore opportunities to provide enhanced risk assessment competency among dam safety practitioners through its existing collaboration with the US Society on Dams, establishing a cooperative arrangement with the Association of State Dam Safety Officials, or potentially involving academic institutions to support or deliver training.
- The Agencies establish historical incident (precursor) analysis and data-driven performance assessment programs to improve their analytical models and calibrate subjectively estimated model parameters and probabilities.
- The Agencies continue to share their RIDM experiences and practices with each other and the broader dam safety community, through professional organizations and international partnerships.

Agency Governance

For the purposes of this report, the term "governance" refers to agency processes of high-level policy setting, management operationalization of the policy to deliver its objectives, and associated checking and feedback to ensure required management actions are completed or refined and policy adjusted, if necessary.

The governance processes implemented by the Agencies appear generally reasonable based on the high-level IEPR and the charge of this Panel. Although our review did not result in developing specific recommendations regarding agency governance, the Panel found that, in general, agency governance could be adjusted to better support continued improvement in public stakeholder communications and risk assessment training to consultants working for FERC licensees.

Agency Methodology

The Panel recommends the following regarding Agency methodologies:

- The Agencies link terms such as *uncertainty* and *confidence* to definitions commonly used in probability, statistics, and decision science. Clear definitions are essential for moving the methodology forward and for communicating among its users.
- The Agencies consider the follow revisions to the PFMA process:
    1. Broaden the scope of the PFMA to include consideration of non-reservoir-release events that have significant consequences.
    2. Standardize and better define the criteria for eliminating PFMs from detailed consideration.
- The agencies improve their methods for dealing with probabilistic dependence, specifically the dependence among failure modes (e.g. multiple seismic pier failures), and the dependence among the uncertainties on probabilities.
- Risk assessments include consideration of the potential impacts of human error; ageing; malfunctioning of monitoring, remote or automated electro-mechanical control equipment; and other events that may not result in dam failure but could lead to casualties, major environmental damage, or significant operational costs.

- Historical data be used to estimate failure rates in lieu of subjective probability, where possible. In practice, these statistical databases are sometimes inadequate, and development of such databases should be seen as both a research and a practical need for the industry.
- The Agencies collaborate in developing a set of practical, state-of-practice guidelines to improve the credibility and quality of probability estimates elicited from subject matter experts.
- The Agencies develop a graded human reliability analysis method for use in dam safety risk assessments, leveraging methods developed by other industries, particularly the nuclear power, petrochemical, and aviation sectors.
- The continued use and improvement of the Agencies' life loss estimation models, USACE's HEC-LifeSim and Reclamation Consequences Estimating Methodology (RCEM), with scrutiny for uncertainties in life loss estimates and the level of detail needed for decision-making.
- The Agencies clearly explain the differences between the two types of probability and life loss charts, f-$\overline{N}$ and F-N charts, in their guidelines, and develop clear protocols for consistently plotting the life loss charts. While the two risk charts look identical, they are fundamentally different. Decision-makers should be aware of these differences to avoid confusion and error.
- The Agencies review use of truncations within their societal risk guidelines and adjust if it is then considered necessary.
- The Agencies more clearly define their objectives related to risk communication and develop practical guidelines for achieving those objectives.
- The Agencies review their policies concerning public access to dam failure inundation (flood) maps and articulate their decisions. Without access to inundation maps, it is impossible for third parties to establish whether they have an interest in a dam's safety.

TABLE OF CONTENTS

SECTION **PAGE NO.**

APPENDICES

TABLES

FIGURES

1 INTRODUCTION

The Independent External Peer Review (IEPR) objective was to examine and comment on the adequacy and provide recommendations to improve dam safety risk policies and methodologies used by United States Army Corps of Engineers (USACE), US Department of the Interior, Bureau of Reclamation (Reclamation); and Federal Energy Regulatory Commission (FERC) (the Agencies).

1.1 Authorization

The USACE, Reclamation, and FERC dam safety risk policies and methodologies IEPR was authorized in two separate contracts: Contract No. W912QR-16-D-0018 / Delivery Order No. W912QR19F0309 to RJH Consultants, Inc. (RJH); and Contract No. W912QR-16-D-0005 / Delivery Order No. W912QR19F0296 to Wood Environment and Infrastructure Solutions, Inc. (Wood). The following five individuals (the Panel) were engaged to complete the review through these contracts:

John W. France, PE, D.GE, D.WRE - Panel Chair

Gregory B. Baecher, PhD

Rudolf B. Jongejan, PhD

Shane G. McGrath, FIE Aust, CPEng, NER

Ali Mosleh, PhD

The Panel was selected to provide a broad range of experience and expertise. It includes individuals from the United States, the Netherlands, and Australia, and includes academicians and private sector consultants. Expertise represented includes civil engineering, dam and levee engineering, dam and levee risk analysis, risk analysis practice in other industries, dam ownership, and dam safety regulation.

Resumes for the individual panel members are provided in Appendix A.

1.2 Purpose

The IEPR objective was to examine and comment on the adequacy of and provide recommendations to improve dam safety risk policies and methodologies used by USACE, Reclamation, and FERC, collectively referred to in this report as the Agencies. The Panel was asked to focus its review on the following “Charge” questions:

1. Given each agency’s missions, are there any governance, critical policies, or methodologies missing from the agencies?
2. Is the overall direction of each of the three agencies appropriate?
3. Is there anything the Panel would like the agencies to consider?

After discussion, the Panel divided the first charge question into Questions 2 and 3 and reordered the questions for clarity, resulting in the following four questions:

1. Is the overall direction of each of the three agencies appropriate?
2. Given each agency’s mission, are there any governance items missing from the Agencies?
3. Given each agency’s mission, are there any critical policies or methodologies missing from the Agencies?
4. Is there anything the panel would like the Agencies to consider?

The charge instructions also encouraged the Panel to include observations beyond the scope of the charge questions to bring important issues to decision makers’ attention, and to identify any lessons learned in the Risk Informed Decision Making (RIDM) process and / or design and construction.

The IEPR is being completed in response to legislative direction from the US Congress, at least in part in response to the 2017 Oroville Dam, California spillway incident. The language in the legislation was:

> The Corps, in cooperation with the Federal Energy Regulatory Commission and the Bureau of Reclamation, shall contract with an independent peer review organization to conduct a comprehensive Independent External Peer Review (IEPR) of risk-informed dam safety practices in these three federal agencies with the intent to inform improvements broadly in national dam safety practices. The Corps is directed to contract with an independent peer review organization in accordance with its current review policy and the National Academy of Science IEPR process. The IEPR shall also consider how dam safety practices are affected by human factors, as well as how risk informed analysis in other industries may be applicable to dam safety practices.

The Panel interprets this language as not being solely focused on the practices of the Agencies, but also on how those practices influence or could influence broader dam safety practices in the United States.

1.3 Report Organization

The remainder of this report is organized in the following sections:

Section 2 - Panel Methodology: A discussion of the methods and approach used by the Panel in completing the review.

Section 3 - Background: A discussion of the dam safety framework in the United States and the history of risk analysis and risk-informed decision making in the Agencies.

Section 4 - Agency Policy: The Panel’s comments on the overall direction of the Agencies’ programs and discussion of selected aspects of agency policy.

Section 5 - Agency Governance: A review of the dam safety program governance in each agency.

Section 6 - Agency Methodology Part 1: Risk Estimation: Discussions of selected aspects of the methodology used by the Agencies to estimate risks.

Section 7 - Agency Methodology Part 2: Risk Assessment and Communication: Discussions of selected aspects of how the Agencies evaluate estimated risks and communicate the risks.

Section 8 - Summary of Findings and Recommendations: A summary of the significant findings and recommendations from the Panel’s work.

The report sections above are supported, as judged appropriate, by appendices with more detailed information. An Executive Summary is provided at the beginning of the report.

2 PANEL METHODOLOGY

The methodology used for this review included:

1. Initial kickoff meeting
2. Initial documents review
3. Briefings from the three agencies
4. Further document review and research
5. Panel discussions
6. Draft report
7. Out brief to the Agencies
8. Final report.

Each of these items is discussed briefly below.

2.1 Kickoff Meeting

The kickoff meeting was held by teleconference, and it was attended by the Panel members, Agencies representatives, and RJH and Wood representatives. The study purpose and the scope were discussed, and the Panel members asked clarifying questions. The schedule for the agency briefings was also established.

2.2 Initial Documents Review

The Agencies provided documents related to their dam safety risk programs, as listed in Appendix B-1. The Panel completed an initial review of those documents in advance of the agency briefings. During the initial review effort, the Panel compiled a list of items that they would like to have addressed in the agency briefings and provided that list to the Agencies.

2.3 Briefings from the Three Agencies

An agency briefing meeting was held in Lakewood, Colorado, 16 through 20 December 2019. Each agency briefed the Panel on its dam safety program history and current application of risk analysis. The agency briefings addressed the items the Panel identified during its initial document review. The Panel asked questions of the Agencies representatives, and the discussions among the Panel and the Agencies were open and far ranging. The Panel, after the briefing, compiled a list of additional information requested for the review, and the Agencies subsequently provided the requested information, as listed in Appendix B-2.

2.4 Further Document Review and Research

The Panel completed further review of the initially-provided information as well as the additional information provided by the Agencies after the briefing. The Panel also reviewed and researched other documents related to risk analysis practices in other industries and other countries, as it judged appropriate. In addition, the Panel interviewed several individuals from the Agencies as part of this review.

2.5 Panel Discussions

The Panel conducted numerous internal discussions regarding the review throughout the process. Face-to-face discussions were held during the week of the agency briefing meeting, and all other Panel discussions were held by teleconference, Web meeting, or email. The Panel was, through these discussions, able to reach the consensus findings ultimately presented in this report.

2.6 Draft Report

The Panel compiled a draft of this report presenting findings from the review. The draft report was submitted to the Agencies for review and comment. The agency review was not intended to change Panel opinions and findings, but rather to provide an opportunity for the Agencies to provide clarifications or corrections for any misinterpretations or misunderstandings that the panel may have developed.

2.7 Outbrief to the Agencies

The final work item for this Task Order is an outbrief from the Panel to the Agencies management personnel in Washington, DC. At the time of report submission, the outbrief has not been scheduled. The current period of performance for this Task Order extends until November 30, 2020.

2.8 Final Report

The Panel compiled this final report of findings after resolving agency comments. The report contents represent the Panel's independent findings and have not been unduly influenced by the Agencies.

3 BACKGROUND: HISTORY OF DAM SAFETY RISK ANALYSIS AND RISK INFORMED DECISION MAKING IN THE UNITED STATES

According to the National Inventory of Dams (NID), managed by the USACE, there are more than 91,000 dams in the United States, as of the latest reporting period in 2018. A dam must meet at least one of the following four criteria to qualify for inclusion in the inventory:

1. **High hazard potential -** loss of human life is likely if the dam fails.
2. **Significant hazard potential -** no probable loss of human life, but dam failure can cause economic loss, environmental damage, lifeline facilities disruption, or impact other concerns.
3. **Equal or exceed 25 feet in height and exceed 15 acre-feet in storage**.
4. **Equal or exceed 50 acre-feet storage and exceed 6 feet in height**.

About 17 percent of the dams in the NID are classified as high hazard potential dams.

Only about 4 percent of the dams in the NID are owned by the US federal government, with the remainder owned by state or local governments, utilities, or private entities. The largest ownership category by far is private entities at 70 percent. Although the percentage of federally-owned dams in the NID is small, federally-owned and federally regulated dams constitute a large portion of dams over 100 feet tall, with about 31 percent of those dams being federally-owned and about 19 percent being federally regulated: a total of 50 percent being either federally-owned or federally regulated. Federal dam owners, such as USACE, Reclamation, US Department of the Interior Bureau of Indian Affairs, US Fish and Wildlife Service, and US National Park Service are self-regulated. The federal dam-owning agencies are not subject to external regulation, but rather each agency establishes policies and procedures to manage their dams' safety in accordance with Congressional authorization and overall federal guidance (Federal Emergency Management Agency [FEMA], 1979). States have their own dam safety programs, and at present all states except Alabama have dam safety programs in place. The state programs vary widely regarding legislative authorities and resources, both workforce and financial. All of the non-federally-owned dams fall under the state programs jurisdiction.

FERC regulates non-federally-owned dams included in hydroelectric power projects. These dams would normally fall within the state dam safety programs jurisdiction. Some states cede the hydropower dams safety regulation to FERC and do not actively regulate them. Other states assert hydropower dams regulation in parallel with FERC.

3.1 History of Dam Safety Risk Analysis at US Federal Agencies

Dam safety risk analysis at US federal agencies has its roots in work by Von Thun (Bureau of Reclamation, 1985) following the Teton Dam failure in 1976. This work advocated:

> an assessment of the historical record of [dam] failure […] as a function of a number of parameters (type of dam, size of dam, date of construction, location, etc.) rather than applying a single estimate of the rate of dam failure for all types, heights, and ages of dams throughout recent history.

This was intended to replace the aggregate rate of modern dam failure of between 10^{-4} to 10^{-5} per annum with something more granular. The purpose was "to determine the number of dam failures that have occurred for each category so that a rate of failure could be determined." This was the seminal paper at the beginning of dam safety risk analysis at Reclamation. USACE still relied on a standards-based approach at the time, and would not begin to apply risk analysis in its dam safety program until the early to mid-2000s.

3.1.1 Standards-based approach

Prior to development of risk-informed approaches, dam safety engineering practice at the federal agencies and across United States practice focused on evaluating dams through visual inspections and comparison of analysis results with standards-based criteria.

Visual inspections would be conducted periodically, with dam safety professionals examining a dam and its appurtenant structures for signs of potential distress, e.g., unexpected seepage, settlement, deformation, or structural deterioration. Analyses would be completed to evaluate the dam and appurtenant structures for various loading conditions: normal operations and normal pool loading, flood loading, earthquake loading, etc. Data on design, composition, and construction of dams constructed before the modern dam safety era (before about 1980) were limited or even non-existent. Analyses would be completed in such cases based on conservative estimates of material property and dam configurations. Further investigation and data collection would be completed if these initial analyses indicated concern, and the analyses would be refined. However the analyses were completed, the results would be compared to deterministic criteria. Some representative examples of such criteria were:

- Spillway capacity would be compared to the ability to safely pass an inflow design flood, which typically ranged from the 100-year flood to the probable maximum flood (PMF), depending on the hazard (potential downstream consequences) posed by the dam.
- Calculated stability factors of safety would be compared to recommended or required minimum factors of safety, for example 1.3 for end of construction, 1.5 for normal operation, 1.2 or 1.3 for rapid drawdown, and 1.0 to 1.2 for earthquake loading. Different organizations established different factor of safety criteria for the various loadings.
- Stresses in structures were compared to allowable or ultimate strengths of the materials composing the structure.

Although there was general consensus in the profession, with some variability, regarding criteria for spillway capacity, stability factors of safety, and structural stresses, there was significant divergence concerning criteria for seepage and internal erosion. There were differing opinions as to whether specific seepage gradients could be judged as safe. Further, although many practitioners held a view that internal erosion was often episodic and thought observation of clear seepage was not proof that internal erosion was not occurring, this view was by no means universal. As a result, conclusions concerning spillway capacity, stability, and stresses were based on definitive numbers, while conclusions concerning seepage and internal erosion were based principally on qualitative professional judgments.

3.1.2 Potential failure modes analysis

Following Von Thun’s innovations (Bureau of Reclamation, 1985), the Agencies evolved toward what is now called potential failure modes analysis (PFMA). This was a modification of failure modes and effects analysis common in reliability engineering and dating to post World War II military standards (U.S. Department of Defense, 1949). As early as 1972 (Dyer et al., 1972), National Aeronautics and Space Administration (NASA) had recommended to US Geological Survey (USGS) that this methodology might be useful in regulating safety in the offshore oil and gas industry. PFMA changed the basic thought process in dam safety engineering from one of evaluating dams against a set of deterministic standards to one of assessing the ways a dam could fail. Probabilities of the different failure mechanisms and their associated consequences would be added.

The steps in the process include:

- Assemble and review available information about the dam, including design and construction records, performance records, instrumentation data, analyses, and photographs (including construction photographs).
- Compile a list of possible ways the dam could fail, known as potential failure modes (PFMs). This list should be compiled initially without consideration of the probability of failure for each failure mode. Note that in reliability engineering these would be called failure mechanisms or failure causes.
- Screen the PFMs to identify which are credible or most significant and document the reasons why some PFMs are not considered further.
- For those PFMs considered further, 1) compile lists of factors making the PFM more likely and factors making the PFM less likely, 2) identify surveillance and instrumentation methods that can be used to detect PFM initiation and progression, 3) identify measures that could reduce the PFM probability, and 4) identify missing data or analyses required to evaluate the PFM probability.
- Compile a list of findings and understandings that came to light during the process.

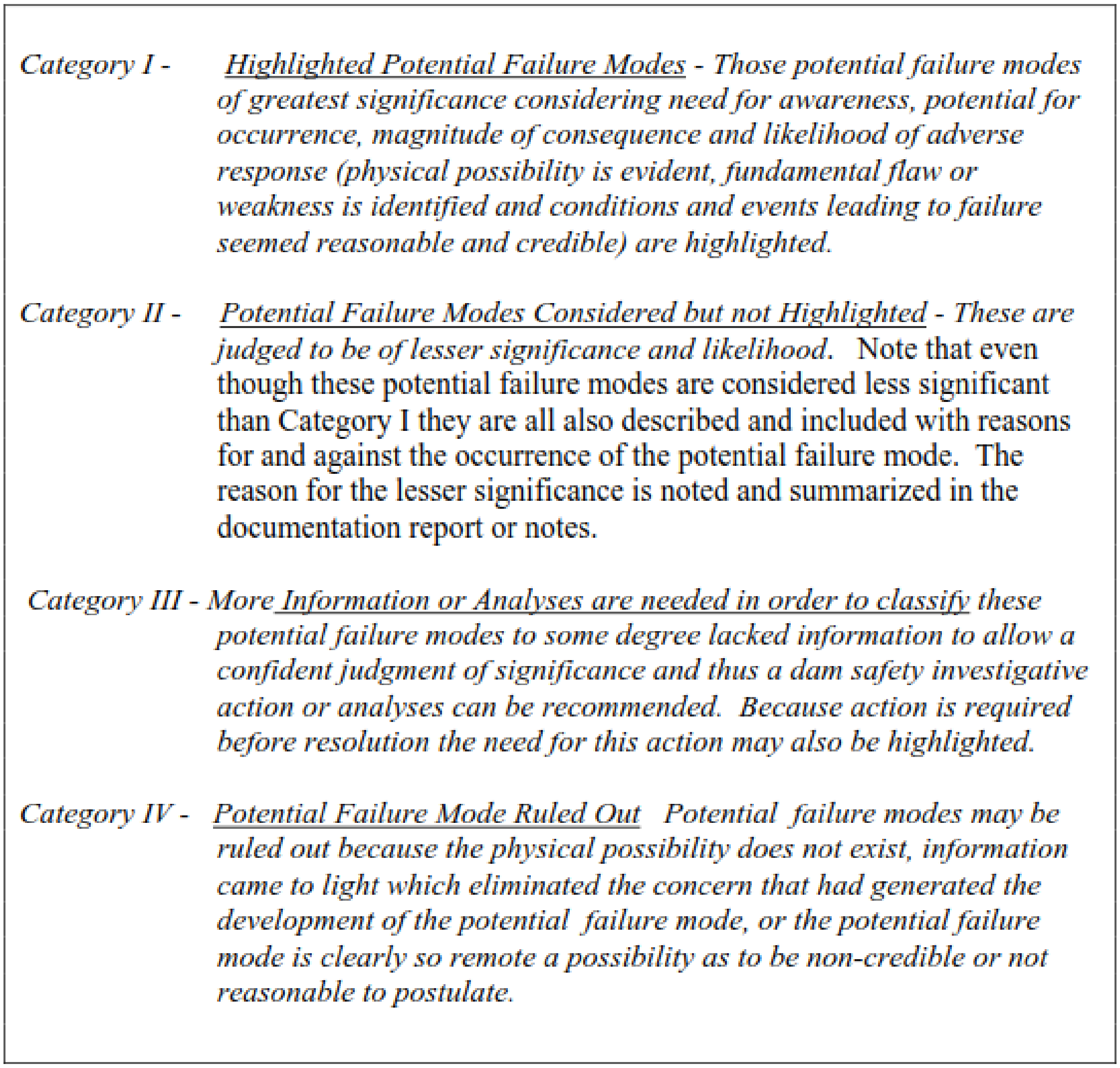
Application of the methodology became significantly more widespread after FERC added a requirement for PFMAs as part of its 5-year independent consultant inspection program for hydropower dams within its regulatory jurisdiction, circa 2000 (FERC, 2003). The FERC PFMA process results in each failure mechanisms being assigned to one of the categories described in Figure 1.

As Reclamation, and later USACE, implemented more explicit risk estimation methods, as described below, the PFMA has remained a key component in any risk analysis, typically as the first task in a quantitative or semi-quantitative risk analysis.

3.1.3 Risk analysis

Reclamation, beginning in the 1990s, evolved its methodology from PFMA to Quantitative Risk Analysis (QRA). Quantitative estimates of risks provide input into Reclamation's decisions concerning whether dam safety modifications are needed, and, if so, the urgency of addressing dam safety concerns and the relative priority of addressing concerns at different dams. USACE began to migrate to a risk-informed approach in the early 2000s, an effort that was accelerated following Hurricane Katrina with its impacts on the New Orleans levee system in 2005. FEMA published *Federal Guidelines for Dam Safety Risk Management (FEMA P-1025)* in 2015 (FEMA, 2015). This document continues to provide impetus to expand use of risk analysis in dam safety.

FERC published *Interim Guidance, Risk-Informed Decision Making (RIDM), Risk Guidelines for Dam Safety* in 2016 (FERC, 2016), to provide guidance for its licensees to apply risk analysis approaches. A number of risk pilot projects are underway or planned by hydropower dam owners in response to the new FERC guidance.



Category I - Highlighted Potential Failure Modes - Those potential failure modes of greatest significance considering need for awareness, potential for occurrence, magnitude of consequence and likelihood of adverse response (physical possibility is evident, fundamental flaw or weakness is identified and conditions and events leading to failure seemed reasonable and credible) are highlighted.

Category II - Potential Failure Modes Considered but not Highlighted - These are judged to be of lesser significance and likelihood. Note that even though these potential failure modes are considered less significant than Category I they are all also described and included with reasons for and against the occurrence of the potential failure mode. The reason for the lesser significance is noted and summarized in the documentation report or notes.

Category III - More Information or Analyses are needed in order to classify these potential failure modes to some degree lacked information to allow a confident judgment of significance and thus a dam safety investigative action or analyses can be recommended. Because action is required before resolution the need for this action may also be highlighted.

Category IV - Potential Failure Mode Ruled Out Potential failure modes may be ruled out because the physical possibility does not exist, information came to light which eliminated the concern that had generated the development of the potential failure mode, or the potential failure mode is clearly so remote a possibility as to be non-credible or not reasonable to postulate.

Figure 1. FERC PFM Classifications (FERC, 2017)

It became apparent, as risk analyses evolved into more common use in dam safety evaluations, that guidance was required to assist dam safety practitioners in applying these emerging methods. Reclamation developed a set of documents and presentations called, *Best Practices in Dam Safety Risk Analysis* in the early 2000s. Later, after USACE began to apply risk analysis to its dam and levee safety programs, Reclamation and USACE worked together to revise and update this document, which is now titled *Best Practices in Dam and Levee Safety Risk Analysis* (Bureau of Reclamation & USACE, 2019) (Best Practices). The current version contains eight parts, addressing a range of topics related to dam and levee safety risk analysis. Recognizing that dam and levee safety risk analysis practice would be evolving quickly, Reclamation and USACE chose not to publish Best Practices as a printed document, but to make it available on the internet (https://www.usbr.gov/ssle/damsafety/risk/methodology.html).

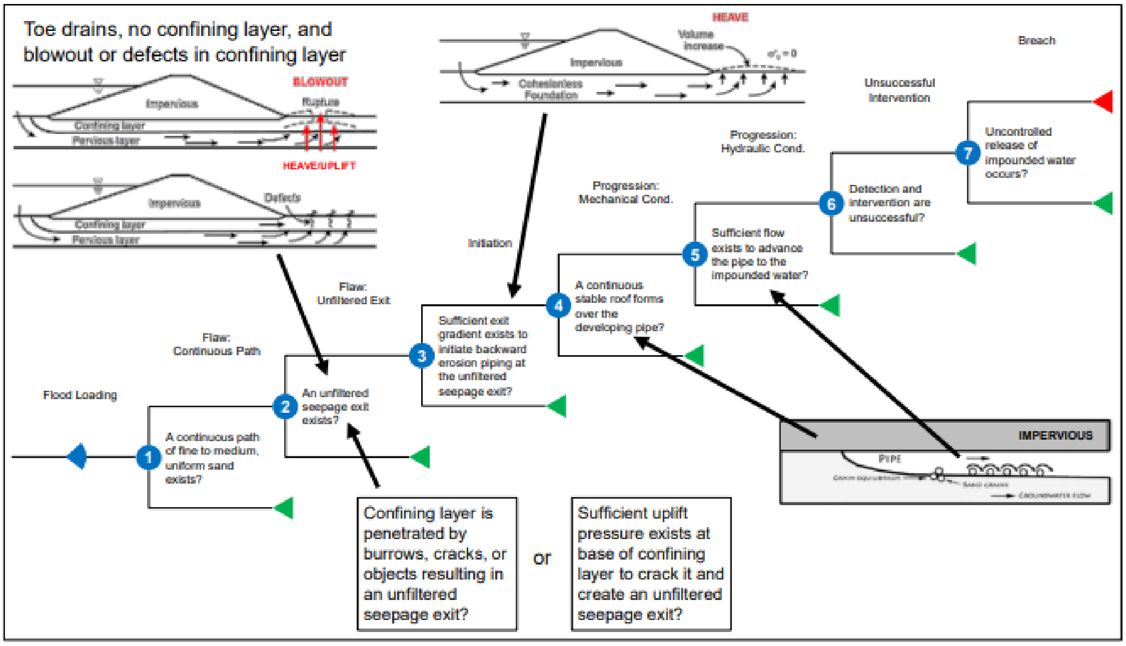
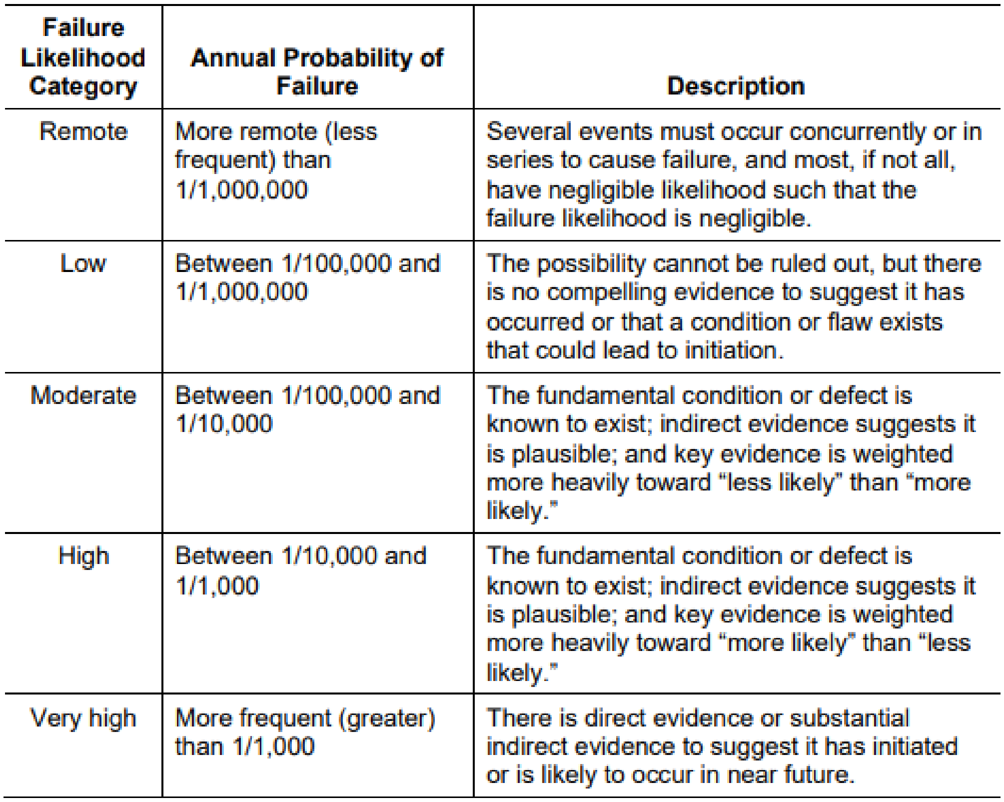


Figure 2. Example Internal Erosion Event Tree (Bureau of Reclamation/USACE, 2019)

QRAs are completed using event trees, an example of which is given in Figure 2. The probabilities for the individual events are estimated based on statistical data or expert opinion elicitation (degree-of-belief estimates or subjective probability estimates by a group of subject matter experts). Detailed quantitative risk analyses are typically used to support decisions to complete more detailed investigations or to implement risk reduction measures. Such analyses have also been used to evaluate risk reduction effectiveness for dam modification alternatives and to evaluate risk during construction of a dam safety modification.

Best Practices also contains guidance for semi-quantitative risk analyses (SQRAs). Subjectively-assigned probabilities and consequence categories are used in place of detailed quantitative estimates of probabilities of failure and consequences in this approach. Examples of these categories are shown on Figure 3 and Figure 4. SQRAs are no longer used often in the USACE and Reclamation program, but they are a part of the FERC guidelines and they are used by other organizations.

SQRAs are sometimes used for portfolio risk analyses for a group of dams as a prioritization tool, to determine which dams or PFMs should be addressed first. However, full quantitative risk analyses can also be used for portfolio risk analyses. SQRAs are typically less expensive and quicker, but do not provide a sound quantitative value for comparison to published risk guidelines (described below). Instead, SQRAs provide a value that can be used as a relative comparison among a set of dams and a general indication of the level of risk a dam or PFM poses.



Failure Likelihood Category	Annual Probability of Failure	Description
Remote	More remote (less frequent) than 1/1,000,000	Several events must occur concurrently or in series to cause failure, and most, if not all, have negligible likelihood such that the failure likelihood is negligible.
Low	Between 1/100,000 and 1/1,000,000	The possibility cannot be ruled out, but there is no compelling evidence to suggest it has occurred or that a condition or flaw exists that could lead to initiation.
Moderate	Between 1/100,000 and 1/10,000	The fundamental condition or defect is known to exist; indirect evidence suggests it is plausible; and key evidence is weighted more heavily toward "less likely" than "more likely."
High	Between 1/10,000 and 1/1,000	The fundamental condition or defect is known to exist; indirect evidence suggests it is plausible; and key evidence is weighted more heavily toward "more likely" than "less likely."
Very high	More frequent (greater) than 1/1,000	There is direct evidence or substantial indirect evidence to suggest it has initiated or is likely to occur in near future.

Figure 3. Likelihood Categories for Dams for SQRA (Bureau of Reclamation & USACE, 2019)

3.1.4 Risk evaluation guidelines

It became apparent, as QRA became more common in dam safety, that guidelines were needed to help evaluate the analyses results. The first such guidelines to be published in the United States were *Interim Guidelines for Achieving Public Protection in Dam Safety Decision Making* (Bureau of Reclamation, 1997). This document was subsequently finalized as *Guidelines for Achieving Public Protection in Dam Safety Decision Making* (Bureau of Reclamation, 2003). Reclamation, after almost a decade of additional experience with risk analysis, updated its guidelines with publication of *Interim Dam Safety Public Protection Guidelines, A Risk Framework to Support Dam Safety Decision Making* (Bureau of Reclamation, 2011b) and a companion document, *Rationale Used to Develop Reclamation's Interim Dam Safety Public Protection Guidelines* (Bureau of Reclamation, 2011c).

Subsequently, USACE published tolerable risk guidelines in *Safety of Dams - Policy and Procedures, Regulation No. 1110-2-1156* (USACE, 2014b), a document originally published in 2011 and updated in 2014. Part of the USACE tolerable risk guidelines are illustrated on Figure 5.

Reclamation's guidelines primarily consist of two measures: annualized failure probability and average annualized life loss (Bureau of Reclamation, 2011b). USACE and FERC also consider these two measures, plus individual incremental life safety flood risk, probability distribution of potential life loss due to dam breach, and consideration of economic risks, and environmental impacts, and other non-monetary risks (FERC, 2016; USACE, 2014b).

Level	Life Loss	Economic Loss
1	Average life loss is less than 1. Although life-threatening flooding occurs, direct loss of life is unlikely due to severity or location of the flooding or effective warning and evacuation.	Average economic loss is less than $10 million. Limited property and/or environmental damage is likely.
2	Average life loss is in the range of 1 to 10. Some direct loss of life is likely, related primarily to difficulties in warning and evacuating small population centers.	Average economic loss is in the range of $10 million to $100 million. Moderate property and/or environmental damage is likely.
3	Average life loss is in the range of 10 to 100. Large direct loss of life is likely, related primarily to difficulties in warning and evacuating small population centers or difficulties evacuating large population centers with significant warning time.	Average economic loss is in the range of $100 million to $1 billion. Significant property and/or environmental damage is likely.
4	Average life loss is in the range of 100 to 1,000. Extensive direct loss of life can be expected due to limited warning for large population centers and/or limited evacuation routes.	Average economic loss is in the range of $1 billion to $10 billion. Extensive property and/or environmental damage is likely.
5	Average life loss is greater than 1,000. Extremely high direct loss of life can be expected due to limited warning for very large population centers and/or limited evacuation routes.	Average economic loss is greater than $10 billion. Extremely high property and/or environmental damage is likely.

Figure 4. Consequence Categories for SQRA (Bureau of Reclamation & USACE, 2019: Table A-4.2)

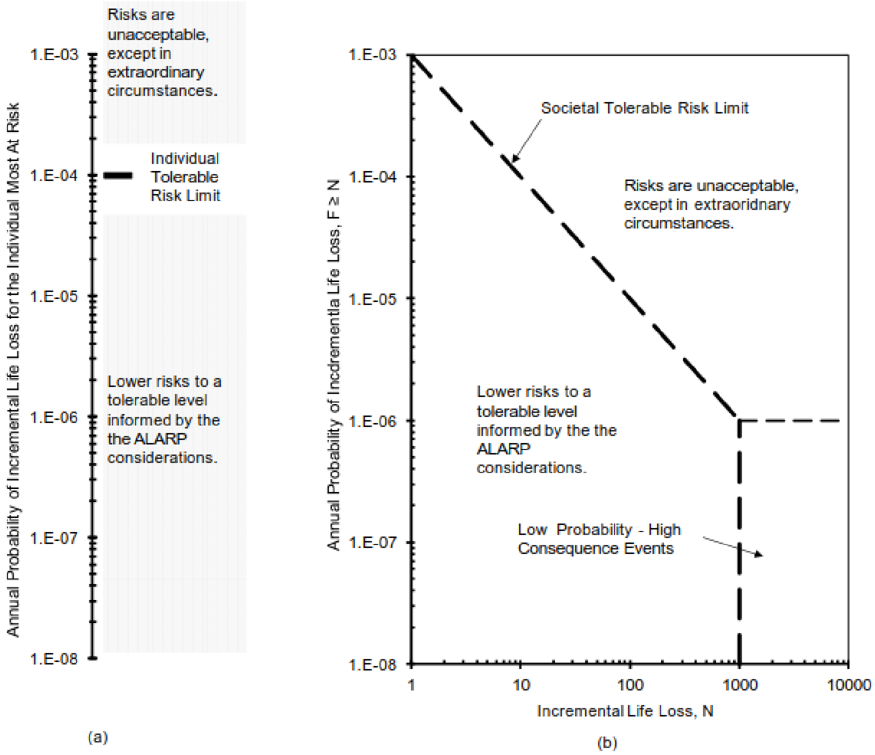


Figure 5. USACE Tolerable Risk Guidelines (USACE, 2014)

Although Reclamation, USACE, and FERC have each published quantitative guidelines for evaluating risk, the Agencies recognize QRA results are imprecise, and, in fact, involve significant uncertainty. As such, the guidelines are not “bright lines in the sand,” but rather a tool to facilitate broader informed decision making. This approach to the quantitative risk estimates is described in the Preface to Best Practices as follows:

> The manual is reviewed and updated periodically to stay current with the state of the practice; however, this guidance is not a prescriptive approach. Procedures and data for quantitative dam and levee safety risk analysis do not provide accurate or precise numerical results. Risk analysis teams must balance calculations with judgment to estimate the risks and to build a case for the risk-driving influences. The numbers are less important than the identification, understanding, and documentation of the major risk contributors.

3.2 RIDM in Dam Safety Practice

Risk informed decision-making for dam safety is similar in important ways to risk analysis performed in other industries, most notably for other large infrastructure projects. However, it is unique - or nearly so - in important ways. Each of the following aspects of risk in dam safety can be found in the practice of other industries, but their combination sets dam safety somewhat apart.

3.2.1 Quantifying risks for unique technological systems

The unique nature of each dam (its foundations, structure, hydrologic and seismic setting) means that risks for dams cannot be quantified simply using failure rates and mean-time-to-failures for standardized components. While probability distributions for hydrologic hazards can be developed largely on the basis of historical data, this is difficult or impossible when it comes to evaluating dam capacity or failure consequences.

Probability estimates can be obtained from expert elicitation when statistical observations or adequate failure mechanism models are unavailable. The Agencies have developed their own expert elicitation procedures, centered on consensus-seeking dialogue under an experienced facilitator’s guidance.

3.2.2 Dependence on subjective probability

Many or most of the important uncertainties in dam safety risk analysis are epistemic. That is, they have to do with lack of information, knowledge, or understanding rather than the stochastic or aleatory frequencies commonly encountered in risk analyses in other industries. As a result, the probabilities in dam safety risk analysis are principally degrees of belief rather than rates in time or space. In response, risk analysis in dam safety relies more heavily on subjective probabilities and their expert elicitations.

Expert elicitation, nonetheless, does play an important role in other industries. For instance, the US Nuclear Regulatory Commission (USNRC) has developed expert elicitation procedures for use in seismic risk analyses (Budnitz et al., 1998; USNRC, 2012). The fact that subjective probabilities are subjective means they may vary from person to person. Various methods have been developed for aggregating or combining expert opinions (see e.g., Bedford & Cooke, 2001). Opportunities for learning are discussed in greater detail in Section 4.3 and Appendix E.

3.2.3 Importance of natural, site-specific uncertainties

The dominant uncertainties in dam safety risk analysis are mostly related to natural processes, specifically those related to hydrology and geology. The corresponding dominant uncertainties in risk analyses applied to chemical plant processes, manufacturing, aerospace engineering, and many other engineering fields attend to facilities, equipment, human factors, and operations. The characteristics of natural versus engineering uncertainties differ in many ways. The basis of uncertainty in natural processes pertains primarily to lack of information, imprecise understanding of natural phenomena, and the potential nonstationary nature of those process.

3.2.4 Long design lives

Dam projects and even dam modifications typically have long design lives. Design lives of 50 years are usual and of 100 years or more are practically realistic. This introduces important questions of how to discount for potential adverse outcomes decades in the future. Should potential loss of life 50 or 75 years in the future be treated at its nominal value, or should it be discounted in the way financial consequences are discounted (Lind, 2007; Rackwitz, 2004)? What would be the appropriate discount rate if the latter is the case? Another consideration of long design life is whether risks are stationary. Do annual probabilities of certain natural processes change with time? Do the consequences associated with failures or other adverse performance change with time? Do public preferences concerning risks (e.g., risk aversion) change with time?

3.2.5 Low-probability, high-consequence risks

Dam failures are low-probability events with potentially catastrophic consequences. Public preferences concerning low-probability, high-consequence risks are difficult to characterize, and are influenced by factors such as dread, control, and equity (P Slovic, 1987). Low-probability, high-consequence risks are not unique to dam safety. The risks from nuclear and chemical installations and certain other technological systems (e.g., planes) are similarly low-probability, high-impact risks.

3.3 Risk to Life Safety

In line with the guiding principle from the Federal Guidelines of Dam Safety Risk Management that life safety is paramount (FEMA, 2015), the Agencies focus on loss of life when evaluating dam safety risks. Risk evaluations are primarily framed with respect to individual and societal risk guidelines, although the Agencies also consider other consequence types and risk metrics when designing modifications or deciding on other risk management actions. Individual and societal risk guidelines are widely used for evaluating dam safety risks. They feature prominently in Australian dam safety regulations and industrial safety regulations in the United Kingdom and the Netherlands. Decisions related to low-probability, high-consequence risks to the public are rarely reduced to simple cost-benefit tests.

There are, however, differences between the tolerable risk guidelines used by the Agencies and those used in other industries and countries. First, the use of f-$\overline{\text{N}}$ for portraying societal risk is less common in other industries; F-N charts are more widely used. Second, the Agencies' interpretation and use of the five factors to be considered in an ALARP (as low as reasonably practicable) assessment appears to differ from Australian dam safety practice and nuclear safety practice.

3.4 Tolerable Risk in Public Policy Decision Making

Decisions about dam safety and its attendant risks are questions of public policy. They are not private sector internal decisions of corporations or organizations. Tolerable levels of risk for dam safety applications, therefore, are questions of public preference. They are fundamentally political decisions, not the decisions of engineers or scientists. In a democracy the level of risk(s) deemed to be tolerable should be the result of a political process, and levels of tolerable risk need to be coordinated among different sectors of the economy. This has not been the situation to date with the tolerable risk guidelines used by the Agencies, whether or not they are considered as criteria.

3.5 Staffing

Successful application of RIDM is not merely about risk analysis methodology and management. It is also about people and culture. RIDM requires staff that is comfortable with the idea that perfect safety cannot be guaranteed, familiar with basic concepts in probability and statistics, and willing to express beliefs in subjective probabilities.

While this is different in some industries and other parts of the world, most civil engineers in the United States are trained in using standards-based approaches. Given this background, many are uncomfortable with the idea that that risks cannot be reduced to zero.

Given the scarcity of risk analysis expertise and experience, Reclamation and USACE invest heavily in staff training, as illustrated by Best Practices. Training is important for maintaining a culture that fosters critical thinking and open dialogue. Such a culture is essential for high-quality risk assessments and for avoiding individual risk estimates or risk attitudes driving decisions. The risk to RIDM caused by staff turnover and new staffing needs (FERC) is discussed in greater detail in Section 4.3.

3.6 RIDM in Broader Dam Safety Practice in the US

Dam engineering and dam safety practitioners in the United States have looked to USACE and Reclamation, the two largest federal dam-owner agencies, for technical guidance and leadership. Hence, it is not surprising that adoption of RIDM by these two agencies has had influence outside of the agencies. While the application of RIDM to the broader dam safety community has been limited, it is growing. It seems likely that FERC's adoption of the methodology will hasten this broader adoption.

Other federal entities have incorporated some degree of RIDM into their programs. Within the Department of the Interior, Bureau of Indian Affairs, US Fish and Wildlife Service, and US National Park Service are applying RIDM approaches within the limitations of their funding resources, though not to as detailed an extent as Reclamation. The Tennessee Valley Authority has begun to incorporate RIDM into its dam safety program within the past few years. The US Department of Agriculture National Resources Conservation Service includes a limited consideration of risk in the programs it administers for the sponsors that own the dams constructed by the agency.

Application of RIDM outside of the 4 percent of dams in the United States owned by the federal government has been limited. As discussed earlier, a few of these dams are being evaluated using RIDM approaches through pilot studies in the FERC program. There have been only a few applications in the broader state-regulated arena. The state dam safety programs in Washington and Montana introduced considerations of annual exceedance probabilities into their program requirements for seismic and hydrologic loading in the 1990s. The Colorado State Dam Safety Program (Hunyadi et al., 2016) has more recently begun using simplified SQRA in its program, and the New Mexico State Dam Safety Program has begun a similar application. The California legislature directed the state's Division of Safety of Dams to incorporate RIDM into its program after the Oroville Dam spillway incident (France et al., 2018). That transition is now underway.

A number of other organizations have begun to use various forms of dam safety risk analysis for evaluating individual dams or portfolios of dams, in many cases recognizing that dam safety RIDM is part of a broader business risk management or asset management strategy. Examples include Denver Water (France & Martin, 2012), PacifiCorp (Raeburn et al., 2012, 2015), and Southern California Edison (Von Gersdorff et al., 2014).

Most other organizations have based their risk approaches on the information published by Reclamation and USACE, sometimes with adaptions to better suit the organization's specific needs.

4 AGENCY POLICY

The directions of the dam safety programs in the Agencies with regard to RIDM are appropriate and sound in the Panel's view. The Agencies' RIDM implementations are consistent with federal guidance provided in the *Federal Guidelines for Dam Safety* (FEMA, 2004) (published in 1979, reprinted in 2004) and the *Federal Guidelines for Dam Safety Risk Management* (FEMA, 2015). However, the levels of maturity in applying RIDM in the three agencies are quite different.

4.1 Agency Program Direction

Reclamation has the most mature RIDM program among the three agencies, having been using quantitative risk analyses to inform dam safety decisions for about 25 years. All of Reclamation's 370 high- and significant-hazard dams have been through multiple rounds of risk analyses as part the agency's comprehensive review process. Detailed issue evaluation and corrective action risk analyses have been completed for most Reclamation dams. Risk informed decisions are fully integrated into the dam safety program for decisions related to investigations and analyses, corrective actions, and performance monitoring. Dam safety actions are prioritized based on the agency's assessment of relative risks among the different dams in its portfolio, with intent to use available funding to address highest risk first. Although Reclamation's implementation of RIDM in the 1990s met with some initial cultural resistance within the agency, RIDM has now become fully embedded in the organizational culture at all levels in the dam safety program.

USACE has the next most mature RIDM program among the Agencies, with more than a decade of experience in using risk analyses to inform dam safety decisions. The agency was able to leverage Reclamation's experience to help in early development of its program in the mid-2000s. All 715 dams in the program have been subject to screening-level risk evaluations, and about two-thirds of the dams have had subsequent team risk analyses as part of the USACE's periodic assessment program. Many dams have been subject to more detailed quantitative risk analyses to evaluate issues of concern and develop dam safety modifications. Dams that have been modified have been subject to post-implementation risk analyses to verify risk reduction was achieved. As with Reclamation, the USACE's dam safety actions are prioritized based on the agency's assessment of relative risks among the different dams in its portfolio, with intent to address highest risk first. RIDM is generally accepted and well supported within the agency. Some within the organization have not fully accepted the approach, but acceptance continues to grow.

Although FERC has been using qualitative risk considerations, in the form of PFMA, for almost two decades, it has only more recently begun to move into more rigorous application of RIDM, with publication of its interim guidance in 2016 (FERC, 2016). FERC licensees presently have the option to continue to use previous FERC standards-based dam safety requirements or the new RIDM guidelines. FERC applications of RIDM to date are considered pilot studies, with two studies completed, ten in progress, and five proposed. These pilot studies represent a small part of the more than 1,000 high- and significant-hazard dams FERC regulates. FERC has also completed internal screening level portfolio risk analyses using an internally-developed tool. It appears there is much interest in the application of RIDM within the FERC organization, but limited experience at this time. It is not yet clear how many licensees will consider employing RIDM under the current optional arrangement. The Panel understands FERC is considering requiring some form of RIDM for its regulated projects, but that decision has not yet been made. Among the challenges FERC faces in implementing RIDM are 1) the need for developing capability not only within its own organization, but also within its licensees and the private sector consulting community that supports the licensees, and 2) highly variable levels of sophistication and resources among its community of licensees. These challenges are unlike those faced by Reclamation and USACE as self-regulated dam-owning agencies.

4.2 Applications of ALARP in RIDM

The ALARP principle originates in the United Kingdom (UK). Ale et al. (2015) trace its origins as far back as the Leeds Act of 1848, and states that much contemporary commentary suggests the term has been enshrined in the UK case law since the case of *Edwards v. National Coal Board (1949).*

4.2.1 ALARP Background

The Edwards case is cited by the UK Health and Safety Executive (HSE) (Health and Safety Executive, 2001) as being particularly important in interpreting “so far as is reasonably practicable” (SFAIRP), a concept the HSE considers equivalent to its interpretation of ALARP (Health and Safety Executive, 2002). The HSE does state, however, that SFAIRP and ALARP are not always interchangeable depending on the relevant legislation.

The HSE (Health and Safety Executive, 2001) states the Edwards case established that a computation must be made in which the quantum of risk is placed on one scale and the sacrifice, whether in money, time, or trouble, involved in the measures necessary to avert the risk is placed on the other; and that, if there is a gross disproportion between them, the risk being insignificant in relation to the sacrifice, the person upon whom a duty is laid discharges the burden of proving compliance was not reasonably practicable.

The Australian National Committee on Large Dams (ANCOLD) has also concluded its application of ALARP is equivalent to SFAIRP, since the guidance provided by ANCOLD is similar to that in Australian Legislation for SFAIRP, and because ANCOLD does not have a target level of risk below which ALARP is considered to have been achieved. This latter aspect is sometimes pointed to by some authors in postulating the terms are not the same (Robinson & Francis, 2014).

It generally appears that, in commonwealth countries with legal systems founded on the UK law, health and safety statutes have given expression to the common law doctrine of duty of care (Health and Safety Executive, 2001). Certainly, in Australia the guidance provided to employers to meet their obligations to reduce risk SFAIRP share similarities to the legal considerations for meeting a duty of care.

4.2.2 ALARP in United States practice for other industries

ALARP has not been widely used in the United States, although the related concept ALARA ("as low as reasonably achievable") arose in earlier decades in consideration of nuclear safety and radiation protection, and remains a regulatory requirement regarding legal dose limits for workers (Federal Register, 1987). ALARA is based on the linear hypothesis that no level of radiation exposure is without harm. The concept of “good engineering practice” in United States legal actions is more common in other fields such as civil infrastructure.

The US Department of Energy (USDOE) and USNRC adopted ALARA in 1972 and in 1977, for regulating radiation exposures under 10 Code of Federal Regulations (CFR) Part 835 and 10 CFR Part 20. ALARA is defined in these federal regulations as “making every reasonable effort to maintain exposures to radiation as far below the dose limits in this part as is practically consistent with the purpose for which the licensed activity is undertaken, taking into account the state of technology, the economics of improvements in relation to state of technology, the economics of improvements in relation to benefits to the public health and safety, and other societal and socioeconomic considerations, and in relation to utilization of nuclear energy and licensed materials in the public interest.”

The California Public Utilities Commission (CPUC) has explored ALARP in developing rate cases for public safety expenditures following a gas pipeline explosion in San Mateo, California (Haine, 2015). The concept was equated to benefit-cost analysis in that implementation. The CPUC cites USACE use of ALARP in Best Practices as a reason for their consideration.

4.2.3 Agency approaches to ALARP

Including ALARP in the Agencies' RIDM programs is consistent with federal guidelines for dam safety management (FEMA, 2015) which state:

> The primary purpose of a dam safety program is to identify the risk to life and property posed by dams. While dam safety risks cannot be eliminated, they should be reduced to a level that is as low as reasonably practicable.

Also, principle three of the five principles that "apply to the overall objectives of these guidelines" states*:*

> Identify and reduce the risk to life and property posed by dams and reduce those risks as low as reasonably practicable.

Reclamation

The current Reclamation guidelines focus ALARP on the "area bounded by the 1 x 10^{-6} annualized failure probability on the top and 1,000 fatalities on the left." (Bureau of Reclamation, 2011b). In this report, the Panel will refer to this as the "bounded area."

However, in the Public Protection Guidelines Examples (Bureau of Reclamation, 2011a) it is stated that:

> ALARP considerations apply when societal risks are estimated in the range where the Annualized Failure Probability is less than 10^{-6} and the estimated loss of life is greater than 1,000. They can also apply to other situations. For example, it may be possible to reduce risk to just below the guidelines for a given cost, but to get it comfortably below the guidelines (e.g., an order of magnitude) could require a substantial increase in cost. ALARP considerations could be used to decide whether that extra cost is justified.

This indicates ALARP considerations must apply in the "bounded area," but could also apply at any point below the quantitative risk guideline and Reclamation has advised the Panel that the ALARP principle has been applied to actions under corrective action studies.

In relation to individual risk (annualized failure probability for Reclamation), Reclamation's Rationale document (Bureau of Reclamation, 2011c) states:

> More recent evaluations by Foster et al [9] and Douglas et al [10] in 1998 indicate failure rates for dams that survived their first five years have reduced somewhat, to about 0.8 x 10^{-4} per annum for both concrete and embankment dams, indicating that the number of dam failures is decreasing as the number of successful dam years of operation is increasing. Continuing to strive to reduce dam failure rates as far below 10^{-4} as reasonably practicable will help ensure that this will continue to be the case.

This implies ALARP considerations may also apply to individual risk.

The briefing presentation provided for this review by Reclamation included the following statement (Bureau of Reclamation, 2019b):

> ALARP Question. Our goal is to achieve a situation where the risks are as-low-as-practicable. We are prioritizing our efforts and may not be taking action now but that could change in the future to reach ALARP.

From the information set out above, and documentation provided to it by Reclamation the Panel concludes Reclamation applies ALARP considerations to risk within the bounded area, and has applied ALARP to actions under corrective action studies, but does not routinely apply ALARP for risks that plot below the quantitative guidelines.

USACE
Engineer Regulation (ER) ER-1110-2-1156 (USACE, 2014b) states the USACE may choose to upgrade a dam to achieve an ALARP position or may take an approach to reduce the risks to below the limit of tolerability without ALARP considerations (this is sometimes referred to in the industry as "progressive improvement"). The Panel assumes that ultimately, in accordance with Sections 5.2.5 and 5.3 of the regulation, it is intended that all dams achieve risks that are ALARP.

FERC
It is clear from Sections 1.1.5, 2.5.5.1 and 3.3.2.2 of the FERC "Risk-Informed Decision-Making Guidelines" (FERC, 2016) that risks should be reduced to ALARP. An extract from Section 1.1.5 follows:

> **Principles**
> The following fundamental principles apply to the overall objectives of these Risk Guidelines:
>
> 1. Life safety is paramount.
> 2. 'Do no harm' must underpin all actions intended to reduce dam safety risk.
> 3. Risk should inform the decision process. Decisions are not 'risk-based'.
> 4. Identify and reduce the risk to life and property posed by dams and reduce those risks to as low as reasonably practicable (ALARP).
> 5. The urgency of completing dam safety actions should be commensurate with the level of risk.
> 6. Dam safety inspections, surveillance and monitoring, emergency action plans and testing, owners dam safety plan, Part 12D Reports, training and other routine dam safety activities are all essential parts of an effective dam safety risk management program.
> 7. Risk communication must be well planned, timely, and involve all parties potentially affected by the decision or a failure of the dam.

4.2.4 Application of the ALARP principle

The Agencies have similar guidance for applying the ALARP principle. As mentioned earlier, the guidance for ALARP that has been developed and proposed by regulatory authorities for ALARP and SFAIRP appears to be based on the legal test to demonstrate that a hazardous facility owner has met its "duty of care" to potentially affected communities and their assets. There is no formula to decide that risks have been reduced to ALARP, and it will ultimately be for a court to decide whether it was achieved by a dam owner. The common factors the Agencies propose are:

- Cost-effectiveness of the risk reduction measures;
- Level of risk in relation to the quantitative or tolerable risk guidelines;
- The disproportion between the sacrifice (money, time, trouble, and effort) in implementing the risk reduction measures and the subsequent risk reduction achieved;
- Any relevant recognized good practice; and
- Societal concerns as revealed by consultation with the community and other stakeholders.

These factors are the same as those proposed by ANCOLD and the Victorian dam safety regulator in Australia. However, little guidance is provided by those organizations on how to make a balanced judgment from consideration of the five factors. Dam owners in Australia have been grappling with the issue since the ANCOLD Guidelines were published in 2003. From briefings received from the Agencies, the Panel understands that guidance for the application of the factors to be considered in ALARP decision making is developing.

The HSE and Office of Nuclear Regulation (ONR) provide guidance for use in their regulatory role for industrial and nuclear safety in the UK. This guidance is consistent in pointing to recognized good practice for a new facility as being the benchmark for assessing an existing facility, and then the reasonable practicality of implementing the required improvements are to be assessed. The regulators point to important considerations in making this assessment, for example for ONR (ONR, 2019):

1. ALARP demonstrations should consider first and foremost factors relating to engineering, operations and the management of safety. These expectations are often referred to by the general term "relevant good practice."
2. For higher consequence situations the consequence should weigh more heavily than the frequency estimates.
3. Furthermore, thought should be given to the robustness of the conclusions with respect to uncertainties and to any assumptions employed in the demonstration … sensitivity analysis should be provided … assume that precautions should be taken unless there is a good reason to think that the risk is insignificant.
4. For long term risk assessment, the uncertainty of how future generations will view the risks left to them (and indeed the uncertainty of any benefits further into the future) argues for a precautionary approach (R2P2 paras 89-93) and hence a particularly stringent demonstration of the ALARP principle ….. Given uncertainties in estimating long term further risks, good practice and application of the Engineering Key Principles Hierarchy with an emphasis on control of hazard are likely to be much more important than numerical risk estimates and CBA (cost-benefit analysis) in establishing the way forward.

The UK public safety regulators provide guidance for using disproportion factors for CBA, depending upon the level of risk and uncertainty in the risk estimate.

Application of the ALARP principle in Australia has developed recently to generally align with the approach described by the UK regulators. Put simply, the approach by some dam safety practitioners in Australia takes into account the following matters:

- Good practice, consisting of the safety controls that would be incorporated in the design for a new dam, is the starting point for the evaluation;
- The reasonable practicability of retrofitting the safety controls to achieve the equivalent to that of a new dam is evaluated by considering:
  - Whether it is practical to retrofit the control (can it be physically done?);
  - Whether it is safe to retrofit the control (would the community be exposed to unacceptable risk during the retrofit?);
  - The costs of implementation versus the improvement in safety through direct comparison of the costs and risk benefits; and
  - The cost effectiveness of the control options (cost to save a statistical life).
- Where the societal concern (the reaction to dam failure) would clearly be very high, then good practice considerations will weigh more heavily on the decision than cost effectiveness.

An example of an approach used in Australia for application of the ALARP principle is outlined in Robilliard and Sih (2018). However, ALARP remains a very difficult issue for dam owners and there is no single approach that will suit all circumstances.

The Panel fully supports the Agencies use of the ALARP principle and the five factors to be considered in making a judgment. It understands that there is little guidance available on how to balance these factors and while no dam safety practitioners in Australia would claim to have a solution; it may be helpful for the Agencies to consider recent developments there.

4.2.5 Disaggregation

A further consideration relates to the need for a "disaggregation" approach to evaluating ALARP to ensure all reasonably practicable measures have been taken to reduce risk at each step in the potential failure mechanism. This can be achieved through carefully examining each branch of the event tree or fault tree, or by the PCM (prevention, control, and mitigation) concept (similar to the 'bow tie' concept). The bow tie methodology is used commonly in other hazardous industries, such as oil, gas, and chemical, for safety cases to demonstrate that risk have been reduced ALARP or SFAIRP through explicit display and quantification of "barriers."

International Committee on Large Dams (ICOLD, 2017) reports that under a PCM approach:

> As a result of this analysis of failure scenarios, the risk management measures or barriers can be divided in three main groups:
>
> - Prevention - Activities necessary to ensure the structural safety of the dam - Design and calculations applying to new or existing dams, dam construction or rehabilitation, dam deconstruction. These activities aim to minimize the risk by optimal design.
> - Control - Operational activities during all of the dam life to maintain the safety of the dam: surveillance (monitoring, visual inspection, equipment tests), flood routing, maintenance, activities ensuring public safety.
> - Mitigation - Emergency preparedness and response

The safety case then contains an evaluation to demonstrate ALARP for all types of controls (risk management measures) and a 'safety management system' is required to ensure controls are maintained to operate reliably and continue to meet the required performance.

The USACE (Snorteland, 2019) appears to have taken a PCM approach with the four-part tolerability of risk guidelines:

1. Understanding risks surrounding dams and levees.
2. Building risk awareness.
3. Fulfilling daily responsibilities
4. Continually considering actions to reduce risk.

Similarly, Reclamation and FERC have requirements for a PCM approach. However, the Panel has not seen information that indicates the Agencies apply the ALARP principle to the control and mitigation activities.

4.2.6 The safety case

The safety case concept evolved primarily in the UK in the aftermath of the Piper Alpha offshore tower failure and other serious accidents of the time. Its application to dam safety is innovative and reasonable. In concept, the safety case presents the logical reasoning leading to a conclusion that a system is safe for a particular use. It must be sufficiently documented and argued that anyone reviewing it is able to understand the chain of thinking that led to the conclusion. It is widely used in the oil and gas industry, aviation and transportation, medical devices, and many other technical enterprises. It identifies hazards and risks, describes how the risks are controlled, and describes the management system intended to ensure the controls are effective.

The Best Practices guidance (Section A-9) describes the dam safety case as:

> […] a logical and objective set of arguments used to advocate a position that either additional safety-related action is justified, or that no additional safety-related action is justified. The safety case should cite the most compelling information and evidence that supports the risk estimates and the overall findings. Confidence and uncertainty identified in the risk analysis should also be discussed, including identifying the sources of uncertainty, describing actions that could be taken to reduce uncertainty, and addressing the level of confidence in all three components of the risk estimate (load probability, structural response likelihood, and consequences).

This is a solid description that seems to capture the important essence of the risk analysis and evaluation. In policy, the various components of the risk analysis lead to this critical summary. However, the Panel understands that it is referring to the prevention component of the PCM approach, and not specifically to control and mitigation.

The safety case as practiced in many industries is also a systems engineering concept that is linked to operations and incident reporting. In contrast, the dam safety case in Best Practices is described as "arguments [that] combine together key evidence regarding the three basic risk components (load probability, response likelihood, and consequences) in order to support decisions related to a dam's or levee's existing condition or ability to withstand future loading." There is nothing wrong with this description, but it is static. It is the set of arguments, data, and thinking leading to a present decision, not to ongoing safe operation of the dam.

The Panel understands that the Agencies have detailed processes to manage all aspects of the PCM approach, are thorough in application, and the Panel strongly supports these efforts. However, to assist with the demonstration of dam safety, the Agencies may find it helpful to establish a safety case description that encompasses the three aspects of prevention, control, and mitigation, and application of the ALARP principle for each.

4.2.7 Panel recommendations

The Panel recommends the Agencies continue to develop their use of the ALARP principle in dam safety decision making and consider whether the recent developments in Australia for applying the factors used in an ALARP judgment could be of benefit.

The Panel recommends the Agencies evaluate whether the safety case approach used to demonstrate ALARP for each of the 'prevention, mitigation, and control' aspects of safety in other hazardous industries would improve the demonstration of dam safety.

4.3 Staff Development, Training, and Continuous Learning

The USACE and Reclamation use the comprehensive Best Practices to provide guidance to staff performing risk analyses. This document is supported by a four-day introductory training course. The methods and procedures set out in the document are used to undertake risk analyses. This is supported through direct involvement of risk analyst specialists, review by specialists, and review by senior personnel.

A brief summary of the USACE, Reclamation, FERC training and review practices is provided below to give an insight into the level of effort those Agencies consider required to try to assure risk analyses results are consistent and of suitable quality for decision making.

4.3.1 USACE training and review

In its publication, Safety of dams: policies and procedures (USACE, 2014a), the USACE set out an extensive program for training personnel in all matters relating to its mission in water resource development. The training covers Dam Safety, Conferences, Exchange, Operations and Maintenance, Risk Assessment, and Consequences. Risk assessment training is provided by the Risk Management Center. Consequence training is provided by USACE Hydrologic Engineering Center (HEC) and through PROSPECT (USACE Proponent Sponsored Engineer Corps Training).

The Risk Management Center consists of risk and reliability experts; dam safety modification centers; mapping, modeling, and consequence center; risk cadres; and a policy and procedures team (N. Snorteland, 2019). The risk cadres are technical specialists and analysts that work with the USACE districts to perform PFMA and risk estimates. The Risk Management Center monitors and provides quality assurance to the dam safety program. It also ensures risk analyses are technically acceptable prior to Dam Senior Oversight Group review.

4.3.2 Reclamation training and review

Reclamation has training goals to inform the organization on current best practices and methodology and ensure consistency and staff development (Major, 2019). This is achieved through on-the-job training, scheduled training (e.g., Best Practices), technical training, and review of case histories, some of which are done in cooperation with USACE.

A risk cadre was initially formed in the 1990s to develop risk methodologies and procedures. The original risk cadre was disbanded in the early to mid-2000s, but it was later re-established sometime before 2008. Later, sometime after 2011, the risk cadre was replaced with a three-person risk advisory team (RAT). A risk cadre was re-established via a formal project management plan in anticipation of the retirements of two RAT members. Currently, the risk cadre consists of six to seven core team members with collateral duties and a dedicated team lead, designated as the risk advisor. The risk cadre is responsible for coordinating and implementing training in dam safety and risk assessment, revising guidelines and documentation for risk assessment, providing outreach and other advisory activities, and supporting Dam Safety Advisory Team (DSAT) meetings and reviews. Additional tasks are delegated to a broader group of participants termed the project cadre. The project cadre also plays a role in succession planning in that its members can be transitioned into the core risk cadre, as needed. Risk assessment documentation is produced by the Technical Service Center (TSC) and is reviewed by DSATs (internal) and consultant review boards (external).

4.3.3 FERC training

FERC has provided PFMA training. Completion of this training has been a requirement to be certified as an approved PFMA facilitator. A three-day, generic SQRA training workshop was presented in October 2019 in collaboration with the United States Society on Dams (USSD). It will be repeated in the future.

Training initiatives in development or proposed include:

- A web-based overview of the FERC RIDM guidelines;
- Web-based training on conducting a screening-level risk assessment using a FERC developed tool; and
- Training workshops on policy and requirements for FERC SQRAs.

4.3.4 Training needs

It is clear from the Reclamation and USACE practices that emphasis is placed on training and governance arrangements to ensure risk analyses quality and consistency. This is partly because the process requires elicited subjective probabilities from subject matter experts. The pool of these resources is small, and methods must be developed whereby initial work may be undertaken by less experienced individuals, but there is an assurance overlay where results are reviewed and adjusted as necessary.

FERC has prepared interim risk-informed decision-making guidelines for its licensees. The guidelines are currently in draft form for use with the intent to finalize them after a trial period. The process will require private sector consultant resources competent in risk analysis to undertake the studies and make informed recommendations to the licensees.

The Division of Dam Safety and Inspections (D2SI) (Boyer, 2019) has an oversight role to:

- Observe and ensure risk analysis sessions are conducted in accordance with general provisions and established processes in the Engineering Guidelines; and
- Observe and ensure the risk analysis team's treatment of the factual project information; engineering evaluations and interpretations; and the resulting findings, conclusions, and recommendations are based on sound engineering judgment and are congruent.

In order for D2SI to fulfil its role and avoid becoming a part of the risk analysis team, it will be essential that:

- D2SI staff are trained in risk analysis processes;
- A pool of trained and experienced private sector consultants and subject matter experts is available to undertake the risk analyses; and
- Trained and experienced risk analysis facilitators are available.

These skilled resources are also essential to ensure the risk analysis is fit for purpose in that there is an understanding of the level of effort appropriate for the decision to be made. Although FERC is proposing to communicate with licensees to provide some support in working with its guidelines, it would be difficult, and undesirable, to be engaged in every step of the process to fill knowledge gaps caused by insufficient training.

A large pool of trained and experienced private sector consulting resources is not currently available for licensees to engage to undertake risk analyses for their dams. So, the question arises, how can this resource deficiency be addressed?

Obvious sources for training in risk analysis would be the USACE and Reclamation. However, the Panel understands Reclamation is currently prevented from providing training for a fee to anyone outside government due to legal guidance, and the level of effort to provide such training gratis cannot be justified by the agency (B. Becker, personal communication, April 30, 2020). Although USACE can deliver paid training, it can only be done through the USACE Learning Center and the training has to be offered to everyone in USACE before opportunities can be offered to the private sector (N. Snorteland, personal communication, January 1, 2020).

Developing risk analysis skills within the private sector consulting businesses is a challenge that will need to be resolved if FERC is to fulfill its ambition for licensees to take advantage of RIDM. Neither Reclamation nor USACE can practically provide the training at the moment.

The Association of State Dam Safety Officials[2] (ASDSO) has a well-regarded training program for the industry that uses experienced dam safety practitioners to deliver courses in classroom, webinar, or

---

[2] The ASDSO vision (ASDSO, 2020.) is "A future where all dams are safe" and its mission is "Improve the condition and safety of dams through education, support for state dam safety programs and fostering a unified dam safety community. A supporting activity is to "advance and expand the technical expertise of dam and levee safety practitioners."

workshop format. It has presented some existing webinars on risk analysis / assessment topics. There may be an opportunity for FERC to leverage its existing collaboration with USSD or to also collaborate with ASDSO (and / or an academic institution) to formulate a range of training for industry.

4.3.5 Panel recommendations

The Panel recommends FERC explore opportunities to provide enhanced risk assessment training to dam safety consultants through its existing collaboration with USSD, establish a cooperative arrangement with ASDSO, or potentially involve academic institutions to support or deliver aspects of the training.

4.3.6 Risk precursor and continuous learning programs

Many major accidents are preceded by precursory events that, although observable, are not recognized as an early indication and forerunner of a catastrophic event until after the fact. Precursor events or conditions provide evidence of the presence of failure mechanisms in the system that pose a significant degree of risk. In one interpretation, any off normal condition, such as the degraded state of a structural element or component of the system, is a sign of reduced margin of safety (increased risk), and should be viewed as a precursor to more severe conditions, deserving some level of evaluation and analysis from a risk and reliability perspective. This is the core concept behind a variety of processes and programs in place in other industries, including Failure Reporting and Corrective Action System, continuous improvement / lessons learned programs, and safety management/monitoring systems. The scope of such programs and systems vary greatly, and the degree of rigor and formalism can also be very different, but they all provide a vehicle for continuous learning.

One of the most formalized programs of this kind is what is generally referred to as accident precursor or “near miss” analysis program, typically designed and implemented in the context of RIDM. The oldest of such programs is the Accident Sequence Precursor (ASP) Program established by the USNRC in 1979 with the following objectives: a) provide a comprehensive, risk-informed view of nuclear power plant operational experience and a measure for trending nuclear power plant core damage risk, b) provide a partial check on dominant core damage scenarios predicted by probabilistic risk analysis (PRA), and c) provide feedback to regulatory activities.

In the USNRC program (USNRC, 2008), the most important precursor events are mapped to generic or site-specific event tree models to quantify the conditional probability of reactor core damage events given the precursor occurrence. This conditional probability is used as a metric of the precursor event risk-significance. Mapping precursors to PRA models also helps in improving the risk models, and in calibrating the calculated probabilities. The USNRC also uses the ASP Program to monitor performance against the safety goal established in the agency’s strategic plan. NASA established a similar program in 2011 (NASA, 2011). The purpose of NASA’s Accident Precursor Analysis program is “to identify and characterize potential sources of safety risk for which indications are received in the form of anomalous events which, although not necessarily presenting an immediate safety impact, may indicate that an unknown or insufficiently understood potential risk-significant condition exists in the system.”

The most recent program of this type was established by the Bureau of Safety and Environmental Enforcement (BSEE) in 2015 (BSEE, 2015) as a near-miss reporting system that includes a confidential collection of equipment failure reports, in an effort to further improve offshore oil and gas operations safety. The data are analyzed by the US Department of Transportation, Bureau of Transportation Statistics to explore safety trends.

Establishing similar precursor analysis and data-driven performance assessment programs, jointly or separately by the Agencies, can provide comparable benefits in the area of dam safety. Such programs can provide factual evidence for postulated as well as previously-unidentified failure scenarios and be used to improve the analytical models and calibrate subjectively estimated model parameters and probabilities.

4.3.7 Panel recommendations

The Panel recommends the Agencies establish precursor analysis and data-driven performance assessment programs to improve the analytical models and calibrate subjectively- estimated model parameters and probabilities.

4.4 Inter-agency Collaboration and Partnerships

The Agencies have on-going collaborations and partnerships with a number of national and international organizations, and they actively participate in the professional community.

4.4.1 Agencies' avenues for collaboration and exchange

The Interagency Committee on Dam Safety (ICODS), chaired by FEMA, is a permanent forum of federal agencies that advises FEMA on institutional, managerial, technical, legislative, and policy issues related to dam safety. Among other efforts, ICODS encourages coordination and exchange of information among federal agencies and publishes guidelines and technical manuals for dam safety (e.g., ICODS, 1979). Members include representatives from the USACE, Reclamation, and FERC, as well as other federal agencies.

The National Dam Safety Review Board monitors dam safety in the United States, monitors state implementation of national dam safety guidelines, and advises FEMA on national dam safety policy. Members include FEMA (chair), representatives from four federal agencies that serve on ICODS, five state dam safety officials, and one private sector member. Representatives from the USACE, Reclamation, and FERC serve as review board members.

The Agencies maintain close collaborative ties at a working level. This is reflected by joint publications such as Best Practices. It is also reflected by the strong similarities among the risk guideline frameworks developed by the USACE, Reclamation, and FERC. While their terminologies differ, the Agencies' approaches to risk analysis and risk assessment are very similar, owing to inter-agency exchanges of information on risk management practices and methodology.

The Agencies have strategic national partnerships with USSD and ASDSO. FERC actively collaborates with the National Hydropower Association. The Agencies are also active members of the international dam safety community:

- The USACE has strategic international partnerships with the Center for Energy Advancement through Technological Innovation (CEATI), the Environment Agency (UK), Rijkswaterstaat (the Netherlands), the Spanish Ministry of Agriculture and Fishing, Food and Environment, Water Services Association of Australia, Australian National Committee on Large Dams (ANCOLD), and ICOLD. Furthermore, the USACE is active in numerous countries where it provides assistance and technical support in the field of dam safety risk management.
- Reclamation has strategic international partnerships with CEATI, ICOLD, ANCOLD, and the Next Generation Liquefaction Project of the Pacific Earthquake Engineering Research Center. It also provides training to various countries within the context of its International Affairs Program.
- FERC has strategic international partnerships with CEATI (Dam Safety Interest Group) and ICOLD.

The USACE, Reclamation, and FERC share their technical reference manuals and risk management guidelines online. They also frequently publish scientific papers on dam safety and risk analysis methodology, thereby opening themselves to external scrutiny and contributing to the national and international states of practice.

4.4.2 Panel recommendation

The Panel recommends that, for identifying learning opportunities and for contributing to industry best practice, the Agencies continue to share their experiences and practices with each other and the broader dam safety community, through professional organizations and international partnerships.

5 AGENCY GOVERNANCE

The General Charge Guidance for this review of risk methods and processes includes the following question for consideration:

> Given each agency's missions, are there any governance, critical policies, or methodologies missing from the agencies?

This section of the report deals with the review of governance

While the word "governance" has historically been used in relation to the processes to control Government or corporations, the word is now also commonly used to refer to control of programs or projects.

The Agencies have clarified that "governance" in the context of the General Charge Guidance means governance at the agency and program level. Therefore, the focus of this section of the report is on examining the Agencies' practices at these levels.

5.1 Introduction

The following broad aspects of governance are set out in this introduction for the purpose of clarifying the review:

1. Defining governance.
2. United States federal guidance and risk governance.
3. The context within which governance is exercised.
4. Program governance.

5.1.1 Defining governance

The Merriam-Webster dictionary defines governance as "government," with an example of use being "the challenges of national governance," and the English language learner's definition of "the way that a city, company, etc., is controlled by the people who run it"[3].

The Oxford English dictionary defines governance as "The office, function, or power of governing; authority or permission to govern."[4]

Governance is a critical factor in managing corporations. A definition with the United States context has been provided by the Committee of Sponsoring Organizations of the Treadway Commission (COSO - "a joint initiative of the American Accounting Association, American Institute of Certified Public Accountants, Financial Executives International, The Association of Accountants and Financial Professionals in Business, and The institute of Internal Auditors and is dedicated to providing thought leadership through the development of frameworks and guidance on enterprise risk management, internal control and fraud deterrence")[5]:

> **Governance** is the act or process of providing oversight or authoritative direction or control. One author defines it as the allocation of power among the board, management and shareholders. It is often applied to describing what the board of directors and executive management does in providing direction and oversight to the organization's affairs. Corporate

---

[3] "Governance." *Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/governance. Accessed 13 Feb. 2020.

[4] https://www.oed.com/

[5] https://www.coso.org/Pages/default.aspx

> governance is typically the domain of the board of directors and refers to the framework of rules and practices by which a board oversees strategy setting and the management of the organization. Effective governance ensures accountability, fairness and transparency in the organization’s relationships with its various stakeholders, e.g., shareholders, lenders, customers, suppliers, employees, governments, regulators and the communities in which it operates.

The Australian Institute of Company Directors quotes the Australian Securities Exchange in defining corporate governance as:

> The framework of rules, relationships, systems and processes within and by which authority is exercised and controlled in corporations. It encompasses the mechanisms by which companies, and those in control, are held to account.

Although the Agencies are not corporations with boards, it is reasonable to expect that guidance established for good governance of an organization can broadly apply within each agency, subject to its internal rules and reporting structure. In any case, such guidelines can provide a framework against which effectiveness of governance over internal functions can be measured.

5.1.2 United States federal guidance and risk governance

Levels of governance within the United States’ federal structure supervise or provide oversight over entities with federal responsibilities for risk management policies. The most important of these are the Office of Management and Budget (OMB) of the Executive Branch and the Government Accountability Office (GAO) of the Legislative Branch. Regarding OMB:

> […] OMB’s mission is to assist the President in meeting his policy, budget, management and regulatory objectives and to fulfill the agency’s statutory responsibilities. […] OMB measures the quality of agency programs, policies, and procedures to see if they comply with the president's policies and coordinates inter-agency policy initiatives.

Regarding GAO:

> The Government Accountability Office is a legislative branch government agency that provides auditing, evaluation, and investigative services for the United States Congress. It is the supreme audit institution of the federal government of the United States.

Under the George W. Bush Administration, OMB attempted to impose government-wide policies and approaches for scientific risk analysis procedures applied especially to environmental regulatory functions of the Administration. This proposal was reviewed by the National Academy of Sciences and eventually abandoned (National Research Council (U.S.), 2007). Under the Obama Administration, enterprise risk management guidance was subsequently introduced as OMB Circular No. A-123, “Management's Responsibility for Enterprise Risk Management and Internal Control,” but this deals principally with financial risk.

The US Congress has long been concerned about the governance of risk-based regulation, particularly in establishing environmental and health and safety standards. This resulted in a long series of reports typified by the National Research Council (NRC) (NRC, 1983, 1993, 2009, 2010). However, these reports do not carry the force of regulation or policy.

5.1.3 The context within which governance is exercised

Through the time of the Carter Administration, the US Water Resource Council (USWRC) provided coordination among the 17 US agencies with responsibility for water resources projects. In principle, this would have included risk governance, although risk-informed decision making in the dams and water resources sector was uncommon at the time. The Reagan Administration de-funded USWRC and, while it still exists administratively, it is not operational. Subsequently, FEMA was assigned responsibility for the National Dam Safety Program (NDSP), which is:

> A partnership of states, federal agencies and other stakeholders to encourage and promote the establishment and maintenance of effective Federal and state dam safety programs to reduce the risk to human life, property, and the environment from dam related hazards.

The NDSP has published a number of technical guidance documents on dam safety, including *Federal Guidelines for Dam Safety Risk Management* (FEMA, 2015):

> This document provides guidelines for implementing risk-informed decision making in a dam safety program. The intended audience is federal agencies that own or regulate dams. The guidelines could also be applied to non-federally-owned or -regulated dams that can impact federally-owned or -regulated facilities; however, this would require cooperation and involvement of the non-federal dam owner.

These guidance documents primarily reflect consensus opinions of the water resources agencies rather than oversight or authoritative direction or control.

5.1.4 Program governance

For the purposes of reviewing agency programs within a governance framework, the Panel chose to categorize the functions within the Agencies in typical board and management roles. For example:

- the level within the Agencies that has responsibility for setting strategy, policies, and objectives could be viewed as having a 'board' type of function;
- the levels within the Agencies that have responsibility for planning and implementation could be viewed as having a 'management' type of function.

In ICOLD Bulletin 154, Dam Safety Management: Operational Phase of the Dam Life Cycle (ICOLD, 2017), "Elements of a Management System" provides a useful way of defining the role of the board versus the role of management in the implementation of a management system to achieve organizational goals (Figure 6). This system is based on the continuous quality improvement principle of the International Organization for Standardization (ISO): plan, act, monitor, improve. It is also Clause 4 in the ISO 31000 guidance on managing risk.

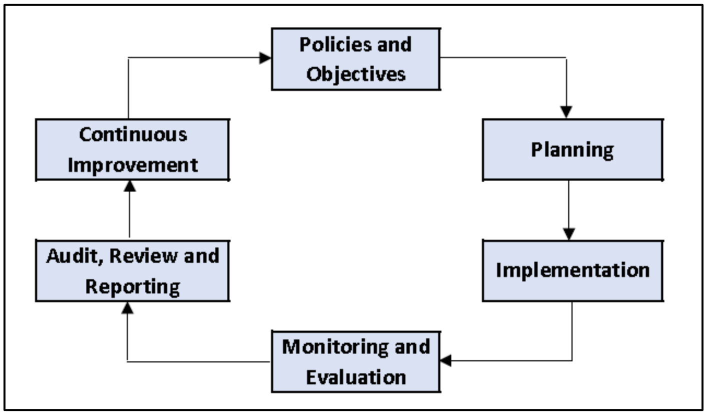


Figure 6. Elements of a Management System (ICOLD, 2017 based on ISO, 2014).

In this model we can see that the ‘board’ function resides in setting the ‘policies and objectives,’ then ultimately reviewing management performance through its reporting. Management is responsible for the other five activities of planning; implementation; monitoring and evaluation; audit, review and reporting; and continuous improvement.

Taking into account the general charge guidance, clarification of the scope from the Agencies, and the discussion on governance set out above, the Panel concluded the review of governance should focus on:

- The board functions of program approval and oversight; and
- The management functions established to deliver the program.

Review of the details and effectiveness of the risk management processes themselves are dealt with in other sections of this report.

5.2 Method

Building on the elements of a management system as defined by ICOLD, the following key activities within each element have been identified. These activities broadly align with ISO 55001:2014 Asset Management - Management Systems—Requirements (ISO, 2014).

Table 1. Governance Elements and Key Activities

Program Governance Element	Key Activities - Summary Description
Policies and Objectives	• Clear direction set through policies and objectives (or equivalent) • Systematic approach to establish the policies and objectives • Major stakeholders other than the decision makers (those who approved the program) involved in establishing the policies and objectives • Decision makers understand, tested, and endorsed the policies and procedures • A process for decision makers to review the performance of the policies and objectives and, if necessary, revise them
Planning	• Program objectives and task targets have been established consistent with the policies and objectives • A program risk management plan is established • Performance standards are in place • The required technical competence and functional understanding has been identified and assured • A communications plan catering for internal and external stakeholders • Outsourcing plan including controls • Activities are contained within an overall program plan
Implementation	• A delivery structure (with roles and responsibilities) and resource plan is established to achieve the policies and objectives, internal program objectives, and task targets • Procedures in place to assure that the responsibilities are met • Risk management plan monitored, updated, and revised as required
Monitoring and Evaluation	• A systematic process to measure the program implementation outcomes are meeting the program objectives, task targets, and performance standards • Outcomes and performance documented
Audit Review and Reporting	• Process to review the monitoring and evaluation results • Systematic internal audit of performance against the program plan • Systematic external audit of performance against the program plan • Process for reporting to the decision makers regarding performance against the established policies and objectives
Continual Improvement	• Implementation of recommendations from audit, review, and reporting • Change management plan

5.3 Agency Governance Evaluation

A detailed program governance evaluation is attached as Appendix C. It was undertaken using the key activities identified in Section 5.2, and the evaluation observations were based on information extracted from the documentation provided by the Agencies, briefings in Denver, and responses provided by each agency where required.

Overall, and noting that the evaluation has been undertaken at a high level, the governance processes implemented by the Agencies is judged by the Panel to be good practice. While the level of the evaluation completed for this review does not directly reveal obvious areas for improvement, other information available to the Panel has led to reservations regarding public stakeholder communications effectiveness and potential challenges in providing risk assessment training to private sector consultants working for

FERC's licensees. These issues are addressed in Section 4.3 Staff Development, Training, and Continuous Learning and Section 7.3 Risk Communication.

Some aspects of FERC's governance are not yet in place, as its risk-informed decision-making program is still being developed, but the agency has indicated these aspects are planned for development and implementation.

6 AGENCY METHODOLOGY PART 1: RISK ESTIMATION

This chapter summarizes and provides comments on the Agencies' methodology for performing risk estimation. It principally relies on the key risk analysis documents for each agency, and on the joint Reclamation-USACE Best Practices guidance.

The Agencies' methodology for risk-informed decision-making practice is principally summarized in:

> Bureau of Reclamation (2011b). *Dam safety public protection guidelines*. US Bureau of Reclamation, Denver.
>
> USACE (2014a). *Safety of dams: policies and procedures*. US Army Corps of Engineers, Washington DC.
>
> FERC (2016). *Risk-Informed Decision Making (RIDM) Risk Guidelines for Dam Safety - Interim Guidance*. Federal Energy Regulatory Commission, Office of Energy Projects, Division of Dam Safety and Inspections.

Training materials are provided in a joint agency website containing narrative primers and presentations:

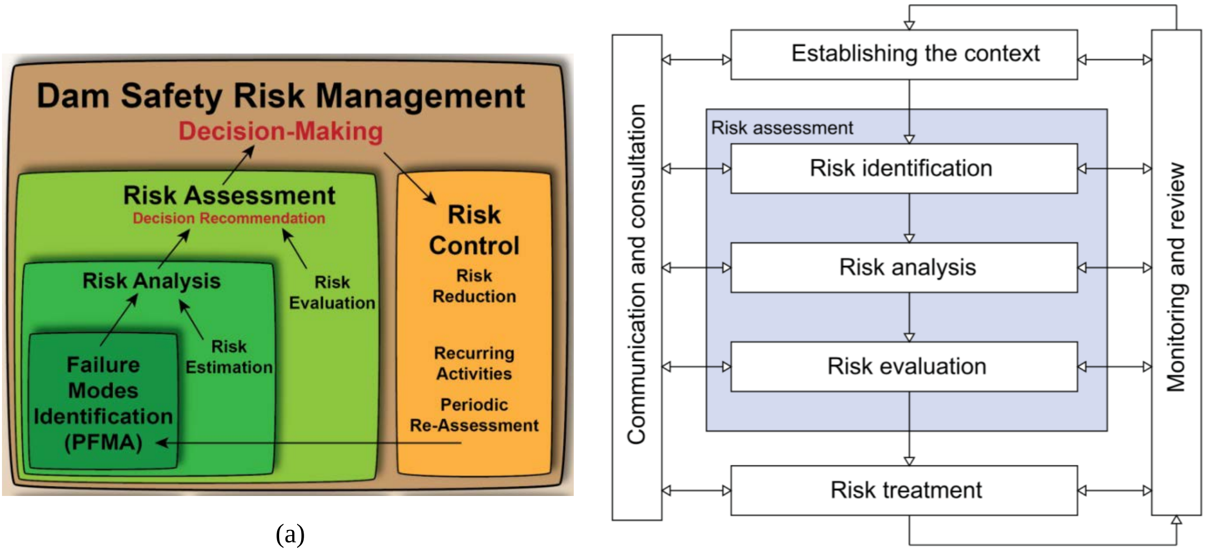
> Bureau of Reclamation and USACE (2019). *Best Practices in Dam and Levee Safety Risk Analysis, v4.1*. US Bureau of Reclamation, Denver.

Section numbers referred to here are taken from the most recent update of the Best Practices, dated July 2019.

The Panel did not review individual engineering models of failure mechanisms, for example, for the reliability of gate structures, geotechnical slope instability, or seepage and piping failure in embankment dams. These were not briefed to the panel, and the panel provides no endorsement or critique of them.

6.1 Risk Management

The risk management process at all three agencies is illustrated by Figure 7. Even though ISO 31000 is not cited in the FEMA document, the process, including monitoring and review, is essentially identical (ISO, 2009). This ISO 31000 process has been cited in the state of Victoria (Australia) dam safety guidance (DELWP, 2015), the Oroville Recovery Project (Craddock & Duval, 2018), the Central Water Commission of India's guidelines for dam safety (Central Water Commission, Government of India, 2019), and used for geotechnical risk management in Sweden (Spross et al., 2018) and elsewhere (Cruz & Rodovalho, 2019).



(a)

(b)

Figure 7. Elements of a Risk Management Process: (a) USACE-USBR process (FEMA, 2015); (b) ISO process (ISO 2014)

6.2 Risk Analysis

The approaches to risk analysis for dam safety used by the Agencies have evolved over a span of approximately 30 years. This history is summarized in Section 3.1. The developments at the Agencies are not independent, and as a result the approaches are much alike, although differing in sometimes subtle ways.

Development of this risk analysis approach has generally followed contemporary methods in other branches of civil, environmental, and mechanical engineering. However, specialized practices and methodologies have arisen, which make the RIDM approaches to risk analysis for federal dam safety programs somewhat unique.

The terms used by the Agencies in referring to the components of dam safety risk management are shown in Figure 7. The same terms are used here. *Risk Analysis* is taken to mean the process of developing quantitative appraisals of the probabilities of hazard loadings, fragility of the dam to those loadings, and consequences should the dam perform adversely. *Risk Assessment* is taken to be the process of evaluating the risk analysis results and other factors, and, if judged necessary, recommending corrective actions.

6.3 Probabilistic Approach

The important change to earlier standards-based practice that was introduced by the Agencies' movement to RIDM was the realization that, no matter how well-designed the structure or how carefully the structure is operated, there is always some chance of adverse performance or even failure, although that chance may be small. This was a fundamentally different concept, requiring methodologies that, up to that time, had not been widely used in dam engineering practice. Foundational among these new methodologies was the probabilistic approach. The probabilistic approach increased consistency in safety across different hazards, different sites, and different aspects of dam performance. It has led to an approach to quantify the priority and urgency of risk reduction action.

6.3.1 Decision confidence

Confidence is a term common in frequentist statistics. In the literature it refers to the probability that sampling results (observations) coincide with the true value of a population parameter. This generates a range (confidence interval) of the population parameter that is consistent with the observations. This is not to be confused with a probability distribution on the true parameter value, but rather it is a range that could have generated the observations with reasonably high probability, typically more than $p=0.05$. The use of the term *confidence* is confusing in the Agencies dam safety risk analysis, because the latter is usually based on degree-of-belief (subjective) probability (Vick, 2002). Degree-of-belief probability assigns a probability distribution directly to the true value of parameter(s), whereas frequentist statistics does not. This is not a simple distinction between aleatory and epistemic probability. In dam safety risk analysis, degree-of-belief probability commonly is used for both, and is the intuitive way of thinking about probability to which most engineers are accustomed.

Within the context of dam safety risk analysis, it seems more useful to speak of "being confident in making the right decision without further study." This seems to be the way the term is used by the Agencies, based on our briefings. Note that this is a different concept from statistical confidence as used in the first paragraph. *Being confident* then means that the probability of making a different decision following further study is small. The smaller this probability, the more confident a decision-maker. Such an interpretation distinguishes *confidence* from *uncertainty*, another term used by the Agencies. While uncertainty has to do with estimation (i.e., risk analysis), *being confident* has to do with decision-making (i.e., risk evaluation). To avoid confusion or misinterpretation the Panel suggests introducing the term *decision confidence*.

6.3.2 Uncertainty in risk analyses

The *uncertainty* related to risk estimates and the suggested interpretation of *being confident* are closely related. The concept of "being confident in making the right decision without further study" presupposes that the uncertainty related to risk estimates is uncertainty that can be reduced through further study. Epistemic uncertainty can be reduced, but aleatory uncertainty is a property of nature and cannot be. New information will never allow a better-informed decision if there is no reducible uncertainty. In this circumstance, no decision-maker could have done better in hindsight, even if the decision turns out to be unfortunate. Decision confidence depends on what is known at the time a decision is made, and how the associated uncertainties are balanced against possible consequences. How reality turns out is simply a gamble.

It may be that *prior* uncertainty related to dam failure or to the severity of loss indicates that further study may be useful. That is, current probability ranges for event tree nodes can be thought of as ranges of potential *posterior* probabilities following further study. This is explained in greater detail in Appendix D. One would expect these changes from *prior* to *posterior* probabilities would narrow with further study (although that is an expected but not a necessary outcome of more study). Most of the nodal probability estimates in dam safety event trees can only be interpreted as degrees of belief (not frequencies), since they concern epistemic uncertainties. These ranges of epistemic probability are expected to become tighter with more study.

If the probability ranges shown on risk plots were interpreted as statements about frequency combined with associated epistemic uncertainty in the probability parameters of those frequencies, these bandwidths would be a combination of natural frequency (which is a property of nature and irreducible) with reducible parameter uncertainty, only the latter of which could change with more study. Some of the uncertainty would be reducible and some would not be. This could be confusing to decision-makers, because the bandwidths would be a combination of two things.

6.3.3 Panel recommendation

The Panel recommends the Agencies link terms such as *uncertainty* and *confidence* to definitions commonly used in probability, statistics, and decision science. The resulting definitions should be mutually agreed to by the three agencies. Clear definitions are essential for moving the methodology forward and for communicating among its users. The term *confidence* is used in Agency practice to mean something akin to, “the probability that a decision would change upon further analysis.” This is a useful definition, but the word “confidence” might be confusing to non-Agency stakeholders, risk analysts, and statisticians. A better term with some history of use in other industries might be *decision confidence* (Dan & Fleming, 2018).

6.4 Potential Failure Modes Analysis

PFMA is the initial process in risk analyses completed by the USACE and Reclamation, whether SQRA or QRA. In the FERC interim RIDM guidelines (FERC, 2016), PFMA is similarly identified as the precursor for SQRA or QRA. In addition, FERC requires PFMAs for all high- and significant-hazard dams in its jurisdiction, even if the licensees choose not to use RIDM, as has been the case since 2002.

The term PFM as used to date in the Agencies’ PFMA practice is a chain of events leading to uncontrolled release of stored water. This use of the term *failure mode* in federal dam safety practice differs from the use of the term in many other hazardous industries, where a *failure mode* is thought of as a systems state. An example of a system state would be unavailability of a vertical-lift gate. A *failure cause* or *mechanism* in those other industries is the process by which the system comes to be in that state. For example, a high-inflow event leads to floating debris, which bypasses a log boom and jams the gate. The PFM construct in federal dam safety practice would more likely be called a *failure cause* or *failure mechanism* in many other industries.

PFMA in dam safety practice commonly involves a team with diverse technical expertise and a senior facilitator. A PFM consists of:

- An initiating event or circumstance (a hazard),
- A failure mechanism, and
- An uncontrolled reservoir release.

The typical hazards considered in PFMAs are normal static loading (“sunny day” conditions), extreme hydrologic loadings, and seismic loadings. Maintenance issues (e.g., deterioration) and operational issues (e.g., instrumentation, automatic controls, and human error) are not expressly considered as hazards, but may be considered as steps or events or contributing factors in the sequence of events that define the failure mechanism. PFMs are identified by judgment and summarized in a narrative scenario[6] (Blackett, 2018), for example:

> During normal maximum reservoir elevation, a continuing increase in uplift pressure on the shale layer slide plane initiates sliding of the buttresses. Major differential movement between two buttresses takes place causing the deck slabs to become unseated from their simply-supported condition on the corbels. Two bays quickly fail followed by failure of adjacent buttresses due to lateral water load resulting in an uncontrolled release of the reservoir.

---

[6] More recently, some practitioners have begun to use a bullet form listing of the sequence of events in a PFM instead of the narrative description.

The PFMA process begins with identifying an initial list of PFMs, often compiled through a brainstorming process. The initial list of PFMs can be quite long, and typically is reduced to a smaller list for detailed consideration. There appear to be different processes used for winnowing the initial list. PFMs can be removed from further consideration because they are judged after consideration to 1) not be credible, 2) be extremely unlikely [often referred to as “remote”], or 3) pose negligible consequences. The rationale for these judgments is normally documented in the PFMA. The USACE reduces the initial list by identifying those PFMs the team considered to be “risk driving” PFMs. It is not clear to the Panel if identifying risk driving PFMs could leave some PFMs that might have risk near to guidelines, but less than the risks of the risk driving PFMs.

Lists of *more likely* and *less likely* specific key factors for development of the PFM are compiled for each PFM selected for detailed consideration. This has the same features as a pros-and-cons list. In a QRA, the factors are typically listed separately for each step in failure mechanism or node in the resulting event tree, rather than in aggregate for the entire PFM. For the FERC the PFMA process completed outside of the RIDM process, more likely and less likely factors are typically listed in aggregate for the entire failure mode sequence.

The focus on PFMs being a sequence of events that leads to uncontrolled release of the reservoir can lead to eliminating from detailed consideration a failure mechanism that might initiate, stop short of reservoir release, but still have significant economic and societal impacts. This was identified in the investigation of the 2017 Oroville Dam spillway incident (France et al., 2018). The event that initiated that incident, failure of the service spillway concrete chute, was identified in a PFMA completed before the incident, but it was not carried forward because the team judged that the service spillway chute failure itself would not likely lead to a reservoir release. That judgment may well have been correct; however, the failed chute compromised the owner’s ability to manage the reservoir level during significant inflow to the reservoir. The service spillway’s compromised condition ultimately led to flow over the emergency spillway, erosion immediately downstream of that spillway, and evacuation of nearly 190,000 downstream residents. The incident cost the owner more than $1 billion in direct cost for response and repair and substantial damage to its reputation with the public, and produced significant stress and trauma to the downstream residents. The owner also faces lawsuits related to alleged indirect damages. Although the initial spillway chute failure does not meet the definition of a PFM in the current dam safety PFMA process, it clearly had significant consequences, though thankfully no loss of life.

Other non-reservoir-release events that could have significant impacts include component or system failure (e.g., gate failures, operating equipment failures) and human errors. Excluding such events from the current PFMA process represents a shortcoming in dam safety risk management programs in the Panel’s opinion.

Recommendations - The Panel recommends the Agencies consider the following revisions to the PFMA process:

1. Broaden the PFMA scope to include consideration of non-reservoir-release events that have significant consequences.
2. Standardize and better define the criteria for eliminating PFMs from detailed consideration.

6.5 Quantitative Risk Analysis

The overall methodological approach to risk analysis adopted by the Agencies conforms at a high level to modern international practice. It uses what is called in more generic practice a threat-vulnerability-consequence model (NRC, 2010). In dam safety terms these would more usually be described as hazard-fragility-consequence, but the meaning is the same.

6.5.1 Hazards

The primary hazards used to initiate event trees are hydrological (reservoir level) and seismological (earthquake-induced ground shaking). For both hazards there are well-known and widely-used probabilistic models the Agencies employ, and the Agencies have helped develop. Application of these models by the Agencies is at least state-of-practice and, in many cases, state-of-the-art.

Other hazard models such as those affecting gate performance, floating debris, ice loads and icing, and those affecting rip rap and surface erosion such as wave runup, seem to be less frequently used by the Agencies, but are nevertheless also state-of-practice.

Hazards associated with human performance and operational risk seem for the most part to appear only infrequently in the Agencies' guidance and practice.

6.5.2 Reliability analysis and fragility

The principal approach for calculating derived distributions and probabilities in the Agencies' methodology seems to be analyzing statistical data on the hazard-side, and mostly subjective expert opinion elicitations (EOE) on the capacity side. There are, however, differences in approach across the different disciplines of structural, hydrology & hydraulics, geotechnical, and mechanical engineering. This is not necessarily poor practice, but it does mean probabilities generated by the different disciplines may need to be interpreted in different ways.

The geotechnical applications of uncertainty analysis appear to be those most uniquely dependent on EOE. At an earlier time in the history of probabilistic methods at the Agencies, geotechnical uncertainty analysis to a large extent used physics-of-failure reliability models, the parameters of which were assessed judgmentally or on the basis of modest amounts of data (USACE, 1995). Calculation of the probability of adverse performance or failure was made using approximations such as the first-order-second-moment method, and expressed either as probabilities of failures or as reliability indices (USACE, 1995, 1999). This seems to have changed over the past decade. Geotechnical probability calculations currently tend to be based on subjective EOEs informed by engineering analysis. That is, engineering analysis is performed and its results are subjectively interpreted to arrive at a probability. The Monte Carlo method is then used as a numerical way to combine the expert elicitations into derived probabilities or probability distributions. The use of Monte Carlo codes for combining probability distributions is common among essentially all industries using risk analysis.

The structural and mechanical engineering applications of uncertainty analysis appear to use physics-of-failure reliability models, parameters of which are taken from published data or are assessed judgmentally. However, as far as the Panel understands, there is currently little structural reliability USACE guidance since previous documents have been rescinded (Schaaf, 2005).

Arguably, this migration of geotechnical reliability practice from physics-of-failure models to judgmental probabilities might be traced to the importance of internal erosion and piping causes of failure in dam safety. The geotechnical profession lacks agreed-upon physics-based models with which to describe and analyze internal erosion failure mechanisms. As a result, descriptive phenomenological models have been used, for which numerical values are assigned intuitively. These phenomenological models are to a reasonable extent informed by the historical performance of dams, but that historical record is sparse given the relatively limited numbers of failures and the many degrees of freedom that attend each failure. Knowing a base failure rate is helpful when using direct elicitation. Validation of subjective assessments of the probabilities of seepage and piping failures are needed.

As a rough approximation, if the base rate of failure of modern, well-engineered dams is presumed to be about $p=10^{-4}$, and the proportion of dams that fail due to internal erosion is about 40 percent, then the default base rate would be about $p=4x10^{-5}$ (Foster et al., 2000). However, it is easy to be overconfident that "this case is different," without the help of models that can point to inconsistencies in our probabilistic thinking. It is good practice to check elicited internal erosion failure probabilities against base failure rates. They should be roughly the same on average. A coherent explanation needs to be made for the discrepancy if they are not.

6.5.3 Dependence modeling

The Agencies generally assume that uncertain variables and events are independent when estimating failure probabilities. This applies most visibly to handling failure modes and to handling uncertainties on nodal probabilities.

The occurrence of a potential failure mechanism may depend on uncertain variables or processes that also play a role in occurrence of other failure modes, such as hydrologic or seismologic loading, or operational and organizational factors. Accounting for the dependences among failure mechanisms may not strongly change a risk estimate when a single failure mechanism dominates the probability of dam failure, or the joint occurrence of multiple failure mechanisms does not lead to substantially greater consequences. This may explain why dependence only appears to be considered for multiple, equally-reliable components, such as multiple gate or pier failures.

The common cause adjustment procedure used by the Agencies for dealing with correlations among failure modes is approximate at best. It requires experts to estimate conditional probabilities that are already implicit in the underlying data. Accounting for the dependence between component failures can be done more simply and accurately by specifying dependence among the uncertain variables that underlie the failure probabilities for the individual components. Doing so would also eliminate the risk of internal inconsistencies.

In agency EOE practice, nodal probabilities are subject to uncertainty bandwidths. These uncertainties are treated as independent, yet randomly combining nodal probability estimates under the assumption they are independent may lead to results that violate the experts' overall degree-of-belief probabilities. Experts may agree on the probability of dam failure, but differ in their opinions on the probabilities of underlying causal factors. Randomly combining nodal probability estimates across experts will indicate there is considerable uncertainty, while, in fact, there is none. Even for a single expert, the uncertainties for different nodal probabilities may not be independent. Assuming they are may portray the expert as more uncertain than he or she really is. Furthermore, nodal probabilities are often elicited for distinct load intervals. Random sampling across load intervals may yield failure probabilities that are non-increasing with demand.

Recommendations - Considering the above, the Panel recommends the Agencies improve their methods for dealing with dependence, more specifically:

1. The dependence among failure modes (e.g. multiple seismic pier failures), and
2. The dependence among the uncertainties on nodal probabilities.

6.5.4 Systems and operational risks

A dam system comprises the structure of the dam itself along with various waterways past the dam (turbines, spillways, conduits, and others), and mechanical and electrical equipment for on-site operations. Mechanical and electrical equipment increasingly includes real-time sensors and supervisory control and data acquisition (SCADA) systems. In a broader context, the dam system may be considered to include the reservoir and its surrounding drainage, communication links, and the human organization responsible for operating the system, including on-site operators, the dispatch center, and agency policy makers. If the dam exists in a series of structures along a river course, then the entire cascade might be considered part of the "system."

The state and nature of these many components of a dam system do not remain static during the lifetime of the system, because of wear, aging, and maintenance, as well as changes to the surrounding infrastructure and to the society being served. The various components also interact to create demands on the system, e.g., hydraulic loads on spillway gates may be generated by interaction of many factors other than just reservoir inflows.

The unwanted consequences of dam operations, especially incidents rather than loss-of-containment failures, are often the result of unusual combinations of usual events rather than of extreme events (King, 2020). Extreme events have been the focus of traditional design and risk analysis methods. Operational risks have to do with the vagaries of unexpected coincidences that happen during operations. Some of these have to do with human factors; some with instrumentation and SCADA systems; and some with floating debris, ice, or other disturbances. Many or most are not predicted ahead of time, and many or most involve an improbable chain of interacting events. At least some number may lead to fatalities, if only of small numbers in individual incidents (Pritchard, 2014).

Learning from operational incidents and failures (see also Section 4.3.6 "Risk precursor and continuous learning programs") is an important part of a dam safety program. In the United States, there have been notable operational incidents and failures that should inform and improve dam safety management. Examples are the Taum Sauk Dam failure and the Oroville Dam spillway incident. At Taum Sauk, a complex system interaction between human actions, monitoring and control equipment, and civil infrastructure led to failure of the dam, and it was fortunate that no lives were lost. The Oroville Dam did not fail as a result of the spillway chute failure, but the incident resulted in massive social dislocation and repair costs.

One can accommodate an unusual sequence of operational events in an event tree analysis, but it is uncommon in practice. Probabilistic Risk Analysis (PRA) suffers the limitation that only specifically enumerated chains of events enter the analysis. An unforeseen or unusual combination not specifically identified does not. As a result, availability on-demand, reliability, durability, resilience, and maintainability of most dam systems remain uncertain. Maintenance policies change over the life of a dam. This makes it difficult to integrate flow control into static, risk-based approaches for assessing dam safety, assess their ability to meet performance goals, and find good corrective measures where inadequacies exist.

Difficulties remain despite some progress with consideration of operational incidents and failures, and the current guidance of Best Practices, Section H-1 provides little help for the analyst attempting to model or even appraise operations risks. Though human factors have been recognized as a significant hazard in dam safety for many years, they are still not usually integrated into dam safety risk analyses. The same observation can be made for monitoring and data transmission systems. The complex interactions between the various influencing factors have generally not been modelled.

Recommendations - The Panel recommends risk assessments include considering potential impacts of human error and organizational factors; ageing; malfunctioning of monitoring, remote, or automated electro-mechanical control equipment; and other events that may not result in dam failure, but could lead to casualties, major environmental damage, or significant response or repair costs.

6.5.5 Gate performance

Hydraulic gates are mechanical-electrical systems. As such, they lend themselves to analysis by fault trees, which is a common approach in dam safety risk analysis. The output of a fault tree analysis in the form of a failure probability is common, and is used as input to larger event tree models, which combine hazards, fragility, and consequences.

The unreliable performance of hydraulic gates is not uncommon, whether caused by the gate itself, its structural support, its hoist systems, or SCADA and automatic controls. Examples are documented in Best

Practices and by the National Dam Performance Program, although the statistical rates of such unreliability are poorly documented across the industry. The USACE Asset Management Program has collected data for many years on locks, gates, and their component reliabilities, especially in the navigation program (e.g., Table G-4-2 of Best Practices). This is an excellent start, and the data are badly needed, but the current compilation is not yet a mature source of such data.

Current federal practice treats reliability analysis of Tainter (radial) gates (Best Practices Section G-1) differently from the reliability of vertical lift, miter, and other gate systems (Best Practices Section G-2), and differently yet again from the mechanical and electrical components included in these gate systems (Best Practices Section G-4). Why this should be the case is unclear, and there seems a need to integrate these three sections of Best Practices into one document. Event trees apparently alone are used to analyze radial gates reliability, while fault trees combined with event trees are used to analyze the other gate systems. Best Practices Section G-4 addressing dam gate systems mechanical or electrical system components failure takes yet another approach by modeling these components in fault trees with traditional Weibull and Dormant Weibull models of physical equipment.

Recommendations - The Panel recommends that, since analytical methods used to appraise gate systems reliability seem well suited to the task, actuarial data can be used to estimate failure rates in lieu of subjective probability, where possible. In practice, these statistical databases are often inadequate, and developing such databases should be seen as both a research and a practical need for the industry.

6.6 Expert Elicitation

The Panel agrees with the Agencies that quantification of expert opinion elicitation (EOE) is critical to the enterprise of dam safety risk analysis. Our historical database is poor for many of the aleatory uncertainties affecting dam safety. Furthermore, many of the most important uncertainties affecting dam safety are epistemic, and, necessarily, based on the experts' degree-of-belief. This is consistent with modern risk analysis practices across a wide spectrum of industries. Ideally, such estimates are in part based on phenomenological models, if available, or explicitly supported by all available verifiable evidence used by the experts in developing their subjective estimates.

Recent decades have seen significant advances in EOE theory and practice. Use of formal protocols and recognition of cognitive biases in the quantification of subjective probabilities is critical to obtaining reliable numbers. The Agencies are strongly encouraged to adopt modern, state-of-the-art protocols for EOE. Quantified probabilities obtained are surely to be suspect absent such evidence-based protocols.

Experience in many fields has shown that use of physics-of-failure models with uncertainties on inputs and model outputs provides more calibrated probabilities than does direct assessment by experts. Such reliability models are commonplace in geotechnical engineering (slope stability, finite-element stress-strain, settlement prediction), mechanical engineering (gate reliability), structural engineering (reliability of concrete monoliths), hydraulic engineering (cavitation of waterways), and other fields. Probability estimation based on physics-of-failure models with uncertainties on model inputs and outputs is generally the preferred approach in many reliability settings. Among others, it is more verifiable and less prone to error caused by heuristics and biases. Only when such models are unavailable, should direct elicitation be the chosen alternative.

6.6.1 Elicitation of expert opinion in dam safety

The elicitation of expert opinion in the form of subjective probabilities is both necessary in dam safety risk analysis and prone to error and unreliability. Extensive and detailed literature has arisen on EOE art and practice over the past 30 years. This body of knowledge needs to be incorporated in the agency approaches. As Cooke (1991), among the pre-eminent researchers in this field, has observed,

> The most obvious way to elicit a subjective probability from someone is to ask them. While this method is surely the most common, it is equally surely the worst […]. Considerable thought must be given to the way in which probabilities are elicited.

Historically, decision-makers have used expert judgment to supplement data or analysis needed to inform their decisions (a longer discussion is found in Appendix E). Cases involving new engineering designs, rare events, and situations that are beyond our direct experience call for the use of expert opinion as a surrogate source of information. In the absence of adequate scientific information, decision-makers have to rely on their own intuition or on expert opinion. Therefore, experts can extensively influence key decisions in vital matters such as politics, economy, science, and engineering. Use of expert opinion is inevitable in almost all risk assessments. This includes judgments made by the risk analysts in modeling, gathering, and analyzing various types of information, and in assessing the model parameters. In all these steps, limitations or lack of established methods, scarcity of required data, and resources are some of the reasons for relying on subject matter experts.

An expert is an individual with specialized knowledge or skill in some specific domain. While in principle any degree of knowledge of a subject qualifies one as an expert to that degree, a person is called an expert only when he or she is believed to be much more knowledgeable than a layperson on the subject of interest. Expert opinion can be viewed as expression of the judgment of an expert on a subject or issue. An opinion is usually regarded as an impression, personal assessment, or subjective estimation of a quality or quantity of interest. Expert opinion, in contrast with factual information, is a judgment or a belief that, at least in the mind of the receiver of the opinion, is based on uncertain information or limited knowledge. It is a tool of last resort for exploring unknown issues that are otherwise inaccessible.

Despite the clear need for expert opinion, it is always applied with much caution. This is because an opinion is not a fact or verified by experiment. It is a person's subjective assessment about a specific subject. The obvious concern is the degree to which an expert’s opinion correlates with objective reality. In risk analysis, where probabilities of rare events are often the subject of expert elicitation, probability underestimation or overestimation could result in costly decisions.

While there is a significant body of research on ways to improve the use of expert opinion, other factors such as limited familiarity with relevant research results and time and budget constraints have impacted the quality and credibility of using of expert judgment in practice. This section aims to offer an assessment of the current state of the practice and summarize important methodological findings that could help in improving use of expert opinion in risk studies.

6.6.2 Elicitation of expert opinions

Elicitation of expert opinions is the processes of selecting and eliciting the opinion of the experts, including selection criteria, expert panel size and composition, and elicitation procedures and protocols.

The available research results indicate the methods by which expert opinions are elicited can have a significant effect on the resulting estimates accuracy. The methods include a set of steps that may or may not be formally and explicitly followed in practice. Categorized into *formal* and *informal* methodologies, expert elicitation methods may range from simple to complex processes. Formal expert elicitation refers to a structured procedure designed to obtain opinions of experts including, for example, training material to debrief the experts, documented criteria for selecting experts, and defined roles and sets of steps and protocols for information exchange and interactions. Compared to an informal expert elicitation process, formal expert elicitation can increase results quality and credibility, and the defensibility of the judgments, in part due to documentation of each expert’s rationale. It also enhances communication of the results. Formal elicitations are generally viewed to produce superior results compared to those informally gathered (Dewispelare et al., 1995).

Informally elicited judgments are obtained through unstructured approaches that lack adherence to established protocol or scientific principles. Examples of these types of judgments include intuition, arbitrary guesses, and gut feelings (Armstrong, 1985, p. 73). Unfortunately, the most commonly used method for decisions under time pressure is unaided judgment. This is not surprising, as unaided-judgment forecasts can often be derived quickly and cheaply. The simplest and most common of these methodologies entails merely asking for an individual's judgment. Most people have poor intuitions regarding numerical probabilities. Consequently, this inquiry yields the least reliable expert performance results, especially for persons unfamiliar with probability concepts (Cooke, 1991).

6.6.3 Post-elicitation processing and utilization of expert opinions

Post-elicitation processing and utilization of the expert opinions is the process of relating the elicited expert opinions to the unknown of interest, including how to use the information / estimates provided by the experts, and also information about the experts, in order to estimate the unknown quantity; and, in case of multiple experts, how to aggregate the opinions.

Based on earlier discussions, we know that some modification to the estimates from experts may be needed to improve accuracy - for instance to adjust for potential bias or to compensate for overconfidence. This can be done by simple "ad-hoc" methods, or in a formal and mathematically structured way via Bayes' theorem. Either way, some level of judgment is needed on the part of the analyst (or decision maker) on the degree and form of the adjustments (post elicitation "calibration") to the expert estimate. Also, when the opinions of several experts are obtained, an aggregation is needed to form one estimate ("aggregated opinion") considering uncertainties associated with the potential diversity of opinions.

6.6.4 Improving the quality of expert judgment

The following list is a summary of several practical recommendations that can be distilled from the literature on the use of expert opinion, particularly in risk analysis of complex systems and processes.

1. Select good domain experts, train them on normative aspects.
2. Aggregation of opinion of multiple experts tends to give more accurate results than the opinion of a single expert. An optimum size of the panel on any single subject is three to five.
3. Mathematical methods of aggregation are generally preferable to behavioral methods for reaching consensus.
4. Quality of judgments can be substantially improved by following a formal and structured elicitation process.
5. Quality of expert opinion can be substantially improved by decomposing the problem into a number of more elementary problems.
6. There is a significant improvement in the overall results if the initial problem definition and decomposition is done with care and in consultation with the experts.
7. Experts' opinions are subject to bias and overconfidence. Effective techniques to reduce overconfidence are: a) using calibration techniques, and b) encouraging experts to actively identify evidence that tends to contradict their initial opinions.
8. Sources of strong dependencies among experts need to be identified. Weak dependence does not seem to have a major impact on the value of multiple expert judgment.

Current Reclamation and USACE guidelines (Bureau of Reclamation & USACE, 2019) adequately cover some but not all of the above insights and recommendations. In particular items 1, 2, 3, 7, and 8 deserve closer attention and more detailed up to date coverage in agency guidelines.

6.6.5 Panel recommendations

The Panel recommends the Agencies continue to rely on expert elicitation of risk estimates when alternative estimation methods are unavailable, which is consistent with modern risk analysis practices across a wide spectrum of industries. In addition, the Panel recommends the Agencies collaborate in developing a set of practical, state-of-practice guidelines to improve credibility and quality of risk estimates elicited from subject matter experts. Such guidelines could include:

1. Process and criteria for expert selection, including guidance on size and composition of a panel of experts.
2. Elicitation protocols and procedures.
3. Expert calibration and training on normative precepts, such as fundamental concepts of probability and uncertainty quantification.
4. Methods for achieving consensus or generating results representative of the spectrum of opinions, with explicit consideration of potential for bias and overconfidence.

6.7 Human Factors in Risk Management

The core concepts and overview of the state-of-the-art in human reliability analysis is more fully discussed in Appendix F. According to various statistics (see for example Kariuki & Löwe, 2007; Wiegmann & Shappell, 2012), human error and organizational factors dominate the causes or contributing factors of accidents and incidents in complex engineered systems. This may be attributable to the fact that humans play a central role in all phases of any project's life cycle.

6.7.1 Human error as a contributor to risk

Alvi (2013) offers the following general observations regarding past failures of dams (and other complex systems) from the perspective of human-system interactions:

> […] major failures are usually preceded by a series of steps involving physical and human factors interacting over a relatively long period of time, often years or even decades.
>
> The interactions among physical and human factors are often not simple and linear. Instead, they may be complex and involve nonlinear relationships, feedback loops, causes having multiple effects, effects having multiple causes, and a lack of distinct root causes or dominant contributing factors.
>
> Interactions among physical and human factors usually generate warning signs [that] are not recognized, or not sufficiently acted upon, prior to the failure.
>
> Physical processes deterministically follow physical laws, with no possibility of physical "mistakes." Therefore, failures - in the sense of human intentions not being fulfilled - are fundamentally due to human factors, as a result of human efforts individually and collectively "falling short" in various ways. A story of *why* a failure happened therefore cannot be complete without reference to contributing human factors.
>
> A natural tendency is for systems to move towards disorder and failure, in line with the concept of increasing "entropy" in physics. Therefore, systems such as dams are typically not inherently safe, and continual human effort is needed to maintain order and prevent failure.

The existing guidelines on risk analysis by the Agencies do not however include any discussion of human error as a contributing factor to dam safety risk. This means any impact of human error on estimated risk levels is at best accounted for only implicitly through the values of other parameters of the risk model. Appropriate delineation of causal factors is essential to fully benefit from risk analysis and effectively apply RIDM, as it provides a basis for risk reduction and cost benefit tradeoffs through consideration of removing or reducing likelihood of the identified risk factors. This obviously applies to human and organizational causes of failure.

There are currently no databases of dam failure events where contributions from human error are systematically gathered and characterized. However, examples can be found in failure investigation reports, where human errors, often combined with organizational and process deficiencies, are mentioned as contributing causes.

Peck (1973), France, et al.(2018), and Alvi (2018) include cases that can be broadly characterized as design error and operational deficiencies, which in common taxonomies of human reliability are associated with human failure in technical judgment, conducting analysis, following procedures, or executing actions.

6.7.2 Human reliability analysis

Human reliability analysis (HRA) involves qualitative and quantitative assessments of human error (or human failure events) in the context of probabilistic risk analysis. The estimated human error probabilities are entered into the PRA models, such as event trees, system failure logic models (e.g., fault trees), process models, and barrier / control models. HRA relies on knowledge of the design and operation of the system with which the human interacts. HRA as a modeling and analysis activity within a PRA generally requires tight interaction with other PRA modeling and analysis tasks, to ensure appropriate representation of potential human errors in the risk models.

Despite the significance of human error in accident causation, many risk assessments of complex systems do not include human error in the risk models. Nuclear power, aerospace, and process industry are exceptions. In fact, the majority of methods for performing HRA have been developed for use in nuclear power plant PRAs. There are currently more than 40 HRA methods, although some are essentially variations of the same approach. Human reliability generally involves three major activities: 1) human error identification, which concerns identifying erroneous human action in the context of a given scenario and a set of human tasks, 2) human error quantification, which aims to assess likelihood of errors occurrence, and 3) identification of the causes and context of the identified errors, to support development of preventive or mitigating measures to reducing the error likelihood (Kirwan, 1994).

HRA is a challenging task for two principal reasons. First, there is no consensus methodology that applies to the full range of human actions of concern in PRA. Second, conducting a detailed HRA for a risk-significant human action could be resource and time consuming. Such a specialized task could involve forming a dedicated team of experienced analysts. Therefore, the HRA task needs to strike a good balance between adequate consideration of the state-of-the-art in HRA methodology and the importance of human actions in the risk models.

6.7.3 Panel recommendations

The Panel recommends the Agencies develop a graded human reliability analysis method for use in dam safety risk assessments, leveraging methods developed by other industries, particularly the nuclear power, petro-chemical, and aviation sectors.

6.8 Life Loss Estimation

This section presents an overview and discussion of the methods the Agencies use for estimating life loss. Life loss estimation typically involves a series of steps:

1. Breach development modeling,
2. Inundation modeling,

3. Warning and evacuation modeling, and
4. Life loss modeling.

Each of these steps builds on the results of previous steps. For instance, the rate of breach development influences the characteristics of the flood wave, which impacts the flood wave's arrival time, flow velocity, and depth. The arrival time affects the probability of timely warning and evacuation, which, together with the hydraulic conditions, influences the resulting life loss estimate.

6.8.1 Overview of USACE practice

In its periodic risk assessments, issue evaluation studies, and modification studies, USACE uses its HEC River Analysis System (HEC-RAS) for breach development and inundation modeling (USACE, 2016). The HEC-RAS outputs are then the inputs for HEC-LifeSim, a tool developed by USACE for evacuation and life loss modeling (USACE, 2020).

The simulation methods included in HEC-RAS are in line with the international state-of-practice in breach development and inundation modeling. HEC-RAS is also used outside USACE, nationally and internationally, for mapping potential inundation zones and estimating the potential consequences of flooding.

HEC-LifeSim is an agent-based model for estimating life loss. It includes a traffic simulation model that simulates how people try to leave the area at risk over the road network upon receiving a flood warning. When the flood wave reaches an individual and immobilizes a person, loss of life is evaluated on the basis of the person's vulnerability (e.g., on the road, in a building, on foot) and the hydraulic conditions (depth and velocity) at that location. The backbone of HEC-LifeSim simulations is a warning and mobilization timeline that determines how the road network is loaded. It is determined by the following parameters:

1. **Imminent hazard identification time:** the time it takes before the hazard is identified.
2. **Hazard communication delay:** the time it takes before the hazard is communicated to the emergency management agency (EMA) responsible for issuing a public notification.
3. **Warning issuance delay:** the time it takes before the EMA issues a warning based on the information it has received.
4. **First alert diffusion delay:** parameter related to the speed with which the warning spreads through the community.
5. **Protective action initiation delay:** parameter related to speed with which individuals respond upon receiving warning.

The estimates of HEC-LifeSim input parameters are adjusted to local circumstances, based on interviews with local EMAs and other parties involved in emergency management. A standardized questionnaire is used for generating the warning and mobilization timeline. While local stakeholders are typically not further involved in the life loss estimation process, the final results are later presented to them. It is understood this often stimulates EMAs to think about emergency preparedness and evacuation plans. Here, it helps that that HEC-LifeSim not only provides numerical outputs, but also visualizes what could happen if a dam were to fail.

The validation of life loss models is notoriously difficult due to the limited number of well-documented historic flood events. For validating the present version of HEC-LifeSim (version 2.0) the USACE has compared model predictions to the actual life loss for three well-documented case histories: the New Orleans east levee failure in Hurricane Katrina (August 2005, 68 deaths), the Kelly Barnes Dam failure (November 1977, 39 deaths) and the Joso levee failure (September 2015, 2 deaths) (Needham et al., 2020; Risher et al., 2017). While the model predictions for the first two cases are in line with the actual life loss, life loss was overestimated by an order of magnitude for the Joso levee failure. This can largely be explained by the fact that rescue is not modeled explicitly in HEC-LifeSim. A recent evaluation based on data from the Oroville spillway incident concluded the modeling of warning issuance delay, first alert diffusion delay, and protection action initiation delay is adequate for short-duration events, but further work is needed for events in which warning and protective action delay exceed 4 hours (Sorensen et al., 2018).

The uncertainties related to breach development and resulting flood wave are accounted for by carrying out simulations for different starting conditions, such as different pool levels. Life loss is then estimated for each of these scenarios in HEC-LifeSim. Nonetheless, other uncertainties such as breach location, shape, dimensions, and development time are at least as important. The current methods of assigning values to these parameters remain primitive, so characterization of uncertainty remains crude.

The uncertainty related to the life loss for a given inundation scenario is modeled using Monte Carlo simulation. Uncertainty is placed on the parameters that determine the warning and mobilization timeline, as well as variables that determine individual behaviors and vulnerabilities. The uncertainty related to potential model structure errors (model uncertainty) is not accounted for explicitly.

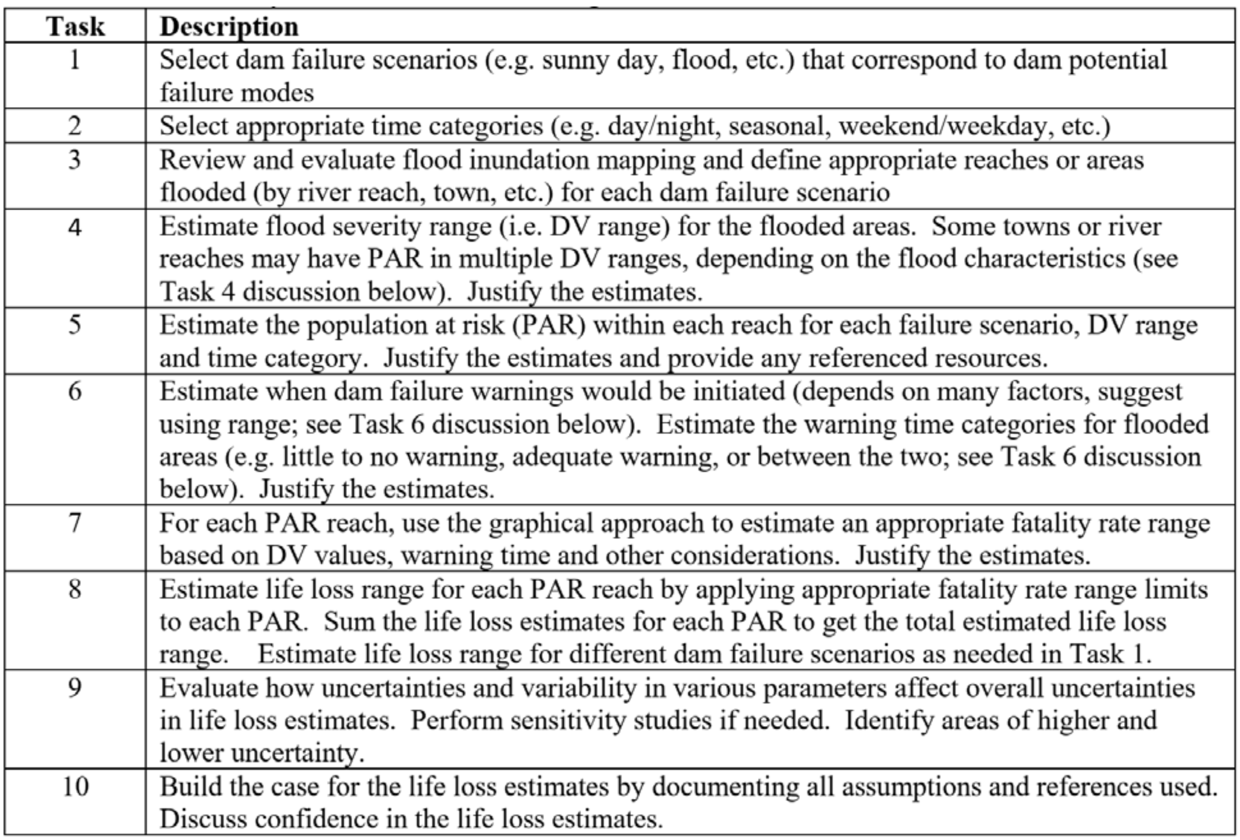
The Panel understands uncertainties on the input parameters of HEC-RAS and HEC-LifeSim are such that the uncertainty related to the estimated life loss for a dam failure is typically over an order of magnitude: often two, sometimes more (J. Needham, personal communication, March 17, 2020). The same uncertainties related to potential life loss are plotted on F-N and F-$\overline{\text{N}}$ charts. This is discussed in greater detail in sub-section 6.8.2.

6.8.2 Overview of Reclamation practice

Reclamation uses MIKE software, developed by the Danish Hydraulic Institute (DHI), for inundation modeling. These are state-of-the art models used worldwide. Reclamation primarily uses the Reclamation Consequences Estimating Methodology (RCEM) (Bureau of Reclamation, 2015) for estimating life loss for inundation scenarios. RCEM builds on Dam Safety Office (DSO) DSO-99-06, a procedure developed in 1999 for estimating life loss (Graham, 1999). RCEM, like its predecessor, is an empirical method based on dam failure case histories and some large flood events (e.g., flooding of New Orleans due to hurricane Katrina).

An overview of the RCEM process is given in Figure 8 below. The RCEM fatality rate charts relate products of depth and velocity (referred to as flood intensity or DV) to fatality rates. These charts differentiate between cases with little or no warning, cases with partial warning, and cases with adequate warning. This explains why the fatality rates from Task 7 of the RCEM process are applied to the pre-evacuation population at risk.

Recognizing the limitations of the RCEM fatality rate charts for high-consequence dams, Reclamation has twice used the Life Safety Model (LSM) for estimating fatality rates for use within the RCEM process. LSM is an agent-based model that represents people's interactions with a flood event (Assaf & Hartford, 2001; Lumbroso et al., 2011), similar to HEC-LifeSim. It was originally developed by BC-Hydro, a Canadian power utility based in the province of British Columbia that operates several high-consequence dams. BC-Hydro later entered into a research and development agreement with HR Wallingford, where the model is currently available for download. The Panel understands Reclamation will move from LSM to HEC-LifeSim for estimating fatality rates when the RCEM fatality rate charts are considered insufficiently applicable, as the agency now considers HEC-LifeSim superior to LSM (B. Feinberg & J. Major, personal communication, April 1, 2020). Unlike the USACE, Reclamation only uses agent-based modeling for special cases, and it only does so for estimating fatality rates. These fatality rates are then used in the broader RCEM process to estimate life loss. The RCEM process thus forms the basis for life loss estimation for all of Reclamation's dams, for consistency in its portfolio of dams.



Task	Description
1	Select dam failure scenarios (e.g. sunny day, flood, etc.) that correspond to dam potential failure modes
2	Select appropriate time categories (e.g. day/night, seasonal, weekend/weekday, etc.)
3	Review and evaluate flood inundation mapping and define appropriate reaches or areas flooded (by river reach, town, etc.) for each dam failure scenario
4	Estimate flood severity range (i.e. DV range) for the flooded areas. Some towns or river reaches may have PAR in multiple DV ranges, depending on the flood characteristics (see Task 4 discussion below). Justify the estimates.
5	Estimate the population at risk (PAR) within each reach for each failure scenario, DV range and time category. Justify the estimates and provide any referenced resources.
6	Estimate when dam failure warnings would be initiated (depends on many factors, suggest using range; see Task 6 discussion below). Estimate the warning time categories for flooded areas (e.g. little to no warning, adequate warning, or between the two; see Task 6 discussion below). Justify the estimates.
7	For each PAR reach, use the graphical approach to estimate an appropriate fatality rate range based on DV values, warning time and other considerations. Justify the estimates.
8	Estimate life loss range for each PAR reach by applying appropriate fatality rate range limits to each PAR. Sum the life loss estimates for each PAR to get the total estimated life loss range. Estimate life loss range for different dam failure scenarios as needed in Task 1.
9	Evaluate how uncertainties and variability in various parameters affect overall uncertainties in life loss estimates. Perform sensitivity studies if needed. Identify areas of higher and lower uncertainty.
10	Build the case for the life loss estimates by documenting all assumptions and references used. Discuss confidence in the life loss estimates.

Figure 8. RCEM Procedure (Bureau of Reclamation, 2015)

The RCEM fatality rate charts are based on 60 case histories. From these 60 case histories, 80 data points (DV and fatality rates) have been estimated; since two of these data points plot outside the charts, only 78 of these appear on the RCEM charts (Bureau of Reclamation, 2014a). While the principal interest is direct life loss (i.e., life loss from exposure to dam failure floodwater), the fatality rates from case histories may sometimes be influenced by recordings of indirect life loss (i.e., life loss from stress-induced medical conditions, water-borne sicknesses and infections, etc.). The uncertainty on life loss estimates is based on judgment. It is typically an order of magnitude, up and down (B. Feinberg & J. Major, personal communication, 1 April 2020).

6.8.3 Overview of FERC practice

FERC, as a regulator, only runs inundation and life loss models for internal use. Its licensees are responsible for preparing inundation maps. Licensees can use any model they want, since FERC is not allowed to prescribe particular models. However, licensees have to lay out an explanation for all the choices and assumptions they make, and the results need to be acceptable to FERC. In practice, most licensees now use HEC-RAS for inundation modeling. Occasionally, they use other models such as DHI's MIKE models. FERC licensees have conducted less than 10 life loss studies so far. Most of these were performed using HEC-LifeSim. RCEM is currently being applied to a low-consequence dam (E. Gross, personal communication, 9 April 2020).

FERC's draft guideline on life loss estimation (FERC, 2014) is outdated (e.g., it references the DSO-99-06 method, rather than RCEM), and it lacks the level of detail needed for guiding licensees with little background in life loss estimation. This does not appear to be an issue at present. The risk analyses overseen by FERC are still pilot projects, carried out voluntarily by licensees committed to doing it right. Things may change if FERC licensees take up RIDM in greater numbers. Given consequence estimates sensitivity to minor modeling details, a tried and tested consequence modeling guideline seems essential before FERC can use RIDM more broadly for regulating the safety of dams.

6.8.4 Discussion: comparing life loss estimation practices

Internationally, there are different models available for estimating life loss from floods, see e.g., (Jonkman et al. 2008, 2016) for an overview and comparison. The methods used by the USACE and Reclamation sit on opposite sides of the spectrum. USACE's HEC-LifeSim is an agent-based model, in which estimates of life loss are obtained from aggregating the outcomes for individuals whose behaviors are modeled individually. Reclamation's RCEM is an empirical model that applies fatality rates to populations at risk. While uncertainty due to potential model structure errors (model uncertainty) is arguably smaller in agent-based models with numerous input parameters (explanatory variables), parameter uncertainty is greater. Overall, both approaches may yield comparable results when accounting for model and parameter uncertainties.

The accuracy of an empirical model depends crucially on the degree to which the historical cases on which it rests are representative of the dams being analyzed. The RCEM charts seem less suitable for very high-consequence dams, located upstream of densely populated urban areas. Only 10 percent of the RCEM case histories have a total population at risk of at least 50,000. Additionally, about 80 percent of the RCEM case histories are over 30 years old, while developments in early warning and evacuation have not stood still. The Panel understands that Reclamation will move towards the use of HEC-Life for supporting the estimation of fatality rates for such cases (B. Feinberg & J. Major, personal communication, April 1, 2020). The Panel also understands that the limited number of recent case histories for dams located upstream of large population centers has led the USACE, which manages a portfolio with many of such dams, to move away from its former empirical approach (a modified version of DSO-99-06), and adopt HEC-LifeSim for life loss estimation (Bureau of Reclamation & USACE, 2019; J. Needham, personal communication, 17 March 2020).

The accuracy of an agent-based model depends crucially on the degree to which it realistically models the actions of individuals and organizations. As indicated by the Joso levee failure validation study, the potential for rescue appears to be an area in which HEC-LifeSim could be further improved. Accepting model uncertainty and modeling it explicitly could be an alternative to adding detail.

Validation studies could serve as a starting point for quantifying model uncertainties, since the differences between well-documented historic events and simulation results are largely due to model uncertainty, not parameter uncertainties. The results of validation studies can be used for identifying opportunities for model improvement and /or for quantifying model uncertainties. Accepting (and quantifying) model uncertainty may sometimes be a cost-effective alternative to building an ever more refined model. Methods have been introduced in recent years on how to incorporate results of validation studies and other types of evidence about accuracy and credibility of models in assessing model uncertainty. When such data and evidence are limited, Bayesian methods for model uncertainty quantification are most appropriate. Examples can be found in (Mosleh et al, 1995; Mosleh and Wood, 2014; and Droguett and Mosleh, 2014).

HEC-LifeSim and RCEM both have their pros and cons. While RCEM is relatively simple to use, it relies heavily on judgment and only provides useful guidance for dams that closely resemble RCEM's case histories. For high-consequence dams, RCEM should be used with caution or supplemented with other estimation methods, as Reclamation does. HEC-LifeSim is more demanding, but more broadly applicable. Its visual outputs can also help create awareness and stimulate discussions with EMAs on improving disaster preparedness.

Considering the above, the Panel is not in a position to judge in favor of HEC-LifeSim or RCEM, or any other life loss estimation model. What works best will depend on local circumstances and the level of experience within each agency or organization.

6.8.5 Panel recommendations

The Panel recommends continued use and improvement of HEC-LifeSim and RCEM, with an eye for uncertainties in life loss estimates and the level of detail needed for decision-making. More specifically, the Panel recommends that the Agencies:

1. Continue to conduct validation studies if useful cases present themselves. Such cases are not limited to dam failures that cause life loss. Events that trigger mass evacuations may offer valuable opportunities for validating and improving evacuation models, as illustrated by the study into the Oroville spillway incident (Sorensen et al., 2018).
2. Perform benchmark studies to identify opportunities for mutual learning, i.e., applying different models to identical cases and comparing findings. Such benchmark studies do not have to be limited to comparisons of life loss estimates, but also encompass comparisons of the estimates for underlying parameters such as warning times and evacuation rates. Only a few examples of such comparative analyses exist to date (see e.g., Kolen et al., 2016; Perdikaris and Zhou, 2018). The Panel understands that USACE is teaming with USSD and other federal agencies to host a consequences benchmarking workshop at the 2022 ICOLD conference. The Panel fully supports this effort.
3. Include model uncertainty in HEC-LifeSim uncertainty analyses. HEC-LifeSim currently accounts for uncertainties on input and model parameters, but not for uncertainty related to the model structure itself. Even if all model inputs were known with perfect certainty, evacuation rates predictions could still differ from actual realizations. The differences between simulation results and the actual life loss for well-documented historic flood events (e.g., Needham et al., 2020) could serve as a starting point for estimating model uncertainty.

In addition, the Panel recommends that FERC update its guidelines on life loss estimation. An up-to-date guidance document is indispensable for moving RIDM into the mainstream of dam safety regulations.

7 AGENCY METHODOLOGY PART 2: RISK ASSESSMENT AND COMMUNICATION

This section covers the assessment of risk based on the risk analysis methodology discussed above and the communication of risk results to stakeholders.

7.1 Portraying Societal Risk

The Federal Guidelines on Dam Safety Risk Management present two types of charts for portraying societal risk. These are commonly referred to as f-N and F-N charts. The f-N chart is referred to hereafter as an f-$\overline{\text{N}}$ chart, per USACE's ER 1110-2-1156 (USACE, 2014a) and FERC´s interim RIDM guidelines for dam safety (FERC, 2016). While the f-$\overline{\text{N}}$ and F-N charts have a similar appearance, they differ in important ways. Decision makers should be aware of these differences to avoid misinterpretation affecting decisions.

7.1.1 Overview of the basic properties of $f - \overline{\text{N}}$ and $F - N$ charts

USACE, Reclamation, and FERC portray societal risk using f-$\overline{\text{N}}$ charts (Figure 9). In such a chart, f is the probability of a failure event (*e.g.*, a specific PFM or a dam failure due to any PFM) and $\overline{\text{N}}$ is the expected (mean) value of life loss[7] (N), given that failure event. This means that $\overline{\text{N}}$ is the probability-weighted sum over all possible fatality numbers that could result from the failure event: $\overline{\text{N}}$ is a conditional expectation. The overbar is widely used in statistics to denote an expectation. An f-$\overline{\text{N}}$ plot is *not* a probability mass function (*i.e.* not an f-N plot), although these plots may look identical when each failure event has a single-valued consequence.

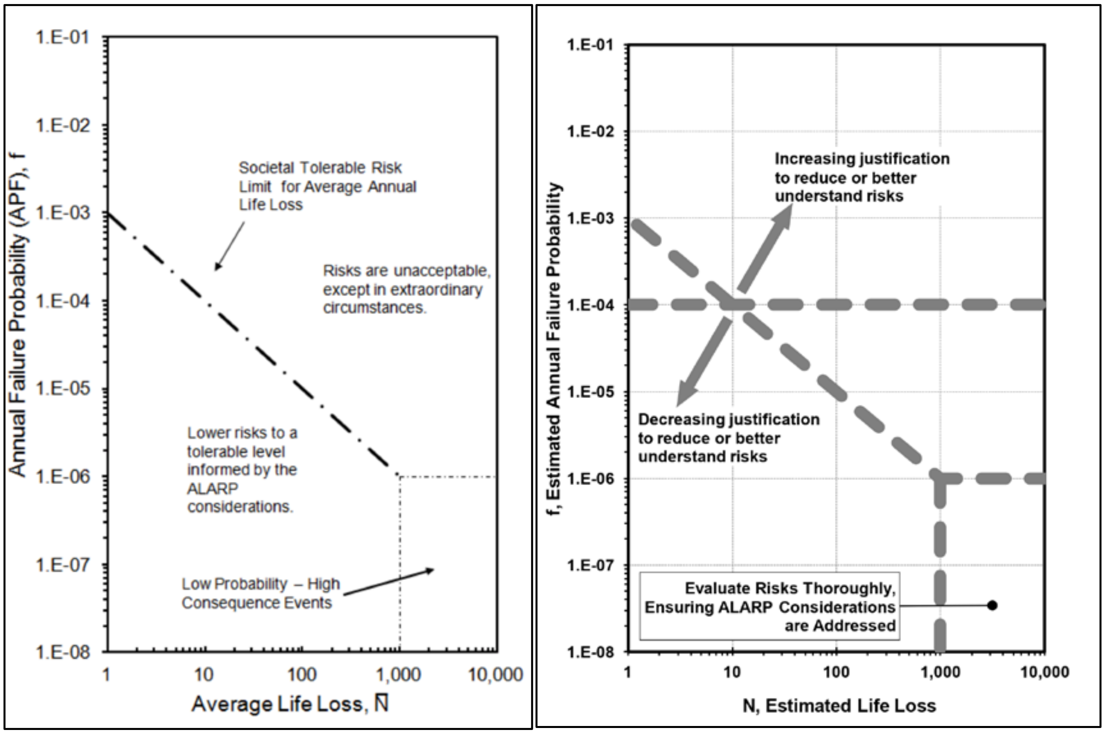


Figure 9. The f-$\overline{\text{N}}$ Chart as Portrayed in USACE (2014) (left) and Bureau of Reclamation (2011) (Right) Publications. Note, the overbar is not shown on the right chart.

---

[7] USACE considers only *incremental* life loss which is the life loss due to dam or levee breach over and above that which occurs under operations as expected (non-failure consequences).

The product of the annual probability of a failure event (f) and the expected number of fatalities conditional on that failure event ($\overline{N}$) is the Average Annualized Life Loss (AALL), *i.e.* AALL = f × $\overline{N}$. That means that (f, $\overline{N}$) pairs with the same AALL can be plotted as downward sloping 45º lines on log-log f-$\overline{N}$ charts (Figure 10.). Any point plotted on such a chart can be meaningfully compared to an AALL contour line. For any point on the chart, the probabilities (on the vertical axis) and conditional expectations (on the horizontal axis) must relate to the same event. This will *only* be the case when the life loss values that are plotted along the horizontal axis are expectations *conditional on the events whose probabilities are plotted along the vertical axis*.

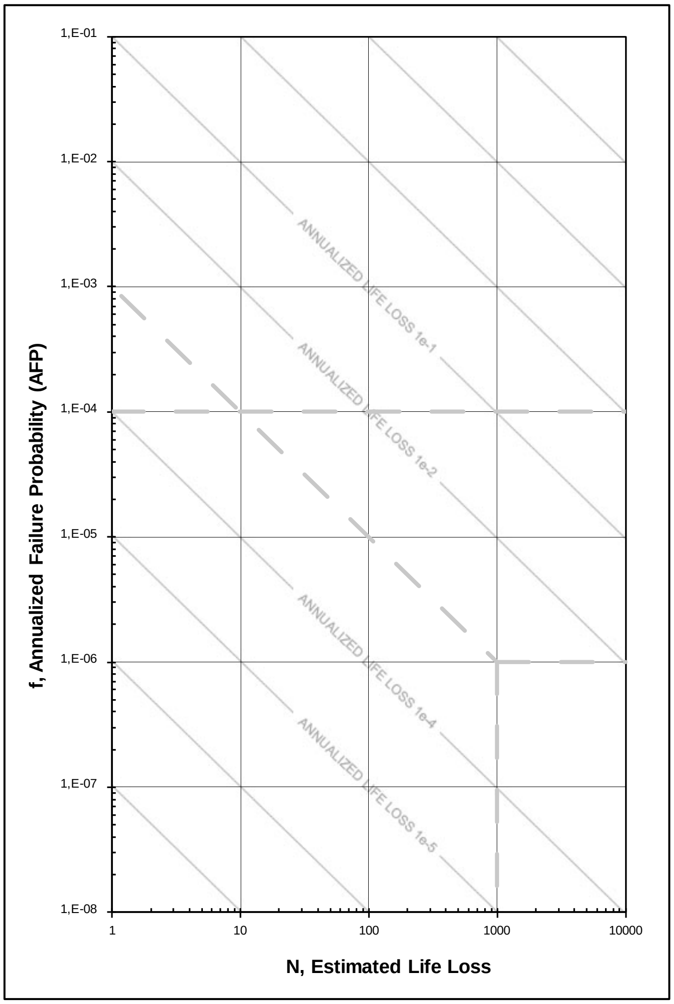


Figure 10. The f-$\overline{N}$ Chart with Diagonal AALL-Contour Lines from an Excel-Sheet that is Commonly Used by Reclamation. Note, the overbar is not shown on the chart.

In addition to the f-$\overline{N}$ charts of Figure 9 and Figure 10, the F-N chart of Figure 11 is also used by USACE and FERC to portray societal risk (FERC, 2016; USACE, 2014b). Here, F is the annual exceedance probability of life loss N. Note, this N is not a conditional expectation. The use of the F-N chart instead of the f-$\overline{N}$ chart is consistent with international practice in Australia, the Netherlands, Canada, and elsewhere (e.g., ANCOLD, 2003; Ball & Floyd, 1998; Jonkman et al., 2011; New South Wales Government Dams Safety Committee, 2006).

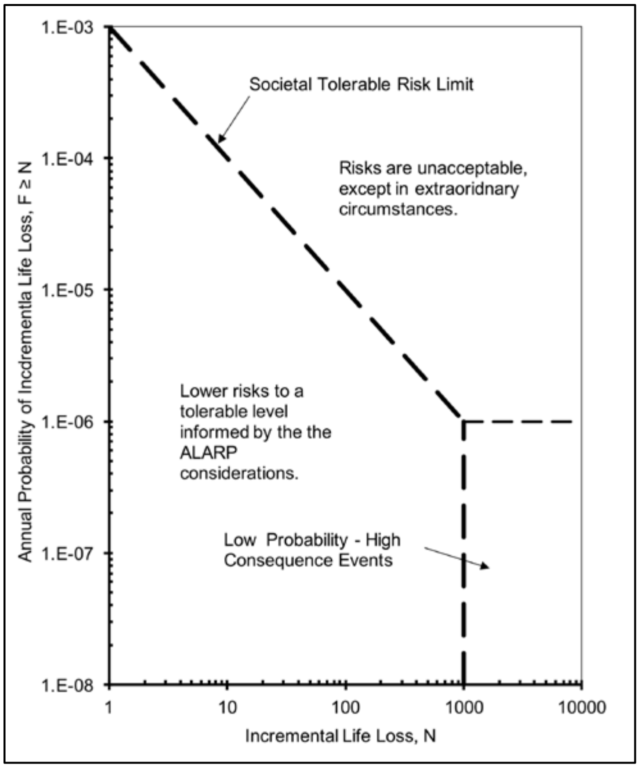


Figure 11. The F-N Chart as Portrayed in (USACE, 2014a) and (FEMA, 2015)

An F-N plot portrays the full range of potential life loss; f-$\overline{N}$ plots only do so when the failure events for which the risks are plotted are so narrowly defined that each failure event yields a unique life loss. When the life loss for each failure event is not only uniquely defined but also different from the life loss for all other events, the f-$\overline{N}$ plot becomes indistinguishable from an f-N plot. As an example, Figure 12 shows an F-N plot and a series of f-$\overline{N}$ plots for the same hypothetical dam where two potential modes could lead to ten different values of life loss depending on such factors as warning time or time of day. It is assumed for simplicity that no uncertainty is to be portrayed separately, e.g. in the form of bandwidths (see Section 7.1.3).

As shown by Figure 12, the information contained in an f-$\overline{N}$ plot depends on how ‘narrowly’ the events are defined for which the risks are shown. The narrower the events are defined, the more information the f-$\overline{N}$ plot will contain, but the harder it will be to relate the positions of the risk markers to Agency guidelines. This is because the Agency guidelines concern the risk for the dam as a whole (*i.e.* the total risk from all potential failure events).

To portray the risk related to a specific PFM on an F-N chart, the F-N plot would have to be made for that particular PFM (not shown in Figure 12).

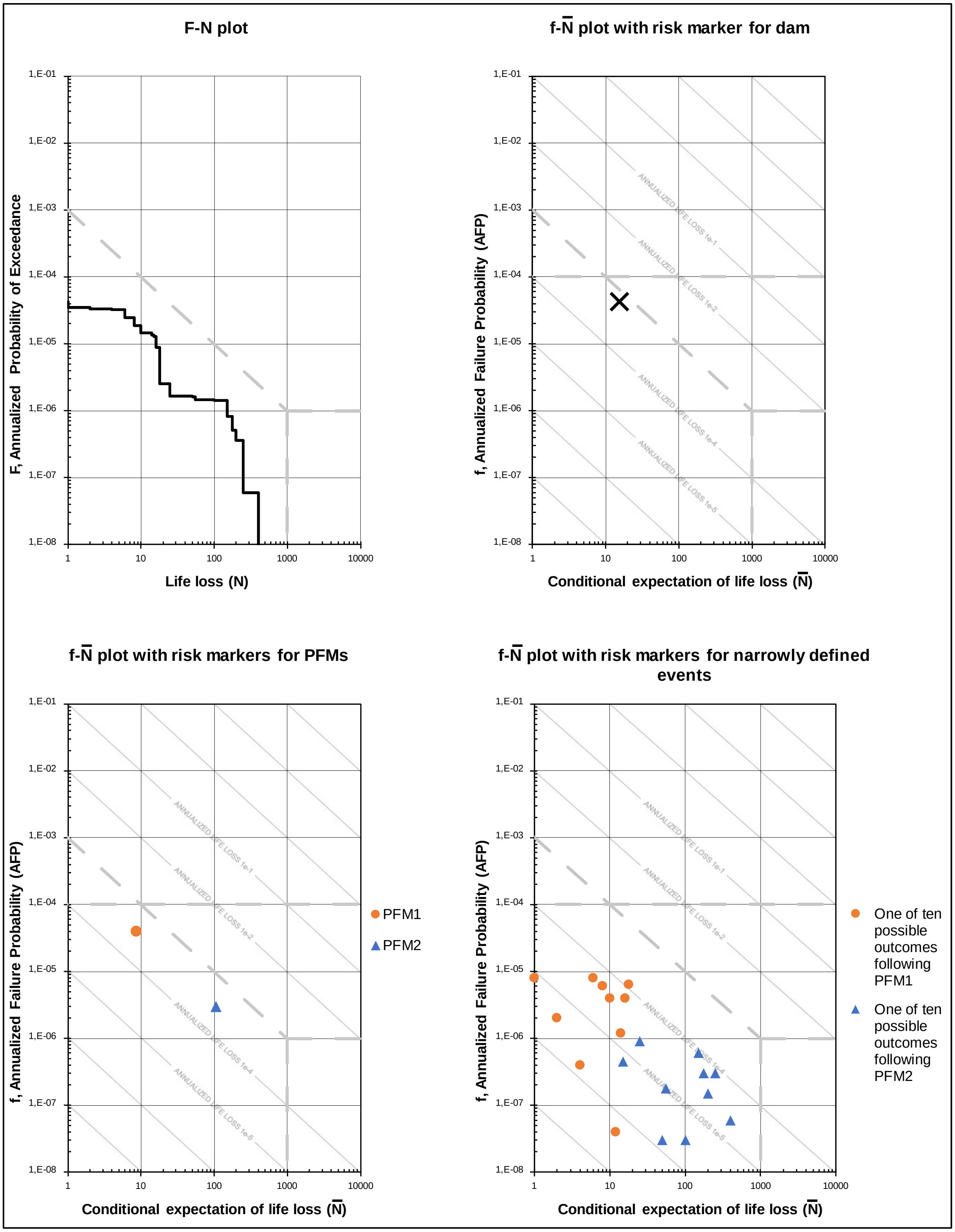


Figure 12. An F-N Plot (top left) and three f-N̄ Plots for a dam with two Potential Failure Modes that could each cause ten different amounts of life loss (hypothetical dam; no uncertainty).

Note that it would be incorrect to interpret the cloud of points for a set of narrowly defined events related to a PMF as a visualization of the uncertainty on the AALL for that PFM. As illustrated by Figure 12, the position of a risk marker for a narrowly defined event related to a PFM cannot be interpreted as a potential position of the "total" risk marker for that PFM. For instance, the probability of PFM1 equals the *sum* of the probabilities of the ten mutually exclusive events that may follow PFM1. This is why the risk marker for PFM1 plots above the cloud of risk makers for the more narrowly defined events related to PFM1. Furthermore, the AALL for PFM1 is the *sum* of the AALLs for the narrowly defined events related to PFM1. The risk marker from PFM1 thus plots closer to the Average Annualized Life Loss Guideline than any of the risk markers for the more narrowly defined events related to PFM1.

7.1.2 Plotting societal risks on $f - \overline{\mathrm{N}}$ and F-N charts: theory and Agency practice

Uncertainty on the probability and life loss estimates from a probabilistic risk analysis can easily be included in any plot by repeating the same plotting procedure for alternative pairs of probability and life loss estimates. The uncertainty can then be portrayed as a cloud of such points, a family of curves, or by means of an uncertainty bandwidth or whisker. With or without uncertainty, the underlying logic stays the same. The fact that the data that are to be portrayed might be subject to uncertainty is therefore irrelevant to the discussion of plotting procedures in the remainder of this section.

Figure 13 shows the f-$\overline{\mathrm{N}}$ and F-N plots for a single PFM that could cause a life loss of 10 with an annualized probability $9\text{x}10^{-6}$ and a life loss of 100 with an annualized probability of $1\text{x}10^{-6}$. For simplicity, assume this is the only PFM so the risk estimate for the dam equals the risk estimate for this PFM. The probability of dam failure thus equals $9\times10^{-6} + 1\times10^{-6} = 1\times10^{-5}$. Since both dam failure scenarios lead to loss of life, the probability of dam failure (the value of f on the f-$\overline{\mathrm{N}}$ chart) equals the probability of at least one fatality (the intersection of the F-N plot with the vertical axis). The AALL (i.e., the expectation of N) equals $(9\times10^{-6} \times 10) + (1\times10^{-6} \times 100) = 1.9\times10^{-4}$. The conditional expectation ($\overline{\mathrm{N}}$) thus equals $1.9\times10^{-4} / 1\times10^{-5} = 19$.

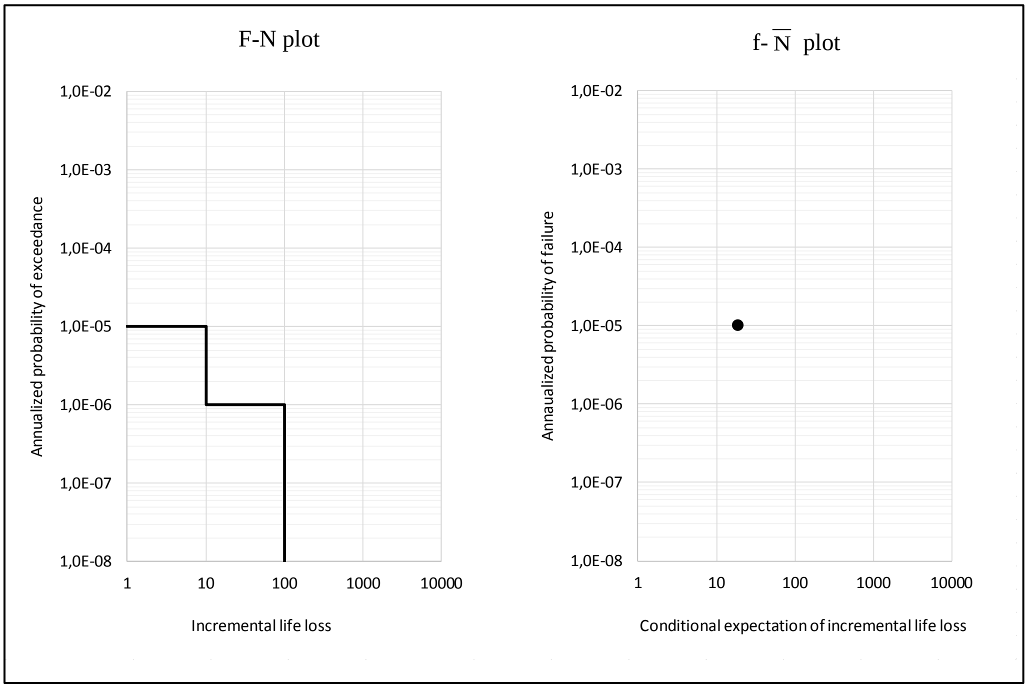


Figure 13. Example of an F-N plot (left) and an f-$\overline{\mathrm{N}}$ Chart (right) and) for the Same Potential Failure Mode (No Uncertainty)

The above example is conceptually similar to the example from Best Practices Chapter A-9 (Risk Guidelines) (Bureau of Reclamation & USACE, 2019). Note, however, that the example from Best Practices shows an F-N-curve that incorrectly displays the underlying data (see also Galic, 2018). The F-N curve from Best Practices is shown in Figure 14 together with the correct F-N curve.

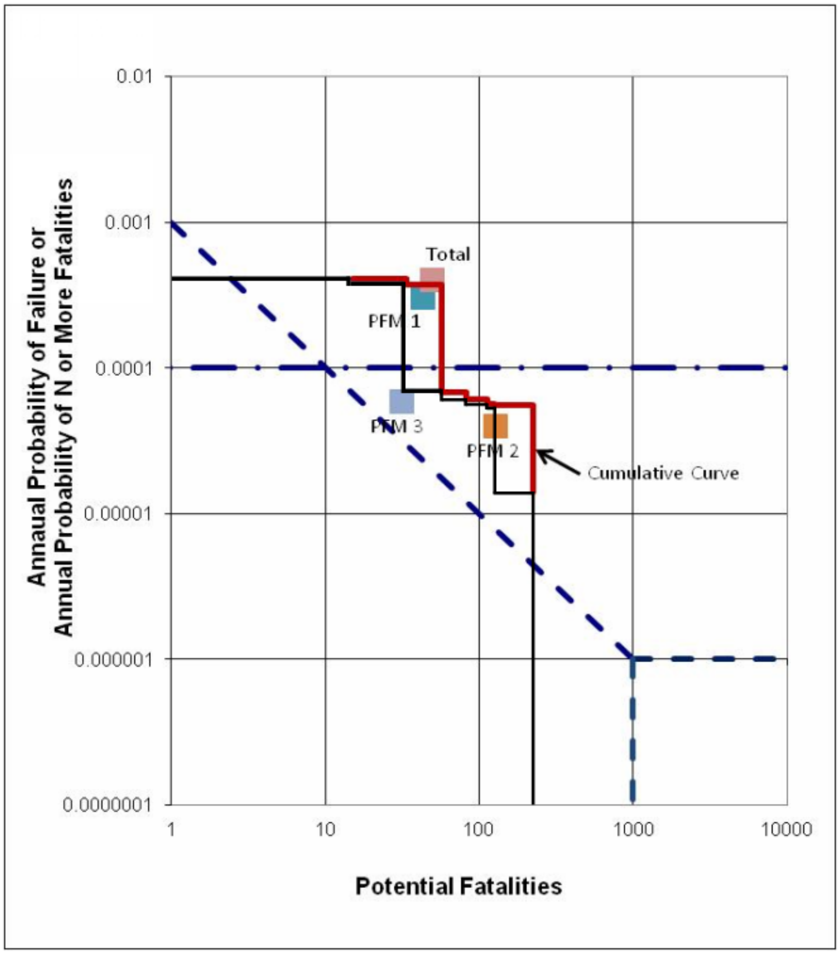


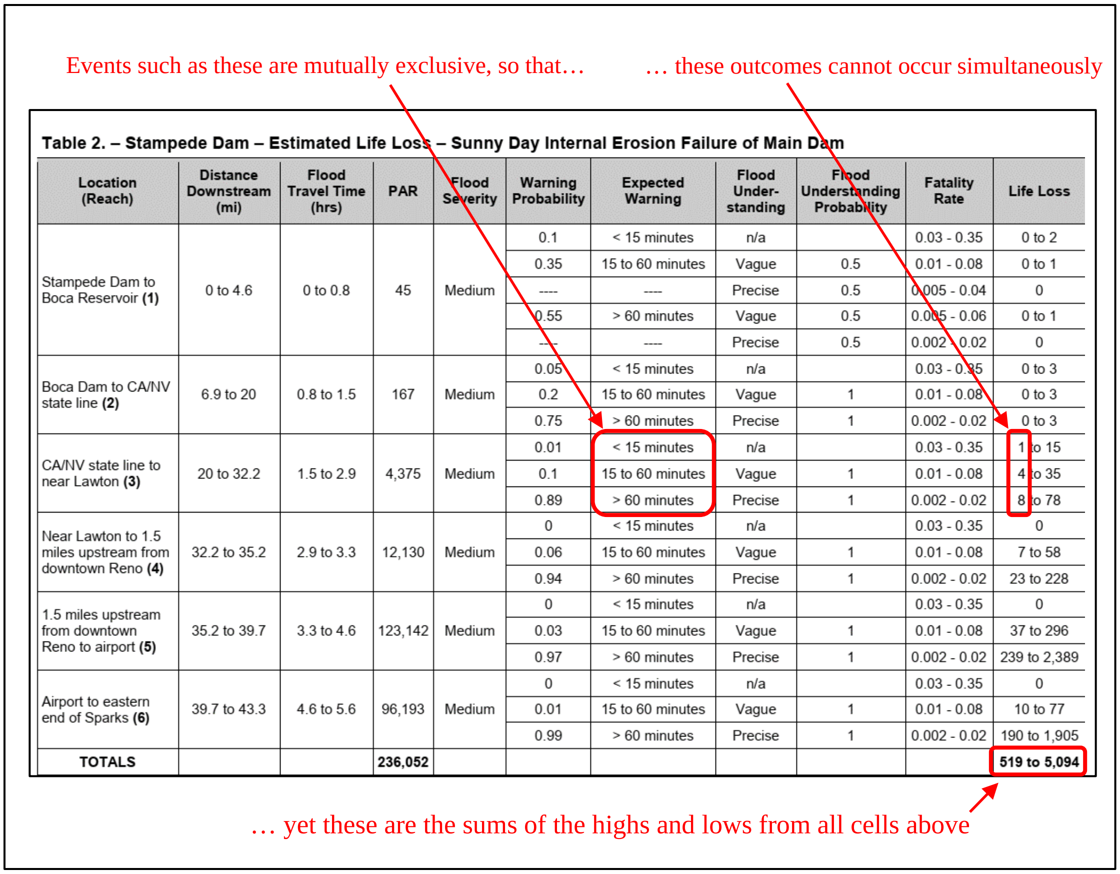
Figure 14. The F-N Plot as Shown in the Best Practices in Dam and Levee Safety Risk Analysis (red) and the F-N Plot that Correctly Displays the Underlying Data (black)

An f-$\overline{N}$ chart can be used to portray the probabilities and associated conditional expectations of life loss for different failure events. As in the example from Best Practices, the Agencies appear routinely to plot risk estimates for individual PFMs and the dam as a whole (*i.e.*, the occurrence of *any* of the identified PFMs) on f-$\overline{N}$ charts, see for example the Federal Guidelines for Dam Safety Risk Management, Reclamation's Public Protection Guidelines, and the various risk analysis reports received by the Panel. Of course, it is entirely possible also to plot the annual probabilities and conditional expectations for more narrowly defined failure events, such as 'a dam failure due to PFM1, at night, with moderate warning time, and no traffic congestion', see also Section 7.1.1.

From a meeting with Reclamation staff, the Panel understands that Reclamation does not plot f-$\overline{N}$ charts in accordance with the example from the Best Practices. However, the Agency's Public Protection Guidelines do not detail how Agency practice differs from this (otherwise correct) example. Also, it is understood from Agency comments that the method for portraying uncertainty on f-$\overline{N}$ charts from the RCEM Examples-of-Use (Bureau of Reclamation, 2014b) is not followed in current practice. This is confusing since the first line of the introduction of the document states: “This document is intended to be used in conjunction with two other companion documents dealing with the estimation of life loss resulting from dam failure, and the document discusses the generation and presentation of uncertainties.” These companion documents are the RCEM Guidelines for Estimating Life Loss for Dam Safety Risk Analysis, and the RCEM Dam Failure and Flood Event Case History Compilation. The Panel concludes that a clear written description of Reclamation’s practice for plotting f-$\overline{N}$ charts is missing.

The Panel further understands that it is incumbent on the risk analysis teams to decide how best to portray probability and consequence estimates on f-$\overline{N}$ charts. Although judgment is required in developing event probability and life loss estimates, no judgment is required for computing AALL estimates from probability and life loss estimates, and the same applies to plotting AALL estimates on f-$\overline{N}$ charts in the form of probabilities (f) and conditional expectations of life loss ($\overline{N}$). An f-$\overline{N}$ plot (or F-N plot) is merely a visualization of the potentially uncertain probability and consequence estimates from a quantitative risk analysis. Plotting procedures should not alter the estimates that are being portrayed. Plotting based on judgement can lead to an incorrect visualization of the probability and life loss estimates, and potentially incorrect understanding of absolute and relative risks within a portfolio of dams. Two examples of incorrect visualizations are discussed below.

First, in the example risk analysis report for Stampede Dam provided to the Panel by Reclamation (Bureau of Reclamation, 2012), the consequences of dam failure depend on, among other things, the warning time. As shown in Figure 15, probabilities have been attributed to warning times of <15 minutes, 15-60 minutes, and >60 minutes. The life loss portrayed on the f-$\overline{N}$ chart for the different PFMs has been computed by taking the sum over all potential consequences per PFM. Yet taking the sum over the consequences of mutually exclusive events yields a fatality number that is neither a conditional expectation nor a possible value. The warning time is either <15 minutes, 15-60 minutes, or >60 minutes; it can never simultaneously be <15 minutes, 15-60 minutes, and >60 minutes.



Events such as these are mutually exclusive, so that…

… these outcomes cannot occur simultaneously

Table 2. – Stampede Dam – Estimated Life Loss – Sunny Day Internal Erosion Failure of Main Dam

Location (Reach)	Distance Downstream (mi)	Flood Travel Time (hrs)	PAR	Flood Severity	Warning Probability	Expected Warning	Flood Under-standing	Flood Understanding Probability	Fatality Rate	Life Loss
Stampede Dam to Boca Reservoir **(1)**	0 to 4.6	0 to 0.8	45	Medium	0.1	< 15 minutes	n/a		0.03 - 0.35	0 to 2
					0.35	15 to 60 minutes	Vague	0.5	0.01 - 0.08	0 to 1
					----	----	Precise	0.5	0.005 - 0.04	0
					0.55	> 60 minutes	Vague	0.5	0.005 - 0.06	0 to 1
					----	----	Precise	0.5	0.002 - 0.02	0
Boca Dam to CA/NV state line **(2)**	6.9 to 20	0.8 to 1.5	167	Medium	0.05	< 15 minutes	n/a		0.03 - 0.35	0 to 3
					0.2	15 to 60 minutes	Vague	1	0.01 - 0.08	0 to 3
					0.75	> 60 minutes	Precise	1	0.002 - 0.02	0 to 3
CA/NV state line to near Lawton **(3)**	20 to 32.2	1.5 to 2.9	4,375	Medium	0.01	< 15 minutes	n/a		0.03 - 0.35	1 to 15
					0.1	15 to 60 minutes	Vague	1	0.01 - 0.08	4 to 35
					0.89	> 60 minutes	Precise	1	0.002 - 0.02	8 to 78
Near Lawton to 1.5 miles upstream from downtown Reno **(4)**	32.2 to 35.2	2.9 to 3.3	12,130	Medium	0	< 15 minutes	n/a		0.03 - 0.35	0
					0.06	15 to 60 minutes	Vague	1	0.01 - 0.08	7 to 58
					0.94	> 60 minutes	Precise	1	0.002 - 0.02	23 to 228
1.5 miles upstream from downtown Reno to airport **(5)**	35.2 to 39.7	3.3 to 4.6	123,142	Medium	0	< 15 minutes	n/a		0.03 - 0.35	0
					0.03	15 to 60 minutes	Vague	1	0.01 - 0.08	37 to 296
					0.97	> 60 minutes	Precise	1	0.002 - 0.02	239 to 2,389
Airport to eastern end of Sparks **(6)**	39.7 to 43.3	4.6 to 5.6	96,193	Medium	0	< 15 minutes	n/a		0.03 - 0.35	0
					0.01	15 to 60 minutes	Vague	1	0.01 - 0.08	10 to 77
					0.99	> 60 minutes	Precise	1	0.002 - 0.02	190 to 1,905
TOTALS			**236,052**							**519 to 5,094**

… yet these are the sums of the highs and lows from all cells above

Figure 15. Estimated Life Loss per Location (Reach) for a PFM from the Stampede Dam Risk Analysis Report (Bureau of Reclamation, 2012).

The conditional expectation of the life loss for this PFM can be obtained by taking the probability-weighted sum of the mutually exclusive outcomes. Due to the uncertainty on the life loss estimates in the final column, these probability-weighted sums are subject to uncertainty. The probability-weighted sum ranges from 451 (i.e. 0.1×0 + 0.35×0.5×0 + 0.55×0.5×0 + 0.55×0.5×0 +…+ 0.01×10 + 0.99×190 = 451) to 4507 (i.e. 0.1×2 + 0.35×0.5×1 + 0.55×0.5×0 + 0.55×0.5×1 …+ 0.01×77 + 0.99×1905 = 4507). Fortunately, these numbers are quite close to the numbers (519 and 5094) in the lower right-hand corner of Figure 15 so that the risk marker for this PFM is not too far from its correct position on the f-$\overline{N}$ chart (the same applies to the other PFMs). However, it can easily be verified that this would have been very different had the warning probabilities for <15 minutes and >60 minutes been reversed. In that case, the conditional expectations per PFM (and the dam as a whole) would have been overestimated by an order of magnitude.

Second, it is understood from the meeting with Reclamation staff that the minima and maxima of N (*i.e.* the potential life loss for best case and worst case events) might be treated as the minima and maxima of $\overline{N}$. Examples of treating the uncertainties on N as the uncertainties on $\overline{N}$ can also be found in the RCEM Examples-of-Use (Bureau of Reclamation, 2014b). Yet the uncertainties related to N and $\overline{N}$ are not the same. To see why, consider again the case of Stampede Dam. From the table shown in Figure 15, the lowest possible number of fatalities in case of a sunny day internal erosion failure of the main dam can be obtained by taking sum of the minima of the potential number of fatalities with a non-zero probability per reach. Since the numbers in the final column of the table shown in Figure 15 are ranges, the lowest potential number of fatalities for this PFM could be treated as uncertain. It ranges from 55 (i.e. min(0,0,0,0,0) + min(0,0,0) + min(1,4,8) + min(7,23) + min(37,239) + min(10,190) = 55) to 449 (i.e.

min(2,1,0,1,0) + min(3,3,3) + min(15,35,78) + min(58,228) + min(296,2389) + min(77,1905) = 449). Similarly, the highest potential number of fatalities for this PFM ranges from 460 to 4605. This means that the life loss in case of a sunny day internal erosion failure of the main dam could be as low as 55 and as high as 4605. Yet the conditional expectation of the life loss for this PFM lies somewhere between 451 and 4507.

Treating the uncertainty on N as if it were the uncertainty on $\overline{\mathrm{N}}$ can lead to substantial overestimation of the uncertainty related to Average Annualized Life Loss (since AALL = $\mathrm{f} \times \overline{\mathrm{N}}$). It also leads to uncertainty bandwidths whose end points cannot be meaningfully compared to AALL-contour lines. This is because the product of the probability of a (*any*) dam failure with the life loss for a *worst case* dam failure is neither the AALL for a (*any*) dam failure, nor the AALL for a *worst case* dam failure. As stated in section 7.1.1, the life loss values that are plotted along the horizontal axis of an f-$\overline{\mathrm{N}}$ chart should be expectations conditional on the events whose probabilities are plotted along the vertical axis. This applies not only to the risk markers shown on f-$\overline{\mathrm{N}}$ charts, but also to the end points of uncertainty bandwidths.

Uncertainties related to N do not necessarily lead to uncertainty related to $\overline{\mathrm{N}}$. Such is the nature of averaging, that the uncertainty related to $\overline{\mathrm{N}}$ (a probability-weighted sum of N-values) will typically be smaller, possibly much smaller, than the uncertainty related to N. This does not mean that “uncertainty gets lost when averaging.” It means $\overline{\mathrm{N}}$ should be thought of as a variable in its own right, with its own uncertainty bandwidth. Note that a conditional expectation can only be uncertain when some uncertainties are excluded from the act of averaging (*i.e.* left outside the probability-weighted sum). This raises the question of what those uncertainties are or should be. This issue is further explored in Section 6.3.2 and Appendix D.

7.1.3 Portraying the uncertainty related to societal risk on $f - \overline{\mathrm{N}}$ charts

Reclamation portrays the uncertainty related to societal risk estimates on f-$\overline{\mathrm{N}}$ charts using vertical and diagonal whiskers (Figure 16). The vertical whiskers portray uncertainty related to the probability of dam failure, the diagonal whiskers portray uncertainty related to the dam’s AALL. There are no horizontal whiskers to portray uncertainty related to the conditional expectation of life loss.

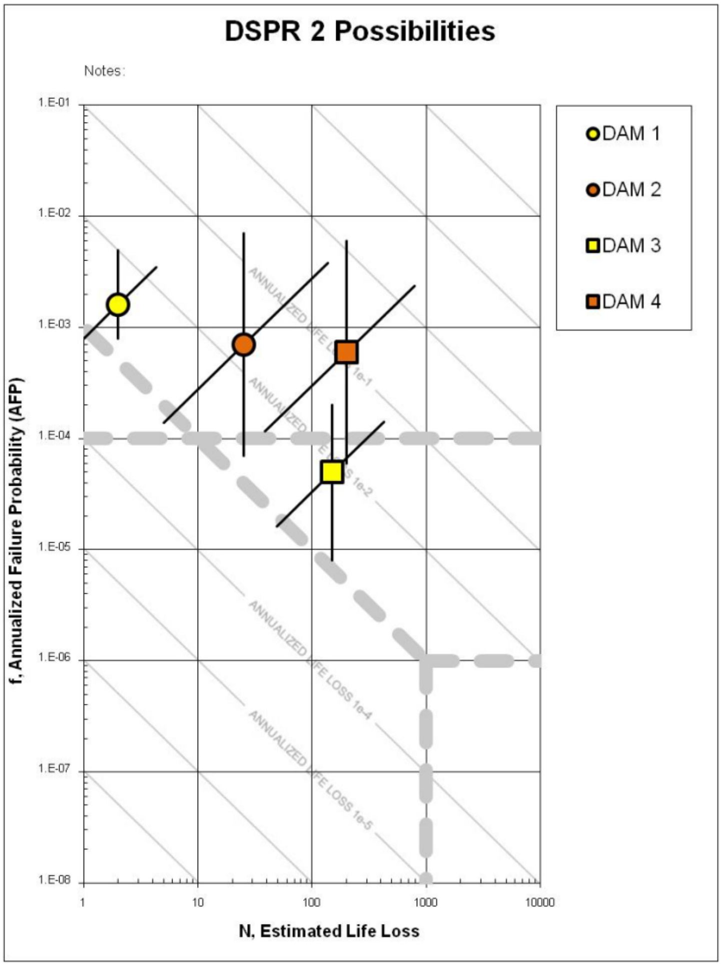


Figure 16. Example of the Bureau of Reclamation's Approach to Portraying the Uncertainty Related to Societal Risk Estimates on f-$\overline{N}$ Plots (from: Bureau of Reclamation, 2011a) (note: the overbar is not shown on this chart)

Casual observers may incorrectly interpret the horizontal range depicted by diagonal whiskers on an f-$\overline{N}$ chart as the uncertainty related to the conditional expectation of life loss. Yet a diagonal whisker could extend over a considerable horizontal distance in the absence of any uncertainty related to the life loss or its conditional expectation. This is illustrated in Figure 16. The figure shows the f-$\overline{N}$ plot with uncertainty bands corresponding to the data from Table 2. In this hypothetical example, potential life loss is assumed to be perfectly known, *i.e.* N=$\overline{N}$=500. The end points of the diagonal whiskers are based on the calculated high and low values of the annual expected life loss of 10^{-5} and 10^{-3}. If we read the values corresponding to these end points from the horizontal axis, we may erroneously conclude that life loss can vary from 158 to 1581, into the bounded area.

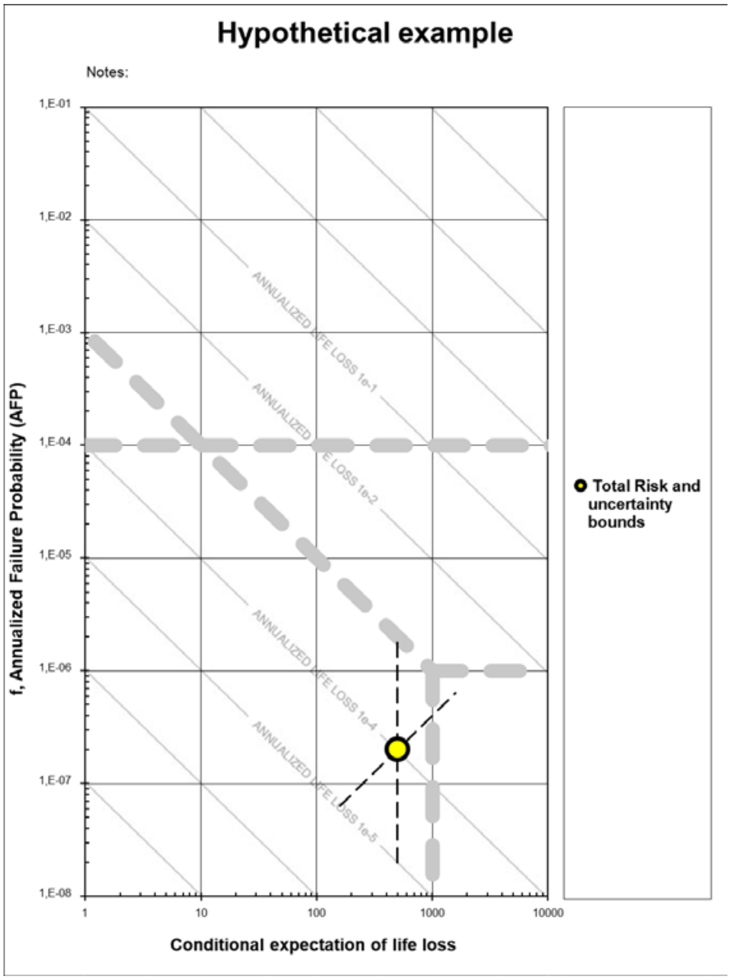


Figure 17. The f-$\overline{N}$ Plot with Uncertainty Bands Corresponding to the Data from Table 2

Table 2. Annualized Failure Probability and Life Loss

PFM	Annualized Probability			Incremental Life Loss		
	Minimum	Mean	Maximum	Minimum	Mean	Maximum
1	2.10^{-8}	2.10^{-7}	2.10^{-6}	500	500	500

As illustrated above, the end points of the diagonal whisker on an on f-$\overline{N}$ chart should not be interpreted as points showing the low and high-end estimates of the dam failure probability together with the low and high-end estimates of the conditional expectation of life loss. The endpoints can only be interpreted in terms of annualized expected life loss. This is not a problem, as long as it is understood.

7.1.4 Panel recommendations

Considering the above discussion, the Panel recommends:

1. The Agencies more clearly explain the differences between f-$\overline{N}$ plots and F-N curves in agency guidelines. The life loss (N) on an F-N chart is not the same as the conditional expectation of the

life loss ($\overline{N}$) on an f-$\overline{N}$ chart. These variables represent different things, and the uncertainties on these variables are not the same. To avoid confusion, the Panel recommends the Agencies consistently use an overbar when referring to the conditional expectation, i.e. $\overline{N}$ instead of N, per USACE's ER 1110-2-1156 (USACE, 2014a) and FERC´s interim RIDM guidelines for dam safety (FERC, 2016).

2. The Agencies correct the F-N plotting example from Best Practices and include a step-by-step procedure for plotting F-N curves.
3. The Agencies develop detailed guidance on recommended practice for portraying risks on f-$\overline{N}$ charts. To ensure that any point that is plotted on such a chart (including the end points of uncertainty whiskers) can be meaningfully compared to an AALL contour line, the life loss values that are plotted along the horizontal axis should be expectations conditional on the events whose probabilities are plotted along the vertical axis. It could be left to risk analysis teams to decide which failure events and uncertainty bandwidths are to be included in f-$\overline{N}$ charts. However, to avoid confusion and error, the procedure for plotting on f-$\overline{N}$ charts should not be based on judgment, but on clear definitions and the rules of probability calculus.
4. Reclamation more clearly explain in its guidelines that the horizontal distance represented by the diagonal whiskers on its f-$\overline{N}$ plots should not be interpreted as life loss ranges. Adding a horizontal whisker to indicate uncertainty related to the conditional expectation of the life loss may help avoid confusion, although clear explanations may also suffice.

7.2 Evaluating Societal Risk: Risk Guidelines

The *risk guideline*[8] USACE and Reclamation use for evaluating societal risk on f-$\overline{N}$ charts is a line on a log-log grid with slope of -1 for a conditional expectation ($\overline{N}$) between 1 and 1,000 (Figure 9), in line with the AALL guideline of 0.001. For higher values of $\overline{N}$, risks should be, as a minimum, as low as reasonably practicable (ALARP) (see Section 4.2). USACE and FERC use a similar societal risk guideline on F-N curves (Figure 11). It, too, is a line on a log-log grid with slope of -1 for the cumulative probability of more than N fatalities between 1 and 1,000. The implications of these two guidelines are not the same (Section 7.2.1).

7.2.1 Comparing the risk guidelines on $f - \overline{N}$ and F-N charts

The societal risk guidelines on the f-$\overline{N}$ and F-N charts look identical (see Figure 9 and Figure 11) despite the fact that different quantities are plotted along the horizontal and vertical axes on these charts. In both cases the societal risk guideline is a downward sloping line (slope -1) for a life loss (N) or the conditional expectation thereof ($\overline{N}$) ranging from 1 to 1,000, starting at an annualized failure probability (f) or an annualized exceedance probability (F) of 0.001. While the Agencies also plot the risk for individual failure modes on f-$\overline{N}$ charts, the risk guideline on an f-$\overline{N}$ chart only applies to the risk for the dam (i.e., all failure modes combined).

A societal risk guideline with a slope of -1 on an f-$\overline{N}$ chart is a line that shows all combinations of probabilities and conditional expectations that yield the same expectation. This is because the product of the probability of dam failure and its conditional expectation is, by definition, an expectation. Such a societal risk guideline is risk neutral. A risk neutral decision maker is solely guided by expected values when making decisions under uncertainty.

A societal risk guideline with a slope of -1 on an F-N chart will only yield identical tolerability judgments to a similar-looking guideline on an f-$\overline{N}$ chart when there is no uncertainty related to the number of

---

[8] Different terms are used in different publications and figures, such as "Tolerable Risk Guideline," "Societal Risk Reference Line," and "Societal Risk Limit."

fatalities in case of dam failure (i.e., the life loss is identical across failure modes, and known perfectly). The exceedance probability of the number of fatalities shown on the F-N chart is then identical to the probability of this number of fatalities shown on the f-$\overline{N}$ chart. In all other cases, there will be differences due to the fundamental differences between the variables plotted along the horizontal and vertical axes of these charts. This is illustrated in Figure 18 below. The figure shows seven different F-N curves for dams with an annualized probability of failure of 1/10,000 and an AALL of 0.001. The F-N curve in red applies to a case without uncertainty related to the number of fatalities given dam failure. Since the annualized probability of failure (f) and the AALL are the same for each dam, the f-$\overline{N}$ charts for these dams look identical (not shown here). The total risk markers on the f-$\overline{N}$ charts plot on top of the risk guideline, just like the F-N curve in red (for this case, F=f=1/10,000 and N=$\overline{N}$=10). Yet all other F-N curves plot at a small distance from the risk guideline, as shown Figure 18.

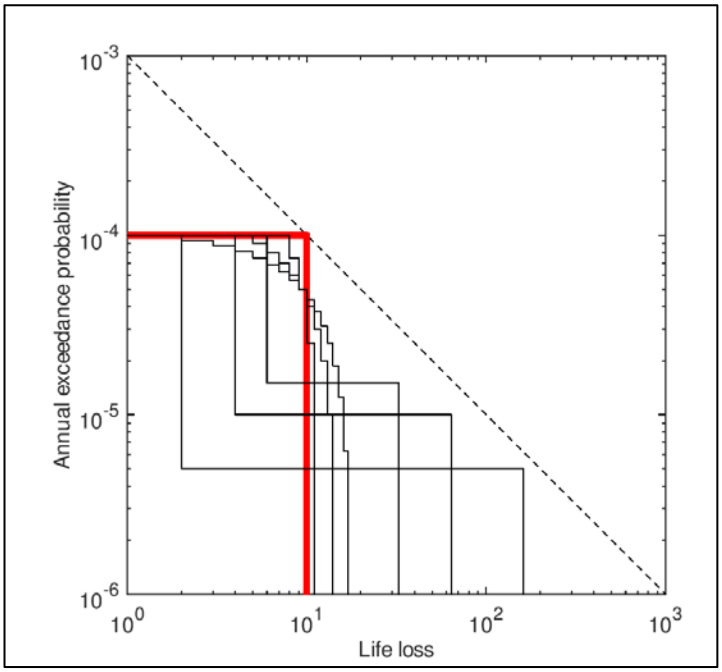


Figure 18. Different F-N Curves with the Same Expected Life Loss in Case of Flooding

While a risk-neutral individual would treat the red and black F-N curves from Figure 18 all the same since their expectations (AALLs) are identical, the red curve touches the societal risk guideline while the other curves plot slightly lower, making them seem preferable. A guideline with a slope of -1 on an F-N chart may thus yield different tolerability judgments than a similar-looking guideline on an f-$\overline{N}$ chart. The downward sloping risk guideline on the F-N chart will typically be less strict than the seemingly identical AALL risk guideline on the f-$\overline{N}$ chart. It is impossible for an F-N curve to plot above $F=10^{-3}/N$ without the annualized expectation of life loss exceeding the AALL guideline of 10^{-3}. In short: while the guidelines on the F-N and f-$\overline{N}$ charts look identical, they are fundamentally different.

While the risk guidelines on the F-N and f-$\overline{N}$ charts from Figure 9 and Figure 11 are fundamentally different, the practical implications of these differences may be insignificant within the context of risk-informed decision making. This is because an F-N curve cannot plot far below a societal risk guideline with slope -1 when the f-$\overline{N}$ estimate plots above a similar-looking guideline (the opposite is impossible).

There is an important caveat, however. The F-N and f-$\overline{N}$ charts contain bounded areas for values of N or $\overline{N}$ exceeding 1,000, yet this threshold means entirely different things on F-N and f-$\overline{N}$ charts. An F-N curve is more likely to end up in the bounded area than an f-$\overline{N}$ plot with risk markers for individual failure modes and the dam as a whole. The probability of a life loss greater than 1,000 could be non-zero, while the conditional expectation $\overline{\mathrm{N}}$ stays far below 1,000. This applies not only to the value of $\overline{\mathrm{N}}$ for the dam as a whole, but also to the values of $\overline{\mathrm{N}}$ for the individual PFMs, see also Figure 12 in section 7.1.1. An F-N plot and an f-$\overline{\mathrm{N}}$ plot will only show the same maximum number of fatalities when the f-$\overline{\mathrm{N}}$ chart contains a risk marker for the worst case event (*i.e.* the narrowly defined event for which $\overline{\mathrm{N}}$ equals the maximum value of N). Yet this does not appear to be standard practice.

7.2.2 Risk Guideline Truncations

The term, "risk guideline truncations" is defined as horizontal and / or vertical boundaries on plots of societal risk guidelines. A horizontal truncation sets the societal risk guideline to a single probability value between specific life loss values. A vertical truncation sets the societal risk guideline to a single value of life loss between specific values of probability. Truncations are included in the Agencies societal risk guidelines shown in Figure 9 and Figure 11 in Section 7.1.1.

The Agencies' tolerable risk guidelines are similar to the approach taken in other jurisdictions that have proposed societal and individual risk guidelines for dam safety. Ball and Floyd (Ball & Floyd, 1998) provide a useful comparison of public safety risk criteria used in various industries and jurisdictions. There are many variations to the slopes and truncations of the risk criteria, but Ball and Floyd concluded, "(…) there is a surprising degree of consistency amongst FN-based societal risk criteria developed at different times, in different places, for different purposes and by different routes."

In Australia ANCOLD suggests an order of magnitude lower tolerability threshold for new dams compared to existing dams. This is stated to be (ANCOLD, 2003) "The higher limit of tolerability for existing dams than for proposed dams, recognizes that many existing dams were constructed decades ago, with much poorer technology than is now available. Also, the marginal cost of achieving lower risk levels is generally much less for proposed dams than for existing dams. Regulatory requirements have accepted higher risks for existing installations than for proposed facilities (…)." In the United Kingdom, the HSE also may accept risks associated with existing facilities to be higher than those that would be achieved by applying good practice at a new facility (Health and Safety Executive, 2003).

ANCOLD's societal risk guidance (ANCOLD, 2003) also has a horizontal truncation at probability (F) of 10^{-5} and fatalities (N) of 100 for existing dams and at F of 10^{-6} and N of 100 for new dams. This represented ANCOLD's judgment at that time of the lowest risks that could be realistically assured in light of the then knowledge and dam technology and methods available to estimate the risks. ANCOLD is currently revising its Guidelines on Risk Assessment, and the truncations are likely to be removed. Although this position has not yet been formally adopted by ANCOLD, the logic of the working group (draft) is as follows:

> As discussed in ANCOLD (2003) Commentary on Guideline G-10 (pages 113, 114) the horizontal truncation of Figures 7.4 and 7.5 in the ANCOLD (2003) Guideline were without precedent at the time of drafting the Guideline, but represented ANCOLD's judgement at that time of the lowest risks that could be realistically assured in light of:
>
> - Present knowledge and dam technology.
> - Methods available to estimate the risks.
>
> ANCOLD now believes that present knowledge and modern dam technology is capable of design and construction of dams and dam upgrades which can achieve the required tolerable risk position without the need for a horizontal truncation. To achieve this requires a thorough knowledge of the dam and its foundation determined by experienced practitioners.

> ANCOLD also believes that the state of practice in dams risk analysis has improved sufficiently since the ANCOLD (2003) Guidelines were published that the second point no longer applies.
>
> (…)
>
> It has also been found that in practice ALARP has been difficult to define and apply, and in some cases appears to have not been considered, at least in the manner the ANCOLD (2003) guideline intended. This has resulted in dams [that] plot just above or below the truncation being considered as representing a tolerable or close to tolerable risk position while strict application of ALARP might indicate otherwise.
>
> ANCOLD also believes it is more consistent with risk management practice in other industries and in dams internationally to not have the horizontal truncation.
>
> It is noted that USACE (2014b) has a horizontal truncation at 1.0E-06. They require that the tolerability of risk must be based on an official review of the benefits and risks as described in the "Except in Extraordinary Circumstances" section of their regulation:
>
> - for Low Probability-High Consequence Events where the life loss is equal to or exceeds 1000 lives; or if
> - the probability per year of potential life loss is less than the 1.0E-06 for an estimated life loss in the range of 1000 or greater. This section of the USACE document is taken from ANCOLD (2003)
>
> Taking account of these factors ANCOLD now believes the horizontal truncation is not appropriate and has adopted the revised guideline. It is anticipated that the "except in exceptional circumstances" qualification may be invoked by some owners whereas in the past it has seldom if ever been used. The wording of the Guideline in Figures G7.4 and G7.5 has been modified from ANCOLD (2003) to better reflect the fact that ANCOLD (2020) is a guideline, not a standard, and hence words implying mandatory application have been modified.
>
> It is expected that for some existing extreme consequence dams it may be very difficult to design and construct upgrades sufficient to satisfy the tolerable risk criterion. To attempt to do so may also in some cases result in unacceptable risks during construction of upgrades. These cases may be considered under the "except in exceptional circumstances" requirements in G 9-10. This is more likely to apply for PLL [Potential Loss of Life] ≥ 100. ANCOLD believes that consultation with State and or Federal Government is advisable at these levels of PLL in any case.
>
> It is anticipated that there will be a greater emphasis on developing warning and evacuation measures to manage the risk as the PLL increases.

The vertical truncation that was in place in the guidelines published by the Dams Safety Committee of New South Wales in Australia is no longer in use, since the Dams Safety Committee was dissolved and a new organization, Dams Safety NSW, established under a new Act and Regulations. The law now requires dam owners to comply with a societal risk guideline that is similar to the ANCOLD (2003) guidance for new and existing dams, except that the horizontal truncations have been removed.

The discussion on the adoption of the societal risk criteria by the USACE (N. J. Snorteland, 2019) indicates the Dams Safety Committee approach was considered in evaluating the criteria. In a presentation made to a May 2008 tolerable risk workshop sponsored by the Agencies, a member of the Dams Safety Committee provided the following information in relation to the vertical truncation:

- The NSW Department of Planning guiding FN Chart has a vertical truncation at N = 1,000,
- The Department of Planning will not consider any proposal with a potential for loss of life over 1,000 lives,
- This vertical truncation is not feasible for dams - some have over 50,000 persons at risk,
- All existing dams with a potential for loss over 1,000 lives were approved by the Government of the day (though in a different societal context), and
- It was clear the ANCOLD horizontal truncation would be untenable in NSW.

This information indicates the principal reason for the vertical truncation and "bounded area" being adopted by the Dams Safety Committee was as a compromise to the position of another government agency that had previously adopted truncated societal risk guidance for different public policy purposes.

Public safety risk guidance is rightly a matter for governments that have the right and power to decide what risks shall be borne by citizens within their jurisdiction. The properly informed views of governments or those acting on their behalf are a relevant consideration, especially when it comes to allocating resources across the various public health and safety sectors.

Where there is no specific government policy, and public agencies develop their own public safety risk guidance, it is generally to inform decisions on the priority and urgency of investments in dam safety. In this context, and where dam owners can be held legally liable in the case of dam failure, risks should ultimately be reduced in accordance with the principles of ALARP and to meet a common law duty of care. Where these principles are applied, but the risk of dam failure remains above the public safety guidance, then the dam must be decommissioned or the government must make a decision (probably on the balance of risk versus benefit) that the dam remain in service.

7.2.3 Panel recommendations

Decisions pertaining to risk tolerability are fundamentally subjective. The Panel therefore considers the choice of f-$\overline{N}$ or F-N charts, and the choice of risk guidelines for evaluating societal risk to be agency or policy decisions on which the Panel has little say. However, given the opportunities for misinterpretation and error discussed above, the Panel provides the following considerations and recommendations:

While the risk guidelines used by the Agencies on f-$\overline{N}$ and F-N charts look identical, they are fundamentally different. Decision makers should be aware of these differences to avoid confusion and error. First, the downward sloping guideline on the f-$\overline{N}$ chart is a perfect representation of the Average Annualized Risk Guideline. The downward sloping guideline on the F-N chart is not. Second, the F-N and f-$\overline{N}$ charts both contain bounded areas for values of N or $\overline{N}$ exceeding 1,000. These similar-looking bounded areas may yield different results due to the difference between N and $\overline{N}$. An F-N curve may plot into the bounded area while the corresponding f-$\overline{N}$ plot does not. The Panel recommends the Agencies more clearly explain these differences in agency guidelines.

In addition, the Panel suggests that the Agencies review use of truncations within their societal risk guidelines and adjust if it is then considered necessary.

7.3 Risk Communication

7.3.1 Risk communication guidelines

USACE's ER 1110-1156 defines risk communication as follows: "Risk communication is the open, two-way exchange of information and opinion about hazards and risks leading to a better understanding of the risks and better risk management decisions" (USACE, 2014: pp 2-3). This definition is in line with a widely cited definition of risk communication as "any purposeful exchange of information about health or environmental risks between interested parties" (Covello et al., 1986: p 172). The USACE's definition of risk communication has also been adopted by FERC (FERC, 2016).

The Federal Guidelines for Dam Safety Risk Management (FEMA, 2015) present risk communication as a critical component of RIDM that must be integrated into all aspects of dam safety risk management. They distinguish between different types of communication or target audiences:

1. internal communication within a dam safety organization,
2. communication with dam owners and stakeholders such as local emergency management authorities, and
3. communication with the organizations and public impacted by the dam.

The Federal Guidelines for Dam Safety Risk Management (FEMA, 2015) covers a wide range of risk management tasks aimed at informing or educating stakeholders, enhancing disaster preparedness, providing emergency information, and making better-informed decisions.

The fundamentals of the principles and guidelines for risk communication developed by the Agencies are broadly in line with those developed elsewhere (e.g., DELWP, 2015; EPA, 1988; NRC, 1988; Renn, 1992, 1998; Zimmerman, 1987), yet putting theory into practice for risk communication is notoriously difficult, not merely in the field of dam safety. This is also true of other fields, such as nuclear power generation and managing risks from induced seismicity (e.g., Fischhoff et al., 1983; Sintubin, 2018). Experts and laypeople may define and perceive risks differently; people may have difficulty understanding technical or probabilistic information; and people may not share common interests and values (e.g., Fischhoff et al., 1984; Slovic, 1987, 1999). Hence, while risk communication is widely seen as an integral part of risk management, getting it right often presents serious challenges to organizations charged with managing the risks to the public from complex technological systems such as dams. The Panel has identified two areas of potential improvement, which are discussed below.

7.3.2 Engaging the public on RIDM

When it comes to engaging the public on RIDM through "open, two-way exchange of information and opinion," the Agencies fall short of the level of ambition set forth by the Federal Guidelines for Dam Safety Risk Management. Risk assessments results are mostly communicated *to* parties that have an interest in dam safety rather than communicated *with* them. There appears to be no systematic effort to engage the public on risk analyses or dam safety decision-making, yet understanding societal concerns is essential in relation to assessing ALARP, see also Section 4.2.

While Reclamation's 2019 Annual Dam Safety Program Assessment Report (Bureau of Reclamation, 2019a) discusses internal communication, interagency communication, and communication aimed at corrective action and emergency management, it does not deal with public engagement on RIDM. The example Issue Evaluation reports, periodic assessment reports, and decision documents the USACE and Reclamation provided to the Panel do not include signs of public consultation or involvement in risk analyses or risk assessments, emergency management apart. The USACE works with local EMAs and other parties for evaluating early warning and evacuation effectiveness (see also Section 6.8 on life loss modeling). Reclamation does not involve third parties in life loss estimation, but engages staff that works with local EMAs. Only the USACE's example Issue Evaluation report (USACE, 2018) deals with risk communication, yet its discussion of the dam's risk communications plan is concerned with explaining USACE risk management practices, creating public awareness and understanding, and notifying selected parties about progress and decisions, i.e., one-way rather than two-way communication.

7.3.3 Public access to inundation maps: balancing safety and security

Inundation maps are important tools for communicating dam safety risks. Without public access to inundation maps, individuals and firms may have difficulty finding out whether they are at risk and have an interest in a dam's safety. Also, people may unwillingly put themselves at risk when they unknowingly move into a potential inundation zone. Yet, at present, the inundation maps developed by the Agencies are not publicly available. Only the USACE has expressed its intent to publish its inundation maps online.

There could be legitimate reasons for not releasing particular types of risk information to the public. The Federal Guidelines for Dam Safety Risk Management (FEMA, 2015: p 37) states that "Decisions on the level of information to share should balance legitimate security concerns with the benefits of creating a public awareness of potential dam safety issues." While security concerns may limit the level of information that is shared with the public, the decision to refrain from communicating potential inundation zones to the public should not be taken lightly. The fact that the USACE has expressed its intent on making inundation maps publicly available suggests that security concerns may not be, or may not always be, sufficient reason for restricting public access.

7.3.4 Panel Recommendations

Considering the above discussion, the Panel recommends:

1. The Agencies engage parties with an interest in dam safety on RIDM, as stipulated by the Federal Guidelines for Dam Safety Risk Management (FEMA, 2015). To this end, the Panel recommends the Agencies more clearly define their objectives related to risk communication, and that they develop practical guidelines for achieving those objectives. Risk communication is not merely a means to improve the *public's* understanding of dam safety risks and agency decisions; it could also be a means to improve the *Agencies'* understanding of risk and public opinion.
2. The Agencies carefully review their policies concerning public access to inundation maps and clearly articulate their decisions.

8 FINDINGS AND RECOMMENDATIONS

The Agencies have developed a coherent and progressive approach to analyzing, assessing, and managing risks associated with dam safety. The directions of the Agencies dam safety RIDM programs are solid and consistent with federal dam safety guidance published by FEMA.

The dam safety RIDM programs in the three agencies are at different levels of maturity. Reclamation, having embraced the RIDM approach earliest, has the most mature adoption of a RIDM program, with RIDM being almost fully integrated into the organizational culture. The USACE, having begun its RIDM program about 10 years after Reclamation, has matured its program substantially, but is still working to fully embed the program into the agency culture. FERC has only more recently begun to adopt RIDM for its dam safety program, and it is just at the beginning of the program's implementation. FERC, as a regulatory agency, faces some hurdles to implementation that are not faced by the other two agencies, which are self-regulated dam owners. FERC must not only develop a RIDM culture within its own organization, but also within the organizations of its licensees - the dam owners. In addition, the success of FERC's implementation of RIDM depends on dam safety risk expertise availability within the private sector, which at present is not sufficient for a full implementation of RIDM to the entire FERC-regulated portfolio of dams.

Moving forward with the programs in the Agencies there are issues of concern that warrant attention, as discussed in the remainder of this section. Many of the following recommendations address details of the RIDM process, but the Panel considers them important details. Completing risk analyses, risk assessments, and dam safety decisions as correctly as possible and continually improving the process are, in the Panel's view, essential to the credibility of the Agencies' dam safety RIDM programs in the minds of decision makers, government officials, and the public. The recommendations presented below are offered in the spirit of continuous improvement.

The findings and recommendations provided in the remainder of this section are grouped under the following three headings: Agency Policy, Agency Governance, and Agency Methodology.

8.1 Agency Policy

The Panel supports the USACE and Reclamation in continued use and development of RIDM in managing the safety of the dams owned and operated by the agencies. The RIDM programs of these two agencies have substantially improved prioritization of risk reduction works according to risks presented by the different dams in the agencies' portfolios. As RIDM has been applied to the agencies' dam safety programs, both agencies have reviewed policies and methodology with an eye toward continuous improvement of the RIDM programs. The Panel encourages both agencies to continue these continuous improvement efforts, and the recommendations offered in this report are directed toward that end.

The Panel recommends FERC continue its movement toward applying RIDM in regulating the dams under its purview. FERC will need to further develop RIDM staff capability and staff resources within its organization and develop or support development of increased RIDM capacity in the private sector, see Section 8.1.4**.** As the first dam safety regulatory agency in the United States to embrace RIDM, FERC's success with this effort will influence the potential application to broader dam safety practice in the country, potentially serving as a guide for state regulatory programs that may be interested in applying RIDM.

8.1.1 As low as reasonably practicable

The Panel recommends the Agencies clarify and broaden their use of the ALARP (risks reduced "As Low as Reasonably Practicable") principle in dam safety decision making, taking into

consideration whether the recent developments in Australian dam safety practice for applying the factors used in an ALARP judgement could be of benefit.

While the federal and agency guidelines on dam safety risk management stress the importance of ALARP, the Panel did not see evidence of consistent and detailed application of ALARP in dam safety decisions.

In contrast, it is the Panel's understanding that in Australian dam safety practice ALARP is a common consideration in dam safety decisions. This consideration includes evaluation of what would be needed to bring the dam into full ALARP compliance, and structuring staged dam safety improvements so they do not preclude long term completion of modifications to comply with ALARP.

The ALARP principle for safety is not widely used in United States federal practice, but is more used within British and commonwealth countries. Nonetheless, the Panel considers the concept to have merit and to warrant broader application in dam safety decision making.

There are challenges that arise with the prospect of using the ALARP principle, as it is not an exercise in mathematics, but involves carefully weighing of several measures to reach a judgment. For the Agencies to advance their practice in this area they will need to consider what is contemporary good practice in the design of a dam, and how to balance the cost and practicality of applying that practice to an existing dam against the benefits that would be realized, and a range of other considerations. If benefit cost analysis is used as a consideration, thought will need to be given to the use of measures such as the value of statistical life saved and the role of "disproportionality."

8.1.2 Safety case

The Panel recommends the Agencies evaluate the safety case approach using 'prevention, mitigation and control' (or bow tie) methods employed in other hazardous industries to help ensure all reasonably practicable measures have been taken to reduce risk for all potential failure mechanisms, including surveillance, maintenance and other activities that control dam safety.

Current Agencies' practices include "making the dam safety case" in presenting risk evaluations to decision makers. As documented in Best Practices, this includes presenting the arguments supporting the recommended decision, based principally on consideration of life safety risks.

The Agencies could derive benefit from exploring inclusion in their programs of the broader safety case approach, as it is employed in several other major hazards industries. The UK Ministry of Defence (Ministry of Defence, 2007: p 9) defines the safety case concept as "a structured argument, supported by a body of evidence, that provides a compelling, comprehensible and valid case that a system is safe for a given application in a given environment." While this approach has been unusual in international dam safety practice, it has been widely used in chemical, oil and gas, offshore, and military applications. The Panel believes this approach is worth further development for dam safety programs. Of important consideration, however, is that the safety case approach needs to apply to the life cycle of a facility, not just to the risk analysis phase, and it needs to encompass such things as human factors, facility monitoring, and maintenance strategies.

A broader safety case approach would also provide benefit by establishing a methodology for disaggregating failure modes and barriers, so ALARP can be applied transparently to each prevention and mitigation safety control.

8.1.3 Agency policy vs. public policy

Based on the information provided and discussions with the Agencies, the Panel believes RIDM policies currently in use in the dam safety programs, including the quantitative risk guidelines

(USACE and FERC refer to the quantitative risk guidelines as *tolerable risk guidelines*), are best described as agency policy rather than public policy. The policies, including the quantitative risk guidelines, seem not to have been subject to any political or public process of agreement or review by federal agencies such as OMB or GAO. The Agencies appear to have been thoughtful and diligent in developing their guidelines, by benchmarking to dam safety practices in other countries and to other daily risks experienced by the US population, but the results cannot be considered public policy without public or political process. The Panel therefore supports the view held by the Agencies that decisions should not be based solely on where risk estimates plot compared to their risk guidelines. This also underscores the importance of ALARP as a guiding principle in dam safety decision-making.

The core question is, who is to decide the levels of risk to life and limb that are tolerable, and who is to decide trade-offs of cost vs. life safety. These are inherently political questions - not engineering questions.

The Panel understands FERC is considering the possibility of incorporating RIDM as a requirement in its dam safety program, which FERC believes may require a public rule making, even though FERC is an independent commission in the sense of Federal Law. If this were to occur, the resulting policies could become public policy and both Reclamation and USACE might find themselves subject to the resulting guidance.

8.1.4 Workforce, training, and continuous learning

The Panel recommends FERC explore opportunities to provide enhanced training in risk assessment to dam safety consultants through its existing collaboration with USSD, establishment of a cooperative arrangement with ASDSO, or potentially involving academic institutions to support or deliver aspects of the training.

The available pool of trained dam safety risk analysts is inadequate for the nation’s need. Today, most of these trained personnel reside within federal agencies. With the potential expansion of dam safety risk analysis to FERC-regulated private sector hydropower facilities, the inadequacy of the pool of trained personnel will become more severe. Training is needed to expand this pool, especially among private sector practitioners.

The Panel recommends the Agencies establish precursor analysis and data-driven performance assessment programs to improve the analytical models and calibrate subjectively estimated model parameters and probabilities.

Many major accidents are preceded by precursory (or near miss) events that, although observable, are not recognized as an early indication and forerunner of a catastrophic event until after the fact. This core concept underpins a variety of processes and programs in industries such as nuclear, petrochemical, and the space program. For example, the most important precursor events can be mapped to events trees to quantify the conditional probability of failure given the precursor, or near miss and equipment failure reports can be used to inform safety improvements.

8.1.5 Inter-agency collaboration and partnerships

For identifying learning opportunities and for contributing to industry best practice, the Panel recommends the Agencies continue to share their experiences and practices with each other and the broader dam safety community, through professional organizations and international partnerships.

USACE, Reclamation and FERC maintain close collaborative ties. Furthermore, each Agency has strategic national and internal partnerships. They all share their technical reference manuals and risk management guidelines online, and frequently publish scientific papers on dam safety and risk analysis methodology. The Agencies thereby open themselves to external scrutiny while contributing to the national and international states of practice.

8.2 Agency Governance

The governance processes implemented by the Agencies appear generally reasonable based on the high-level IEPR and the charge of this Panel.

While the level of the evaluation does not directly reveal obvious areas for improvement, other information available to the Panel has led to reservations regarding effectiveness of public stakeholder communications and potential challenges in the provision of risk assessment training to consultants working for FERC's licensees. These issues are addressed in Section 8.3.11 Risk Communication and Section 8.1.4 Training.

For the purposes of this report the term "governance" refers to the organizational processes of high-level policy setting, management operationalization of the policy to deliver its objectives, and the associated checking and feedback loops to ensure the required management actions are completed or refined and that policy is adjusted if it is necessary.

8.3 Agency Methodology

In the Panel's view the RIDM directions of the dam safety programs in the Agencies are appropriate and sound. The Agencies' implementations of RIDM are generally consistent with federal guidance provided in FEMA, 1979 (reprinted in 2004) and FEMA, 2015. However, the levels of maturity in the application of RIDM in Agencies differ. In each agency, details of implementation would benefit from refinement.

8.3.1 Consistent terminology

The Panel recommends the Agencies link terms such as *uncertainty* and *confidence* to definitions commonly used in probability, statistics, and decision science. Clear definitions are essential for moving the methodology forward and for communicating among its users.

Uncertainty is an umbrella term, and its usage in risk analysis can be many-faceted. It could be useful to establish a fuller taxonomy of the sources and types of uncertainty in dam safety risk studies than the taxonomy that now exists in the Best Practices guidance. Among the uncertainties encountered in dam safety are the historical frequencies of naturally random events (e.g., flood or earthquake frequency), often called aleatory. Uncertainties in the statistical models of these frequencies (e.g., annual exceedance curves and their parameters) are due to limited sample sizes and are calculated using statistical principles. They are epistemic but not subjective. Uncertainties about physical processes such as the physics of failure are usually epistemic and may be based either on validation experiments or subjective degrees-of-belief.

Confidence is a sampling theory (frequentist statistical) term associated with estimates of the frequencies of naturally random events based on finite numbers of observations. The term has well-defined meaning within traditional statistical practice, and should be used with caution when the intent is other than this historical use. In statistics, *confidence*, is only used in relation to the variability of a sample and of the summary statistics calculated from it.

In Agency practice, the term *confidence* is used to mean something akin to, "the probability that a decision would not change upon further analysis." This is a useful definition, but confusing to non-Agency risk analysts and statisticians. A better term with some history of use might be *decision confidence* (Dan & Fleming, 2018).

8.3.2 Potential failure mode analysis

The Panel recommends the Agencies consider the follow revisions to the PFMA process:

1. **Broaden the scope of the PFMA to include consideration of non-reservoir-release events that have significant consequences.**
2. **Standardize and better define the criteria for eliminating PFMs from detailed consideration.**

The PFMA process used to date by the Agencies, and the broader dam safety community, is limited to consideration of PFMs defined as a sequence of events leading to uncontrolled release of the reservoir. This excludes from consideration events that may have significant consequences, even though they do not ultimately cause a reservoir release, such as the 2017 Oroville Dam spillway incident. The Panel believes this is a shortcoming for dam safety management programs.

It is the Panel's understanding that the criteria used to screen the initial long lists of PFMs to the shorter list to be considered in detail vary among organizations and teams. Clarification of these criteria would improve consistency.

8.3.3 Dependence

The Panel recommends the Agencies improve their methods for dealing with dependence, more specifically:

1. **The dependence among failure modes (e.g. multiple seismic pier failures), and**
2. **The dependence among the uncertainties on nodal probabilities.**

The Agencies generally treat uncertain variables as independent. Stochastic dependence only appears to be considered for some failure modes. The common cause adjustment procedure used by the Agencies for dealing with correlations between failure modes is approximate at best. Simpler and more accurate methods are available in the literature and used elsewhere.

The assumption that uncertainties are independent across nodes and load intervals can, at least in theory, yield probabilities that are not monotonically increasing with the loading on a dam and/or probabilities that violate experts' overall degree-of-belief probabilities. The Agencies are therefore encouraged to carefully review and improve their methodologies for dealing with dependence.

8.3.4 Systems and operational risk

The Panel recommends risk assessments include consideration of the potential impacts of human error; ageing; malfunctioning of monitoring, remote or automated electro-mechanical control equipment; and other events that may not result in dam failure but could lead to casualties, major environmental damage, or significant repair costs.

A dam safety management program should incorporate learning from dam failures and incidents and, where appropriate, measures taken to improve the program accordingly. Two well-known examples where important lessons have been learned in the United States are the Taum Sauk dam failure and the Oroville Dam spillway incident. At Taum Sauk, a complex system interaction between human actions, monitoring and control equipment, and civil infrastructure lead to dam failure. It was fortunate that no lives were lost. The dam at Oroville did not fail as a result of the spillway chute failure, but the incident resulted in massive social dislocation and repair costs.

At present, the Agencies' approach to dam safety risk is to focus on catastrophic loss of containment failures. These studies are sometimes limited to the hazards posed by extreme hydrologic or seismic events and internal erosion. Spillway gate availability is sometimes included. This practice fits the purpose intended at present. However, as QRA for dam safety evolves and matures it should increasingly incorporate systems engineering considerations, as is now common in other hazardous industries, including offshore operations, chemical plant safety, and nuclear power safety. Such an approach has also been found important to safety case propositions.

8.3.5 Gate systems reliability

The Panel recommends that, since analytical methods used to appraise gate systems reliability seem well suited to the task, actuarial data be used to estimate failure rates in lieu of subjective probability where possible. In practice, these statistical databases are evolving, and continued development of such databases is seen as both a research and a practical need for the industry.

The unreliable performance of hydraulic gates is not uncommon, whether caused by the gate itself, its structural support, its hoist systems, or SCADA and automatic controls. Examples are documented in Best Practices and by the National Dam Performance Program, although the statistical rates of such unreliability are poorly documented across the industry. The USACE Asset Management Program has collected data for many years on locks, gates, and their component reliabilities, especially in the navigation program (e.g., Best Practices Table G-4-2), and these data are considered by the USACE in routine periodic assessments and in fault tree analyses for QRAs. This is an excellent start, and the data are badly needed across the industry. Current efforts at compilation of such data should continue to be a research priority.

8.3.6 Expert opinion elicitation

The Panel recommends the Agencies collaborate in developing a set of practical, state-of-practice guidelines to improve the credibility and quality of risk estimates elicited from subject matter experts.

EOE guidelines should be consistent with modern good practice of subjective probability elicitation as reflected in the current EOE literature. They could include a) process and criteria for experts selection, including guidance on size and composition of a panel of experts, b) elicitation protocols and procedures, c) expert calibration and training on normative precepts such as fundamental concepts of probability and uncertainty quantification, and d) methods for achieving consensus or generating results representative of the spectrum of opinions, with explicit consideration of potential for bias and overconfidence.

Developing such guidelines will also provide an opportunity to bring in lessons learned from past applications of EOE and evaluate, for example, whether the results of past expert elicitations are broadly calibrated to experienced failure rates and the differences among past expert opinion elicitations.

8.3.7 Human error

The Panel recommends the Agencies develop a graded human reliability analysis method for use in dam safety risk assessments, leveraging methods developed by other industries, particularly the nuclear power, petrochemical, and aviation sectors.

Human error during design, construction, operation, and maintenance can be an important contributor to dam safety risk. Yet a human reliability analysis method does not currently exist for dam safety applications. Such a method can improve risk estimates accuracy. In addition, the systematic analysis of the causal factors of human failure can aid the search for effective risk reduction measures.

8.3.8 Life loss estimation

The Panel recommends continued use and improvement of HEC-LifeSim and RCEM, with scrutiny for uncertainties in life loss estimates and the level of detail needed for decision making. More specifically, the Panel recommends the Agencies:

1. **Continue to conduct validation studies when useful cases present themselves. Events such as the Oroville Dam spillway incident may offer valuable opportunities for validating and improving the warning and evacuation models in HEC-LifeSim,**
2. **Perform benchmark studies to identify opportunities for mutual learning,**
3. **Include model uncertainty in HEC-LifeSim uncertainty analyses. Methods have been introduced in recent years on how to incorporate results of validation studies and other types of evidence about accuracy and credibility of models in assessing model uncertainty.**

In addition, the Panel recommends FERC develop up-to-date guidelines on life loss estimation for use by its licensees.

While Reclamation’s RCEM is relatively simple to use, it relies heavily on judgment and only provides useful guidance for dams that closely resemble RCEM’s case histories. For high-consequence dams, RCEM should be used with caution or supplemented with other estimation methods, as done by Reclamation. The USACE’s HEC-LifeSim is more demanding, but more broadly applicable. Its visual outputs can also help create awareness and stimulate discussions with EMAs on improving disaster preparedness. Considering both methods have their pros and cons, the Panel refrains from recommending use of either model. What works best will depend on local circumstances and the level of experience within each agency.

8.3.9 Portraying societal risk

The Panel recommends the Agencies more clearly explain the different charts they use for portraying societal risks. More specifically, the Panel recommends that:

1. **The Agencies more clearly explain the differences between f-$\overline{N}$ plots and F-N curves in agency guidelines. The incremental life loss (N) on an F-N chart is not the same as the conditional expectation of the incremental life loss ($\overline{N}$) on an f-$\overline{N}$ chart.**
2. **The Agencies correct the F-N plotting example from Best Practices, and include a step-by-step plan for plotting F-N curves.**
3. **The Agencies develop detailed guidance on recommended practice for portraying risks on f-$\overline{N}$ charts. To ensure that any point that is plotted on such a chart (including the end points of uncertainty whiskers) can be meaningfully compared to an AALL contour line, the life loss values that are plotted along the horizontal axis should be expectations conditional on the events whose probabilities are plotted along the vertical axis. It could be left to risk analysis teams to decide which failure events and uncertainty bandwidths are to be included in f-$\overline{N}$ charts. However, to avoid confusion and error, the procedure for plotting on f-$\overline{N}$ charts should not be based on judgment, but on clear definitions and the rules of probability calculus.**
4. **Reclamation more clearly explain that the horizontal distance range of the diagonal whiskers on its f-$\overline{N}$ plots should not be interpreted as life loss ranges in its guidelines. Adding a horizontal whisker to indicate the uncertainty related to the conditional expectation of the incremental life loss may also help to avoid confusion.**

USACE and FERC use what are designated in their guidelines as the F-N and f-$\overline{N}$ charts, while Reclamation uses only the f-$\overline{N}$ charts. The latter type of chart is also sometimes referred to as an f-N chart, such as in the Federal Guidelines for Dam Safety Risk Management (FEMA, 2015). This is confusing to non-Agencies stakeholders, since the incremental life loss and its conditional expectation are different quantities. This difference has important ramifications. A thorough understanding of the backgrounds of F-N and f-$\overline{N}$ plots is essential for avoiding misinterpretation and error.

8.3.10 Evaluating societal risk: risk guidelines

While the risk guidelines used by the Agencies on f-$\bar{N}$ and F-N charts look identical, they are fundamentally different. Decision-makers should be aware of these differences to avoid confusion and error. The Panel recommends the Agencies more clearly explain these differences in agency guidelines.

While the downward sloping guideline on the f-$\bar{N}$ chart is a perfect representation of the Average Annualized Risk Guideline, the downward sloping guideline on the F-N chart is not. Furthermore, the bounded areas on an f-$\bar{N}$ and F-N charts (for N or $\bar{N}$ greater than 1,000) mean different things. An F-N curve is more likely to plot into the bounded area than an f-$\bar{N}$ plot due to the difference between N and $\bar{N}$. Because of this, the risk guidelines on f-$\bar{N}$ and F-N charts may yield different judgements in some (presumably rare) cases.

The Panel recommends the Agencies review use of truncations within their societal risk guidelines.

Truncations are sometimes added to societal risk guidelines because of government policy or practical considerations relating to risk estimation. The Agencies have adopted truncations for their societal risk guidelines, and it appears that this, at least partially, may be related to the use of truncations in Australia. The practice in Australia relating to truncations is changing and in this light the Panel suggests that the Agencies review the use of truncations within their societal risk guidelines and adjust if it is then considered necessary.

8.3.11 Risk communication

The Panel recommends the Agencies engage parties outside of the Agencies with an interest in dam safety regarding the Agencies' use of RIDM for dam safety, as stipulated in the Federal Guidelines for Dam Safety Risk Management. To this end, the Panel recommends the Agencies more clearly define their objectives related to risk communication, and that they develop practical guidelines for achieving those objectives.

The Panel recommends the Agencies carefully review their policies concerning public access to inundation maps and clearly articulate their decisions. Without access to inundation maps, it is impossible for third parties to establish whether they have an interest in a dam's safety.

The Federal Guidelines for Dam Safety Risk Management (FEMA, 2015) presents risk communication as a critical component of RIDM that must be integrated into all aspects of dam safety risk management. According to these guidelines, risk communication is not merely a means to improve the *public's* understanding of dam safety risks and agency decisions, but also a means to improve the *Agencies'* understanding of risk and public opinion. In practice, the results of risk assessments are mostly communicated *to* parties that have an interest in dam safety, rather than communicated *with* them, insofar as results are made public at all.

While security might be a reason for restricting public access to some types of information, such as inundation maps, it cannot justify the gap between the level of ambition set forth by the Federal Guidelines for Dam Safety Risk Management and Agencies' practice. Despite the vast literature on the topic, effective risk communication is notoriously difficult, and it often requires expertise and skills that are not necessarily the same as those required for conducting risk assessments.

REFERENCES

Ale, B., Hartford, D., & Slater, D. (2015). ALARP and CBA all in the same game. *Safety Science*.

Alvi, I. (2013). *Human Factors in Dam Failures*. ADSO Annual Conference, Providence, Rhode Island.

Alvi, L. (2018). *Human Factors in the Oroville Dam Spillway Incident*. ADSO Webinar, Lexington, Kentucky.

ANCOLD. (2003). *Guidelines on Risk Assessment*. Australian National Committee on Large Dams.

Armstrong, J. S. (1985). *Long-range forecasting: From crystal ball to computer* (2nd ed). Wiley.

ASDSO. (2020). *ASDSO Mission*. Retrieved March 30, 2020

Assaf, H., & Hartford, D. N. D. (2001). Physically-Based Modelling of Life Safety Considerations in Water Resource Decision-Making. *Bridging the Gap*, 1-10.

Ball, D., & Floyd, P. (1998). *Societal Risks*. Health and Safety Executive.

Becker, B. (2020, April 30). *Three Agency Review—Governance* [Personal communication].

Bedford, T., & Cooke, R. (2001). *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press.

Blackett, F. (2018). *Identifying, Describing, and Classifying Potential Failure Modes*. Adapted from Bureau of Reclamation Best Practices Presentation.

Boyer, D. (2019, December 18). *FERC Overview Presentation Tri-Agency Dam Safety Risk Review Meeting*.

BSEE. (2015). *BSEE launches near-miss reporting system at OTC*. Bureau of Safety and Environmental Enforcement.

Budnitz, R. J., Apostolakis, G., Boore, D. M., Cluff, L. S., Coppersmith, K. J., Cornell, C. A., & Morris, P. A. (1998). Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis*. *Risk Analysis, 18*(4), 463-469.

Bureau of Reclamation. (1985). *Application of Statistical Data from Dam Failures and Accidents to Risk-Based Decision Analysis on Existing Dams*. U.S. Department of the Interior, Bureau of Reclamation.

Bureau of Reclamation. (1997). *Interim Guidelines for Achieving Public Protection in Dam Safety Decision Making*. U.S. Department of the Interior, Bureau of Reclamation.

Bureau of Reclamation. (2003). *Guidelines for Achieving Public Protection in Dam Safety Decision Making*. U.S. Department of the Interior, Bureau of Reclamation.

Bureau of Reclamation. (2011a). *Dam Safety Public Protection Guidelines—Examples of Use*. U.S. Department of the Interior, Bureau of Reclamation, Dam Safety Office.

Bureau of Reclamation. (2011b). *Dam Safety Public Protection Guidelines, A Risk Framework to Support Dam Safety Decision-Making—Interim*. U.S. Department of the Interior, Bureau of Reclamation, Dam Safety Office.

Bureau of Reclamation. (2011c). *Rationale Used to Develop Reclamation's Dam Safety Public Protection Guidelines—Interim*. U.S. Department of the Interior, Bureau of Reclamation, Dam Safety Office.

Bureau of Reclamation. (2012). *Stampede Dam—Final Design Modifications—Risk Analyses*. U.S. Department of the Interior, Bureau of Reclamation, Dam Safety Office.

Bureau of Reclamation. (2014a). *RCEM - Reclamation Consequence Estimating Methodology—Dam Failure and Flood Event Case History Compilation—Interim—Draft*.

Bureau of Reclamation. (2014b). *RCEM - Reclamation Consequence Estimating Methodology—Examples of Use—Interim—Draft*.

Bureau of Reclamation. (2015). *RCEM - Reclamation Consequence Estimating Methodology—Guidelines for Estimating Life Loss for Dam Safety Risk Analysis—Interim*.

Bureau of Reclamation. (2019a). *Annual Dam Safety Assessment Report—Program Evaluation Fiscal Year 2018*. U.S. Department of the Interior, Bureau of Reclamation, Dam Safety Office.

Bureau of Reclamation. (2019b, December). *Bureau of Reclamation Dam Safety Program Policies and Practices*. Tri Agency Briefing, Denver, Colorado.

Bureau of Reclamation, & USACE. (July 2019). *Best Practices in Dam and Levee Safety Risk Analysis*.

Capka, D. (2020, April 14). *Three Agency Review—Governance* [Personal communication].

Central Water Commission, Government of India. (2019). *Guidelines for Assessing and Managing Risks Associated with Dams* (Doc. No. CDSO_GUD_DS_10_v1.0; p. 426). Central Dam Safety Organization.

Cooke, R. M. (1991). *Experts in uncertainty: Opinion and subjective probability in science*. Oxford University Press.

Covello, V. T., von Winterfeldt, D., & Slovic, P. (1986). Risk Communication: A Review of the Literature. *Risk Abstracts*, *3*, 171-182.

Craddock, T., & Duval, D. (2018). *2017 Oroville Response and Recovery*. State Water Project Oroville Recovery & Asset Management.

Cruz, C. O., & Rodovalho, E. da C. (2019). Application of ISO 31000 standard on tailings dam safety. *REM - International Engineering Journal*, *72*(1 suppl 1), 47-54.

Dan, B., & Fleming, S. M. (2018). Distinct Encoding of Decision Confidence in Human Medial Prefrontal Cortex. *Proceedings of the National Academy of Sciences*, *115*(23), 6082-6087.

DELWP. (2015). *Engaging communities on dam safety—A guide for dam owners*. State Government of Victoria, Department of Environment, Land, Water and Planning.

Dewispelare, A., Herren, L., & Clemen, R. (1995). The Use of Probability Elicitation in the High-Level Nuclear Waste Regulation Program. *International Journal of Forecasting*, *11*(1), 5-24.

Droguett, Enrique L, and A. Mosleh. (2014). *Bayesian treatment of model uncertainty for partially applicable models*. Risk Analysis 34.2: 252-270

Dyer, M. K., Little, D. G., Hoard, E. G., Taylor, A. C., & Campbell, R. (1972). *Applicability of NASA Contract Quality Management and Failure Mode Effect Analysis Procedures to the USGS Outer Continental Shelf Oil and Gas Lease Management Program*.

EPA. (1988). *Seven Cardinal Rules of Risk Communication*. U.S. Environmental Protection Agency, Office of Policy Analysis.

Evans, A. W., & Verlander, N. Q. (1997). What Is Wrong with Criterion FN-Lines for Judging the Tolerability of Risk? *Risk Analysis*, *17*(2), 157-168.

Federal Register. (1987). *Radiation Protection Guidance to Federal Agencies for Occupational Exposure; Approval of Environmental Protection Agency Recommendations*. Federal Register, Part II.

Feinberg, B., & Major, J. (2020, April 1). *Life Loss Modeling, U.S. Bureau of Reclamation* [Personal communication].

FEMA. (2004). *Federal Guidelines for Dam Safety*. Interagency Committie on Dam Safety, U.S. Department of Homeland Security, Federal Emergency Management Agency.

FEMA. (2015). *Federal guidelines for dam safety risk management* (P-1025; p. 49). Federal Emergency Management Agency.

FERC. (2003, April). *Draft FMA guidance 2001, Chapter 14, Rev0, FERC presentation, slide 58*.

FERC. (2014). *FERC Engineering Guidelines, Risk-Informed Decision Making, Chapter 22, Estimation of Life Safety Consequences, draft*.

FERC. (2016). *Risk-Informed Decision Making (RIDM) Risk Guidelines for Dam Safety—Interim Guidance* (Version 4.1). Federal Energy Regulatory Commission, Office of Energy Projects, Division of Dam Safety and Inspections.

FERC. (2017). *Dam Safety Performance Monitoring Program: Vol. Chapter 14*. Federal Energy Regulatory Commission.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1983). "The Public" Vs. "The Experts": Perceived Vs. Actual Disagreements About Risks of Nuclear Power. In V. T. Covello, W. G. Flamm, J. V. Rodricks, & R. G. Tardiff (Eds.), *The Analysis of Actual Versus Perceived Risks* (pp. 235-249). Springer US.

Fischhoff, B., Watson, S. R., & Hope, C. (1984). Defining risk. *Policy Sciences*, *17*(2), 123-139.

Foster, M., Fell, R., & Spannagle, M. (2000). The statistics of embankment dam failures and accidents. *Canadian Geotechnical Journal*, *37*(5), 1000-1024.

France, J. W., Alvi, I. A., Falvey, H. T., Rigbey, S. J., & Trojanowski, J. (2018). *Independent Forensic Team Report, Oroville Dam Spillway Incident*.

France, J. W., & Martin, J. (2012). Risk Analysis Provides Perspective for Antero Dam. *Annual Conference of the Association of State Dam Safety Officials*.

Galic, D. (2018). The big F-N plotting style: Right for every organization? *U.S. Society on Dams (USSD) Annual Meeting*.

Graham, W. J. (1999). *A Procedure for Estimating Life Loss Caused by Dam Failure—DSO-99-06*. Bureau of Reclamation, Dam Safety Office, Sedimentation & River Hydraulics.

Gross, E. (2020, April 9). *Life loss Modeling, FERC* [Personal communication].

Haine, S. (2015). *Safety and Enforcement Division Staff White Paper on As Low As Reasonably Practicable (ALARP) Risk-informed Decision Framework Applied to Public Utility Safety* (p. 50) [Staff Paper]. California Public Utilities Commission.

Health and Safety Executive. (2002). *Risk Management: Principles and guidelines to assist HSE in its judgements that duty holders have reduced risk as low as reasonably practicable*.

Health and Safety Executive. (2001). *Reducing Risks, Protecting People HSE's decision making process*. Health and Safety Executive.

Health and Safety Executive. (2003). *Assessing compliance with the law in individual cases and the use of good practice*.

Hunyadi, J., Perry, M., France, J. W., & Williams, J. L. (2016). Comprehensive Dam safety Evaluations: Colorado Expands Risk Considerations in its Dam safety Program. *Dam Safety 2016, 33rd Annual Conference of the Association of State Dam Safety Officials*. Dam Safety 2016, 33rd Annual Conference of the Association of State Dam Safety Officials, Philadelphia, Pennsylvania.

ICOLD. (2017). *Dam Safety Management: Operational phase of the dam life cycle. Bulletin 154*. International Commission on Large Dams (ICOLD).

ISO. (2009). *ISO 31000:2009, Risk management—Principles and guidelines*. International Standards Organization.

ISO. (2014). *ISO 55001: Asset management - Management systems - Requirements.* International Organization for Standardization.

Jonkman, S. N., Maaskant, B., Kolen, B., & Needham, J. T. (2016). Loss of life estimation - Review, developments and challenges. *E3S Web of Conferences*, *7*, 06004.

Jonkman, S. N., Vrijling, J. K., & Vrouwenvelder, A. C. W. M. (2008). Methods for the estimation of loss of life due to floods: A literature review and a proposal for a new method. *Natural Hazards*, *46*(3), 353-389.

Jonkman, Sebastiaan N., Jongejan, R., & Maaskant, B. (2011). The Use of Individual and Societal Risk Criteria Within the Dutch Flood Safety Policy-Nationwide Estimates of Societal Risk and Policy Applications: The Use of Individual and Societal Risk Criteria Within the Dutch Flood Safety Policy. *Risk Analysis*, *31*(2), 282-300.

Kariuki, S. G., & Löwe, K. (2007). Integrating human factors into process hazard analysis. *Reliability Engineering & System Safety*, *92*(12), 1764-1773.

King, L. M. (2020). *Using a Systems Approach to Analyze the Operational Safety of Dams, thesis submitted in partial fulfillment of the requirements for the degree Doctor of Philosophy, Graduate Program in Civil and Environmental Engineering*. University of Western Ontario.

Kirwan, B. (1994). *A guide to practical human reliability assessment*. Taylor & Francis.

Kolen, B., Maaskant, B., Jonkman, S. N., & Needham, J. T. (2016). Comparison of evacuation methods used in the Netherlands and the USACE Dam and Levee Safety Programs for the Natomas Basin (CA). *E3S Web of Conferences*, *7*, 19007.

Lind, N. (2007). Discounting risks in the far future. *Reliability Engineering & System Safety*, *92*(10), 1328-1332.

Lumbroso, D. M., Sakamoto, D., Johnstone, W. M., Tagg, A. F., & Lence, B. J. (2011). Development of a life safety model to estimate the risk posed to people by dam failures and floods. *Dams and Reservoirs*, *21*(1), 31-43.

Major, J. (2019, December 17). *Tri-agency IEPR Risk Cadre Overview*. Tri Agency Risk Review Briefing, Denver, Colorado.

Mosleh, A., Siu, N., Smidts, C., & Liu(Eds.), C. (1995). *Model Uncertainty: Characterization and Quantification*. CRR Publication, University of Maryland.

Mosleh, A., Wood (Eds), J. (2014), *Second International Workshop on Model Uncertainty*, CRR Publications, University of Maryland.

NASA. (2011). *NASA Accident Precursor Analysis Handbook*. National Aeronautics and Space Administration, Office of Safety and Mission Assurance.

National Research Council (U.S.). (2007). *Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget*. The National Academies Press.

Needham, J. (2020, March 17). *Life Loss Modeling, U.S. Army Corps of Engineers* [Personal communication].

Needham, J., Morrill-Winter, J., Beam, B., Owen, S., & Fields, W. (2020). *Validating HEC-LifeSim 2.0: Lessons learned from application on historic events*. USSD 2020 Annual Conference.

New South Wales Government Dams Safety Committee. (2006). *Risk Management Policy Framework for Dam Safety*.

NRC. (1983). *Risk assessment in the federal government: Managing the process*. National Academy Press.

NRC. (1988). *Improving Risk Communication.* National Academies Press.

NRC. (1993). *Science and Judgment in Risk Assessment.* National Academies Press.

NRC. (2009). *Science and Decisions: Advancing Risk Assessment*. The National Academies Press.

NRC. (2010). *Review of the Department of Homeland Security's Approach to Risk Analysis*. (U.S.)National Reserach Council.

NSW Government. (2019). *Dams Safety Regulation 2019—NSW legislation*. NSW Government New Legislation.

ONR. (2019). *Guidance on the Demonstration of ALARP (As Low As Reasonably Practicable)* (NS-TAST-GD-005 Revision 10). Office for Nuclear Regulation.

Peck, R. (1973). Influence of Nontechnical Factors on the Quality of Embankment Dams. In *Embankment-Dam Engineering: Casagrande Volume* (pp. 201-208). Wiley.

Perdikaris, J., & Zhou, R. (2018, June). *Consequences of Flooding: Methods for Estimating Loss of Life*. Advances in the Assessment of Loss of Life Workshop, Toronto, Ontario.

Pritchard, S. (2014, March 13). Public safety around dams: Is it taken seriously enough? - International Water Power. *Water Power & Dam Construction*.

Rackwitz, R. (2004). Optimal and Acceptable Technical Facilities Involving Risks. *Risk Analysis*, *24*(3), 675-695.

Raeburn, R., Williams, J., & Blackett, F. (2012). *Risk Informed Decision Making Influences on the Ashton Dam Remediation Project Design*. 32nd Annual United States Society of Dams Meeting and Conference, New Orleans.

Raeburn, R., Williams, J., & Davidson, R. (2015). Ashton Dam - Risk Guided Rehabilitation. *The Journal of Dam Safety*, *13*(5).

Renn, O. (1992). Risk communication: Towards a rational discourse with the public. *Journal of Hazardous Materials*, *29*(3), 465-519.

Renn, O. (1998). Three decades of risk research: Accomplishments and new challenges. *Journal of Risk Research*, *1*(1), 49-71.

Risher, P., Ackerman, C., Morrill-Winter, J., Fields, W., & Needham, J. (2017, September). *Levee Breach Consequence Model Validated by Case Study in Joso, Japan*. ASDSO Dam Safety, San Antonio, TX.

Robilliard, K., & Sih, K. (2018). Proceedings of the 2018 ANCOLD Conference, Melbourne, Australia

Robinson, R., & Francis, G. (2014). *SFAIRP vs ALARP*. Conference on Railway Excellence, Adelaide, Australia.

Schaaf, D. M. (2005). *Engineering Reliability Guidance for Existing USACE Civil Works*.

Sintubin, M. (2018). The Groningen Case: When Science Becomes Part of the Problem, Not the Solution. *Seismological Research Letters*, *89*(6), 2001-2007.

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280-285.

Slovic, Paul. (1999). Trust, emotion, sex, politics, and science: Surveying the risk-assessment battlefield. *Risk Analysis*, *19*(4), 689-701.

Snorteland, N. (2019, December 17). *Tri Agency Risk Review Policies and Procedures*. Tri Agency RIsk Review Briefing, Denver, Colorado.

Snorteland, N. (2020, January 1). *Two Other Questions* [Personal communication].

Snorteland, N. (2020, April 15). *Governance Evaluation Table Draft—USACE* [Personal communication].

Snorteland, N. J. (2019). *Rationale behind the U.S. Army Corps of Engineers Tolerable Risk Guidelines* (RMC-TR-2019-02). U.S. Army Corps of Engineers.

Sorensen, J. H., Mileti, D. S., Richards, G. L., & Pope, J. M. (2018). *Warning issuance, diffusion and public protective action initation during the February 2017 Oroville dam event*.

Spross, J., Olsson, L., & Stille, H. (2018). The Swedish Geotechnical Society's methodology for risk management: A tool for engineers in their everyday work. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, *12*(3), 183-189.

U.S. Department of Defense. (1949). *Procedures for performing a failure mode effect and criticality analysis*. Department of Defense.

U.S. Code of Federal Regulations, Pub. L. No. Title 18.

U.S. Code Title 16 Conservation, 16 Conservation § Chapter 12, Subchapter 1.

USACE. (1995). *Introduction to probability and reliability methods for use in geotechnical engineering* (ETL 1110-2-547; p. 15). US Army Corps of Engineers.

USACE. (1999). *An Overview of Probabilistic Analysis for Geotechnical Engineering Problems* (ETL 1110-2-556; p. 23). US Army Corps of Engineers.

USACE. (2014a). *Safety of dams—Policies and procedures* (ER 1110-2-1156). US Army Corps of Engineers.

USACE. (2014b). *Engineering and Design Safety of Dams—Policy and Procedures* (ER 1110-2-1156).

USACE. (2016). *HEC-RAS River Analysis System—Hydraulic Reference Manual* (Version 5.0). US Army Corps of Engineers.

USACE. (2018). *Phase 2 Issue Evaluation Study Report—Proctor Lake Dam (TX00010), Leon Riverm Comanchy County, Texas*. US Army Corps of Engineers.

USACE. (2020). *HEC-LifeSim Life Loss Estimation, Technical Reference Manual, draft*. US Army Corps of Engineers (USACE), Hydrologic Engineering Center's (HEC).

USNRC. (2008). *United States Nuclear Regulatory Commission Accident Sequence Precursor (ASP) Program Summary De-scription*. U.S. Nuclear Regulatory Commission, Division of Risk analysis, Office of Nuclear Regulatory Research. NRC-2008

USNRC. (2012). *Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies*. U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.

Vick, S.G. (2002). *Degrees of Belief: Subjective Probability and Engineering Judgment*. ASCE.

Von Gersdorff, N., Yen, J., & Cruz, M. (2014). Establishing a Risk Assessment Program Within the Current Regulatory Framework. *The Journal of Dam Safety*, *12*(4).

Wiegmann, D. A., & Shappell, S. A. (2012). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Ashgate.

Zimmerman, R. (1987). A Process Framework for Risk Communication. *Science, Technology & Human Values*, *12*(3/4).

APPENDIX A

PANEL RESUMES

JOHN W. FRANCE, P.E., D. GE, D.WRE

GEOTECHNICAL AND DAM RISK EXPERT; PANEL LEAD

BACKGROUND

Mr. France has more than 45 years of experience in engineering consulting and design. Most of his technical work for the past 35 years has focused on dams and water retention structures. This experience includes dam safety inspections and analyses, risk analyses, detailed geotechnical and geological field and laboratory investigations, hazard classification, seepage and static stability analyses and evaluations, seismic stability/seismic deformation analyses, conceptual and final designs of new structures, rehabilitation of existing structures, and consultation during construction.

Mr. France He has served on several Consultant Review Boards and Senior Technical Advisory Panels for federal agencies, including the Bureau of Reclamation (Reclamation), Federal Energy Regulatory Commission (FERC), and U.S. Army Corps of Engineers (USACE), to review dam safety programs and modification designs. He has served on two Advisory Boards for BC Hydro for dam safety studies and modifications of five of its dams. He has also instructed courses for seepage analysis and rehabilitation, slope stability analysis, emergency action plans, and embankment dam design for the Association of State Dam Safety Officials (ASDSO) and USACE.

He is an active member of the ASDSO and U.S. Society on Dams (USSD). He is a past Vice president and Ex-Officio Member of the Board of Directors for USSD and a past chairman of the Affiliate Member Advisory Committee for ASDSO. Mr. France also served for six years as the private sector member on the National Dam Safety Review Board. In 2010, he was the recipient of the prestigious President's Award from ASDSO for his contributions to dam safety. He has published several papers and presentations for ASDSO, USSD, and the American Society of Civil Engineers.

EDUCATION

M.S., Civil Engineering, Cornell University, 1976

B.S., Civil Engineering, Cornell University, 1972

PROFESSIONAL REGISTRATION

Registered Professional Engineer: Colorado, Massachusetts, North Carolina, Oklahoma

RELEVANT EXPERIENCE

MOSUL DAM, IRAQ, FOR U.S. ARMY CORPS OF ENGINEERS. Facilitated a semi-quantitative risk analysis for Mosul Dam in support of the USACE's role as engineer of record for the re-initiation of foundation at this critical dam in Iraq. The three-week-long effort considered the full range of potential failure modes (PFMs) for this dam founded on rock strata with soluble gypsum layers. Also provided senior technical review and advice for the grouting program.

OROVILLE DAM SPILLWAY INCIDENT FORENSIC INVESTIGATION TEAM, BUTTE COUNTY, CALIFORNIA. Leader of a six-member team charged with investigating the February 2017 Oroville Dam spillway incident, which consisted of failure of a concrete spillway chute and erosion of a natural hillside downstream of the emergency spillway crest structure, ultimately resulting in the temporary evacuation of almost 190,000 downstream residents. The team developed opinions on causes of the incident, considering both physical factors and human factors.

REVIEW OF DAM SAFETY RISK ANALYSIS PRACTICES. Chair of a five-person team tasked with reviewing the dam safety risk analysis practices of the USACE, U.S. Department of the Interior, Reclamation; and FERC. This effort was undertaken in response to a directive from the U.S. Congress, after the Oroville Dam spillway incident.

HERBERT HOOVER DIKE, FLORIDA, FOR U.S. ARMY CORPS OF ENGINEERS. Served as a subject matter expert and estimator on a team that completed a potential failure mode analysis and detailed quantitative dam safety risk assessment for Herbert Hoover Dike, FL, which is a Dam Safety Action Class (DSAC) 1 (highest risk) facility in the USACE's inventory of dams. Potential dam safety concerns for this 150-mile-long embankment structure centered on seepage and internal erosion PFMs and overtopping and overwash failure modes. This was a four-year long risk analysis effort.

ISABELLA DAM INDEPENDENT EXPERT PANEL, KERN COUNTY, CALIFORNIA, FOR U.S. ARMY CORPS OF ENGINEERS. Served on a team for an independent expert panel review of design and construction of dam safety modifications for Isabella Dam, CA, which is a DSAC 1 facility. The modifications are being designed to address seismic stability and spillway capacity concerns for this existing facility.

WOLF CREEK DAM TECHNICAL ADVISORY PANEL, RUSSELL COUNTY, KENTUCKY, FOR U.S. ARMY CORPS OF ENGINEERS. Chairman of a Technical Advisory Panel reviewing design and construction of major dam safety modifications for Wolf Creek Dam, which has a storage capacity of more than 6 million acre-feet and is a DSAC 1 facility – the class of highest dam safety concern for USACE. The modifications were completed to address seepage concerns in the karstic foundation of the embankment section of the dam. The solution included a concrete diaphragm seepage barrier wall of unprecedented proportions.

DAM SAFETY INDEPENDENT REVIEW PANEL, FOR BUREAU OF RECLAMATION. For five years, served as one of three members of an expert panel charged with providing an annual review of Reclamation's Dam Safety Program. The panel met twice per year, for one week at each meeting, to review Reclamation's dam safety program and provide findings and recommendations, as judged appropriate by the panel. A principal focus of the panel's activities was a detailed use of Reclamation's application of risk analysis and risk-based dam safety decision making.

MORMON ISLAND AUXILIARY DAM AND OTHER EMBANKMENT DAMS ASSOCIATED WITH THE FOLSOM DAM PROJECT, PLACER COUNTY, CALIFORNIA, FOR BUREAU OF RECLAMATION. Member of Consultant Review Boards which provided senior technical review of dam safety evaluations, dam modification designs, and construction for one of the embankment dams that impound Folsom Lake, California. The principal dam safety issues are embankment and foundation seepage and piping, seismic stability concerns and inadequate spillway capacity.

GREGORY B. BAECHER, PH.D., NAE

GEOTECHNICAL RISK EXPERT

BACKGROUND

Dr. Gregory B. Baecher is a Glenn L Martin Institute Professor of Engineering at the University of Maryland. He works principally in risk and reliability of geotechnical and water resource infrastructure. He has authored five books and 250+ publications on these topics.

Dr. Baecher previously was a Tenured Professor of Civil Engineering at the Massachusetts Institute of Technology from 1975 to 1988 and then served as Chief Executive Officer of ConSolve Incorporated from 1988 to 1995. He has been a professor at the University of Maryland since 1995.

EDUCATION

Ph.D., Geotechnical Engineering, Massachusetts Institute of Technology, 1972

M.S., Geotechnical Engineering, Massachusetts Institute of Technology, 1972

B.S., Civil Engineering, University of California (UC) Berkeley, 1968

RELEVANT EXPERIENCE

RECENT PROJECTS RELEVANT TO DAM SAFETY, RELIABILITY, AND RISK MANAGEMENT:

- Oak Ridge National Laboratory, Tenn., Hydropower template for non-hydro dams (2020-Present)
- California Department of Water Resources, Aqueduct Subsidence Program consulting board (2019-Present).
- TransAlta Generation, Calgary, Spillway system risk analysis program (2019-Present).
- CEATI International, Inc., Research needs for floating debris management at dams (2019-2020).
- USACE, Seminal Concepts in Risk, senior risk cadre training program (2017-2020).
- U.S. Nuclear Regulatory Commission and ORNL, State of practice in dam safety risk analysis (2017-2019).
- International Commission on Large Dams, Committee to Update ICOLD Bulletin 130, “Risk analysis in dam safety management” (2014-2019).
- Panama Canal Authority. Development of enterprise risk management guidance for natural and engineering risks (2011-2016).
- BC Hydro, Ontario Power Generation, Elforsk, USACE. Joint industry study on operations and flow-control in dam systems safety (2008-2015).
- USACE, Lakes and Rivers Division Development of risk-screening tools for dams and levees (2007-2013).
- Ontario Ministry of Natural resources: Risk-based dam safety regulations (2006-2012).
- USACE, New England District Multi-year IDIQ for geotechnical risk management consulting services (2006-2010).

- California Department of Water Resources Consultant, California Delta Risk Management Study of levee systems (2006-2008).
- Interagency Performance Evaluation Taskforce Evaluation of existing and future risk protection provided by the New Orleans Hurricane Protection System (2005-2009).

RECENT GOVERNMENT AND RELATED SERVICE:

- Science Advisory Board, Chair, The Water Institute of the Gulf, Baton Rouge, Louisiana (2011-present).
- Coastal Master Plan Science and Engineering Board, State of Louisiana (2009-2011).
- Sub-Committee on Planetary Protection, Science Mission Directorate, NASA (2007-2012).
- Chairman, Committee on Geotechnical and Geological Engineering, National Research Council, DC, (2006-2009).
- Member, Board on Earth Science and Resources, National Research Council, DC, (2006-2009).
- Member, Water Science and Technology Board, National Research Council, DC, (2000-2008).
- Member, Board on Infrastructure and the Constructed Environment, National Research Council, DC, (2001-2005).

AWARDS AND SIGNIFICANT LECTURES:

- 59th Terzaghi Lecture, Geo-Institute of ASCE (invited), Dallas (2021)
- 27th Buchanan Lecture, Texas A&M University, College Station (2019)
- 2nd Lacasse Lecture of GEOSnet, ASCE GeoCongress, Denver (2017)
- GEOSnet Distinguished Achievement Award for contributions to geotechnical reliability (2015)
- UC Berkeley Civil Engineering Academy of Distinguished Alumni (2014)
- Panamanian National Award for Science and Technology Innovation (2012)
- Harleman Lecture, Pennsylvania State University, State College (2010)
- Commander's Award for Public Service, HQ US Army Corps of Engineers (2007)
- US National Academy of Engineering (2006)

BOOK PUBLICATIONS

Hartford, D.N.D., Baecher, G.B., Zielinski, P.A., Patev, R.C., Ascila, R., and Rytter, K., *Operational safety of dams and reservoirs*, Thomas Telford, Ltd., London, 400pp, 2016.

Makhutov, N. and Baecher, G.B. (Eds.), Co*mparative Analysis of Technological & Sociological Consequences of Terrorism (NATO Science Series)*, IOS , Amsterdam, 2012.

Frolov, K. and Baecher, G.B. (Eds.), *Protection of Civil Infrastructure from Acts of Terrorism (NATO Science Series)*, Kluwer Academic Publishers, Dordrecht, 2006.

Hartford, D., and Baecher, G.B., *Risk and Uncertainty in Dam Safety*, Thomas Telford, Ltd., London, 416pp., 2004.

Baecher, G.B., and Christian, J.T., *Reliability and Statistics in Geotechnical Engineering*, John Wiley and Sons, London and New York, 605pp., 2003.

RUDOLF BERNARD (RUBEN) JONGEJAN, PH.D.

FLOOD RISK AND RISK ANALYSIS EXPERT

BACKGROUND

Dr. Rudolf Bernard (Ruben) Jongejan is the director and sole employee of Jongejan Risk Management Consulting B.V. Dr. Jongejan is the co-author of Dutch levee safety guidelines, including the basic guideline "Fundamentals on Flood Protection" (2015) and the new Dutch guideline on the design of hydraulic structures (2019). He serves in numerous work groups including the Safety and Flood Risk Approach Workgroup of the Dutch Expertise Network for Flood Protection (ENW). The ENW advises the Dutch national government on matters related to flood safety.

Besides providing engineering support, Dr. Jongejan also acts as a management or policy advisor. For instance, he supports the Groningen Safety Advisory Board on managing the risk from induced seismicity to the building stock in the northern Netherlands. He has also played an active role in the creation and introduction of new flood protection standards in the Netherlands (codified in 2017).

EDUCATION

Ph.D. for risk management and insurance with applications to flood protection and major industrial hazards, Delft University of Technology (Netherlands), 2008

M.A. with Honors, Political Science, Leiden University (Netherlands), 2007

M.S. with Honors, Civil Engineering and Geosciences, Delft University of Technology (Netherlands), 2004

RELEVANT EXPERIENCE

SYSTEM RISK ANALYSIS, DENVER, COLORADO AND PORTLAND, OREGON. Advisor to the USACE on risk analysis methodology. Responsible for five tasks: (1) compare and advise on methods for handling spatial correlations; (2) advise on (2a) defining the system in a risk analysis for the Willamette Valley, (2b) combining risks, and (2c) portraying system risks to the decision-makers; (3) review the best practices in dam and levee risk analysis; (4) evaluate levee screening tool methodology; and (5) evaluate USACE methodology for estimating and portraying hydrologic uncertainty. Developed a prototype Bayesian flood hazard analysis model. Prepared advisory reports for each task.

JONGEJAN RISK MANAGEMENT CONSULTING B.V (2008-PRESENT). As Director, Dr. Jongejan has participated in several professional activities including:

- Advisor to the Groningen Safety Advisory Board (ACVG) on managing the risk from induced seismicity to the building stock in the northern Netherlands (2020-present).
- Member of a Dutch technical advisory body named (AD, formerly KPR) that provides guidance and technical support for the design of primary flood defences in the Netherlands (2014-present).
- Advisor to the Dutch Ministry of Economic Affairs and Climate Policy on managing the seismic risk in the northern Netherlands (EZK) (2018-2020).

- Project leader of the development and application of new methods for assessing the seismic performance of the embankments with sheet pile walls along the Eemscanal. The project was carried out in collaboration with Fugro, NAM and Noorderzijlvest (2017).
- Project leader of the development of a new method for assessing the seismic performance of large earthen levees with an application to the primary flood defense between Eemshaven and Delfzijl (11.7 km). The project was carried out in collaboration with Fugro, Deltares, NAM, Noorderzijlvest and the National Coordinator Groningen (2016-2017).
- Member of the CRISP project team that carried out a quantitative risk analysis for two coastal areas in Sri Lanka (2015-2017).
- Advisor to Deltares on the development of new tools and guidelines for probabilistic and semi-probabilistic levee safety assessments (2011-2016).
- Advisor to the Dutch government on the introduction of new flood protection standards (2010-2016).
- Partner in a project aimed at developing a risk-informed approach for managing the risk of coastline recession in Australia with UNESCO-IHE and the universities of Wollongong, Sydney and Queensland (2011-2012).
- Member of the technical management team of the VNK2-project, a large-scale quantitative risk analysis for all major levee systems in the Netherlands (about 3,300 km of dykes, dunes and structures that protect roughly two-thirds of the country from flooding) (2010-2014).
- Project leader of a quantitative risk analysis for Central Holland, one of the largest levee systems in the Netherlands with close to 4 million inhabitants (2010).

PUBLICATIONS

Roscoe, K., Hanea, A., Jongejan, R., Vrouwenvelder, T. (2020). Levee System Reliability Modeling: The Length Effect and Bayesian Updating. *Safety* 6 (1): 7. https://doi.org/10.3390/safety6010007.

Jongejan, R.B., Diermanse, F., Kanning, W., Bottema M. (2020). Reliability-Based Partial Factors for Flood Defenses. *Reliability Engineering & System Safety*, January, 193: 106589. https://doi.org/10.1016/j.ress.2019.106589.

Jongejan, R., Drosos, V., Giannakou, A., Chacko, J., Tasiopoulou, P., Zuideveld-Venema, N., De Wit, S., Huissoon, H. (2018). Probabilistic Assessments of Flood Defence Performance Subject to Induced Seismicity. *Bulletin of Earthquake Engineering*, November. https://doi.org/10.1007/s10518-018-0521-7.

Ranasinghe, R., Jongejan, R. (2018). Climate Change, Coasts and Coastal Risk. *Journal of Marine Science and Engineering, 6* (4): 141. https://doi.org/10.3390/jmse6040141.

Dastgheib, A., Jongejan, R., Wickramanayake, M., Ranasinghe, R. (2018). Regional Scale Risk-Informed Land-Use Planning Using Probabilistic Coastline Recession Modelling and

Economical Optimisation: East Coast of Sri Lanka. Jou*rnal of Marine Science and Engineering, 6*(4): 120. https://doi.org/10.3390/jmse6040120.

Zimmaro, P., Kwak, D.Y., Stewart, J.P., Brandenberg, S.J., Balakrishnan, A., Jongejan, R., Ausilio, E., Dente, G., Xie, J., Mikami, A. (2017). Procedures from international guidelines for assessing seismic risk to flood control levees. *Earthquake Spectra*, August 2017, 33(3): 1191-1218. https://doi.org/10.1193/072316EQS117EP.

Jongejan, R.B., Bottema, M. (2017). A new partial factor approach for assessing the reliability of flood defenses. *Proc. Seventh International Conference on Flood Management* (ICFM7), 5-7 September, Leeds, UK.

Kwak, D.Y., Jongejan, R., Zimmaro, P., Brandenberg, S.J., Stewart, J.P. (2017). Methods for Probabilistic Seismic Levee System Reliability Analysis. ASCE-Geo-Risk 2017: *Reliability-Based Design and Code Developments*. Geotechnical Special Publication, 283: 140-150.

Jongejan, R., Ranasinghe, R., Wainwright, D, Callaghan, D., Reyns. J. (2016). Drawing the line on coastline recession risk. *Ocean and Coastal Management*, 122: 87-94.

Jongejan, R.B., Maaskant, B. (2015). Quantifying flood risks in the Netherlands, *Risk Analysis*, 35(2): 252-264.

Ter Horst, W.L.A., Jongejan, R.B. (2015). The importance of domino effects in flood risk assessments: a case study from the VNK2 project, *International Journal of River Basin Management*: 1-9.

Jongejan, R.B., Slootjes, N., Kok, M. (2015). Lessons learned from the introduction of fully probabilistic flood protection standards in the Netherlands. Proc. *Canadian Dam Association 2015 Annual Conference*, 5-10 October 2015, Mississauga, Ontario, Canada.

Wainwright, D.J., Ranasinghe, R., Callaghan, D.P., Woodroffe, C.D., Jongejan, R.B., Dougherty, A.J., Rogers, K., Cowell, P.J. (2015). Moving from deterministic towards probabilistic coastal hazard and risk assessment: Development of a modelling framework and application to Narrabeen Beach, New South Wales, Australia, *Coastal Engineering*, 96: 92-99.

Li, F., Van Gelder, P.H.A.J.M., Ranasinghe, R. Callaghan, D.P., Jongejan, R.B. (2014). Probabilistic modelling of extreme storms along the Dutch coast, *Coastal Engineering*, 86: 1-13.

Jongejan, R.B., Maaskant, B. (2013). The use of quantitative risk analysis for prioritizing flood risk management actions in the Netherlands, Proc. *CDA 2013 Annual Conference*, 5-10 October 2013, Montreal, Quebec, Canada.

Jongejan, R.B., Calle, E.O.F. (2013). Calibrating semi-probabilistic safety assessments rules for flood defences. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*. 7(2): 88-98.

Jongejan, R., Stefess, H., Roode, N., Ter Horst, W., Maaskant, B., (2012). The VNK2-project: a fully probabilistic risk analysis for all major levee systems in the Netherlands, Floods: From Risk to Opportunity, *IAHS Red Book*.

Ranasinghe, R. Jongejan, R.B., Callaghan, D., Vrijling, H. (2012). An innovative approach to determine economically optimal coastal setback lines for risk informed coastal zone management, *Proc. 8th International Conference on Coastal and Port Engineering in Developing Countries* (COPEDEC 2012), 20-24 Feb. 2012, IIT Madras, Chennai, India.

Jonkman, S.N., Jongejan, R.B., Maaskant, B. (2011). The use of individual and societal risk criteria within the Dutch flood safety policy: nationwide estimates of societal risk and policy applications, *Risk Analysis*, 31(2), 282-300.

Jongejan, R.B., Helsloot, I., Beerens, R.J.J, Vrijling, J.K. (2011). How prepared is prepared enough?, *Disasters*, 35(1): 130-142.

Jongejan, R.B. (2008). How safe is safe enough? The government's response to industrial and flood risks, PhD-Thesis, ISBN978-90-9023432-8.

Jongejan, R. and Barrieu, P. (2008). Insuring Large-Scale Floods in the Netherlands, *The Geneva Papers on Risk and Insurance*, 33: 250-268.

SHANE MCGRATH, P.E.

EXPERT DAM SAFETY RISK PRACTITIONER

BACKGROUND

Mr. McGrath is a recognized expert in dam engineering and has served an as independent advisor for dam safety and civil performance review groups. He has held leadership roles for several advisory boards throughout Australia. His past professional institution membership includes being Australia's representative on the International Commission on Large Dams, Dam Safety Committee (leading the working group on assessment of dam break consequences); Past Chairman and the Convener for Guidelines on Risk Assessment for the Australian National Committee on Large Dams; and Board Member of Dams Safety New South Wales (NSW).

Mr. McGrath has been awarded the Churchill Fellowship for "International Practice and Use of Risk Assessment in Dam Management," the River Murray Medal from the Murray-Darling Basin Authority, and was the leadership member of two projects awarded Excellence Awards by the Institution of Engineers. He has also participated as a member of the judging panel for the Institution of Engineers, Victoria Division Excellence Awards.

EDUCATION

Bachelor of Engineering (Civil) – Monash University, Melbourne, Australia

Construction/Project Management Program, Institute of Administration, University of New South Wales, Australia

Australian Institute of Company Directors Diploma

PROFESSIONAL REGISTRATION

Registered Professional Engineer of Queensland (RPEQ).

International Professional Engineer (Australia).

RELEVANT EXPERIENCE

DIRECTOR FOR SGM CONSULTING PTY LTD. (2012-PRESENT). Founder and Director for SGM Consulting, which provides practical advisory services to asset owners on project and program delivery for the management, construction, maintenance and operation of dams, water supply and irrigation infrastructure. Projects include:

- Melbourne Water (Victoria):
  - Safety case development for Greenvale Dam.
  - Peer review of Greenvale Dam safety upgrade design and construction.
  - Peer review of Upper Yarra Dam safety upgrade design and construction.
- Snowy Hydro Limited (NSW):
  - Dam safety program review and management system.
  - New emergency plans for all dams and plan exercises.
- Department of Primary Industries, Parks Water and Environment (Tasmania):

  - Review of Dam Safety Regulatory Strategy and Plan.

- Goulburn Murray Water (Victoria):
  - Peer review of detailed risk assessments for its portfolio of dams.
- Department of Water, Land Environment and Planning (Victoria):
  - Benchmarking of dam safety performance of 14 State dam owning authorities.
  - Briefing of water corporation Directors on dam safety obligations and liabilities.
  - Victorian dam safety regulation framework review.
- South Gippsland Water (Victoria):
  - Delivery of a dam decommissioning guide and project plan for the decommissioning of five water supply dams.
  - Dam safety policy and management plan.
- TasWater (Tasmania):
  - Dam safety program review.
  - Dam safety management system.
- Lower Murray Water (Victoria):
  - Peer review of Victorian Murray Floodplain Restoration Project.
- United Nations (FAO):
  - Review of draft dam safety guidelines for Afghanistan.

BOARD MEMBER FOR SEQWATER (2014-PRESENT). Seqwater is one of Australia’s largest water businesses, with more than $10 billion in assets used to capture, store, treat and distribute bulk water for more than 3.6 million people across South-East Queensland. It has responsibility for 50 dams and 25 weirs.

GOULBURN-MURRAY RURAL WATER CORPORATION (G-MW) (1996-2012). General Manager of Infrastructure (2009-2012) for Australia’s largest rural water authority. Responsible for the management of all G-MW’s major assets and major investment programs (both retail and bulk water services), management of GM-W’s dam portfolio and bulk water services, and the provision of property and environmental services ($5 billion in assets). Acting Managing Director (2010-2011) responsible for the leadership and direction of G-MW during a period of major transformation and change.

PUBLICATIONS

Numerous papers published in the Bulletin of the Australian National Committee on Large Dams and in Congress papers of the International Commission on Large Dams on risk assessment, the safety upgrading of dams and dam safety management.

Member of the working group for the Australian National Committee on Large Dams publication, “Risk Assessment for Dams”, 2003.

Member of the working group to write the International Commission on Large Dams Bulletin 154, "Dam safety management: Operational phase of the dam life cycle".

Risk Assessment and Dam Safety. USA Center for Infrastructure Protection and Homeland Security (CIP/HS) "CIP Report" Volume 10, No 4, October 2011, The Dams Sector. George Mason University School of Law.

ALI MOSLEH, PH.D.

RISK AND RELIABILITY EXPERT

BACKGROUND

Dr. Ali Mosleh, Ph.D., is currently employed as a distinguished professor and Evelyn Knight Chair in Engineering at the University of California, Los Angeles (UCLA). He is also Joint Appointment at the departments of Materials Science and Engineering, Civil and Environmental Engineering, Electrical Engineering, and Mechanical and Aerospace Engineering at UCLA. He was employed at the University of Maryland from 1992-to 2014, and was Nicole J. Kim Eminent Professor of Engineering from 1996 to 2014.

Dr. Mosleh's research areas include simulation-based methods for reliability and probabilistic risk assessment of complex systems; hybrid systems risk and reliability analysis methods (identification of failure mechanisms and reliability prediction of hybrid systems of hardware, software, and human); Bayesian methods for system health monitoring and sensor placement (solution algorithms for hybrid dynamic Bayesian belief networks); Bayesian methods for inference with uncertain evidence, use of expert opinion, and causal modeling with Bayesian belief networks; common cause failure analysis (reliability models, data classification, and analysis); risk-based oversight and inspection of infrastructure systems (i.e., civil aviation safety, oil and gas pipeline integrity management); methods for the impact of organizational factors on socio-technical systems risk; human reliability analysis (cognitive modeling, simulation, and experimental validation); methods for assessing model uncertainty and model validation (Bayesian formalism and application to complex computational codes); risk-informed healthcare system management; system reliability growth tracking and prediction methods; software reliability assessment methodology; cyber security risk management; natural hazards risk analysis; and precursor methodology for risk indicators and early warning systems.

Dr. Mosleh has developed three patents: Quantitative Risk Assessment System, Automation of Common Cause Failure Modeling and Quantification, and System and Methods for Assessing Risk Using Hybrid Causal Logic. Throughout his career, Dr. Mosleh has written 600 papers and reports for journals, conference proceedings, and research projects. He has also either written or served as guest editor for 14 guidebooks and source books.

EDUCATION

Ph.D., Nuclear Science and Engineering, UCLA, 1981

RELEVANT EXPERIENCE

OTHER POSITIONS AND RESPONSIBILITIES:

- Board Member, U.S. Nuclear Waste Technical Review Board (Appointed by President G.W. Bush, 2004 and 2008, continued under President Obama through 2012)
- Advisory Committees and Review Panels including
    - President Clinton's Commission on Critical Infrastructure Protection (1998-2000)
    - Board Member, Marine Board, Transportation Research Board of the National Academies (2010-present)

  - Technical Advisory Board, Japan Nuclear Safety Institute (2012-present)
  - External Advisory Board, Idaho National Engineering Laboratory, NST Directorate, Department of Energy (2014-present)
  - Technical Review Board, The B. John Garrick Foundation for the Advancement of the Risk Sciences (2001-present)
  - National Academies of Science and Engineering, numerous panels and committees (2001-Present)
  - National Academy of Engineering, Bernard M. Gordon Prize for Innovation in Engineering and Technology Education (2013-2018)
  - Civil Aviation Authority, Government of The Netherlands, Aviation Safety Advisory Panel (2005-2009)
  - International Atomic Energy Agency (expert missions to various countries, and member of numerous technical committees on risk, 1986-2004)
  - US Department of Defense, Neutralization of Chemical Weapons (Advisory Panel 2002-2004)
  - NASA Space Shuttle and International Space Station risk studies, (1996-2002)
  - National Institute of Standards and Technology, Information Security Risk (1989-1999)
  - National Science Foundation (Review Panels 1984-2004).
  - Electric Power Research Institute (Advisory and Technical Review Panels, 1984-2004).
  - Nuclear Regulatory Commission (Expert Panels, Review Committees, 1988-Present).
- International Chair Professor, Norwegian University of Science and Technology, Norway (2016-Present).
- Strategic Scientist, Wuhan University of Technology, China (2017-Present).
- Board of Directors, Human Reliability Society (2012-present).
- Board of Directors, International Association of Probabilistic Safety Assessment and Management (1994-2004).
- Chair, Engineering Division, Society for Risk Analysis (1995-2005).
- Engineering Editor, International Journal of Risk Analysis (2001-2005).
- Editorial Boards: Reliability Engineering and System Safety (1994-1999), Risk Analysis (1988-1998), Journal of Risk and Reliability (2006-present)
- Chair, Executive Committee, Human Factors Division of the American Nuclear Society (1996-1998).

HONORS AND AWARDS:

- Member, U.S. National Academy of Engineering (Elected, 2010)
- Nicole J. Kim Eminent Professor of Engineering (2005-2014)
- Glenn L. Martin Institute Professor (2010-2014)

- Fellow, American Nuclear Society (Elected 2013)
- Fellow, Society for Risk Analysis (Elected 1999)
- Tommy Thompson Award, American Nuclear Society (2013)
- Gilbert White Fellow, Resources for the Future (2010)
- Evans- McElroy Award, IEEE Reliability Society (2009)
- Rotary National Award for Space Achievement (1998)
- NASA Flight Safety Award (1998)
- Ford Quality Award, Ford Motor Company (1997)
- P.K. McElroy Award, IEEE Reliability Society (1995)
- Numerous Best Paper Awards: Institute of Mechanical Engineers, American Nuclear Society, European Nuclear Society, System Safety Society, IEEE Reliability Society, Society for Risk Analyses, American Society of Mechanical Engineers, International Conference on Nuclear Engineering (ICONE)

APPENDIX B

LISTS OF DOCUMENTS REVIEWED

APPENDIX B-1

DOCUMENTS INITIALLY PROVIDED BY THE AGENCIES

Federal Emergency Management Agency (FEMA) (2004). *Federal Guidelines for Dam Safety (FEMA-93).*

Federal Emergency Management Agency (FEMA) (2015). *Federal Guidelines for Dam Safety Risk Management (FEMA P-1025).*

Federal Energy Regulatory Commission (FERC). *Safety of Water Power Projects and Project Works (18 CFR Part 12).*

Federal Energy Regulatory Commission (FERC) (2016). *Risk-Informed Decision Making (RIDM) Risk Guidelines for Dam Safety*.

Snorteland, N. J. (2019). *Rationale behind the U.S. Army Corps of Engineers Tolerable Risk Guidelines.* United States Army Corps of Engineers (USACE).

United States Army Corps of Engineers (USACE) (2014). *Safety of Dams - Policy and Proceedures (ER 1110-2-1156).*

United States Army Corps of Engineers (USACE) (2018). *Hydrologic Hazard Methodology for Semi-Quantitative Risk Assessments (RMC-TR-2018-03).*

United States Army Corps of Engineers (USACE) (2018). *SQRA Calculation Methodology (RMC-TN-2018-01).*

United States Department of the Interior, Bureau of Reclamation (Reclamation) (2011). *Dam Safety Public Protection Guidelines.*

United States Department of the Interior, Bureau of Reclamation (Reclamation) (2011). *Dam Safety Public Protection Guidelines - Examples of Use.*

United States Department of the Interior, Bureau of Reclamation (Reclamation) (2011). *Rationale Used to Develop Reclamation's Dam Safety Public Protection Guidelines.*

United States Department of the Interior, Bureau of Reclamation (Reclamation), United States Army Corps of Engineers (USACE) (2015). *Best Practices in Dam and Levee Safety Risk Analysis*

APPENDIX B-2

ADDITIONAL DOCUMENTS PROVIDED BY THE AGENCIES

Part 1: Agency Briefing Documents

Baecher, G. B., Abedinisohi, F., & Patev, R. C. (2015). *Societal Risk Criteria for Loss of Life - Concepts, History, and Mathematics.* University of Maryland.

Best Practices in Dam and Levee Safety Risk Analysis [PowerPoint slides]. (2019).

Collaborations and Sharing Between the Agencies [PowerPoint slides]. (2019).

(2018). *Dam Safety Program Independent Review Panel Report Twentieth Periodic Review.*

(2019). *Dam Safety Program Independent Review Panel Report Twenty-First Periodic Review.*

Differences Between the Agencies [PowerPoint slides]. (2019).

Dudley, S. E., & Hays, S. L. (2007). *Memorandum for the Heads of Executive Departments and Agencies: Updated Principals for Risk Analysis.*

Federal Energy Regulatory Commission (FERC). (2019). FERC Overview Presentation Tri-Agency Dam Safety Risk Review Meeting [PowerPoint slides].

Federal Guidelines for Dam Safety (FEMA-93) [PowerPoint slides]. (2019).

Federal Guidelines for Dam Safety Risk Management (FEMA P-1025) [PowerPoint slides]. (2019).

Galic, P.E., D. (2017). *The Common Cause Adjustment in Dam Safety Risk Analysis: Scope, Purpose, and Applicability*. U.S. Society on Dams.

Galic, P.E., D. (2018). *The Big F-N Plotting Style: Right for Every Organization?* U.S. Society on Dams.

History of the Agencies [PowerPoint slides]. (2019).

National Research Council. (2007). *Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget.*

(2018). *Public Law No: 115-270 (10/23/2018) America's Water Infrastructure Act of 2018 (WRDA).*

Steven Haine, P. (2015). *Preliminary SED Staff Whitepaper on As Low As Reasonably Practicable (ALARP) Risk-informed Decision Framework Applied to Public Utility Safety*. California Public Utilities Commission.

(2018). *Template for Dam Safety Program Decisions after Dam Comprehensive Review or Issue Evaluation.*

U.S. Department of the Interior, Bureau of Reclamation. (2019). *Annual Dam Safety Program Accomplishment Report - Fiscal Year 2018.*

U.S. Department of the Interior, Bureau of Reclamation. (2019). *Annual Dam Safety Program Assessment Report - Program Evaluation Fiscal Year 2018.*

U.S. Department of the Interior, Bureau of Reclamation. (2019). Bureau of Reclamation Dam Safety Program Policies and Practices [PowerPoint slides].

U.S. Department of the Interior, Bureau of Reclamation. (2019). Tri-Agency IEPR Methodology Presentation [PowerPoint slides].

U.S. Department of the Interior, Bureau of Reclamation. (2019). Tri-Agency IEPR Risk Cadre Overview [PowerPoint slides].

U.S. Department of the Interior, Bureau of Reclamation. (2019). Tri-Agency Risk Review Introduction [PowerPoint slides].

U.S. Department of the Interior, Bureau of Reclamation; U.S. Army Corps of Engineers (USACE); University of New South Wales (UNSW); United Research Service (URS). (2007). *Risk Analysis for Dam Safety - A Unified Method for Estimating Probabilities of Failure of Embankment Dams by Internal Erosion and Piping Supporting Document.*

U.S. Department of the Interior, Bureau of Reclamation; U.S. Army Corps of Engineers (USACE); University of New South Wales (UNSW); United Research Service (URS). (2008). *Risk Analysis for Dam Safety - A Unified Method for Estimating Probabilities of Failure of Embankment Dams by Internal Erosion and Piping.*

U.S. Society on Dams. (2011). *21st Century Dam Design - Advances and Adaptations.*

United States Army Corps of Engineers (USACE). (2013). *Draft Meeting Agenda for Federal Guidelines for Dam Safety*.

United States Army Corps of Engineers (USACE). (2019). Dam Safety Program History [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). Tri-Agency Risk Review Introduction [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). Tri-Agency Risk Review Legislation and Authorities [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). Tri-Agency Risk Review Policies and Procedures [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). Tri-Agency Risk Review Questions [PowerPoint slides].

Vroman, N. D., Sills, G. L., Cyganiewicz, J., Fell, R., Foster, M., & Davidson, R. R. (2007). *A Unified Method for Estimating Probabilities of Failure of Embankment Dams by Internal Erosion and Piping.*

Part 2 – Documents Provided by Agencies at Panel Request

From Reclamation

(2017). *Dam Safety Program Independent Review Panel Report Nineteenth Periodic Review.*

Feinberg, B. (2017). *Reclamation's Consequence Estimating Methodology - Using a Two-Dimensional Hydraulic Model to Increase Confidence in Fatality Rate Selection.* U.S. Department of the Interior, Bureau of Reclamation.

McFarland, T. S. (2017). *Baseline Risk Analysis for Minidoka Dam - Decision Document and Technical Report of Findings - Minidoka Project, ID.* U.S. Department of the Interior, Bureau of Reclamation.

Rocha, M. (2018). *Comprehensive Review (CR) Report for New Melones Dam - Safety Evaluation of Existing Dams and Review of Operation and Maintenance Programs Central Valley Project, New Melones Unit, California.* U.S. Department of the Interior, Bureau of Reclamation.

U.S. Department of the Interior, Bureau of Reclamation. (1997). *Reclamation Manual - Directives and Standards (FAC 01-06).*

U.S. Department of the Interior, Bureau of Reclamation. (1999). *Reclamation Manual - Directives and Standards (FAC 01-08).*

U.S. Department of the Interior, Bureau of Reclamation. (2004). *Reclamation Manual - Directives and Standards (FAC 06-01).*

U.S. Department of the Interior, Bureau of Reclamation. (2005). *Reclamation Manual - Directives and Standards (FAC 01-04).*

U.S. Department of the Interior, Bureau of Reclamation. (2005). *Reclamation Manual - Directives and Standards (FAC 06-03).*

U.S. Department of the Interior, Bureau of Reclamation. (2012). *Stampede Dam - Final Design Modifications - Risk Analyses (Technical Memorandum No. STM-8130-FD-2011-04).*

U.S. Department of the Interior, Bureau of Reclamation. (2015). *Reclamation Manual - Directives and Standards (FAC TRMR-66).*

U.S. Department of the Interior, Bureau of Reclamation. (2016). *Reclamation Manual - Directives and Standards (FAC TRMR-95).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01 Appendix A).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01 Appendix B).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01 Appendix C).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01 Appendix D).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01 Appendix E).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-01).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-09 Appendix A).*

U.S. Department of the Interior, Bureau of Reclamation. (2017). *Reclamation Manual - Directives and Standards (FAC 01-09).*

U.S. Department of the Interior, Bureau of Reclamation. (2018). *Flow Charts for Dam Safety Program - Overview, Safety Evaluation of Existing Dams Process, ISCA Process, and Post Award Activities.*

U.S. Department of the Interior, Bureau of Reclamation. (2018). *Reclamation Manual - Directives and Standards (FAC 01-07).*

U.S. Department of the Interior, Bureau of Reclamation. (2019). *C.C. Cragin Dam Decision Document Technical Report of Findings (Technical Memorandum No. CCC-8110-IE-2019-3).*

U.S. Department of the Interior, Bureau of Reclamation. (2019). *C.C. Cragin Dam Issue Evaluation Risk Analysis (Technical Memorandum No. CCC-8110-IE-2019-2).*

U.S. Department of the Interior, Bureau of Reclamation. (2019). *DSPR, Risk, and Dam Safety Activities as of September 1, 2019.*

U.S. Department of the Interior, Bureau of Reclamation. (2020). *Bureau of Reclamation Gate Reliability Practices.*

U.S. Department of the Interior, Bureau of Reclamation. (2020). *Bureau of Reclamation Dam Safety Funding Sources.*

U.S. Department of the Interior, Bureau of Reclamation. (2020). *Contact List for Interviewees.*

U.S. Department of the Interior, Bureau of Reclamation. (2020). List of Dams, Annualized Failure Probability, Life Loss, and Annualized Life Loss Total: Creation of the f-N Chart [Excel Document].

U.S. Department of the Interior, Bureau of Reclamation. *Mission Statements for the Department of Interior, Reclamation, and Dam Safety.*

From USACE

Needham, P.E., J., Morrill-Winter, J., Beam, B., Owen, S., & Fields, W. (2020). *Validating HEC-LifeSim 2.0: Lessons Learned from Application on Historic Events.* United States Army Corps of Engineers (USACE).

Sorensen, P. J., & Mileti, P. D. (2018). *Oroville USACE Warning Project List of Appendices.* United States Army Corps of Engineers (USACE).

Sorensen, Ph.D., J. H., & Mileti, Ph.D., D. S. (2018). *Warning Issuance, Diffusion and Public Protective Action Initiation During the February 2017 Oroville Dam Event.* United States Army Corps of Engineers (USACE).

United States Army Corps of Engineers (USACE). (2018). *Nolin River Dam (KY03011), Periodic Inspection No. 11, Periodic Assessment No. 1, Report.*

United States Army Corps of Engineers (USACE). (2018). Proctor Dam (TX00010) Phase 2 - IES [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2018). *Proctor Dam (TX00010) Phase 2 Issue Evaluation Study Appendices.*

United States Army Corps of Engineers (USACE). (2018). *Proctor Lake Dam (TX00010) Phase 2 Issue Evaluation Study Report.*

United States Army Corps of Engineers (USACE). (2019). *Center Hill Dam Appendix A - Risk Assessment.*

United States Army Corps of Engineers (USACE). (2019). *Center Hill Dam Appendix C - Hydrology and Hydraulics.*

United States Army Corps of Engineers (USACE). (2019). Center Hill Dam Project Overview and Risk Management Plans [Publisher Document].

United States Army Corps of Engineers (USACE). (2019). *Center Hill Dam Safety Modification Report.*

United States Army Corps of Engineers (USACE). (2019). Center Hill Dam Safety Modification Study (TN04102) [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). *Memorandum for the Endorsement of Tentatively Selected Risk Management Plan for Center Hill Dam (TN04102).*

United States Army Corps of Engineers (USACE). (2019). *Memorandum for the Review of the Dam Safety Action Classification for Mojave River Dam (CA10021).*

United States Army Corps of Engineers (USACE). (2019). *Memorandum for the Review of the Dam Safety Action Classification for Nolin Lake, Kentucky (KY03011).*

United States Army Corps of Engineers (USACE). (2019). *Memorandum for the Review of the Dam Safety Action Classification for Proctor Dam (TX00010).*

United States Army Corps of Engineers (USACE). (2019). *Memorandum for the Review of the Dam Safety Action Classification for Truscott Brine Lake Dam (NID TX00038).*

United States Army Corps of Engineers (USACE). (2019). Mojave River Dam (CA10021) Periodic Assessment [PowerPoint slides]. United States Army Corps of Engineers (USACE).

United States Army Corps of Engineers (USACE). (2019). *Mojave River Dam (CA10021), Periodic Inspection No. 14, Periodic Assessment No. 01, Appendices.*

United States Army Corps of Engineers (USACE). (2019). *Mojave River Dam (CA10021), Periodic Inspection No. 14, Periodic Assessment No. 01, Report.*

United States Army Corps of Engineers (USACE). (2019). Nolin River Dam (KY03011) Periodic Assessment Findings [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2019). *Truscott Brine Lake (NID TX00038), Periodic Inspection No. 12, Periodic Assessment No. 1, Report.*

United States Army Corps of Engineers (USACE). (2019). Truscott Dam (TX00038) Periodic Assessment Findings [PowerPoint slides].

United States Army Corps of Engineers (USACE). (2020). *HEC-LifeSim Life Loss Estimation Technical Reference Manual.*

From FERC

Federal Energy Regulatory Commission (FERC). (2014). *FERC Engineering Guidelines Risk-Informed Decision Making, Chapter R22 - Estimation of Life Safety Consequences.*

Federal Energy Regulatory Commission (FERC). (2020). *Response for Request for Additional Information.*

Federal Energy Regulatory Commission (FERC). *FERC Dam Safety Decision Process Chart.*

APPENDIX C

PROGRAM GOVERNANCE EVALUATION

The table below sets out the evaluation of program governance by the Agencies. The evaluation observations were based on information extracted from the documentation provided by the Agencies, briefings in Denver, and responses provided by each agency where required.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
1. Policies and Objectives	1.1. Primary documentation of policy and objectives	USACE	• Engineering Regulation - ER 1110-2-1156 (2014) Safety of Dams - Policies and Procedures (USACE, 2014a). • RMC-TR-2019-02 Rationale behind the US Army Corps of Engineers Tolerable Risk Guidelines (N. J. Snorteland, 2019). The purpose of the Dam Safety program is set out in Section 1.1 of the ER as "compliance with the Federal Guidelines for dam safety." For the purposes of this report this is taken to be the equivalent of a policy. In Section 1.11.1 of (USACE, 2014a) it is stated that "A key mission of the USACE dam safety program is to achieve an equitable and reasonably low level of risk to the public from its dams." For the purposes of this report this is taken to be the objective for the policy.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		Reclamation	• Dam Safety Public Protection Guidelines - A Risk Framework to Support Dam Safety Decision Making (Bureau of Reclamation, 2011b) • Rationale Used to Develop Reclamation's Dam Safety Public Protection Guidelines August 2011 (Bureau of Reclamation, 2011c). • [*FAC P02 Decisions Related to Dam Safety Issues.*](B. Becker, personal communication, April 30, 2020) Page 2 of the guidelines states, "Reclamation has established a risk-informed framework to meet the objectives of its program, the Safety of Dams Act, and the Federal Guidelines." For the purposes of this report this is taken to be the equivalent of a policy. Also, on page 2 of the Guidelines the mission of the Reclamation Dam Safety Program is descried as "To ensure Reclamation dams do not present unreasonable risk to people property, and the environment." For the purposes of this report this is taken to be the objective for the policy. [*FAC P02* - *The objective of the Reclamation dam safety program is to ensure that Reclamation facilities do not present unreasonable risks to the public, public safety, property, or the environment*.] (B. Becker, personal communication, 30 April 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		FERC	• US Code Title 16 - Conservation (U.S. Code Title 16 Conservation, 1935) • US Code of Federal Regulations - Title 18 (U.S. Code of Federal Regulations, 1981) • Interim Guidance - Risk Informed Decision Making (RIDM) (FERC, 2016). The Code S 803 (c) requires licensees to …. "make all necessary renewals and replacement…… and shall conform to such rules and regulations as the Commission may from time to time prescribe for the protection of life, health and property." Code of Federal Regulations (CFR) 18 Part 12 (Safety of Water Power Projects and Project Works) S 12.32 includes the requirement for dams to be "… inspected and evaluated … to identify any actual or potential deficiencies … that might endanger public safety." For the purposes of this report this is taken to be the equivalent of a policy. The mission of the Office of Energy Projects is to facilitate potential benefits to the nation through the review of natural gas and hydropower infrastructure proposals and minimize risks to the public associated with FERC jurisdictional energy infrastructure. For the purposes of this report this is taken to be the equivalent of a policy objective. The Division of Dam Safety and Inspections D2SI Mission Statement - The Division of Dam Safety and Inspections (D2SI) is responsible for ensuring the safety of the Commission's hydroelectric projects and implementing the Commission's dam safety, public safety, security and license compliance programs. The safety programs apply advances in technology to address the technical challenges presented by the aging national water resources infrastructure. D2SI's safety related programs have a direct bearing on life, property and the environment.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
	1.2. Systematic approach to development of policy and objectives including major stakeholder involvement	USACE	The Agency has clearly been diligent in carefully developing its approach to risk informed decision making. [*Stakeholders are defined in ER 1110-2-1156 as cost-sharing partners, communities associated with the flood plain, the Administration, and the Congress. Stakeholder involvement is defined and sometimes required throughout the ER. The agency emphasis is on risk communication and the most robust discussion of this is in Chapter 10. When risk reduction actions are being developed, USACE follows NEPA for official engagement of stakeholders*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	The Agency has clearly been diligent in carefully developing its approach to risk informed decision making. The extent of involvement of major stakeholders is not known
		FERC	The Agency has clearly been diligent in carefully developing its approach to risk informed decision making. [*The draft RIDM Guidelines were posted to the D2SI website and made available for public review and comment. The draft RIDM guidelines were independently reviewed by an external panel of risk experts in the US. The RIDM Guidelines are currently being tested through a series of pilot projects and will be made available for public review and comment prior to finalization*.] (D. Capka, personal communication, April 14, 2020)
	1.3. Endorsed	USACE	[*All USACE Engineering Regulations are approved by the Chief of Staff, office of the Chief of Engineer*.] (N. Snorteland, personal communication, April 15, 2020)]

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		Reclamation	The documentation relates the guidelines to the mission of the dam safety program, the Act and the federal guidelines. [*Reclamation's most recent version of the Public Protection Guidelines were formally transmitted in September 2011 from the Director of the Security, Safety and Law Enforcement to the Deputy Commissioners, all Directors, Regional Directors and Area Managers for adoption and utilization. Reclamation policy for dam safety decision making is mentioned along with the delegation contained within the policy*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*Incorporation of RIDM into the dam safety program was established and endorsed by the Agency and included in the Agency's published strategic plan. The RIDM guidelines have followed established Agency practices for the development and review of new dam safety guidelines. The interim RIDM guidelines have been reviewed and vetted internally by senior management*.] (D. Capka, personal communication, April 14, 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
	1.4. Clear direction	USACE	[*ER 1110-2-1156 was intentionally written to both develop the regulations and document some of the rationale behind the significant change from the previous regulation. The current plan is to revise ER 1110-2-1156 to remove the rationale and procedures within the document. The intent is to replace that with a group of documents:* • *ER 1110-2-1156 - The policy that outlines the objectives and requirements for the program.* • *USACE Dam Safety Program Management Plan - Outline how 1156 will be implemented including roles and responsibilities. Each organizational unit in USACE will be required to have one. (Currently in Draft)* • *Standard Operating Procedures for each major phase of risk-related activities - Periodic Assessment, Issue Evaluation Study, and Dam Safety Modification Study - (Each of these are in draft)* • *Quality Management Plan for the dam and levee safety programs. (Currently each organization has an individual QMP)*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	The policy and objective are clear, but it appears there is opportunity for these to be simply articulated to those who are charged with implementing the policy and objectives.
		FERC	The policy and objective are clear, but it appears there is opportunity for these to be simply articulated to those who are charged with implementing the policy and objectives.
	1.5. Review policy performance and revise		Is there a process for review of policy performance at the management level at which policy is approved?

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		USACE	[*The Chief of Engineers has delegated dam safety authority (as the Dam Safety Officer or DSO) to a Senior Executive at HQUSACE. All policy changes are reviewed with the DSO on a frequent basis. However, most policy items are handled by the Deputy DSO at HQ. Additionally, the program:* • *Annually briefs HQ Senior Executives and the Assistant Secretary of the Army for Civil Works on the state of the dam safety program* • *Quarterly briefs the status of the national program and the individual projects to the agency. This includes program financial information and program metrics.* • *Monthly briefs detailed program and project status to the Deputy DSO and DSO*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Reclamation's Policies and Directives and Standards are subject to a minimum 4-year review and revision (if appropriate) cycle to ensure continued effectiveness and applicability to Reclamation's mission and activities.*] (B. Becker, personal communication, April 30, 2020)
		FERC	[*The review and update of the risk policy documents are the responsibility of the Director, Division of Dam Safety and Inspections. Revisions of the risk guidelines are performed under their direction.*] (D. Capka, personal communication, April 14, 2020)
2. Planning	2.1 Objectives and targets established		Are there program objectives and targets developed from the policy and primary objective?

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		USACE	ER 1.11.1 - Meet tolerable risk guidelines ER 1.11.2.1 - Do no harm (do not increase risk) ER 1.11.2.2 - Implement interim risk reduction measures while longer term solutions are pursued ER 1.11.3 - Risk-informed approach, not risk-based ER 1.11.8 - Urgency of action to reduce risks commensurate to risk - DSAC system (3.2) These are the specific risk-assessment related objectives. These are supported by Continuing Evaluation (Chapter 11), Operations and Maintenance (Chapter 12), Emergency Action Planning (Chapter 16) and other aspects of dam safety management. [*There are additional program metrics in the RMC Program Management Plan that are updated quarterly and briefed to the Deputy DSO and DSO*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Public Protection Guidelines are utilized to ensure reasonable risks to the public. They set the risk assessment guidelines for annualized failure probability and annualized life loss.*] (B. Becker, personal communication, April 30, 2020) A Dam Safety Prioritization Rating (DSPR) is used to evaluate the priority and urgency of risk reduction actions. The risk assessment is supported by monitoring and other risk management activities (inspections, monitoring, Emergency Action Plans (EAPs), periodic examinations etc.).

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		FERC	RIDM Guidelines 1.1.5 - Reduce risk to as low as reasonably practicable (ALARP) Do no harm (do not increase risk) Urgency of dam safety actions is commensurate with risk Risk communication required The risk assessment is supported by monitoring and other risk management activities (inspections, monitoring, EAPs, periodic examinations etc.).
	2.2 Program risk management plan		The identification of risks to the successful implementation of the program and risk mitigation measures?
		USACE	[*The section in the RMC Program Management Plan on "Objectives, Priorities and Metrics" provides a framework for how risks to the successful implementation of the program are identified and mitigated.*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Reclamation's risk management process is documented in the "Safety of Dams Project Management Guidelines (January 2018). The risk management process and associated Dam Safety Decision milestones are specified within the Guidelines. The Guidelines cover all activities from initial identification of potential Dam Safety issues through completion of permanent risk reduction actions (as appropriate). Each Dam Safety Decision document indicates the facility can continue to be operated in accordance with the Standing Operating Procedures or directs implementation of appropriate risk reduction measures, as appropriate.*](B. Becker, personal communication, April 30, 2020)
		FERC	[*A programmatic D2SI-wide and regional office-specific risk management approach and plan are currently under development.*] (D. Capka, personal communication, April 14, 2020)

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
	2.3 Performance Standards		To what standards should the work of the program be undertaken?
		USACE	“Best practices in Dam and Levee Safety Risk Analysis.” (Bureau of Reclamation & USACE, 2019) ER-1110-2-1156, in particular Chapter 18. [*The RMC Program Management Plan also defines the standards to which the program is undertaken, along with the DSMS SOP, IES SOP and PA Facilitators Briefcase*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	“Best practices in Dam and Levee Safety Risk Analysis” Dam Safety Public Protection Guidelines - Examples of Use [*Reclamation Guidance For Evaluation of Internal Erosion - April 2017* *Reclamation Consequence Estimation Methodology* *Comprehensive Review Guidelines (available internally)* *USGS Bulletin 17C for hydrology studies* *Reclamation Design Standards* *Reclamation Manual and Directives and Standards as presented at opening in Denver*.] (B. Becker, personal communication, April 30, 2020)
		FERC	“Best practices in Dam and Levee Safety Risk Analysis” FERC Risk-Informed Decision-Making Guidelines
	2.4 Technical Competence and Functional Understanding		Has the Agency identified the skill requirements to do the work and ensure that only competent people undertake the work? Are employees engaged in the work initiated and trained to understand the program?

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		USACE	Chapter 4 of the ER sets out the required competencies of staff engaged in the dam safety program. [*The dam and levee safety programs also have a training plan that is updated annually. It’s not an official document, but it’s a work in progress that is used by the RMC to select facilitators, team members, and reviewers for risk-related activities. There is a missing section of the document that the agency is working on related to qualifications for engineers responsible for design and construction. This is part of a broader effort in E&C at USACE to establish education, experience, and training standards across the enterprise*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	The Denver Briefing presentation included the qualification requirements for key staff involved in the risk assessments. [*Facilitators, Dam Safety Advisory Team members, and Dam Safety Decision-makers*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*Chapter 2 of the RIDM guidelines provide a table listing the experience and qualifications of individuals performing a risk analysis*.] (D. Capka, personal communication, April 14, 2020)
	2.5 Communications Plan (Internal / External)		Is there a plan on how the program will be communicated, both within the Agency and to external stakeholders?
		USACE	Chapter 10 of the ER sets out the communications plan for the program. This is externally focused. Internal communications is mentioned in relation to EAP Exercises (16.6) and Project Delivery (21.3). [*Each project has a risk communication plan and, as mentioned above, routine monthly calls and meetings are held to communicate the program within USACE*.] (N. Snorteland, personal communication, April 15, 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		Reclamation	Best Practices in Dam And Levee Safety Risk Analysis, Part IX - Risk Assessment Management, Chapter IX-1: Risk Guidelines, S 8.0 Risk Communications sets out the need for internal and external communications plans. [*Reclamation utilizes this information to develop specific risk communication plans for use in emergency management exercises, development and implementation of risk reduction measures, etc.* *FAC P02 Decisions Related to Dam Safety Issues identifies requirements to communicate risk within the agency. Decision documents and decision memorandums are the primary mechanism for documenting and internally communicating dam safety risks and decisions.* *Communication with water and power contractors is required and identified in The Safety of Dams Act of 1978, FAC P02 Decisions in related to Dam Safety Issues, and the Directives and Standards that govern the program.* *Communication with the downstream emergency managers complies with FEMA 64 and with Reclamation's Directive and Standard FAC 01-01 Emergency Action Planning Program for High-and Significant- Hazard Dams.* *Communication with the public is governed and complies with the National Environmental Policy Act*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*Chapter 4 of the RIDM Guidelines sets out the framework for dam safety risk communication for FERC licensees, both internally and externally*.] (D. Capka, personal communication, April 14, 2020)
	2.6 Outsourcing Plan (where required)		Where work is to be outsourced, does the Agency have a plan on how to ensure that the competence and program understanding requirements are met? This could be through mandatory requirements for inclusion in tender documentation and tender selection for example.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		USACE	[*Typically, USACE only outsources individual components of risk assessments, engineering work, or design work. Typically, this falls under tasks like data collection, hydrology studies, seismology studies, and design of individual components. USACE retains overall responsibility to ensure individual work items are consistent with program objectives.*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Reclamation generally utilizes the Technical Services Center for Dam Safety related analyses and designs. When necessary, Reclamation utilizes Indefinite Delivery Indefinite Quantity contracts for outsourcing work to Architecture and Engineering firms. Qualification requirements are specified within the contract. All work performed by outside firms is reviewed by technical experts within the TSC and subject to required vetting through the Dam Safety Advisory Team review and Dam Safety Decision-making processes.*] (B. Becker, personal communication, April 30, 2020)
		FERC	All of the risk assessment work is or will be undertaken by consultants. [*FERC has identified minimum qualifications of individuals performing the work (Chapter 2 of the RIDM Guidelines). Those guidelines also provide the requirement for the risk studies to be reviewed by an independent risk review board (RRB). Members of the D2SI RIDM Branch provides licensees and internal D2SI staff with the information needed to understand the program requirements.*] (D. Capka, personal communication, April 14, 2020)
	2.7 Overall Program Plan		Is there an overall plan for implementation of the program?
		USACE	The USACE presentation at the Denver briefing included a 10-year plan.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		Reclamation	[*Reclamation has utilized a Risk Informed Decision Making process to manage Dam Safety Program activities for over 25 years. Program activities, processes, and decision points are specified in the "Dam Safety Project Management Guidelines." Every high-hazard potential dam owned by Reclamation has been subject to at least 4 quantitative risk analyses as part of our routine risk management processes identified in the Guidelines. Reclamation included a presentation of the 10 year plan for implementation of risk reduction actions during the presentation in Denver*.] (B. Becker, personal communication, April 30, 2020)
		FERC	There is no overall program plan, as the implementation is not mandatory.
3. Implementation	3.1 Delivery structure and resource plan		Is there a specific structure established to deliver the plan? Have the required resources been identified?
		USACE	The ER Chapter 4 contains information on structure and responsibilities. The Denver briefing presentation included the management structure.
		Reclamation	This is a long term program that is embedded within the Reclamation organization. Denver briefing presentation included the management structure.
		FERC	[*Within the past several years, the Agency has created a RIDM branch within D2SI to establish policy, procedures, and methodologies for implementing risk concepts in dam safety. The Agency is in the process of developing specific internal processes and structures to develop, track, and manage risk related activities within the dam safety program*.] (D. Capka, personal communication, April 14, 2020)
	3.2 Procedures to ensure responsibilities are met		Are there position descriptions (or equivalent) that describes each program workers' accountabilities and responsibilities?

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		USACE	The ER Chapter 4 contains information on structure and responsibilities. Are these reflected in the individual position description? [*The structure and responsibilities are also further defined in the draft program management plans and SOPs*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Formal Position Descriptions are established and implemented for each program worker.* *Reclamation utilizes several internal and external control measures to ensure Dam Safety responsibilities are met, including:* • *Conducting Annual Dam Safety Reporting Meetings involving the Dam Safety Office, Dam Safety Officer, the Director of Dam Safety and Infrastructure, each Regional Director and each Area Manager* • *Completion of an Annual Dam Safety Accomplishment Report to the Commissioner summarizing accomplishment of all key risk management activities* • *Completion of an Annual Independent Review Panel overview of the Dam Safety Program and specific program activities and initiatives* *Completion of an Annual Dam Safety Assessment Report by the Dam Safety Officer reviewing and assessing program performance and effectiveness*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*Procedures are under development to identify and describe staff accountabilities and responsibilities. Existing RIDM staff have position descriptions that establish these responsibilities. Draft position descriptions have been prepared for future RIDM staff responsibilities*.] (D. Capka, personal communication, April 14, 2020)
	3.3 Risk management plan monitored and updated		Is the risk management plan routinely/periodically monitored an updated?

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		USACE	[*The risk management plan is updated quarterly and reviewed by the Dam Senior Oversight Group. The plan is briefed annually to the DSO, HQ senior leaders, and the ASA(CW).*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Dam Safety risks are evaluated at every high-hazard potential facility at least every 8 years, as part of the Comprehensive Review process. Dam Safety recommendations to better understand risk or to reduce risk are prioritized based upon the Dam Safety Priority Rating for the facility (and other considerations). Work plans and priorities are reviewed and updated at least 4 times annually through a Change Management process.*] (B. Becker, personal communication, April 30, 2020)
		FERC	[*A programmatic D2SI-wide and regional office-specific risk management approach and plan are currently under development.*] (D. Capka, personal communication, April 14, 2020)
4. Monitoring and Evaluation	4.1 Internal program assessment against objectives, tasks and standards		Is there a system within the program to review progress against targets?
		USACE	[*There are enterprise metrics, organizational metrics, and project-specific metrics. Enterprise metrics are contained in the draft dam safety program management plan and the USACE Campaign Plan. These are briefed annually to USACE and ASA(CW) leadership. There are organizational metrics for each District, Division, and Center that are reviewed quarterly. There are project metrics for routine and non-routine dam safety activities that are reviewed either monthly or quarterly.*] (N. Snorteland, personal communication, April 15, 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		Reclamation	[*See response to 3.2. Additionally, Reclamation utilizes several performance metrics to measure progress against objectives, tasks, and standards. These include:* • *Dam Safety Recommendations completed as a percentage of total Dam Safety Recommendations made (Goal of maintaining or exceeding 95% of all recommendations have been completed)* • *Decisions documented within 60 days of recommendations being made* • *Measurement of and reduction in average portfolio risk over time* *These metrics are reported within the Annual Dam Safety Accomplishment Report to the Commissioner*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*A programmatic D2SI-wide and regional office-specific RIDM approach and plan are currently under development*.] (D. Capka, personal communication, April 14, 2020)
	4.2 Internal program documentation/reporting on performance		Is the internal program performance assessment formally documented?
		USACE	[*The internal program performance is documented through quarterly briefs or annual briefs*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*See response to 3.2 and 4.1*] (B. Becker, personal communication, April 30, 2020)
		FERC	[*A preliminary conceptual-level framework for the documentation and tracking of RIDM performance has been developed internally; however, until the plan is fully developed and implemented there is no formal documentation available at this time*.] (D. Capka, personal communication, April 14, 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
5. Audit, Review and Reporting	5.1 Review of monitoring and evaluation		Is there a process for program management to regularly review the monitoring and evaluation data and take action as required?
		USACE	[*There are monthly Project Review Board meetings with the Deputy DSO and quarterly Project Review Board meetings with the agency. Both reviews discuss programs and individual projects*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*Reclamation conducts an annual Independent Review of the Dam Safety Program utilizing a panel of 3 independent consultants. The IRP provides a report documenting their review, findings, and recommendations to the Dam Safety Officer. The Dam Safety Officer utilizes this information to generate specific recommendations for program improvement within the annual Dam Safety Assessment Report to the Commissioner. The Dam Safety Office provides status updates for each IRP and Dam Safety Officer recommendation to the subsequent IRPs. The Dam Safety Officers Assessment report includes the status of all incomplete recommendations*.] (B. Becker, personal communication, April 30, 2020)
		FERC	
	5.2 Internal performance audit against program plan		Is there an internal audit of the program?
		USACE	[*There are no official internal audits*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	There is an annual internal review of the program by the Dam Safety Officer, who is independent from the program delivery arm of the organization.

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		FERC	[*As the RIDM program plan is still under development, there has not been opportunity for review of the RIDM-generated monitoring and evaluation data or to take appropriate action. This important element will be addressed in the programmatic D2SI-wide RIDM plan currently under development*.] (D. Capka, personal communication, April 14, 2020)
	5.3 External performance audit against program plan		Is there an external audit of the program?
		USACE	The ER mentions an external ASDSO audit in the year 2001 (S1.8). [*Every 3 years, an Independent External Peer Review is conducted of the USACE dam safety program. The reports are available upon request. Additionally, the program and aspects of the program have been audited by the Army Audit Agency and the General Accountability Office*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	There is an annual independent review panel that reviews the performance of the program in relation to the Federal Guidelines for Dam Safety. In FY 2018, Reclamation contracted with ASDSO to evaluate its Dam Safety Periodic Facility Review (PFR) process.

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		FERC	[*As the programmatic D2SI-wide and regional office-specific RIDM plan is currently under development and RIDM has not been fully implemented, there has not been opportunity for external review or audit of the program incorporating RIDM approaches. That said, the Agency does perform external reviews and audits of specific elements or of the overall dam safety program as needed. For example, the most recent review of select elements of the dam safety program occurred as a result of the Oroville spillways incident, after which the Government Accounting Office (GAO) performed an external audit of the overall dam safety program and the Agency established a FERC After Actions Panel (FAAP) of independent consultants to review the program elements and provide recommendations relative established dam safety program procedures.*] (D. Capka, personal communication, April 14, 2020)
	5.4 Reporting performance to decision makers		Is there a regular report on overall program performance to senior management (level above program delivery management)
		USACE	[*There is an annual brief to USACE leadership and ASA(CW).*] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[See responses to 3.2, 4.1, and 5.1] (B. Becker, personal communication, April 30, 2020)
		FERC	[*As the programmatic D2SI-wide and regional office-specific RIDM plan is currently under development and RIDM has not been fully implemented, there has not been opportunity for the reporting of program performance to senior management (Director, Office of Energy Projects or the Commissioners) incorporating RIDM approaches.*] (D. Capka, personal communication, April 14, 2020)
6. Continual Improvement	6.1 Process to implement recommendations from audit, review and reporting		Does senior management (level above program delivery management) have a process to implement recommendations for change from audit, review and reporting?

Program Governance Element	**Summary Activity Description**	**Agency**	**Observations and Sources**
		USACE	[*There is no scheduled official process, but leadership responds and makes changes at their discretion*.] (N. Snorteland, personal communication, April 15, 2020)
		Reclamation	[*The annual Independent Review Panel and Dam Safety Officer's assessment reports form the basis for formulation and implementation of recommendations for continuous program and process improvements*.] (B. Becker, personal communication, April 30, 2020)
		FERC	[*Depending on the specific recommendation, senior management has a wide range of authorities and approaches available to implement programmatic recommendations. That said, there are still Agency and Congressional limitations regarding resources and other factors that must be considered in implementing the recommendation*.] (D. Capka, personal communication, April 14, 2020)
	6.2 Change management plan		Was there a change management plan to implement the program plan and to manage modifications to the plan?
		USACE	N/A
		Reclamation	[*The Dam Safety Office works with the TSC Risk Cadre to prioritize and implement program and process improvements*.] (B. Becker, personal communication, April 30, 2020)

Program Governance Element	Summary Activity Description	Agency	Observations and Sources
		FERC	[*Identification of the advantages and desire to incorporation RIDM approaches into the Agency dam safety program occurred through normal internal programmatic reviews. A RIDM development and implementation plan was created and vetted through senior management using established internal procedures. Incorporation of RIDM in the dam safety program was established and endorsed by the Agency and included in the Agency's Strategic Plan. The RIDM guidelines have followed established Agency practices for the development and review of new dam safety guidelines*.] (D. Capka, personal communication, April 14, 2020)

APPENDIX D

UNCERTAINTY RELATED TO PROBABILITY

If a degree-of-belief probability is a measure of uncertainty, how can such a probability be uncertain? This can (only) be the case when this probability is conditional on an uncertain outcome, such as the uncertain outcome of further study. The following stylized example is intended to illustrate the fact that it is possible to distill numerous conditional failure probabilities from a dam safety risk analysis, each with a different uncertainty bandwidth, and a different meaning. This makes it important to clearly specify what is meant by 'the' uncertainty related to a risk estimate.

Consider a limit state function for a failure mechanism with four stochastic (uncertain) variables,

$$Z = \beta - \sum_{i=1}^{4} a_i X_i$$

in which β and a_1, a_2, a_3 and a_4 are constants, $\sum a_i^2 = 1$, and X_1, X_2, X_3 and X_4 are independent standard Normal variables. Every random realization of $\{X_1, X_2, X_3, X_4\}$ is a random realization of the limit state function, Z.

A limit state function is an indicator function that yields a value smaller than zero when load exceeds resistance. While the Agencies do not use limit state functions explicitly in failure probability calculations, the results of event tree analyses essentially translate into such functions.

To make the example less abstract, let us assume that:

X_1	is the aleatory uncertainty in next year's maximum flood flow
X_2	is sampling uncertainty (an epistemic uncertainty)
X_3	is the capacity of the dam based on the 'best estimate' probabilities per load interval (an epistemic uncertainty)
X_4	is the uncertainty related to 'best estimate' capacity (an epistemic uncertainty)

The constants a_1, a_2, a_3 and a_4 could be thought of as weights that indicate how important the above uncertainties are to the probability of dam failure.

Based on the above, one could compute a number of different (conditional) probabilities, each with a different uncertainty bandwidth. Table 3 gives several examples. It is emphasized that the table only shows a selection of possibilities. In this relatively simple example with only 4 stochastic variables, there are already 14 different possibilities.

While every (conditional) probability has the same expectation (always equal to P(Z<0), they all have a different uncertainty bandwidth. The size of this bandwidth and what it stands for depends crucially on what we are conditioning on.

From a mathematical standpoint, 'anything goes' when it comes to conditioning: there is no right or wrong. Yet from a decision-making standpoint, not all options are equally informative. Consider, for instance, a bandwidth that shows the potential effect on the dam failure probability of knowing next year's maximum flood flow. Such a bandwidth would be of little practical use, since there is no way to know what next year's maximum flow will be. The same applies to a bandwidth for sampling uncertainty, since this uncertainty can only be reduced by waiting many years to extend the record length (assuming all other data sources have already been exploited). Delaying a decision for such a long time will typically not be an option. By contrast, a bandwidth for uncertainties that can realistically be reduced through more study could be helpful for evaluating whether further study is worthwhile.

The above discussion is closely related to the logic behind *decision confidence*. If a dam failure probability could change from tolerable to intolerable through further study (or vice versa), we could make a different decision following further study. In the above example, the probability of this happening could be obtained from the uncertainty bandwidths shown in the last column of Table 3.

Table 3. Different ways to establish and interpret the uncertainty related to a dam failure probability estimate (selection).

Notation	Meaning	Uncertainty bandwidth
$P(Z<0)$	This is the probability of dam failure when accounting for all uncertainties related to demand and capacity	There is no uncertainty bandwidth related to $P(Z<0)$. The uncertainties related to all stochastic variables are expressed by $P(Z<0)$. $P(Z<0)$ equals $\Phi(-\beta)$, where Φ is the standard normal cdf.
$P(Z<0 \mid X_2)$ Since X_2 is uncertain, this conditional probability is an uncertain quantity	This is the probability of dam failure conditional on an uncertain realization of sampling uncertainty	$P(Z<0 \mid X_2)$ is uncertain: $P(Z<0 \mid X_2) = \Phi(-\beta^*)$ with $\beta^* = \mathrm{Norm}(\beta/c_2, a_2/c_2)$ and $c_2 = (1-a_2^2)^{1/2}$ Note that β^* is a stochastic variable Some useful statistics of $P(Z<0 \mid X_2)$: Mean: $\Phi(-\beta) = P(Z<0)$ Median: $\Phi(-\beta/c_2)$ 5% value: $\Phi(-(\beta+1{,}645a_2)/c_2)$ 95% value: $\Phi(-(\beta-1{,}645a_2)/c_2)$
$P(Z<0 \mid X_4)$ Since X_4 is uncertain, this conditional probability is an uncertain quantity	This is the probability of dam failure conditional on an uncertain realization of the uncertainty related to 'best estimate' capacity	$P(Z<0 \mid X_4)$ is uncertain: $P(Z<0 \mid X_4) = \Phi(-\beta^*)$ with $\beta^* = \mathrm{Norm}(\beta/c_4, a_4/c_4)$ and $c_4 = (1-a_4^2)^{1/2}$ Note that β^* is a stochastic variable Some useful statistics $P(Z<0 \mid X_4)$: Mean: $\Phi(-\beta) = P(Z<0)$ Median: $\Phi(-\beta/c_4)$ 5% value: $\Phi(-(\beta+1{,}645a_4)/c_4)$ 95% value: $\Phi(-(\beta-1{,}645a_4)/c_4)$
$P(Z<0 \mid X_2, X_4)$ Since X_2 and X_4 uncertain, this	This is the probability of dam failure conditional on uncertain realizations of sampling uncertainty and the	$P(Z<0 \mid X_2, X_4)$ is uncertain: $P(Z<0 \mid X_2, X_4) = \Phi(-\beta^*)$

conditional probability is an uncertain quantity	uncertainty related to ’best estimate’ capacity	with $\beta^* = \mathrm{Norm}(\beta/c_{2,4}, (a_2^2 + a_4^2)^{1/2}/c_{2,4})$ and $c_{2,4} = (1-a_2^2-a_4^2)^{1/2}$ Note that β^* is a stochastic variable Some useful statistics $P(Z<0\|X_2,X_4)$: Mean: $\Phi(-\beta) = P(Z<0)$ Median: $\Phi(-\beta/c_{2,4})$ 5% value: $\Phi(-(\beta+1{,}645(a_2^2+a_4^2)^{1/2})/c_{2,4})$ 95% value: $\Phi(-(\beta-1{,}645(a_2^2+a_4^2)^{1/2})/c_{2,4})$

APPENDIX E

EXPERT OPINION IN RISK ASSESSMENT

E. Use of Expert Opinion in Risk Assessment

E.1 Introduction

Historically, decision-makers have utilized expert judgment to supplement data or analysis needed to inform their decision. Cases involving new engineering designs, very rare events, and situations that are beyond our direct experience, call for the use of expert opinion as a surrogate source of information. In the absence of adequate scientific information, decision-makers have to rely on their own intuition or on expert opinion. Therefore, experts can extensively influence key decisions in vital matters such as politics, economy, science, and engineering. Use of expert opinion is inevitable in almost all risk assessments. This includes judgments made by the risk analysts in modeling, gathering and analysis of various types of information, and in assessment of the model parameters. In all these steps limitations or lack of established methods, scarcity of required data, and resources are some of the reasons for relying on subject matter experts.

An expert is an individual with specialized knowledge or skill in some specific domain. While in principle any degree of knowledge of a subject qualifies one as an expert to that degree, a person is called an expert only when he or she is believed to be much more knowledgeable than a layperson on the subject of interest. Expert opinion can be viewed as expression of the judgment of an expert on a subject or issue. An opinion is usually regarded as impression, personal assessment, or subjective estimation of a quality or quantity of interest. Expert opinion, in contrast with factual information, is a judgment or a belief that, at least in the mind of the receiver of the opinion, is based on uncertain information or limited knowledge. It is a tool of last resort for exploring unknown issues that are otherwise inaccessible.

Despite the clear need for expert opinion, it is always applied with much caution. This is because an opinion is not a fact or verified by experiment; it is a person's subjective assessment about a specific subject. The obvious concern is the degree to which the opinion of an expert correlates with the objective reality. In risk analysis where probabilities of rare events are often the subject of expert elicitation and use, underestimation or overestimation of relative or absolute likelihoods could result in costly decisions such as overdesigning with large margins, or accepting undue risk exposure.

While there is a significant body of research on ways to improve the use of expert opinion, other factors such as limited familiarity with relevant research results and time and budget constraints have impacted the quality and credibility of using of expert judgment is practice. This section aims to offer an assessment of the current state of the practice and summarize important methodological findings that could help in improving the utilization of expert opinion in risk studies.

The literature on the use of expert opinions can be roughly categorized under two topical areas:

> **Elicitation of Expert Opinions**, i.e., the processes of selecting and eliciting the opinion of the experts, including selection criteria, expert panel size and composition, and elicitation procedures and protocols.
>
> **Post Elicitation Processing and Utilization of the Expert Opinions**, i.e., the process of relating the elicited expert opinions to the unknown of interest, including how to use the information/estimate provided by the experts, and also information about the experts, in order to estimate the unknown quantity; and in case of multiple experts, how to aggregate the opinions.

Important aspects of these topics are disused in more details in the following.

E.2 Elicitation of Expert Opinion

The available research results indicate that the methods by which expert opinions are elicited can have a significant effect on the accuracy of the resulting estimates. They include a set of steps which may or may not be formally and explicitly followed in practice. Categorized into *formal* and *informal* methodologies, expert elicitation methods may range from simple to complex processes. Formal expert elicitation refers to a structured procedure designed to obtain opinion of experts including, for example, training material to debrief the experts, documented criteria for the selection of experts, and defined roles and set of steps and protocols for information exchange and interactions. Compared to an informal expert elicitation process, formal expert elicitation can increase the quality and credibility of the results, and the defensibility of the judgments, in part due to documentation of each expert's rationale. It also enhances the communication of the results. Formal elicitations are generally viewed to produce superior results compared to those informally gathered (DeWispelare et al. 1995).

Informally elicited judgments are obtained through unstructured approaches which lack adherence to established protocol or scientific principles. Examples of these types of judgments include intuition, arbitrary guesses, and gut feelings (Armstrong 1985, pg 73). Unfortunately, the most commonly used method for decisions under time pressure is unaided judgment. This is not surprising, as unaided-judgment forecasts can often be derived quickly and cheaply. The simplest and most common of these methodologies entails merely asking for an individual's judgment. Most people have poor intuitions regarding numerical probabilities. Consequently, this inquiry yields the least reliable expert performance results, especially for persons unfamiliar with probability concepts (Cooke 1991).

The earliest methodologies for formal scientific use of expert judgments was developed by the Research and Development (RAND) Corporation in the 1950s for the Air Force to facilitate decision making. The first two methods were the Delphi Approach and Scenario Analysis. It should be noted that Delphi refers to a class of methods for elicitation of expert opinion and that the term has been used for many, often totally different, methods. In essence the Delphi Method is a group forecasting technique that uses estimates from experts and rounds of feedbacks to obtain updated estimates by the experts until a reasonable consensus occurs.

Steps of the Original Delphi Method (Dalkey, 1969) are:

(1) Selection of issues or questions and development of questionnaires.
(2) Selection of experts who are most knowledgeable about issues or questions of concern.
(3) Issue familiarization of experts by providing sufficient details on the issues on the questionnaires.
(4) Elicitation of experts about the issues. The experts might not know who the other respondents are.
(5) Aggregation and presentation of results in the form of median values and an inter-quartile range (i.e., 25% and 75% percentile values).
(6) Review of results by the experts and revision of initial answers by experts. This iterative reexamination of issues would sometimes increase the accuracy of results. Respondents who provide answers outside the inter-quartile range need to provide written justifications or arguments on the second cycle of completing the questionnaires.
(7) Revision of results and review for another cycle. The process should be repeated until a complete consensus is achieved. Typically, the Delphi method requires about two cycles or iterations.
(8) A summary of the results is prepared with argument summary for out of inter-quartile range values

Scenario Analysis also introduced by RAND is a process of analyzing future events by considering alternative possible outcomes (scenarios). The analysis is designed to allow improved decision-making by allowing more complete consideration of outcomes and their implications. We note that scenario analysis in form of event sequence diagram (ESD) or event tree (ET) form the modeling backbone of most modern risk assessments of engineered systems.

E.2.1 Selection of Experts

It is generally believed that formal selection processes and application of objective criteria can safeguard against personal biases and preferences in selecting experts. Some studies have relied on a nomination process to establish a list of candidate experts by consulting archival literature, technical societies, professional organizations, and other knowledgeable experts. Formal nomination and selection processes should establish appropriate criteria for nomination, selection, and removal of experts. An example of selection criteria is found in (NRC, 1997) in the context of seismic hazard analysis:

- Strong relevant expertise through academic training professional accomplishment and experiences, and peer-reviewed publications;
- Familiarity and knowledge of various aspects related to the issues of interest;
- Willingness to act as proponents or impartial evaluators;
- Availability and willingness to commit needed time and effort;
- Specific related knowledge and expertise of the issues of interest;
- Willingness to effectively participate in needed debates, to prepare for discussions, and provide needed evaluations and interpretations; and
- Strong communication skills, interpersonal skills, flexibility, impartiality, and ability to generalize and simplify

It has been suggested that the panel of experts should include the following:

- Experts familiar with the specific technical subject
- Experts in the broader domain of knowledge
- Experts in support areas, and related domains, such as statistics, risk analysis, and decision-making
- Observers, discussion facilitators, and expert opinion elicitation experts

The latter two categories can participate in the elicitation process and can contribute to the discussion, but they cannot provide expert opinion that enters in the aggregated opinion of the experts. The integrators and facilitators are responsible for conducting the expert-opinion elicitation process. They can be considered observers or experts depending on the circumstances and the needs of the process.

In selecting and experts one should note that in the case of probability estimation two measures of the quality of expert opinion are **"substantive goodness"** and **"normative goodness”**. Substantive goodness refers to the knowledge of the expert relative to the problem at hand. Normative goodness, on the other hand, refers to the expert's ability to express that knowledge in accordance with the calculus of probability and in close correspondence with his/her actual opinions. In many cases experts are weaker in normative aspects and can significantly benefit from some level of training and calibration. This is an important issue and a key to ensuring highest possible quality and robustness of the results.

E.2.2 Size of Expert Panel

Speculations made about the correlation between expert accuracy and the number of experts used in a study, lead many to conclude the number of experts is directly proportional to accuracy. Of course, if possible, the panel should be large enough to capture complementary expertise and achieve diversity of opinion to ensure a balanced and broad spectrum of viewpoints, expertise, and technical points of view.

"N-heads rule" (Dalkey, 1977) suggests under certain assumptions, the higher the number of experts the better the result. Questions still remain; is this true? And, with seldom unlimited funding, are there a minimal number of experts needed to obtain optimal accuracy? Hogarth's normative model suggested that maximal accuracy can be obtained with 6-10 experts, and Ashton's (Ashton & Ashton 1985) empirical work as well as many of the studies reviewed showed that between 3 and 6 experts lead to high accuracy levels. Cooke, & Probst, (2006) reported that based on panelists interviewed during an expert judgment policy symposium and workshop, the number of experts for most studies they conducted was "targeted to lie between 6 and 12". Seaver, D. A. (1976) observed that "in general the group judgment will be more accurate than the individual judgments primarily due to a decrease in error variance." Martz, et al (1985) concluded that "Aggregation of expert opinion using group medians does give some improvement in accuracy." Other research suggests, however, that gains in accuracy are attributed to the inter-correlations of the experts, and minimal gain in accuracy is achieved from redundancy in experts (Budescu and Rantilla 2000).

Feldman (2016) conducted a study on the variation in the accuracy of the aggregated estimate, versus the number of experts using a larger date base of expert estimates and realized values. The data sets used was extracts from Delft University of Technology (TUD) expert judgment meta-databases described in (Cooke and Goossens, 2008), and the University of Maryland Center for Reliability and Risk Analysis (UMD) expert judgment meta-data sources. The UMD source was derived from two dissertations addressing expert judgment Shirazi (Shirazi, 2009) and Forester (Forrester, 2005), and related research at UMD. The database includes expert estimates for physical quantities and probabilities, with vast majority belonging in the physical category. The number of experts associated with each of the 1788 records in the database ranges from 1 to 45.

The primary metric of expert performance used is the study was the ratio, r = e/e', of the true value to expert predicted value (median estimate if a range was provided by the expert). A metric called Maximum Multiplicative Error (MME) defined as the maximum of r and its inverse, was used to study the variation in the accuracy of aggregated expert judgments as a function of the size of the expert panel. (use of MME penalizes multiplicative excursions of estimates on either side of the realized value equally).

This study was done using several popular aggregation methods including the simple geometric mean method. The analysis was performed separately for physical and probabilistic data types. Figure E.1 shows an example of the results using Geometric Mean Method of aggregation, where MME is plotted against number of experts for the physical quantity estimates in the database.

Feldman (Feldman 2016) concludes that:

1. Increasing the number of experts from one to two reduces MME by a factor of ten.
2. Based on this broad base of meta-data, largest improvements in average MME are observed between n=1 and n=4 experts.
3. Results for probabilistic data are more dispersed than for physical data; therefore, larger numbers of experts should be employed for the former, to help avert large estimation errors.
4. A plateau is reached at about n=6 or n=7, with little consistent improvement for larger n.
5. The extent of improvement associated with increases in numbers of experts is highly dependent on the domain (field of study in which experts' opinions were elicited)

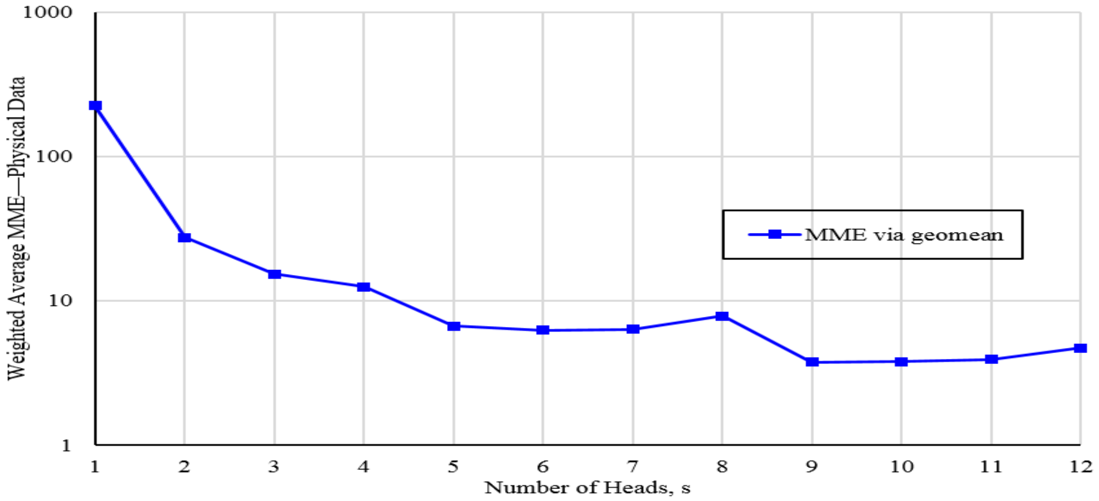


Figure E.1 MME vs Number of Experts for Physical Data (Feldman, 2016)

E.2.3 Preparation for Elicitation

The technical subject of elicitation and questions to be covered should be carefully selected to achieve the stated objectives. The issues should be structured in a logical sequence, starting with background, followed by questions, and then answer selections or answer format, and any needed scales. Such materials can be used to familiarize the experts with the issue of interest.

The following are some guidelines on constructing questions and issues based on social research practices (Ayyub, 2001):

- Each issue can include several questions; however, each question should elicit only one opinion. It is a poor practice to include two questions in one.
- Question and issue statements should not be ambiguous. Also, the use of ambiguous words should be avoided. Special attention should be given to the definitions within the context of each issue or question. The level of wording should be kept to a minimum. Also, word choice might affect the connotation of an issue.
- The use of factual questions is preferred over abstract questions. Questions that refer to concrete and specific matters result in concrete and specific answers.
- Questions should be carefully structured in order to reduce biases of subjects. Questions should be asked in a neutral format; sometimes it is more appropriate not to have lead statements.
- Sensitive topics might require stating questions with lead statements to establish supposedly accepted social norms in order to encourage subjects to answers the questions truthfully.

Another set of guidelines for elicitation is offered by Cooke (1991):

- The issues or questions should be clearly stated without any ambiguity. Sometimes there might be a need for testing the issues or questions to ensure their adequate interpretation by others
- The questions or issues should be stated using appropriate format with listed answers, perhaps graphically expressed, in order to facilitate and expedite the elicitation and scoring processes
- It is advisable to test the processes by performing a dry run
- The analysts must be present during the elicitation and scoring processes
- Training and calibration of experts must be performed. Examples should be presented with explanations of elicitation and scoring processes, and aggregation and reduction of results. The

analysts should avoid coaching the experts or leading them to certain views and answers.
- The elicitation sessions should not be too long. In order to handle many issues, several sessions with appropriate breaks might be needed

Some researchers recommend the use of **decomposition** (breaking a problem into sub-problems) in the problem formulation to improve the overall quality of elicited opinions (see for example Armstrong et al. 1975). Decomposition is one of the most effective ways to utilize an expert's information base and specific evidence. The expert(s) can be asked to respond to questions on each of the parts of a problem. The analyst then synthesizes the responses to construct the forecast. This strategy can be especially useful with a group of experts, each of whom may have information on only part of the problem. The use of decomposition is particularly advantageous when relevant theories exist only for certain aspects of the problem under consideration, or when different experts have information on different aspects of the problem (Armstrong, 1975).

E.2.4 Sources and Treatment of Bias

It has been widely documented that judgmental estimates are subject to a number of possible biases; among the influential early discussions of the subject are those by Tversky and Kahneman[2] and Hogarth. Two biases that are particularly important in the practice of risk assessment are:

(a) The possibility of *systematic* overestimation or underestimation.
(b) *Overconfidenc*e; i.e., the tendency for people to give "overly narrow confidence intervals which reflect more certainty than is justified by their knowledge about the assessed quantities" (Tversky and Kahneman, 1974).

Some causes of bias include:
- <u>Availability,</u> when experts can recall events or situations similar to the event or issue of interest. Accordingly, probabilities of well-publicized events tend to be overestimated, whereas probabilities of events not generally known are underestimated.
- <u>Anchoring,</u> referring to the tendency of people to start with an initial estimate and correct it to the issue at hand. However, the correction might not be sufficient.
- <u>Representativeness,</u> individuals tend to evaluate intuitively the conditional probability P(B|A) by assessing the similarity between A and B. The problem with this assessment is that similarity is symmetric whereas conditional probabilities are not.
- <u>Control factor</u>, refers to the perception of subjects that they can control over outcomes related to the issue at hand.

To this list one can also add a tendency to overestimate extremely small risks and underestimate extremely large risks according to Slovic, et al.

It is clearly important for risk studies based on expert opinion to take the problems of overconfidence and systematic bias (i.e. overestimation or underestimation) into account, and two general approaches have been proposed for doing this. One approach is to correct the expert opinions after they have been elicited to remove overconfidence and systematic overestimation or underestimation; Mosleh and Apostolakis (1986) developed formal Bayesian models for doing this. An alternative approach is to improve the quality of the expert opinions that are elicited in the first place. The latter is explained next, while the former is disused in Section E.3.

Expert **calibration** can help in reducing the effects of bias. A well-calibrated expert can be defined as an individual who would consistently produce an estimate that is in agreement with frequencies observed in the world. Experts can be calibrated by providing them with feedback on their assessments in training-like sessions. The calibration process involves training on probability concepts, error sources, biases, expectation, issue familiarization.

The problem of overconfidence has been extensively studied. For example, Fischhoff, et al. among others document substantial degrees of overconfidence in a wide variety of judgmental tasks involving both almanac questions (i.e., documented facts that are unlikely to be known by the experimental subjects) and also forecasts of future events. The results presented by Lichtenstein and Fischhoff (1977)are the most relevant to our concerns since they document that overconfidence is essentially unaffected by a person's degree of expertise. In other words, an expert does not tend to be any less overconfident than a lay person.

This result has been confirmed by further observations. For example, Fischhoff, et al, show that experts are just as likely as naive subjects to overlook causes of failure that have been omitted from a fault tree. Similarly, in Hynes and Vanmarcke, only three of the confidence intervals provided by a group of seven internationally known geotechnical engineers cover the actual embankment height at which a clay foundation was observed to fail. Based on his review of the literature, Armstrong concludes that feeling of expertise "increase confidence but not accuracy." This suggests that experts may actually be more overconfident than lay persons in some cases. Lichtenstein, et al., review experimental results that test the adequacy of probability assessments, and they conclude that "the overwhelming evidence from research on uncertain quantities is that people's probability distributions tend to be too tight. The assessment of extreme fractiles is particularly prone to bias."

Two of the techniques that have been suggested for improving the quality of elicited opinions are specifically designed to reduce overconfidence. The first of these is calibration training, which involves feedback on the extent of overconfidence that has been exhibited in past assessments (see, for example, Winkler, 1971). The second involves encouraging the expert to actively identify evidence that would tend to contradict his/her initial opinion; see Wason (1968) and also Koriat, et al (1980). Evidence confirming the value of calibration training is provided by Stael von Holstein and Lichtenstein and Fischhoff, among others). Similiarly, Koriat, et al. document the value of active searches for disconfirming evidence. The second technique is to encourage the expert to actively identify evidence that would tend to contradict his/her initial opinion.

Research results indicate that these techniques are moderately effective. In other words, they do not completely eliminate overconfidence, but they do appear to be effective at reducing the extent of the problem. These techniques are particularly valuable in cases for which relevant empirical data are extremely limited or totally lacking for calibration.

E.2.5 Elicitation of Probabilities

Elicitation of probabilities pose special challenges including the fact that technical domain knowledge does not automatically translate into ability to express graded knowledge on the probability scale. In other words, one should not equate "substantive expertise" with "normative expertise". In many cases experts are weaker in normative aspects and can significantly benefit from some level of training and calibration in probability metrics. Notions such as *aleatory* and *epistemic* uncertainties, population variability, and difference between probability, relative frequency, and rate of occurrence of events need to be distinguished as part of elicitation issue definition and in subject matter expert training and orientation. This, along with effective use of calibration and anchoring techniques would improve the accuracy (or at a minimum consistency) of the probability estimates.

Probability elicitation can be facilitated by choosing direct or indirect estimations techniques as appropriate. *Direct Elicitation* of the probability or likelihood of the event of interest can be done in various forms, including:

- Direct assessment without feedback and opportunity for revising the estimate
- Initial direct assessment and one or more rounds of feedback and revision (including knowledge of other experts' estimates when multiple experts are consulted)
- Direct estimation at the end of an interactive nominal group session
- Comparative assessment of the probability of interest in relation to a probability familiar to the expert. For example if the expert is familiar with event A and its occurrence probability P(A) but does not feel as comfortable with assessing the probability of the event of interest B, the expert is asked to assess the relative likelihood, r, of B to A, (e.g., twice as likely). The probability of B can then be estimated as $P(B) = r. P(A)$ subject to compliance with the axioms of probability.

Indirect Elicitation of probabilities entails elicitation of an answer to a question which can be used to obtain the answer to the question of interest, through an explicit or implicit actual or perceived relation. An example is betting rules for developing probability estimates. Another example is estimation of frequency of occurrence based on estimated time to first occurrence of the event.

E.3 Post Elicitation Processing and Utilization of the Expert Opinion

Whether we have estimates from one expert or several experts, the next question to address is how to use the information obtained to develop the estimate to be used estimate in the final analysis or decision making. Based on earlier discussions we know that some modification to the estimates from experts may be needed to improve accuracy (for instance adjusting for potential bias or compensating for overconfidence). Also, when the opinions of several experts are obtained some kind of aggregation is needed to form one estimate ("aggregated opinion").

In the following we will discuss the case of post elicitation calibration of experts, but first it is important to generally characterize the forms in which the elicited quantities estimate may take.

The unknown quantity being estimated could be "single-valued" or "distributed". Estimates of single valued quantities could come in the form of a point estimate, confidence bound, or a distribution (parametric or non-parametric). Estimates of distributed quantities (representing aleatory variability) could also come in form of a point estimate, confidence bounds, a single distribution, or a family of distributions (expressed for instance by uncertainty distribution over the distribution parameters).

E.3.1 Single Expert

Based on earlier discussions we know that some modification to the estimates from experts may be needed to improve accuracy, for instance to adjust for potential bias or to compensate for overconfidence. This can be done by simple "ad-hoc" methods, or in a formal and mathematically structured way via Bayes theorem. Either way, some level of judgment is needed on the part of the analyst (or decision maker) on the degree and form of the adjustments (post elicitation "calibration") to the expert estimate.

Bayesian Methods

In the Bayesian theory expert opinion is treated as evidence that can be used in Bayes' theorem to update the state of knowledge about an unknown of interest:

$$\pi(x \mid x^*) = k^{-1} L(x^* \mid x)\, \pi_0(x)$$

- $\pi_0(x)$ = the analyst/decision maker's "prior" state of knowledge about the unknown quantity x (prior to receiving the opinions of the experts),
- x* = expert's of the value of x,
- L(x*|x) = the likelihood of the evidence x* given that the true value of the unknown quantity is x,
- π(x| x*) = the decision maker's posterior state of knowledge about the unknown quantity, x, given the expert's x*

The problem of expert opinion is thus translated to the assessment of π_0 and L by the decision maker. L(x*|x) dx* is the probability that the expert's estimate will be between x* and (x* + dx*) when the true value is x. In other words, the likelihood function is a measure of the accuracy of the experts estimate, in the eyes of the decision maker. It is the decision maker's model of the expert's ability to correctly estimate the value of the unknown quantity. The shape and functional form of the likelihood may differ from one expert to another for the same unknown quantity. Also, the form of the likelihood, in general, is different for a given expert's assessments of different quantities.

The Additive Error Model (and its logarithmic variation, the Multiplicative Error Model) proposed by Mosleh and Apostolakis (1984) have been used in many applications as models of the likelihood function. In this model the decision maker treats the expert's estimate as a variable, X^*, and thinks of it as the sum of two terms, i.e., $X^* = x + E$, where x is the true value and E is the error term which is assumed to be normally distributed, with mean b (a measure of systematic bias), and standard deviation σ (a measure of range of error). The concept is overconfidence on the part of the expert can be compensated by increasing the value of σ. The two quantities can be estimated from the type of information presented in the following section.

Other Methods

A less formal method is to simply stretch the distribution of the estimate provided by the expert, as proposed by Apostolakis (1978) and applied extensively in PRAs of nuclear power plants. The recommendation was for instance to take 5th and 95th percentiles of the expert provided distribution at 20th and 80th percentiles (reducing 90% confidence to 60%). Some empirical evidence can be cited as a basis for these types of adjustments in a generic way. Among the studies on the accuracy of expert assessments of component failure frequencies is a survey conducted by the U.K. Atomic Energy Authority (Snaith), where observed and predicted reliability parameters for some 130 diverse equipment and systems used in nuclear power plants were evaluated. The predicted values include both direct assessments by experts and also the results of analysis. The objective was to examine correlations between the predicted and observed values.

Figure E.2 shows the cumulative frequency of the ratio (R) of observed to predicted values. As we can see, the majority of the points lie within the dashed boundary lines representing ratios between 0.25 and 4. In fact, 63% of all predicted values are within a factor of 2 from the observed values and 93% are within a factor of 4. The figure also shows that R = 1 is the median value, indicating that there is no systematic bias in either direction. Finally, the linear nature of the curve shows that R tends to be lognormally distributed, at least within the central region.

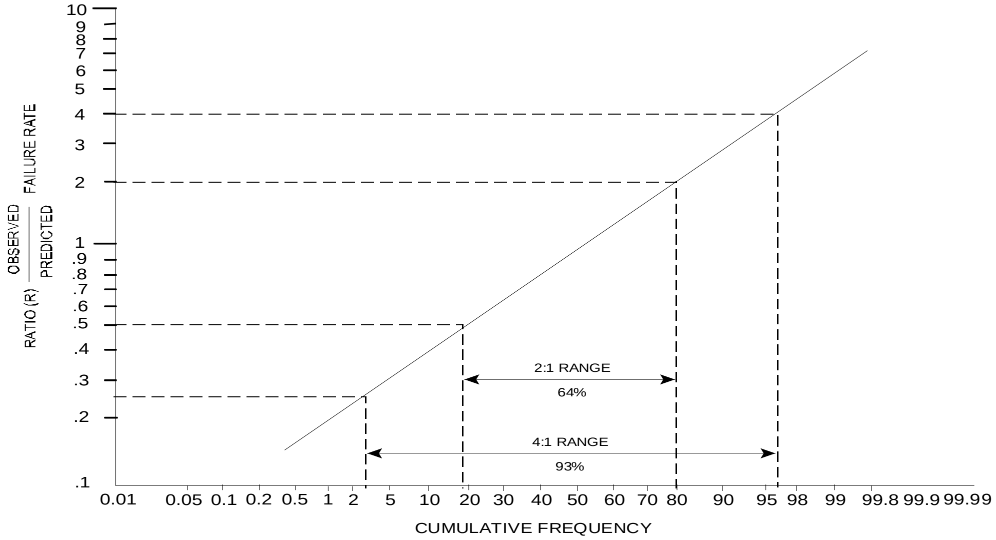


Figure E.2 Cumulative Distribution of the Ratio of Observed to Predicted Estimates for Component Failure Rates

Note that the above information applies only to the assessment of a single unknown quantity rather than a probability distribution. The question of expert performance in assessing the uncertainty about reliability parameters is discussed later in this section.

Feldman and Mosleh (Feldman and Mosleh 2020) used the large database mentioned earlier to explore several questions, including the degree of error (as measures by the distribution of the ratio r≡e/e', where e is the realized value of a quantity, and to e' the elicited prediction). Figures E.3, E.4, and E.5 show the results in a few different forms for both physical quantity and probability estimates. Figure E.3 provides evidence that generically speaking the likelihood of error being outside of the range -2 <Ln (e/e') < +2 is relatively narrow. In the absence of any the other evidence, the distribution shown in Fig 3 can form the basis for assigning effort range to the estimates provided by the experts, or formally used as the likelihood function in the Bayesian formulation mentioned above.

Figures E.4 and E.5 show the overestimation distribution by factor respectively for physical and probabilistic estimates. Comparison of the distribution of r revealed that predictions for probabilistic data were less reliable. For example, estimates against probabilistic variables were roughly twice as likely to overestimate the true value by a factor of ten, compared to physical variables. Consolidating over-and underestimation errors into MME, it was found that MME exceeds various factors such as 2, 5, 10, 100, or 1000 approximately twice as frequently for probabilistic data as for physical data.

Overestimation errors by factors of 2, 5, and 10 are approximately half as likely (to within ten percent) to occur for the full physical data set as for the physical data subset. However, very large overestimations are more likely to occur over the full physical data set. For example, the probability of a factor of 100 overestimation is 1% for the full set, while it is only 0.1% for the subset. The probabilistic data overestimation errors range from 1.5 to 3.5 times more likely to occur, at each factor. The largest discrepancy is at a factor of 1000.

Underestimation probabilities at each factor agree to within approximately 15% with their counterparts for the physical data subset. For example, the probability of a factor-of-two underestimation error is 17% for the subset, and 15% over the full physical data set. The probability of underestimation ranges from 1.5 to 2.5 times more likely for probabilistic data as for physical data over these factors.

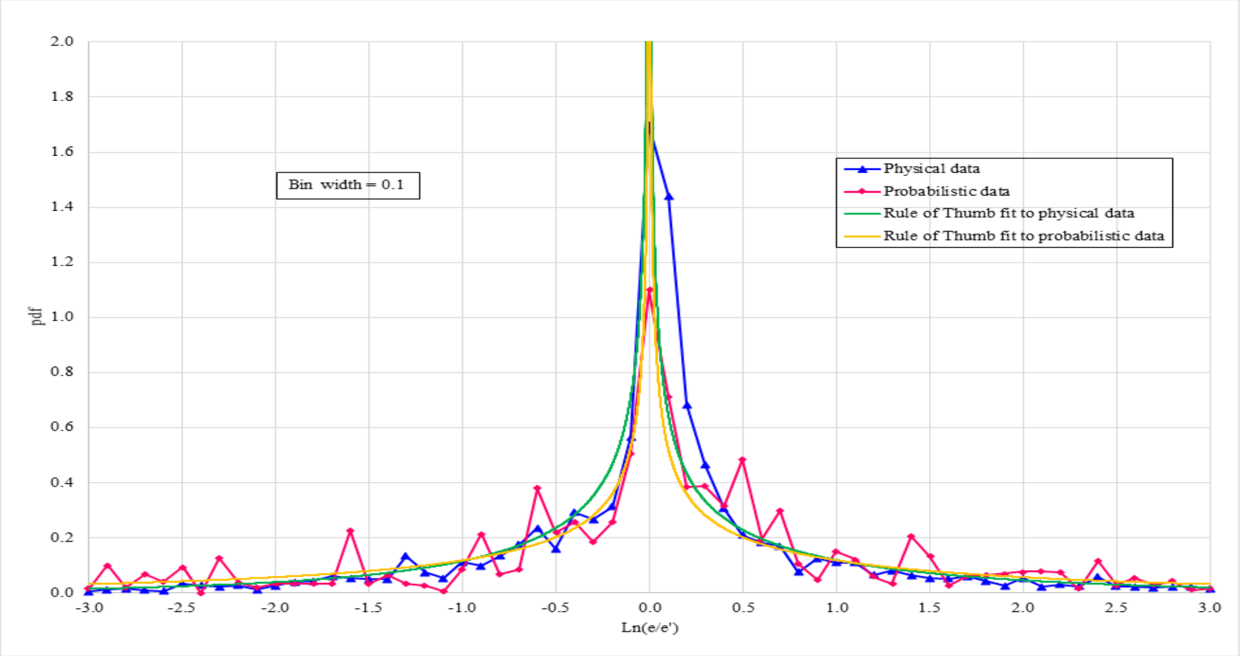


Figure E.3 Density of Ln(e/e') for Physical and Probabilistic Data (Feldman, 2016)

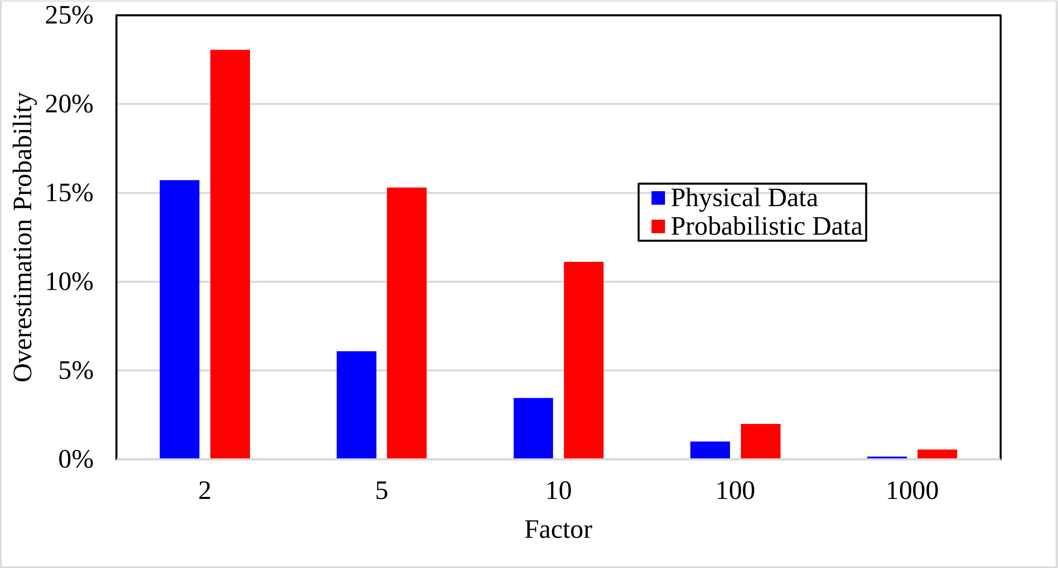


Figure E.4 Overestimation Probabilities by Factor for Physical and Probabilistic Estimates (Feldman, 2016)

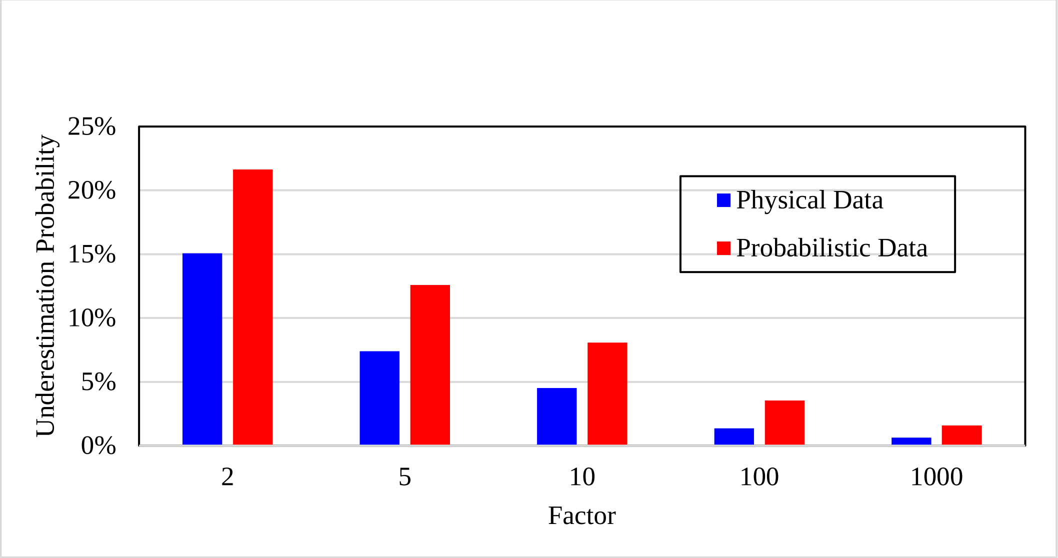


Figure E.5 Underestimation Probabilities by Factor for Physical and Probabilistic Estimates (Feldman, 2016)

Evidence of expert overconfidence can be found in actual probabilistic risk assessments. In several probabilistic risk assessment PRAs of nuclear power plants in which failure rate distributions were developed by updating generic distributions with plant-specific data using Bayesian methods, it was observed that the posterior distributions lay in the right-hand tails of the generic distributions. One explanation for this phenomenon is that perhaps the prior distributions of the Reactor Safety Study (RSS) (US NRC1975) and the ranges of IEEE-500 (EPRI, 1984) were too tight, a result of expert overconfidence. Referring to the RSS, Lichtenstein, et al. state: "The research reviewed here suggests that distributions built from assessments of the 0.05 and 0.95 fractiles may be grossly biased." Commenting on judgmental biases in risk perception, Slovic, et al, state: "A typical task in estimating uncertain quantities like failure rates is to set upper and lower bounds such that there is a 98% chance that the true values lies between them. Experiments with diverse groups of people making different kinds of judgments have shown that, rather than 2% of true values falling outside the 98% confidence bounds, 20% to 50% to do so. Thus, people think that they can estimate such values with much greater precision than is actually the case."

Following these observations, Apostolakis (1978) suggested that the RSS distributions should be broadened by taking the endpoints of the assessed ranges as the 20th and 80th percentiles of lognormal distributions instead of the 5th and 95th. It should be mentioned that the evidence from use of the RSS and IEEE-500 distributions is not conclusive as to the reason for the data falling outside the range; e.g., a bias toward low values, overconfidence, or both. However, failure rate distributions developed using methods that accurately represent the plant-to-plant variability of failure rates, indicate that typical distributions have range factors from 5 for mechanical equipment to 20 for electrical and electronic components. These values are larger than those suggested by the RSS (range factors [1] between 3 and 10).

Further evidence on expert overconfidence was obtained from a comparative analysis of expert judgment and actual data on component maintenance unavailability Mosleh, et.al. Table 1 shows the results of this comparison for distributions is mean maintenance durations for several key components that are typically modeled in PRAs. Listed in the table are the mean values and range factors of the expert-estimated distributions that have been used in several PRAs.

Also listed are similar characteristics of component maintenance duration distributions developed based on detailed review and analysis of component histories at nine operating nuclear power plants. The ratio of the observed to the predicted mean and range factor for each of the component maintenance categories is also given in Table E.1.

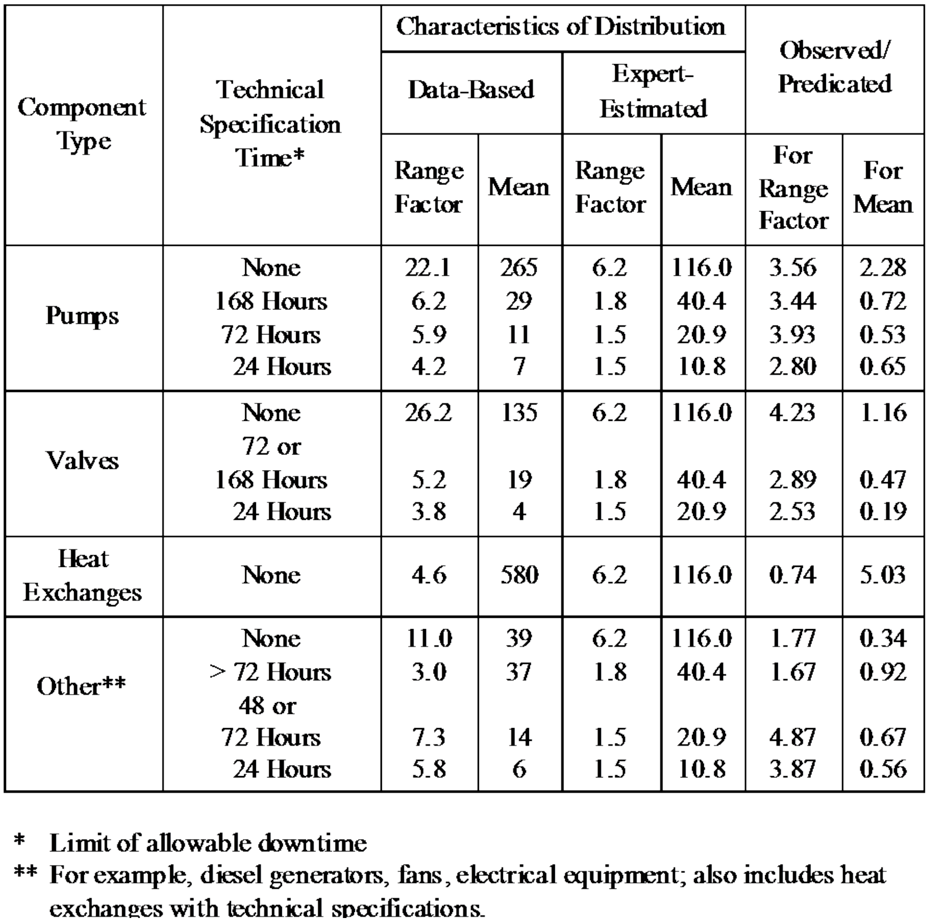
These results indicate that:

- The experts systematically underestimated the actual variation of the mean maintenance duration from one plant or component to another.
- In 11 of the 12 cases listed, the magnitude of the error in estimating the mean value is well within a factor of 4. Also, the experts do seem to be slightly biased toward long durations, as 75% of the estimates are above and 25% below the observed values. These results provide additional data for comparison with the findings of Green and Bourne, and Snaith, as discussed earlier.

---

[1] Range factor is a term used by some authors to refer to the ratio of an upper percentile to the median value of an uncertain quantity which is distributed according to lognormal distribution. Typically, the 95th percentile is used as the upper percentile in this respect. Therefore, a range factor of 2 means $X_{95}/X_{50} = 2$.

Table E.1 Comparison of Data and Expert Opinion on the Distribution of Component Maintenance Time (Pickard, Lowe and Garrick, 1983)



Component Type	Technical Specification Time*	Characteristics of Distribution				Observed/ Predicated	
		Data-Based		Expert-Estimated			
		Range Factor	Mean	Range Factor	Mean	For Range Factor	For Mean
Pumps	None	22.1	265	6.2	116.0	3.56	2.28
	168 Hours	6.2	29	1.8	40.4	3.44	0.72
	72 Hours	5.9	11	1.5	20.9	3.93	0.53
	24 Hours	4.2	7	1.5	10.8	2.80	0.65
Valves	None	26.2	135	6.2	116.0	4.23	1.16
	72 or 168 Hours	5.2	19	1.8	40.4	2.89	0.47
	24 Hours	3.8	4	1.5	20.9	2.53	0.19
Heat Exchanges	None	4.6	580	6.2	116.0	0.74	5.03
Other**	None	11.0	39	6.2	116.0	1.77	0.34
	> 72 Hours	3.0	37	1.8	40.4	1.67	0.92
	48 or 72 Hours	7.3	14	1.5	20.9	4.87	0.67
	24 Hours	5.8	6	1.5	10.8	3.87	0.56

* Limit of allowable downtime

** For example, diesel generators, fans, electrical equipment; also includes heat exchanges with technical specifications.

E.3.2 Aggregating the Opinions of Multiple Experts

There is ample evidence indicating that the aggregation of multiple opinions or forecasts tends to be more accurate than the opinion of a single expert. However, there are important questions about how such aggregations should be performed. This section explores these questions, considering both Mathematical and Behavioral Aggregation

Behavioral Aggregation

Winkler, Stael von Holstein, Rowse, et al, and Brown among others provide evidence indicating that mathematical methods of aggregation generally yield better results than "behavioral" or "consensus" methods, such as the Delphi approach, Dalkey (1969, 1977). This is consistent with the literature on the dysfunctions of interactive group processes, Gustafson, et al.; in particular:

- A "central tendency" effect.
- The tendency for less confident members of the group to limit their participation.
- Group pressures for conformity.
- The strong influence of dominant personalities.

- An investment in maintaining the integrity of the group itself.
- A tendency to reach speedy decisions.

It is worth mentioning that Gustafson, et al., finds that the Estimate-Talk-Estimate method (a structured approach roughly similar to the Delphi technique) is more accurate than the use of unstructured group meetings. Fischer fails to produce a statistically significant replication of this result, but Riggs finds the Delphi technique to be more accurate than use of unstructured group meetings. Based on his review of the available literature in this area, Armstrong concludes that "The traditional group meeting has many problems, not the least of which is the strong pressure to conform ... . The best advice is to use some structured technique."

Mathematical Aggregation

While it is recognized that inputs from a set of subject matter experts conveys more information than inputs from a single expert, there is no consensus on the best way to combine inputs through a mathematical procedure. Methods of expert opinion aggregation vary from simple geometric and, arithmetic aggregations, to weighted tools, to more complex methods that account for dependency. Mosleh (1992) suggested a Bayesian approach to combining expert opinions when there is an inherent variability among the estimates.

Feldman (Feldman 2016) has compared the performance of nine aggregation techniques ranging from simple arithmetic averaging to more complex Bayesian and “classical” (Cooke 1991) methods. The methods are compared in a scorecard considering such factors as accuracy, bounds coverage, bounds width, sensitivity to outliers, and complexity. He concludes that there are no significant differences between simple aggregation schemes (e.g., using equal weights for all experts) and more complicated techniques, such as the use of self-weights, peer-assigned weights, or weights based on past predictive performance. The *median* and the *geometric mean*—are, respectively, the most accurate and in the “middle of the pack” for accuracy. (The arithmetic and harmonic mean, while simple methods of aggregation, were sensitive to outliers, and had poor performance in terms of average MME). Therefore, practical considerations (e.g., cost) lead many researchers to recommend the use of mathematical aggregation with equal weights for all experts.

E.4 Summary of Key Findings for Improving Quality Expert Judgment in Risk Studies

The following list is a summary of several practical recommendations that can be distilled form the literature on the use of expert opinion, particularly in risk analysis of complex systems and processes.

- Select good domain experts, train them on normative aspects
- Aggregation of opinion of multiple experts tends to give more accurate results than the opinion of a single expert. An optimum size of the panel on any single subject is 3 to 5.
- Mathematical methods of aggregation are generally preferable to behavioral methods for reaching consensus.
- Quality of judgments can be substantially improved by following a formal and structured elicitation process
- Quality of expert opinion can be substantially improved by decomposing the problem into a number of more elementary problems
- There is a significant improvement in the overall results if the initial problem definition and decomposition is done with care and in consultation with the experts.
- Experts opinions are subject to bias and overconfidence

- Effective techniques to reduce overconfidence are: (a) use of calibration techniques, and (b) encouraging experts to actively identify evidence that tends to contradict their initial opinions.
- Sources of strong dependencies among experts should be identified. Weak dependence does not seem to have a major impact on the value of multiple expert judgment.

E.5 References

Apostolakis, G., The broadening of failure rate distributions in risk analysis: how good the experts? also Martz, H. F. response to the preceding letter, *Risk Analysis*, **5** (1985), pp. 89-95.
Applied Meteorology, 1 (1978), pp. 751-8

Apostolakis, G. and Mosleh, A., Expert opinion and statistical evidence: an application to reactor core melt frequency, *Nuclear Science and Engineering*, **70** (1979), pp. 135-49.

Armstrong, J**.** S., *Long-range forecasting: From crystal ball to computer*, Wiley, New York, 1985.

Armstrong, J.Scott, William B. Denniston, and Matt M. Gordon. "The Use of the Decomposition Principle in Making Judgments." *Organizational Behavior and Human Performance* 14, no. 2 (October 1975): 257–63.

Ayyub, Bilal M., Elicitation of expert opinions for uncertainty and risk, CRC Press 2001

Ayyub, Bilal, *Methods for Expert-Opinion Elicitation of Probabilities and Consequences for Corps Facilities*, IWR Report -00-R-10 Prepared for U.S. Army Corps of Engineers Institute for Water Resources, Alexandria, VA 22315, December 2000.

Brown, T. A., *An experiment in probabilistic forecasting*, Rand Corporation, 1973.

Budescu, D.V. and Rantilla, A.K. (2000). Confidence in aggregation of expert opinions. Acta Psychologica, 104, 371-398

Clemen RT, Winkler RL. Combining probability distributions from experts in risk analysis. Risk Analysis, 1999;19(2)187-203.

Cooke, R. M., & Probst, K. N. (2006). "*Highlights of the expert judgment policy symposium and technical workshop*" Retrieved on October 19, 2014 from http://www.rff.org/Documents/Conference-Summary.pdf.

Cooke, R. M., &. Goossens L. H. J. (2008). *TU Delft expert judgment data base*. Reliability Engineering and System Safety, 93(5), 657–674. doi:10.1016/j.ress.2007.03.005.

Cooke, R. M., *Experts in Uncertainty*, Oxford University Press, 1991, New York.

Dalkey, N.C (1969). The Delphi Method: An Experimental Study of Group Opinion, prepared for the United States Air Force Project RAND. Rand Corporation, Santa Monica, California.

Dalkey, N. C. (1970). The Delphi Method: An Experimental Study of Group Opinion.
Technical Report RM-5888-PR, The Rand Corporation

DeWispelare, A. R.; Herren, L. T. & Clemen, R. T. (1995). The use of probability elicitation in the high-level nuclear waste regulation program. International Journal of Forecasting 11, 5–24.

Dalkey, N. C., *Group decision theory*, School of Engineering and Applied Science, University of California, Los Angeles, UCLA-ENG- 7749, 1977.

Feldman ES. A meta-data informed expert judgment aggregation and calibration technique [dissertation]. [College Park (MD)]: University of Maryland; 2016, 251 p.

Feldman, E. and A. Mosleh, Expert Judgment Reliability, a Meta-data Informed Comparisons, (submitted) Risk Analysis, 2020

Fischoff, B., Slovic, P. and Lichtenstein, S., Fault trees: sensitivity of estimated failure probabilities to problem representation, *Journal of Experimental Psychology: Human Perception and Performance,* **4** (1978), pp. 330-44.

Fischoff, B., Slovic, P. and Lichtenstein, S., Knowing with certainty: the appropriateness of extreme confidence, *Journal of Experimental Psychology: Human Perception and Performance,* **3** (1977), pp. 552-64.

Forrester Y. The quality of expert judgment: An interdisciplinary investigation [dissertation]. College Park: University of Maryland; 2005. 179. Available at http://hdl.handle.net/1903/3267.

French, S (1985). Group Consensus Probability Distribution: A Critical Survey. In Bayesian Statistics 2nd ed. JM Bernado et al. North Holland, Amsterdam, pp 183-201.

Genest, C.; Zidek, J (1986). Combining Probability Distribution: A Critique and an Annotated Biliography. Statistical Science, 1: 114-148.

Green, A. E. and Bourne, A. J., *Reliability technology,* Wiley, New York, 1972.

Gustafson, D. H., Shukla, R. K. and Delbecq, A., A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups, and nominal groups, *Organizational Behavior and Human Performance,* **9** (1973), pp. 280-91.

Hogarth, R. M., Cognitive processes and the assessment of subjective probability distributions, *Journal of the American Statistical Association,* **70** (1975), pp. 271-89.

Hynes, M. E. and Vanmarcke, E. **H.,** Reliability of embankment performance predictions, *Proceedings of the ASCE Engineering Mechanics Division Specialty Conference,* 1982.

Institute of Electrical and Electronics Engineers, IEEE *guide to the collection and presentation of electrical, electronic, and sensing component reliability data for nuclear-power generating stations,* IEEE Standard 500, 1984.

Koriat, A., Lichtenstein, S. and Fischoff, B., Reasons for Confidence, *Journal of Experimental Psychology: Human Learning and Memory,* **6** (1980), pp. 107-18.

Lichtenstein, S. and Fischoff, **B.,** Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance,* **3** (1977), 159-83.

Lichtenstein, S., Fischoff, B. and Phillips, L. D., Calibration of probabilities: the state of the art to 1980, in *Judgment under uncertainty: heuristics and biases* (D. Kahneman, P. Slovic, and A. Tversky, Editors), Cambridge University Press, Cambridge, England, 1982.

Martz, H. F. and Bryson, M. C., On combining data for estimating the frequency of low-probability events with applications to sodium valve failure rates, *Nuclear Science and Engineering,* **83** (1983), pp. 267-80.

Martz, H. F., Bryson, M. C., & Waller, R. A. (1985). *Eliciting and aggregating subjective judgements – some experimental results*. Los Alamos National Laboratory LU-UR—84-3193.

Mosleh, A, Bayesian Modeling of Expert-to-Expert Variability and Dependence in Estimating Rare Event Frequencies. Reliability Engineering and System Safety, 38. 1992

Mosleh, A. and Apostolakis, G., Combining various types of data in estimating failure rates, *Transactions of the 1983 Winter Meeting of the American Nuclear Society,* San Francisco, California, 1983.

Mosleh, A. and Apostolakis, G., Models for the use of expert opinions, in *Low- probability /high consequence risk analysis; issues, methods, and case studies* (R. A. Waller and V. T. Covello, Editors), Plenum, New York, 1984.

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* 20 (1988) 63-

Pickard, Lowe and Garrick, Inc., *Seabrook Station probabilistic safety assessment*, prepared for Public Service Company of New Hampshire and Yankee Atomic Electric Company, 1983.

Rowse, G. L., Gustafson, D. H. and Ludke, R. L., Comparison of rules for aggregating subjective likelihood ratios, *Organizational Behavior and Human Performance,* **12** (1974), pp. 274-85.

Seaver, D. A. (1976). *Assessment of group preferences and group uncertainty for decision making*. (Defense Technical Information Center Technical Report. Jul 75-Sep 76). Retrieved on August 14, 2018 from http://www.dtic.mil/dtic/tr/fulltext/u2/a033246.pdf.

Shirazi, C. H. (2009). *Data-informed calibration and aggregation of expert judgment in a Bayesian framework* (Doctoral Dissertation). Retrieved on August 14, 2018 from http://hdl.handle.net/1903/9883.

Slovic, P., Fischoff, B. and Lichtenstein, S., Rating the risks, *Environment,* **21**

Snaith, E. R., *The correlation between the predicted and observed reliabilities of components, equipment, and systems,* National Center of Systems Reliability, UK Atomic Energy Authority, NCSR R18, 1981.

Stael von Holstein, C.-A. S., Probabilistic forecasting: an experiment related to the stock market, *Organizational Behavior and Human Performance,* **8** (1972), pp. 139-58.

Tversky, A. and Kahneman, D., Judgment under uncertainty: heuristics and biases, *Science,* **185** (1974), pp. 1124-31.

USNRC, 1997 Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts, Prepared by Senior Seismic Hazard Assessment committee (R. J. Budnitz, G.

Apostolakis, D. M. Boore, L. S. Cluff, K. J. Coppersmith, C. A. Cornell, and P. A. Morris), NUREG/CR-6372, U.S. Nuclear Regulatory Commission, 1997.

US Nuclear Regulatory Commission, *Reactor safety study; an assessment of accident risks in US commercial nuclear power plants,* WASH-1400, NUREG, 75/014, 1975.

Wason, P. C., On the failure to eliminate hypotheses-a second look, *Thinking and reasoning* (P. C. Wason and P. N. Johnson-Laird, Editors), Penguin, Baltimore, Maryland, 1968.

Winkler RL. (1971). Probabilistic prediction: Some experimental results. Journal of the American Statistical Association, 1971;66 (336)675-685. Available at http://www.jstor.org/stable/2284212

APPENDIX F

HUMAN RELIABILITY

F.1 Introduction

Despite the significance of human error in accident causation, many risk assessments of complex systems do not include human error in the risk models. Nuclear power, aerospace, and process industry are exceptions. In fact, the majority of methods for performing what is known as human reliability analysis (HRA) have been developed for use in nuclear power plant probabilistic risk analysis (PRA)s.

The history of HRA can be traced back to 1952, when it was first applied to weapon system feasibility by Sandia National Laboratories. It began to be used in civil applications, with a focus on human-machine system design, in the 1960s. The first formal method for HRA was presented in November 1962 at the Sixth Annual Meeting of the Human Factors Society, followed by a monograph from Sandia National Laboratories outlining its quantification. This method is called Technique of Human Error Rate Prediction (THERP) (A. D. Swain & Guttmann, 1983), Throughout the 1980s the number of HRA methods significantly increased, mostly in response to the fact that the 1979 accident at the Three-Mile Island nuclear power plant was mainly attributed to operator error.

There are currently more than 40 HRA methods, although some are essentially variations of the same approach. Human reliability generally involves three major activities:

1. Human error identification, which concerns identifying erroneous human action in the contest of a given scenario and a set of human tasks;
2. Human error quantification, which aims to assess likelihood of errors occurrence; and,
3. Identification of the causes and context of the identified errors, to support development of preventive or mitigating measures to reducing the error likelihood (Kirwan, 1994).

The core concepts and overview of the state-of-the-art in human reliability analysis is briefly discussed in the following.

F.2 HRA Scope and General Requirements

HRA involves qualitative and quantitative assessments of human error (or human failure events) in the context of probabilistic risk analysis. The estimated human error probabilities are input to the PRA models such as event trees, system failure logic models (e.g., fault trees), process models, and barrier / control models. Performing HRA relies on knowledge of the design and operation of the system with which the human interacts. HRA as a modeling and analysis activity within a PRA generally requires tight interaction with other PRA modeling and analysis tasks, to ensure appropriate representation of potential human errors in the risk models.

HRA is a challenging task for two principal reasons. First, there is no consensus methodology that applies to the full range of human actions of concern in PRA. This is compounded by the fact mentioned earlier that, with few exceptions, almost all HRA methods have been developed for nuclear industry applications. In addition, conducting a detailed HRA for a risk-significant human action could be resource and time consuming. Such a specialized task could involve forming a dedicated team of experienced analysts. Therefore, the HRA task needs to strike a good balance between adequate consideration of the state-of-the-art in HRA methodology and the importance of human actions in the risk models.

F.3 HRA Steps

Despite differences among HRA methods, majority include three distinct analysis phases:

1. Modeling of the potential human errors--involving a) “task analysis” to decompose an overall sequence of human [e.g., system operator] activities into smaller units [subtasks] suitable for analysis, and b) identification of possible human errors [often referred to as human failure event,

or HFE] in conducting the task. There is no universally agreed upon approach or standard for the best level of task decomposition.

2. Identification of the potential contributors to human error--selecting relevant performance shaping factors.
3. Quantification of human error probabilities --the process of using available evidence to estimate human error probability (HEP) for the identified HFEs.

These phases are briefly described in the following.

F.3.1 Modeling of Potential Human Errors

Modeling the potential human errors requires a good understanding of possible human-system interactions. At a high level, the analyst should identify:

- What information the operator receives from the system, in terms of content and timing;
- What actions are the operators required to perform on the system, and through which human-system interface;
- What non-required actions can the operator perform on the system;
- Which and how components, functions, systems could be affected by the operator actions.

Task analysis (TA) is a useful tool for identifying potential human errors. TA was developed by two industrial psychologists, John Annett and Keith Duncan, in the 1960s (Shepherd, 2001). TA has experienced continuous improvement since then to capture factors related to the psychological context of tasks (involving both physical and cognitive activities). Hierarchical task analysis (HTA) is one of the most popular forms of TA.

The key feature of HTA is defining tasks through *goals* rather than actions, and the possibility of analyzing complex tasks through decomposition of the goals into sub-goals. The goals and sub-goals are organized in HTA through *plans*. Plans indicate how subordinate operations are organized to meet their common goal. HTA offers different plans, such as fixed sequence (a specified second operation is carried when a first goal has been successfully attained) and choices (a decision-making activity is first undertaken). A detailed description of the plan concept can be found in (Shepherd, 2001).

The general method of conducting an HTA is:

1. Deciding whether the goal warrants examining; if it does then;
2. Examining the operation by considering how the operator and the system interact, and identifying the goal;
3. Re-describing the goal in terms of its subordinate goals and their plan.

Because it is generic in nature, HTA may be used to analyze any type of task in any domain. In fact, HTA has been used in many applications including personnel selection and staffing, training, workplace layout, equipment design, work organization, and procedure design.

Once TA has been completed, a human error identification process then determines what can go wrong. TA thus allows the analyst to identify which tasks the operator must accomplish for a successful operation, and the human error identification step then analyzes how the operators can fail in completing these tasks. Human error can be generally categorized as *errors of omission* (EOOs)-when an operator fails in performing an expected or prescribed action, and *errors of commission* (EOCs)-when the operator’s action is not prescribed or anticipated and leads to failure or undesired state of the system. Other characterizations of types of error have been introduced by various HRA approaches.

F.3.2 Identification of Potential Contributors to Human Error

In almost all HRA methods the so-called performance influencing factors (PIFs) are used to represent the situational contexts and causes that affect human performance in interacting with the system. PIFs are also named performance shaping factors (PSF)s, error forcing contexts, common performance conditions, or error producing conditions, depending on the HRA method. As with task decomposition, there is no standard list of performance shaping factors, and there is considerable variability between HRA methods. THERP, for instance, categorizes PIFs as *external*, including work environment (e.g., equipment design, written procedures, or oral instructions), and *internal*, including individual characteristics of operators (e.g., skills, motivations, and experience), and psychological and physiological stress, Phoenix Method (Ekanem and Mosleh) makes use of 9 main groups of PIFs: Stress, (covering emotional factors), Knowledge / Abilities and Bias (covering cognitive response), Team Effectiveness, Human System Interface, Task Load, Time Constraint, Resources, and Procedures (covering the physical world). PIFs can also reflect organizational matters such as *safety culture.*

Considering PIFs when analyzing human error allows identifying why operators may fail, rather than simply which errors they can commit. Identification and assessment of PIFs is also an acknowledgement that the reasons for the human error include external factors that may be beyond the abilities of the operator to control, or even know about.

F.3.3 Quantitation of Human Error Probabilities

Different HRA methods feature different approaches to quantification of human error probability (HEP). The quantification determines the likelihood that the particular action identified in the qualitative analysis steps will fail. In almost all methods PIFs are often used to derive or modify HEP values. In some methods PIFs are used as multipliers of a "base HEP." Others use Bayesian Belief Networks to quantify HEPs through probabilistic metrics of the degree of the influence of PIFs on HEPs. Some models allow for including any desired PSF defined by the analyst and calculates. HRA methods also differ with respect to the sources for the values used in HEP calculations.

F.4 HRA BEST PRACTICES

A good reference for HRA best practices is the US Nuclear Regulatory Commission (USNRC) report "Good Practices for Implementing Human Reliability Analysis (HRA)" - NUREG-1792 (Kolaczkowski et al., 2005). It contains discussions and conclusions on general features of HRA methods, and identifies 38 Good Practices, related to various the phases of the HRA (analysis activities). Some of these Good Practices are summarized below:

- The HRA assessment should involve a multi-disciplinary team and should include field observations, review of system documents and talk-troughs with the operators;
- HRA should address human errors both in diagnosis (situation assessment) and response (action execution);
- HRA should account for dependencies among the Human Error Probabilities in any given accident sequence;
- Possibility of recovery from errors needs to be considered;
- HRA documentations make the results traceable and reproducible.

An evaluation of the strengths and limitations of various HRA methods with respect to Good Practices report is provided in "Evaluation of Human Reliability Analysis Methods against Good Practices" - (USNRC, 2006). Review of various is also provided in a report by the Health and Safety Executive - UK (Bell & Holroyd, 2009).

As an example of HRA methods, a summary of the main features of the THERP method is provided here.

F.5 OVERVIEW OF THERP

THERP is known as the first formal HRA method. Its handbook was prepared by Swain and Guttman in 1983 (1983) for the USNRC. The aim of THERP is to calculate the probability of successful performance of activities needed for the execution of a task. The results are presented graphically in an HRA event tree, which is a formal representation of required action sequence, as can be seen in Figure F-1.

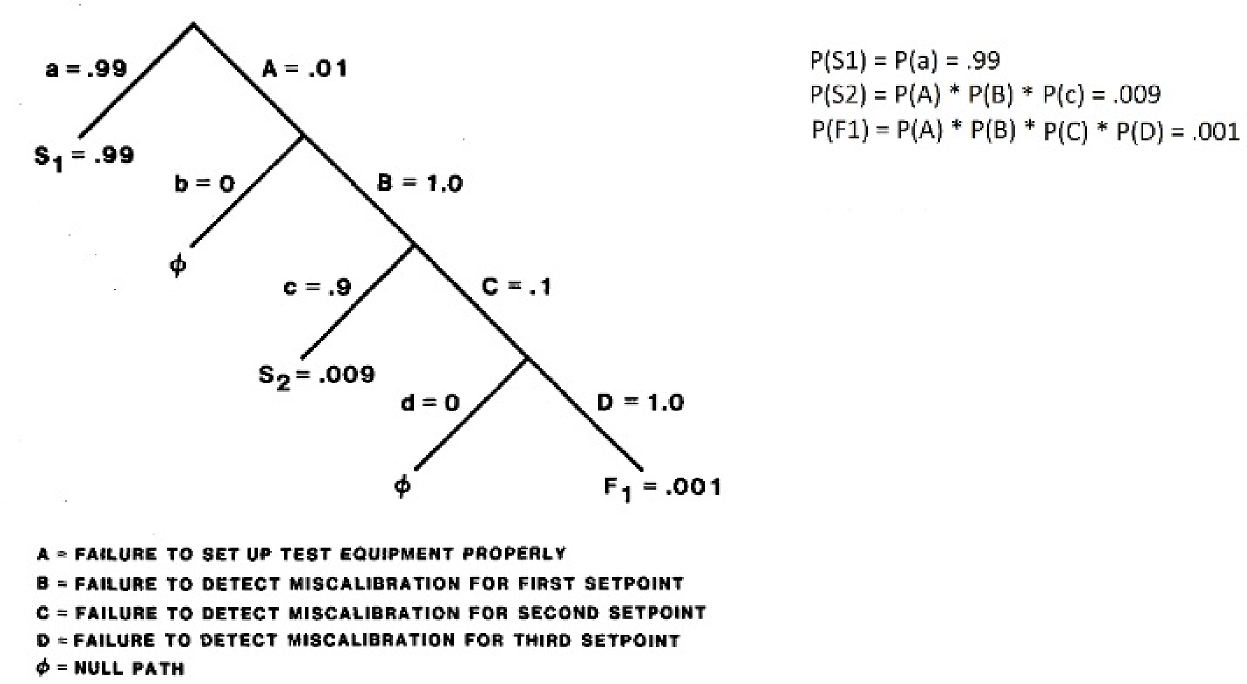


Figure F-1: Example of THERP binary event tree

THERP relies on a human reliability database containing HEPs, drawn from some operational data and expert judgments. The nominal probability estimates from the analysis of the HRA event tree are modified for the effects of sequence-specific PSFs.

THERP provides tables for the following types of errors:

- Screening and detection of system abnormalities
- Diagnosis and identification of the causes of system abnormalities
- Omitted actions, including actions in procedure preparation, use of a specified procedure (i.e., administrative control), execution of a procedure step, and providing an oral instruction
- Writing down incorrect information
- Acting on a wrong object includes reading from an unintended display, acting at an unintended control, and unintended control (e.g., turn a control at wrong direction).

THERP models errors of omission (EOO)s and errors of commission (EOCs). In THERP EOOs relate to omission of a task or a step in a task; whereas EOC refers to a selection error (selecting wrong control, malposition control, issuing wrong command), errors of sequence, time errors (too early or too late), and qualitative errors (too little or too much). Regarding cognitive error modeling, however, THERP uses available time to determine the probabilities of diagnosis failure and does not offer further breakdown in terms of specific cognitive or decision errors. A lack of detailed modeling of cognitive errors is common to first generation HRA methods, thus the primary motivation for the second-generation HRA methods.

Figure F-2 presents THERP correlation between time passed and probability of failure of diagnosis of the event by the operating crew. The event starts with a sign indication of an abnormal event, and the more time the crew would have to diagnose it, the lower the probability of failure in diagnosing.

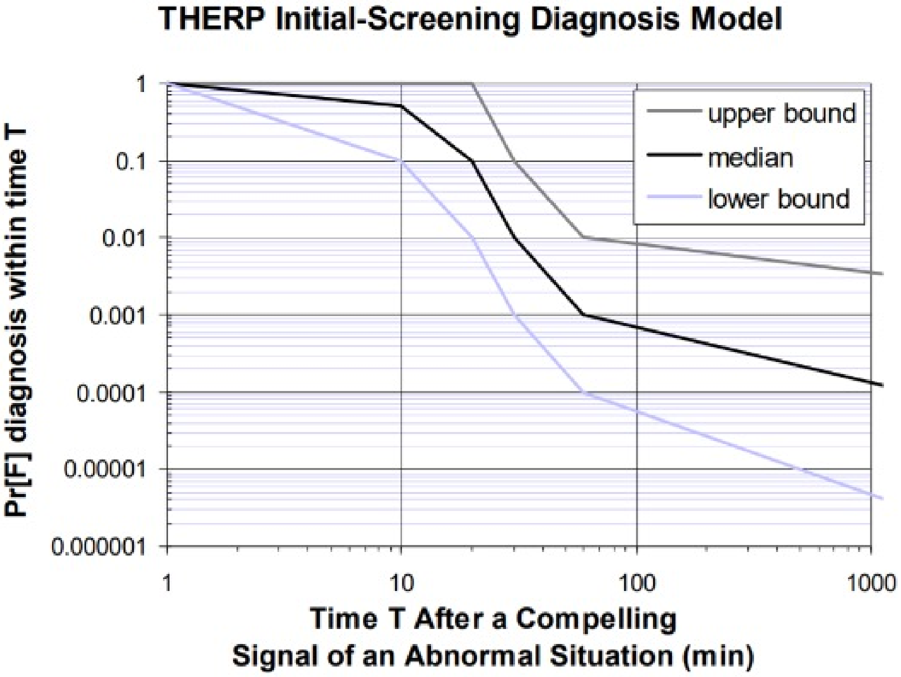


Figure F-2: Initial screening model of estimated HEPs for diagnosis within time T of one abnormal event

THERP divides the PSF into three categories:

1. External PSFs: Factors related to work situations of the operators, technicians, maintenance personnel, engineers, and others who keep the nuclear power plant performing reliably and safely. They can be subdivided into:
    a. Situational characteristics: include PSFs that are often plant-wide in influence or that cover many different jobs and tasks in the plant;
    b. Task and equipment characteristics: include PSFs connected with the instructional materials used in jobs;
    c. Job and task instructions: although they are task characteristics, they are singled out because THERP considers them much more important than is generally believed.
2. Internal PSFs: Factors associated to the operators, such as personality, attitudes, knowledge, stress and physical conditions.
3. Stressors: Factors deriving from work environment in which the demands placed on the operator by the system do not conform to his / her capabilities and limitations. THERP considers stress to be the human response to a stressor, they are subdivided into:
    a. Psychological stressors and
    b. Physiological stressors.

Although THERP provides the list of these PIFs, it does not offer specific rules for causal identification. In addition, it does not provide a specific procedure for identification of error mode.

The quantitative procedure of THERP is based on tables: the handbook presents tabled entries of HEPs that can be modified by the effects of plant specific PSFs, using other tables only three of the PSFs are used in HEP calculation. These are: Tagging Levels (of components or controls), Experience, and Stress.

The application of THERP consists of seven steps:

1. Definition of the system failures of interest. These failures include functions of the system in which human error has a greater likelihood to influence probability of a fault, and those which are of interest to the analyst;
2. Identification and analysis of human operations and their relationship to system tasks and function of interest. For this step, it is necessary to conduct a comprehensive task and human error analysis. For each step of the task, the analyst considers the possible errors, which will be included in the HRA tree. The tree thus shows a number of different paths each of which has an associated end state or consequence (Figure 1)
3. Estimation of human error probabilities (detailed below);
4. Estimation of the effects of human error on the system failure events;
5. Recommendation of changes to the system and recalculation of the system failure probabilities;
6. In this step the analyst can use sensitivity analysis to identify how certain risks may be improved in the reduction of HEPs, and incorporated recovery into the event tree;
7. Review of proposed changes with respect to availability, reliability and cost-benefit.

The Step 3 above - the HEP calculation, consists of:

1. For each branching point of the HRA Tree, the analyst uses the HEP search scheme provided in the handbook to identify the likely human errors and the corresponding nominal HEPs as well as the uncertainty bounds.
2. The analyst identifies the factors / interactions affecting human performance, and assesses the effect of the tagging levels, experience, and stress on the HEPs as well as the uncertainty bounds of the HEPs.
3. The analyst then assesses the levels of task dependencies based on the five-level dependency scale specified by THERP.
4. The analyst assesses the possible recovery branches in the HRA ET and assess the success probabilities (Step 5 of the procedure detailed above).
5. The analyst then determines the success and failure states within the HRA Tree and calculates the HEP of the HRA Tree. The calculated HEP is used in the PRA model (See Figure F.3)

With respect to cognitive error modeling, which is the primary interest of the second generation HRA methods, THERP uses available time to determine the probabilities of diagnosis failure. No further breakdown in terms of specific cognitive or decision errors is offered.

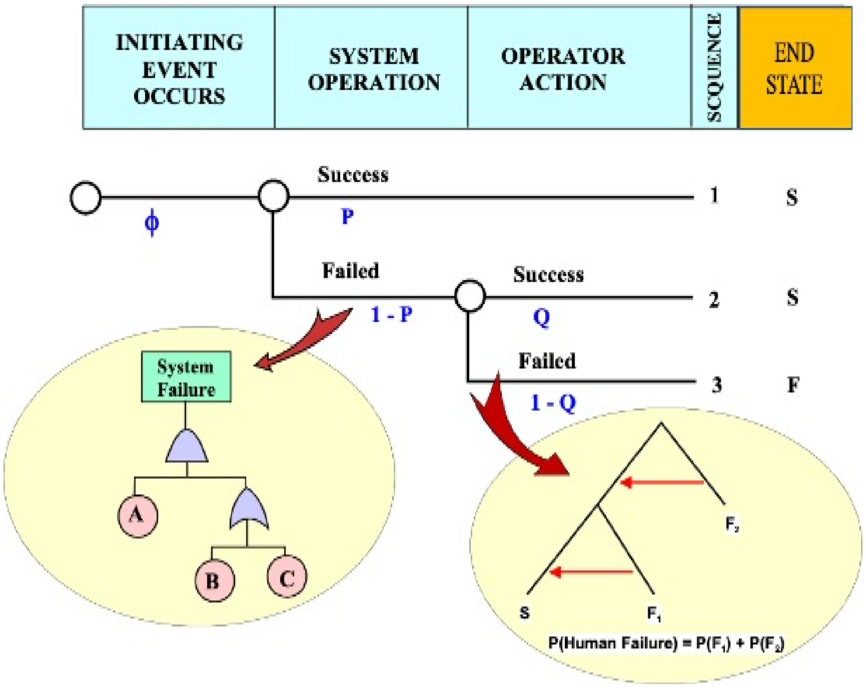


Figure F-3 Inclusion of HRA (THERP) Results in Risk Model ETs and FTs

REFERENCES

Ekanem NJ, Mosleh A. Phoenix - A Model-Based Human Reliability Analysis Methodology: Qualitative Analysis Overview. Reliab Eng Syst Saf 2016;145.

Swain AD, Guttmann HE. Handbook of human reliability analysis with emphasis on nuclear power plant applications, NUREG/CR-1278. vol. null. 1983.

Kirwan B. A Guide to Practical Human Reliability Assessment. CRC Press; 1994.

Kolaczkowski A, Forester J, Lois E, Cooper S. Good practices for implementing human reliability analysis (HRA) (NUREG-1792). Sandia Natl Lab ,US Nucl Regul Comm 2005:110.

USNRC. NUREG 1842 - Evaluation of Human Reliability Analysis Methods Against Good Practices. 2006.

Bell J, Holroyd J. Review of human reliability assessment methods. Heal Saf Lab 2009:78.

Shepherd A. Hierarchical Task Analysis. London: Taylor & Francis; 2001.

Kirwan B. A guide to practical human reliability assessment. London: CRC Press; 1994.