

Technical Basis for the use of Probabilistic Fracture Mechanics in Regulatory Applications

Draft Report for Comment

AVAILABILITY OF REFERENCE MATERIALS IN NRC PUBLICATIONS

NRC Reference Material

As of November 1999, you may electronically access NUREG-series publications and other NRC records at the NRC's Library at www.nrc.gov/reading-rm.html. Publicly released records include, to name a few, NUREG-series publications; *Federal Register* notices; applicant, licensee, and vendor documents and correspondence; NRC correspondence and internal memoranda; bulletins and information notices; inspection and investigative reports; licensee event reports; and Commission papers and their attachments.

NRC publications in the NUREG series, NRC regulations, and Title 10, "Energy," in the *Code of Federal Regulations* may also be purchased from one of these two sources:

1. The Superintendent of Documents

U.S. Government Publishing Office
Washington, DC 20402-0001
Internet: www.bookstore.gpo.gov
Telephone: (202) 512-1800
Fax: (202) 512-2104

2. The National Technical Information Service

5301 Shawnee Road
Alexandria, VA 22312-0002
Internet: www.ntis.gov
1-800-553-6847 or, locally, (703) 605-6000

A single copy of each NRC draft report for comment is available free, to the extent of supply, upon written request as follows:

Address: **U.S. Nuclear Regulatory Commission**
Office of Administration
Digital Communications and Administrative
Services Branch
Washington, DC 20555-0001
E-mail: Reproduction.Resource@nrc.gov
Facsimile: (301) 415-2289

Some publications in the NUREG series that are posted at the NRC's Web site address www.nrc.gov/reading-rm/doc-collections/nuregs are updated periodically and may differ from the last printed version. Although references to material found on a Web site bear the date the material was accessed, the material available on the date cited may subsequently be removed from the site.

Non-NRC Reference Material

Documents available from public and special technical libraries include all open literature items, such as books, journal articles, transactions, *Federal Register* notices, Federal and State legislation, and congressional reports. Such documents as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings may be purchased from their sponsoring organization.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at—

The NRC Technical Library

Two White Flint North
11545 Rockville Pike
Rockville, MD 20852-2738

These standards are available in the library for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from—

American National Standards Institute

11 West 42nd Street
New York, NY 10036-8002
Internet: www.ansi.org
(212) 642-4900

Legally binding regulatory requirements are stated only in laws; NRC regulations; licenses, including technical specifications; or orders, not in NUREG-series publications. The views expressed in contractor prepared publications in this series are not necessarily those of the NRC.

The NUREG series comprises (1) technical and administrative reports and books prepared by the staff (NUREG-XXXX) or agency contractors (NUREG/CR-XXXX), (2) proceedings of conferences (NUREG/CP-XXXX), (3) reports resulting from international agreements (NUREG/IA-XXXX), (4) brochures (NUREG/BR-XXXX), and (5) compilations of legal decisions and orders of the Commission and the Atomic and Safety Licensing Boards and of Directors' decisions under Section 2.206 of the NRC's regulations (NUREG-0750).

DISCLAIMER: This report was prepared as an account of work sponsored by an agency of the U.S. Government. Neither the U.S. Government nor any agency thereof, nor any employee, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product, or process disclosed in this publication, or represents that its use by such third party would not infringe privately owned rights.

Technical Basis for the use of Probabilistic Fracture Mechanics in Regulatory Applications

Draft Report for Comment

Manuscript Completed: January 2021

Date Published: September 2021

Prepared by:

L. Hund¹, J. Lewis¹, N. Martin¹, M. Starr¹,
D. Brooks¹, A. Zhang¹, R. Dingreville¹,
A. Eckert¹, J. Mullins¹, P. Raynaud²,
D. Rudland², D. Dijamco², and S. Cumblidge²

¹Sandia National Laboratories
P.O. Box 8500
Albuquerque, NM 87185

²U.S. Nuclear Regulatory Commission

P. Raynaud, NRC Project Manager

Office of Nuclear Regulatory Research

COMMENTS ON DRAFT REPORT

Any interested party may submit comments on this report for consideration by the NRC staff. Comments may be accompanied by additional relevant information or supporting data. Please specify the report number **NUREG/CR-7278** in your comments, and send them by the end of the comment period specified in the *Federal Register* notice announcing the availability of this report.

Addresses: You may submit comments by any one of the following methods. However, the NRC encourages electronic comment submission through the Federal Rulemaking Website. Please include Docket ID **NRC-2021-0087** in the subject line of your comments. Comments submitted in writing or in electronic form will be posted on the NRC website and on the Federal rulemaking website <http://www.regulations.gov>.

Federal Rulemaking Website: Go to <http://www.regulations.gov> and search for documents filed under Docket ID **NRC-2021-0087**.

Mail comments to: Office of Administration, Mail Stop: TWFN-7-A60M, U.S. Nuclear Regulatory Commission, Washington, DC 20555-0001, ATTN: Division of Resource Management and Analysis Staff.

For any questions about the material in this report, please contact: Patrick Raynaud, Senior Materials Engineer, (301) 415-1987 or by e-mail at Patrick.Raynaud@nrc.gov

Please be aware that any comments that you submit to the NRC will be considered a public record and entered into the Agencywide Documents Access and Management System (ADAMS). Do not provide information you would not want to be publicly available.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

ABSTRACT

This NUREG, on probabilistic fracture mechanics (PFM), is a companion document to Draft Regulatory Guide 1382 (DG-1382) / Regulatory Guide 1.245 (RG-1.245), "Preparing Probabilistic Fracture Mechanics (PFM) Submittals." This NUREG describes a graded approach to developing PFM submittal documentation, provides a generalized technical basis for conducting PFM analyses, and constitutes the technical basis for DG-1382 / RG-1.245.

The graded approach that is outlined below represents a balance between the benefits of clear, consistent, and comprehensive submittals and the need to maintain flexibility for PFM analyses that, by their nature, include many situation-specific aspects. The resulting guidance, provided in DG-1382/RG1.245, outlines a procedure that includes this suggested graded approach for PFM analyses and submittals. The unique characteristics of the underlying regulatory application dictate the breadth and depth of content included in the submission.

This document also describes a hypothetical process for conducting a PFM analysis. This process is aligned with the position on documentation elements given previously in the U.S. Nuclear Regulatory Commission's (NRC's) technical letter report, "Important Aspects of Probabilistic Fracture Mechanics Analyses," issued in 2018. The NUREG provides fundamental background for the concepts and methods introduced in the analysis process. The examples give details for analysts on (nonprescriptive) approaches for PFM analyses.

TABLE OF CONTENTS

1		
2	ABSTRACT	iii
3	LIST OF FIGURES	ix
4	LIST OF TABLES	xi
5	EXECUTIVE SUMMARY	xiii
6	ABBREVIATIONS AND ACRONYMS	xv
7	1 INTRODUCTION	1
8	1.1 Fracture Mechanics Approach to Structural Integrity Analysis	1
9	1.2 Historical Perspective on Probabilistic Fracture Mechanics Analysis of Nuclear	
10	Structures	2
11	1.3 Objective.....	4
12	1.4 Structure of This Document.....	4
13	1.5 References	5
14	2 PROBABILISTIC FRACTURE MECHANICS GRADED APPROACH	7
15	2.1 Background	7
16	2.2 Supporting Information for RG-1.245	8
17	2.2.1 Quantities of Interest and Acceptance Criteria	8
18	2.2.2 Software Quality Assurance and Verification and Validation	8
19	2.2.3 Models.....	9
20	2.3 References	10
21	3 ANALYTICAL STEPS IN A PROBABILISTIC FRACTURE MECHANICS	
22	ANALYSIS	11
23	3.1 Step 1: Translation of Regulatory Requirements into an Analysis Plan	13
24	3.1.1 Step 1: Action 1—Define the Regulatory Context	13
25	3.1.2 Step 1: Action 2—Define the Quantity of Interest and How it Relates to	
26	the Model Output and Acceptance Criteria	13
27	3.1.3 Step 1: Action 3—Determine the Suitability of the Probabilistic Fracture	
28	Mechanics Code for the Specific Application	14
29	3.1.4 Step 1: Action 4—Identify Key Elements of the Problem that Impact	
30	Analysis Choices	16
31	3.2 Step 2: Model Input Uncertainty Characterization	17
32	3.2.1 Step 2: Action 1—Identify Uncertain Model Inputs	18
33	3.2.2 Step 2: Action 2—Specify Probability Distributions on Uncertain Inputs	19
34	3.3 Step 3: Estimation of Quantity of Interest and Associated Uncertainty	20

1	3.3.1	Step 3: Action 1—Select a Sampling Scheme for Sampling Uncertain	
2		Model Inputs.....	21
3	3.3.2	Step 3: Action 2—Assess Sampling Uncertainty: Statistical	
4		Convergence Analysis.....	22
5	3.3.3	Step 3: Action 3—Conduct Sensitivity Analyses to Determine Input	
6		Uncertainty Importance	24
7	3.3.4	Step 3: Action 4—Conduct Output Uncertainty Analysis.....	26
8	3.4	Step 4: Sensitivity Studies to Assess the Credibility of Modeling Assumptions	27
9	3.4.1	Step 4: Action 1—Determine a Set of Sensitivity Studies	28
10	3.4.2	Step 4: Action 2—Conduct Sensitivity Studies and Present Results.....	28
11	3.5	Step 5: Draw Conclusions from Analysis Results.....	29
12	3.5.1	Step 5: Action 1—Interpret Analysis Results.....	29
13	3.5.2	Step 5: Action 2—Iterate on the Analysis Process to Refine Model	
14		Results	29
15	3.6	References	30
16	4	USEFUL METHODS FOR ESTABLISHING CONFIDENCE IN	
17		PROBABILISTIC FRACTURE MECHANICS ANALYSIS.....	33
18	4.1	Useful Methods for Translating Regulatory Requirements into an Analysis Plan	33
19	4.1.1	Separation of Aleatory and Epistemic Uncertainty	33
20	4.2	Methods for Model Input Uncertainty Characterization	39
21	4.2.1	Statistical Distribution Fitting	39
22	4.2.2	Preserving Physical Relationships between Inputs.....	44
23	4.3	Useful Methods for Forward Propagation of Input Uncertainty.....	48
24	4.3.1	Simple Random Sampling.....	48
25	4.3.2	Latin Hypercube Sampling	51
26	4.3.3	Importance Sampling	54
27	4.3.4	First- and Second-Order Reliability Methods	57
28	4.3.5	Convergence Analysis.....	60
29	4.3.6	Closed-Form Metric for Simple Random Sampling Uncertainty in a	
30		Probability Estimate.....	66
31	4.3.7	Statistical Bootstrapping.....	68
32	4.3.8	Global Sensitivity Analysis	72
33	4.3.9	Local Sensitivity Analysis	76
34	4.3.10	Surrogate Models	77
35	4.3.11	Visualizing Output Uncertainty Due to Input Uncertainty	82
36	4.4	Useful Methods for Sensitivity Studies	87

1	4.4.1 Sensitivity Studies	87
2	4.5 References	91
3	5 SUMMARY AND CONCLUSIONS	99
4	GLOSSARY	101
5		

LIST OF FIGURES

1		
2	Figure 3-1	Flowchart Describing the Steps and Actions of a PFM Analysis 12
3	Figure 4-1	The Estimated Probability of Pipe Rupture for the First 50 Epistemic
4		Samples 37
5	Figure 4-2	Histogram of Estimated Probabilities Across 1,000 Epistemic Realizations 38
6	Figure 4-3	Distributions with Nonnegative Input Parameters..... 41
7	Figure 4-4	Heavy-Tailed and Skewed Distributions..... 41
8	Figure 4-5	Sampling Uncertainty in Input Distribution Fits..... 42
9	Figure 4-6	Graphical Diagnostics for Parametric Model Fit 43
10	Figure 4-7	Randomly Sampled Inputs (Left) and Transformed Inputs (Right) 46
11	Figure 4-8	Sampled Inputs from the Joint Distribution of X and Y 47
12	Figure 4-9	Constraining Input Distributions to Ensure that Yield Strength Is Less than
13		Ultimate Tensile Strength (Open Circles Are Not Admissible) 48
14	Figure 4-10	SRS Sample in the Quantile Space for Two Input Variables ($n=10$) 50
15	Figure 4-11	SRS Sample Transformed into the Input Space for Two Input Variables
16		($n=10$) 50
17	Figure 4-12	Example of an LHS in the Quantile Space for Two Input Variables ($n=10$)..... 52
18	Figure 4-13	LHS Transformed into the Input Space for Two Input Variables ($n=10$)..... 53
19	Figure 4-14	Example of Estimating a Probability Using Random Sampling and
20		Importance Sampling 56
21	Figure 4-15	Example of the FORM and SORM Methods in the Standard Normal Space
22		(Following Reference 4-42) 60
23	Figure 4-16	Histogram of Probability Estimates from a Simple Random Sample..... 62
24	Figure 4-17	Confidence Interval used to Assess the Convergence of a Probability
25		Estimate 63
26	Figure 4-18	Estimates of the Probability of Axial Crack for $r=5$ Independent Replications
27		Using the Same Sampling Scheme..... 65
28	Figure 4-19	Prediction Interval Computed from the Five Independent Simulations..... 66
29	Figure 4-20	Visualization of the Steps Taken for the Standard Statistical Bootstrap..... 70
30	Figure 4-21	Bootstrap Sampling Distribution along with a 95-Percent Confidence
31		Interval for the Complex Estimator Example 72
32	Figure 4-22	Scatterplots Showing an Input Without (Left) and with (Right) a Significant
33		Relationship with the Output Variable 74
34	Figure 4-23	Scatterplots Showing Linear (Left), Nonlinear/Monotonic (Middle), and
35		Nonlinear/Nonmonotonic (Right) Relationships between the Input and
36		Output Variables..... 74
37	Figure 4-24	GP Surrogate Fit to Training Data (Black Points) from the True but
38		Unknown Function $f(x)$ 79
39	Figure 4-25	Example of a Generalized Linear Model for Binary Data—Component
40		Failure as a Function of Age 82
41	Figure 4-26	Continuous Output at a Single Time Point (Left) and Over Time (Right) 83
42	Figure 4-27	Failure Probability Over Time when Aleatory and Epistemic Uncertainty are
43		not Separated; Linear Scale (Left) and Log Scale (Right)..... 84
44	Figure 4-28	Frequency over Aleatory Samples at a Single Time Point (Left) and as a
45		Function of Time (Right)..... 85
46	Figure 4-29	Importance-Sampled Distribution (Blue), Simple Random Sample (Black)
47		Distribution, and Reconstructed Unweighted Distribution (Green) for a
48		Probability of Failure..... 86
49		

LIST OF TABLES

1
2
3 Table 2-1 Content Mapping between RG1.245, NUREG Section 3, and EPRI White
4 Paper (Reference 2-2).....8
5
6

EXECUTIVE SUMMARY

This technical basis NUREG document, and the regulatory guide it is associated with, were developed from the concepts for using PFM in support of regulatory applications outlined in the technical letter report, "Important Aspects of Probabilistic Fracture Mechanics Analyses," issued in 2018. In conjunction with the release of the technical letter report, the U.S. Nuclear Regulatory Commission held a series of public meetings to present a general framework of the expected content of a probabilistic fracture mechanics (PFM) analysis. This NUREG further develops the concept of a PFM analysis methodology and outlines important considerations for a high-quality and high-confidence PFM analysis.

This NUREG contains three technical sections: Section 2 presents technical details for the contents of a PFM submittal following a graded approach, Section 3 presents the analytical steps in a PFM analysis, and Section 4 presents methods that can be used in PFM analysis. These three sections are linked together through the development structure, but the technical information provided in each section is geared toward different audiences. Section 2 is intended for applicants of all experience levels. Section 3 is intended to be used by applicants who are familiar with PFM submittals but are seeking additional information regarding the development of an analysis structure or formalism. Section 4 is intended to be used by applicants who are seeking to further understand the theoretical underpinnings of the processes that are used to establish the credibility of a PFM analysis.

The technical background provided for PFM submittal documentation represents a balance between the efficiencies gained by clear, consistent, and comprehensive submittals and the need to maintain flexibility for PFM analyses that by their nature will include many situation-specific aspects. The resulting guidance is provided in DG-1382 / RG 1.245, and outlines a procedure in which a suggested minimum set of documented evidence may be augmented by additional details. As explained in DG-1382 / RG 1.245, the unique characteristics of the underlying regulatory application dictate the breadth and depth of content included in the submission. Thus, the documentation elements that would be expected in an application are explicitly linked to the analysis framework that is described herein.

This NUREG presents a general framework for describing, performing, and evaluating PFM that will produce a high-quality and high-confidence PFM analysis. The important pieces of a PFM analysis that should be considered include models, inputs, uncertainty characterization, probabilistic framework, and PFM outputs:

- Models can be categorized into different types, but in all cases, model verification, validation, and uncertainty quantification are key steps to gain confidence in the adequacy of the models used.
- Treatment of random inputs may consist of constructing probability distributions; determining input bounds if applicable; and quantifying any assumptions, conservatisms, or dependencies among inputs.
- Uncertainty characterization and treatment are at the core of a PFM analysis. In many PFM analyses, separation of epistemic and aleatory uncertainty may be useful. Uncertainty identification, quantification, and propagation are essential elements in describing a PFM methodology or analysis. The proper choice of sampling techniques is also an important step that needs justification. The report discusses concepts and methods to verify and validate a probabilistic framework.

- 1 • Ways to demonstrate PFM convergence include varying sample size and sampling
2 strategy, as well as performing stability analysis. Output uncertainty analysis can take
3 various forms depending on the problem being analyzed. Sensitivity analyses can help
4 to identify the drivers of uncertainty for a given problem or output. Sensitivity studies are
5 useful to understand which parameters drive the issue being investigated, and to show
6 that some expected trends are indeed reflected in the analysis results. The report
7 presents methods to perform such studies.

1

ABBREVIATIONS AND ACRONYMS

2	AIC	Akaike Information Criterion
3	ANS	American Nuclear Society
4	ASME	American Society of Mechanical Engineers
5	BIC	Bayesian Information Criterion CDF cumulative distribution function
6	CFR	<i>Code of Federal Regulations</i>
7	CV	coefficient of variation
8	EPRI	Electric Power Research Institute
9	FAVOR	Fracture Analysis of Vessels—Oak Ridge
10	FORM	first-order reliability method
11	GP	Gaussian process
12	IAEA	International Atomic Energy Agency
13	LHS	Latin hypercube sample/sampling
14	MARS	multivariate adaptive regression splines
15	ML	machine learning
16	MPP	most probable point
17	MRP	Materials Reliability Program
18	NRC	U.S. Nuclear Regulatory Commission
19	NUREG	NRC technical report designation
20	PDF	probability density function
21	PFM	probabilistic fracture mechanics
22	PRA	probabilistic risk assessment
23	QA	quality assurance
24	QoI	quantity of interest
25	SA	sensitivity analysis
26	SORM	second-order reliability method
27	SQA	software quality assurance
28	SRS	simple random sampling
29	V&V	verification and validation
30	xLPR	extremely low probability of rupture
31		

1 INTRODUCTION

2 The purpose of this NUREG is to provide a generalized technical basis for conducting
3 probabilistic fracture mechanics (PFM) analyses and to describe a graded approach for
4 developing submittal documentation. PFM is a subset of fracture mechanics that complements
5 deterministic fracture analysis. Specifically, PFM is based on a deterministic fracture mechanics
6 framework that quantifies crack propagation or damage accumulation while accounting for
7 uncertainty in aspects such as the physical models, physical parameters, geometry, loading,
8 deformation mechanisms, and environmental exposure. Analysis of a PFM framework allows for
9 assessments of the structural integrity of components to enable risk-informed decisions in a
10 regulatory application. PFM allows the direct representation of uncertainties using best estimate
11 models and distributed inputs.

12 1.1 Fracture Mechanics Approach to Structural Integrity Analysis

13 Any fracture mechanics approach (deterministic or probabilistic) to structural integrity analysis
14 quantifies the combination of at least three key elements: (1) the applied stress produced by
15 structural loading, (2) the flaw size, and (3) the fracture toughness. The stress and flaw size
16 provide the driving force for fracture, while the fracture toughness provides a measure of the
17 material's resistance to crack propagation and failure. Techniques for computing fracture driving
18 force range from simple to complex, and the most appropriate methodology depends on the
19 geometry, loading, and materials properties. Flaw size may be determined by nondestructive
20 evaluation of an indication found to exist in the structure. It may represent the size of a flaw that
21 nondestructive evaluation could miss, or it may represent a nominal flaw size agreed to as
22 appropriate for certain types of assessments. The driving force and fracture toughness are
23 compared to assess the likelihood of failure. Environment and time generally complete the list of
24 other elements included in most fracture mechanics analyses. All these variables may or may
25 not evolve with time and spatial location within a component or structure. Fracture mechanics
26 provides mathematical relationships among these quantities.

27
28 There are two general options for performing a fracture analysis (although they can be
29 equivalent in certain circumstances)—the energy balance approach and the stress-intensity
30 factor approach:

- 31 • In the energy balance approach, a fracture mechanics-based failure criterion is
32 considered when the strain energy release rate associated with crack advance matches
33 or exceeds the energy needed to create new crack surfaces, to account for plastic flow,
34 and to account for other types of energy dissipation associated with the degradation
35 mechanisms considered. In this interpretation of fracture mechanics, the crack will grow
36 when the critical energy release rate is exceeded.
- 37 • In the stress-intensity factor approach, a fracture mechanics-based failure criterion
38 considers that the material fails locally at some critical combination of stress and strain
39 for given crack-tip conditions. In the case of a linear elastic body, the classic
40 stress-intensity factor is used; in the case of a nonlinear body (or equivalently an
41 elastic-plastic body under monotonic loading), the J-integral is used.

42 For both the energy balance and stress-intensity factor approaches, the applied load is typically
43 determined either through a finite-element analysis of the actual structure or by a closed-form
44 analysis of a simplified representation of the structure. In the case of linear elastic fracture
45 mechanics, one considers materials under quasistatic conditions, while elasto-plastic fracture

1 mechanics involve consideration of plastic deformation under quasistatic conditions. Dynamic,
2 viscoelastic, and viscoplastic fracture mechanics include time as a variable.

3 **1.2 Historical Perspective on Probabilistic Fracture Mechanics Analysis of** 4 **Nuclear Structures**

5 Historically, most assessments of structural integrity have been performed deterministically; for
6 example, a single value of fracture toughness is used to estimate the failure stress or critical
7 flaw size. This is true for many U.S. Nuclear Regulatory Commission (NRC) regulations. In the
8 past, the NRC has typically regulated the use of nuclear reactor structural materials on a
9 deterministic basis. Consensus codes and standards used for the design and analysis of such
10 structures, such as the American Society of Mechanical Engineers (ASME) Boiler and Pressure
11 Vessel Code, typically rely on conservative fracture models with applied safety factors and
12 conservative bounding inputs to account for the numerous uncertainties that may be present.
13 Improving the reliability of such models by quantifying the impacts of the assumptions and
14 uncertainties becomes difficult because of the conservative nature of the models and inputs and
15 the lack of historical documentation of the basis for safety factors.

16
17 Observations of the character of the three key fracture mechanics elements show that (1) loads
18 exerted on a structure may include random noise, (2) structures contain many flaws with various
19 sizes, orientations, and locations, and (3) fracture toughness data in the ductile-brittle transition
20 region are widely scattered. As such, the reliance on a deterministic basis for engineering
21 designs and regulations has given way to increased use of probabilistic techniques. Many
22 factors support and motivate this evolution:

- 23 • **NRC policy decision.** In the mid-1990s, the NRC issued a policy statement
24 (Reference 1-1) that encouraged the use of probabilistic risk assessments (PRAs) to
25 improve safety decisionmaking and improve regulatory efficiency. This policy statement
26 formalized the Commission’s commitment to the expanded use of PRA, stating in part
27 that “the use of PRA technology should be increased in all regulatory matters to the
28 extent supported by the state-of-the-art in PRA methods and data and in a manner that
29 complements the NRC’s deterministic approach and supports the NRC’s traditional
30 defense-in-depth philosophy.” Since that time, the NRC has made progress in its efforts
31 to implement risk-informed and performance-based approaches into its regulation and
32 continues to revisit and update the approaches on a regular basis. Two notable efforts in
33 PFM include the FAVOR (Fracture Analysis of Vessels—Oak Ridge) (References 1-2, 1-
34 3) and xLPR (extremely low probability of rupture) projects (Reference 1-4).
- 35 • **Factors unanticipated in the design phase or not addressed by codes and**
36 **standards.** There is a fundamental difference between how deficiencies, or potential
37 deficiencies, are addressed when they are discovered during the design and
38 construction of a structure versus when they are revealed later, often after many years
39 or decades of safe service. During design and construction, deficiencies that do not
40 meet specifications are often addressed by repair, replacement, or reconstruction,
41 because the effort to demonstrate the acceptability of the deficiency often exceeds the
42 effort associated with correcting the deficiency. However, once operation begins, repairs
43 that were considered feasible during construction can become cost prohibitive (“cost” in
44 terms of dollars, time, or dose). While the NRC’s primary mission is safety, it is obligated
45 (see Title 10 of the *Code of Federal Regulations* (10 CFR) 50.109(c)(5) and (7)
46 (Reference 1-5)) to assess whether safety benefits justify the attendant cost. PFM
47 assessments are ideally suited to such situations because PFM metrics relate directly

1 and clearly to systems that can challenge safety (i.e., probability of structural failure).
2 Indeed, the Backfit Rule (Reference 1-5) explicitly requires an assessment of risk. PFM
3 also provides more flexible methods to account for factors that occur during service
4 (e.g., new damage mechanisms, unanticipated loadings, aging) that were not considered
5 during design. Especially when such factors are encountered for the first time, the
6 performance of deterministic analyses following the guidelines of codes, standards, and
7 regulations can be difficult because these established procedures may not account for
8 the new factors. Historically, unanticipated material degradation mechanisms have
9 regularly arisen in nuclear power plants (Reference 1-6). Examples include the primary
10 water stress-corrosion cracking aging issue in Alloy 600 and 182/82 welds in
11 pressurized-water reactors (which led in part to the development of the xLPR code), cold
12 head cracking, and control rod drive mechanism thermal sleeve wear.

- 13 • **Need to understand conservatisms.** One of the factors in the evolution of PFM is that
14 bases are needed to understand the level of conservatism in typical deterministic
15 evaluation. PFM is a means to calculate best estimate values and the associated
16 uncertainties and margins, and in turn it is a means to quantify conservatisms. By
17 understanding these conservatisms, analysts can refine the safety requirements.
18

19 Over the years, the NRC has received numerous submittals that contain PFM results, with
20 varying levels of quality. The inconsistency in the contents of the submittals has often led to low
21 efficiency in the reviews and a lack of predictable regulatory outcomes. For example, the
22 Electric Power Research Institute's (EPRI's) Materials Reliability Program (MRP) has submitted
23 to the NRC several reports containing PFM analyses, both for informational purposes and for
24 seeking review and approval. Such efforts include the following:

- 25 • "Materials Reliability Program: Probabilistic Fracture Mechanics Analysis of PWR
26 Reactor Pressure Vessel Top Head Nozzle Cracking (MRP-105)," Report 1007834,
27 issued 2004 (Reference 1-7)
- 29 • "Materials Reliability Program: Alloy 82/182 Pipe Butt Weld Safety Assessment for
30 U.S. PWR Plant Designs (MRP-113)," Report 1009549, issued 2006 (Reference 1-8)
- 31 • "Materials Reliability Program: Probabilistic Risk Assessment of Alloy 82/182 Piping Butt
32 Welds (MRP-116)," Report 1009806, issued 2004 (Reference 1-9)
- 33 • "Materials Reliability Program: Inspection and Evaluation Guidelines for Reactor Vessel
34 Bottom-Mounted Nozzles in U.S. PWR Plants (MRP-206)," Report 1016594, issued
35 2009 (Reference 1-10)
- 36 • "Materials Reliability Program: Topical Report for Primary Water Stress Corrosion
37 Cracking Mitigation by Surface Stress Improvement (MRP-335 Revision 3),"
38 Report 3002007392, issued 2016 (Reference 1-11)
- 39 • "Materials Reliability Program: Reevaluation of Technical Basis for Inspection of
40 Alloy 600 PWR Reactor Vessel Top Head Nozzles (MRP-395)," Report 3002003099,
41 issued 2014 (Reference 1-12)
- 42 • "BWRVIP-05: BWR Reactor Pressure Vessel Shell Weld Inspection Recommendations
43 (BWRVIP-05)," TR-105697, issued 1995 (Reference 1-13)

- 1 • “BWRVIP-241-A: BWR Vessel and Internals Project: Probabilistic Fracture Mechanics
2 Evaluation for the Boiling Water Reactor Nozzle-to-Vessel Shell Welds and Nozzle
3 Blend Radii,” Report 3002013093, issued 2018 (Reference 1-14)
- 4 • “BWRVIP-108-A: BWR Vessel and Internals Project: Technical Basis for the Reduction
5 of Inspection Requirements for the Boiling Water Reactor Nozzle-to-Vessel Shell Welds
6 and Nozzle Blend Radii,” Report 3002013092, issued 2018 (Reference 1-15)

7 **1.3 Objective**

8 The NRC intends this document to provide a generalized technical basis for the following:

- 9 • validating and verifying a PFM capability
- 10 • developing input distributions that feed into the PFM framework
- 11 • characterizing and propagating input and model uncertainties
- 12 • understanding the impacts of problem assumptions on the adequacy of the results
- 13 • choosing a methodology with the appropriate complexity for the intended application
- 14 • properly conducting a PFM analysis
- 15 • correctly interpreting the results of a PFM analysis in a regulatory context
- 16 • documenting the important steps and information relevant to the PFM code and analysis
17 at hand

18 This NUREG describes how to improve confidence in structural analyses performed using PFM
19 by focusing on topics such as problem definition, PFM model development, input definition,
20 uncertainty analyses, probabilistic framework development, and output analysis, including
21 sensitivity analyses (SAs) (to determine impact of uncertainties on result) and sensitivity studies
22 (to determine impact of plausible changes to analysis assumptions). For each of these topics,
23 this NUREG shows how a graded approach for PFM analyses and submittals can improve
24 confidence in the structural analyses performed.

25 **1.4 Structure of This Document**

26 This NUREG has three technical sections. The content provided in all three sections is linked,
27 but an applicant’s experience and familiarity with PFM analyses will determine whether it needs
28 to refer to that content.

29
30 Section 2 provides a tiered framework for a submittal that contains PFM analyses and results
31 and could be used by applicants of all experience levels. This section provides a graded
32 approach for developing PFM analyses and submittals.

33
34 Section 3 provides a framework for performing a PFM analysis. This section could be used by
35 applicants who have used PFM in prior submittals but who are seeking additional information
36 regarding the development of an analysis structure or formalism. This section is not intended to
37 prescribe a linear analysis, since PFM analyses are typically iterative in nature. Furthermore,

1 not every application needs all steps and actions, and the analyst can evaluate the necessity to
2 perform each step and action on a case-by-case basis. Table 2-1 in Section 2 provides a
3 mapping between the analysis actions given in Section 3 and the associated documentation.
4
5 Section 4 details analysis methodologies, including notional examples for context. This section
6 is intended to be used by applicants who are seeking to more fully understand the theoretical
7 underpinnings of the processes that are used to establish the credibility of a PFM analysis. Each
8 subsection is linked to an action that was introduced in Section 3.

9 **1.5 References**

- 10 1-1. "Use of Probabilistic Risk Assessment Methods in Nuclear Regulatory Activities; Final
11 Policy Statement," *Federal Register*, Vol. 60, No. 158, page 42622, August 16, 1995.
- 12 1-2. Williams, P.T., Dickson, T.L., Bass, B.R., and Klasky, H.B., "Fracture Analysis of
13 Vessels—Oak Ridge—FAVOR v16.1, Computer Code: Theory and Implementation
14 Algorithms, Methods, and Correlations," ORNL/LTR-2016/309, Oak Ridge National
15 Laboratory, 2016 (ML16273A033).
- 16 1-3. Dickson, T.L., Williams, P.T., Bass, B.R., and Klasky, H.B., "Fracture Analysis of
17 Vessels—Oak Ridge—FAVOR v16.1, Computer Code: User's Guide,"
18 ORNL/LTR-2016/310, Oak Ridge National Laboratory, 2016 (ML16273A034).
- 19 1-4. Rudland, D.L., and Harrington, C., "xLPR Pilot Study Report," NUREG-2110, U.S.
20 Nuclear Regulatory Commission, May 2012 (ML12145A470).
- 21 1-5. Backfitting, 10 C.F.R. § 50.109(c)(3).
- 22 1-6. Wilkowski, G., Tregoning, R., Scott, P., and Rudland, D., "Status of Efforts to Evaluate
23 LOCA Frequency Estimates Using Combined PRA and PFM Approaches," in
24 *Proceedings of 28th MPA Seminar*, Stuttgart, Germany, October 2002.
- 25 1-7. Electric Power Research Institute, "Materials Reliability Program: Probabilistic Fracture
26 Mechanics Analysis of PWR Reactor Pressure Vessel Top Head Nozzle Cracking
27 (MRP-105)," Report 1007834, 2004 (ML041680489).
- 28 1-8. Electric Power Research Institute, "Materials Reliability Program: Alloy 82/182 Pipe Butt
29 Weld Safety Assessment for US PWR Plant Designs (MRP-113)," Report 1007029,
30 2004 (ML042080193).
- 31 1-9. Electric Power Research Institute, "Materials Reliability Program: Probabilistic Risk
32 Assessment of Alloy 82/182 Piping Butt Welds (MRP-116)," Report 1009806, 2004
33 (ML043200641).
- 34 1-10. Electric Power Research Institute, "Materials Reliability Program: Inspection and
35 Evaluation Guidelines for Reactor Vessel Bottom-Mounted Nozzles in U.S. PWR Plants
36 (MRP-206)," Report 1016594, 2009.
- 37 1-11. Electric Power Research Institute, "Materials Reliability Program: Topical Report for
38 Primary Water Stress Corrosion Cracking Mitigation by Surface Stress Improvement
39 (MRP-335 Rev. 3)," Report 3002007392, 2016 (ML16055A215).

- 1 1-12. Electric Power Research Institute, "Materials Reliability Program: Reevaluation of
2 Technical Basis for Inspection of Alloy 600 PWR Reactor Vessel Top Head Nozzles
3 (MRP-395)," Report 3002003099, 2014 (ML14307B007).
- 4 1-13. Electric Power Research Institute, "BWR Vessel and Internals Project: BWR Reactor
5 Pressure Vessel Shell Weld Inspection Recommendations (BWRVIP-05)," TR-105697,
6 1995 (ML032200246).
- 7 1-14. Electric Power Research Institute, "BWRVIP-241-A: BWR Vessel and Internals Project:
8 Probabilistic Fracture Mechanics Evaluation for the Boiling Water Reactor
9 Nozzle-to-Vessel Shell Welds and Nozzle Blend Radii," Report 3002013093, 2018
10 (ML19297G738).
- 11 1-15. Electric Power Research Institute, "BWRVIP-108-A: BWR Vessel and Internals Project:
12 Technical Basis for the Reduction of Inspection Requirements for the Boiling Water
13 Reactor Nozzle-to-Vessel Shell Welds and Nozzle Blend Radii," Report 3002013092,
14 2018 (ML19297F806).

1 **2 PROBABILISTIC FRACTURE MECHANICS GRADED APPROACH**

2 This section provides background information on developing a graded approach for PFM
3 analyses, and, where applicable, provides explanations and supporting information for portions
4 of new proposed RG-1.245 (Reference 2-1)..

5 **2.1 Background**

6 In the past, the NRC has typically regulated the use of nuclear structural materials on a
7 deterministic basis. Safety factors, margins, and conservatisms were used to account for model
8 and input uncertainty. However, as described in Section 1, the NRC has progressed in its efforts
9 to implement risk-informed approaches into its regulation. In one such effort, the NRC
10 developed guidance on a risk-informed decisionmaking process that is acceptable to use for
11 design-basis changes. This guidance is contained in new proposed regulatory guide, RG-1.245.

12
13 When solving a probabilistic fracture mechanics problem, the level of effort associated with
14 analysis and documentation activities is generally dependent upon the goals of the analysis. In
15 fact, each analysis is usually considered uniquely within its own specific context to make
16 determinations about the expected level of rigor and to demonstrate that the problem is resolved
17 in a satisfactory way. This is particularly true as the safety significance of the analysis
18 application increases, and the consequences of an incorrect decision are more severe. The
19 availability of supplemental evidence to support the decision is also part of the consideration.
20 For example, if inspection data or operational measurements are available in addition to
21 analysis results, the analysis may be viewed as one piece of evidence in a larger context, and
22 the level of rigor may be adjusted accordingly. The guiding principle is that the level of detail
23 should be commensurate with the safety significance of the subject and the complexity of the
24 problem.

25
26 In October 2018, the NRC held a public meeting to discuss a graded approach for PFM codes
27 and analyses for regulatory applications. At the meeting, EPRI presented suggestions for
28 expected content in a PFM submittal. EPRI also submitted a white paper containing additional
29 details and guidelines. The NRC staff concurred that EPRI’s approach constituted a quality
30 basis from which to build further guidance. Consequently, based on a submitted proposal from
31 industry (Reference 2-2), the new proposed regulatory guide, RG-1.245, defines a practical
32 framework for developing the content of PFM submittals that maintains the effectiveness of
33 NRC reviews of such submittals while improving review efficiency.

34
35 Table 2-1 gives the complete mapping between the guidance in new proposed RG-1.245, the
36 NRC’s analytical steps in Section 3, and the item number of the suggested minimum content
37 and considerations of additional content given in Tables 1 and 2 from EPRI’s white paper
38 (Reference 2-2).

1 **Table 2-1 Content Mapping between RG1.245, NUREG Section 3, and EPRI White Paper**
 2 **(Reference 2-2)**

RG-1.245 Section	Content	NUREG PFM Analytical Steps Section 3	EPRI White Paper (Reference 2-2)	
			Suggested Content, Table 1	Additional Considerations, Table 2
2.1	Regulatory Context	3.1.1	-	-
2.2	Information Made Available to NRC Staff		1	-
2.2.1	PFM Software	3.1.3	1.1	1, 4, 11, 12, 13
2.2.2	Supporting Documents	3.1.3	1.2	-
2.3	Quantities of Interest and Acceptance Criteria	3.1.2	8	-
2.4	Software quality assurance and verification and validation	3.1.3	6	1
2.5	Models	3.1.3	2	1, 2, 5, 6, 9, 10
2.6	Inputs	3.2.1 3.2.2 3.3.1 3.4.1	3, 5	3, 4, 5, 6
2.7	Uncertainty Propagation	3.3.1	7	3, 10
2.8	Convergence	3.3.2	4	3
2.9	Sensitivity Analyses	3.3.3	-	-
2.10	Output Uncertainty Characterization	3.3.4	-	-
2.11	Sensitivity Studies	3.4.1 3.4.2	5	1, 2, 11

3

4 **2.2 Supporting Information for RG-1.245**

5 This section presents useful explanations and background for topics addressed in new
 6 proposed RG-1.245.

7 **2.2.1 Quantities of Interest and Acceptance Criteria**

8 The NRC typically approves the acceptance criteria, which may be relative or absolute. Relative
 9 acceptance criteria refer to a relative comparison of probabilistic results under the proposed
 10 approach versus an already acceptable approach. In general, the rigor required in
 11 demonstrating that a relative acceptance criterion is met is lower than that required in
 12 demonstrating that an absolute acceptance criterion is met.

13 **2.2.2 Software Quality Assurance and Verification and Validation**

14 New or more complex PFM codes may warrant a more thorough review than codes that are less
 15 complex or more familiar to the NRC staff. The extent of the differences between a new PFM
 16 code and the codes previously approved by the NRC are likely to have an impact on the
 17 complexity and depth of a given regulatory review. Similar impacts on the scope of review may

1 exist when a code previously reviewed by the NRC is applied in a new way (i.e., outside the
2 previously reviewed range of use for the code). Certain specific applications of the code, such
3 as those involving a high safety significance or if the code is plant specific (versus an intended
4 generic application), may also warrant deeper and more thorough investigations.

5
6 For a code that the NRC has previously approved, the technical basis for using the code is likely
7 well understood, such that supplemental SQA and V&V efforts are unnecessary to understand
8 the credibility of the results. For a code that the NRC has previously approved but that has been
9 modified for the analysis being performed, understanding the technical basis for the
10 modifications is important. For a code that is new and has not been previously approved in any
11 form, understanding the entire technical basis informs the credibility of the results. With this in
12 mind, the set of different analysis codes can be divided into several categories that may warrant
13 different levels of QA and V&V...

14
15 In general, to meet the objectives of a given QA program, V&V may be performed on individual
16 submodels and the unifying framework, or it can be performed directly on the overall code.
17 Some QA programs also allow for checks using alternate calculation methods (e.g.,
18 spreadsheets or alternate implementations). The applicable QA program, plan, or procedures
19 define the supporting documents created in conjunction with PFM analysis code development. A
20 graded approach to QA for software development, with different minimum requirements
21 depending on the software application, such as that outlined in International Atomic Energy
22 Agency (IAEA) Technical Report Series No. 397, "Quality Assurance for Software Important to
23 Safety," issued in 2000 (Reference 2-3), may be useful to reduce unnecessary documentation.
24 Furthermore, as the applicable QA program may depend on the safety significance of the
25 component or system being evaluated, the corresponding rigor of V&V may also vary.

26
27 If a code is used for an application that is different than the one for which it was developed, the
28 existing verification may still be valid, but the validation may need to be extended or redone if
29 the previous validation was specific to a different range of use.

30 **2.2.3 Models**

31 The goal of any engineering assessment methodology is to determine the response of a system
32 to a variety of inputs. The behavior of the system in question can be predicted more accurately
33 when using a set of models that best represents the physical behavior of the system of interest.
34 However, because of analytical and computational limitations, model selection may involve
35 balancing the accuracy and practicality of various mathematical approaches. Whenever a model
36 is constructed, inherent simplifications are injected into the representation to make model
37 evaluation feasible.

38
39 The process of developing a model begins with a conceptual model, which defines the physics
40 to be included. This decision is often aided by a process that defines the most critical physics to
41 capture in the analysis. Then, for each relevant physics, a mathematical model is chosen to
42 represent that physics, and a code is selected or developed to solve the chosen mathematical
43 model. Over the course of an analysis, the model and code may be updated, revised, or
44 calibrated with available data to improve predictive capability and understand how similar the
45 conditions of validation tests are to the application space of interest. Engineering judgment is an
46 inevitable and integral part of model development.

47
48 Another factor that influences model development is computational resources. While a particular
49 approach may be considered the "best estimate," it may not be practical for a PFM analysis

1 given the time and resource constraints imposed on the analyst. The occasional need to choose
2 a model that has less fidelity but is easier to solve due to the solution speed requirements of
3 PFM may affect results. Model choice can be complicated further by the fact that PFM
4 encourages the use of the most accurate deterministic models rather than conservative models
5 (so as to maximize accuracy in estimating quantities of interest and their uncertainty). These
6 more accurate models may require longer solution times.

7 **2.3 References**

8 2-1. U.S. Nuclear Regulatory Commission, DG-1382/RG-1.245: Preparing Probabilistic
9 Fracture Mechanics (PFM) Submittals, Washington, DC, USA: U.S. NRC.

10 2-2. Palm, N., "White Paper on Suggested Content for PFM Submittals to the NRC,"
11 BWRVIP 2019-016, Electrical Power Research Institute, 2019 (ML19241A545).

12 2-3. International Atomic Energy Agency, "Quality Assurance for Software Important to
13 Safety," Technical Report Series No. 397 (TRS-397), Vienna, Austria, 2000.

14

3 ANALYTICAL STEPS IN A PROBABILISTIC FRACTURE MECHANICS ANALYSIS

This section describes a process for conducting a PFM analysis. It is generally assumed that an analysis process is implemented after PFM code quality and credibility have been established through SQA processes and V&V. The process followed in performing analyses for a PFM submittal is not required to be the same as the process outlined here, but it should be structured to address the specific features of the application under investigation.

A generalized PFM analysis process is structured according to five key steps:

- (1) Translate regulatory requirements into an analysis plan.
- (2) Characterize input uncertainty.
- (3) Estimate Qols and their associated uncertainty.
- (4) Conduct sensitivity studies to assess credibility of modeling assumptions.
- (5) Draw conclusions from analysis results.

This section describes each step in the PFM analysis process and its corresponding analyst actions, along with the following information:

- **Purpose.** Motivation for including this step in a PFM analysis.
- **Description.** High-level description of the concept.

These steps and actions are intended to provide a conceptual framework for conducting and presenting the results of a PFM analysis that can be used in a risk-informed regulatory assessment, but they are not intended to be performed in a strictly linear fashion. PFM analyses are typically iterative in nature. Furthermore, not all steps and actions are needed in every application, and the analyst should evaluate the necessity to perform each step and action on a case-by-case basis. Different applications will warrant different levels of analysis complexity and documentation. If separate PFM analyses are conducted for different regulatory contexts or Qols, then these analyses should be documented separately.

Figure 3-1 summarizes the steps and actions and their relationship to one another. This figure also shows the organization of this section and the iterative nature of PFM analyses.

	Step	Action
Plan	1. Translate regulatory requirements into an analysis plan	<ul style="list-style-type: none"> • Define the regulatory context (Sec. 3.1.1) • Define the QoI and how it relates to the model output and acceptance criteria (Sec. 3.1.2) • Determine the suitability of the PFM code for the application (Sec. 3.1.3) • Identify key elements of the problem that impact analysis choices (Sec. 3.1.4)
	2. Characterize model input uncertainty	<ul style="list-style-type: none"> • Identify uncertain model inputs (Sec. 3.2.1) • Specify probability distributions on uncertain inputs (Sec. 3.2.2)
Analyze	3. Estimate QoIs and their associated uncertainty	<ul style="list-style-type: none"> • Select a sampling scheme (Sec. 3.3.1) • Assess sampling uncertainty (Sec. 3.3.2) • Conduct sensitivity analysis (Sec. 3.3.3) • Conduct output uncertainty analysis (Sec. 3.3.4)
	4. Conduct sensitivity studies to assess the credibility of modeling assumptions	<ul style="list-style-type: none"> • Determine a set of sensitivity studies (Sec. 3.4.1) • Conduct sensitivity studies and present results (Sec. 3.4.2)
Synthesize	5. Draw conclusions From analysis results	<ul style="list-style-type: none"> • Interpret analysis results (Sec. 3.5.1) • Iterate on the analysis process to refine model results (Sec. 3.5.2)

Analysis results inconclusive, refinements needed.

1

2

3

Figure 3-1 Flowchart Describing the Steps and Actions of a PFM Analysis

1 A key element of risk-informed decisionmaking is identifying uncertainties that impact the
2 analysis results and subsequent regulatory decision. The proposed steps and actions are
3 intended to reflect sources of uncertainty that are common to all PFM applications, including the
4 following:

- 5 • **Step 2: Input uncertainty.** The specific values of model inputs are typically unknown;
6 this input uncertainty results in uncertainty in the model output, such as the likelihood of
7 an adverse event. Accounting for this uncertainty in model inputs is what distinguishes
8 deterministic and probabilistic fracture mechanics applications.
- 9 • **Step 3: QoI approximation uncertainty.** PFM analyses are based on a finite number of
10 model realizations, resulting in sampling uncertainty. This sampling uncertainty can
11 impact the accuracy of the analysis results.
- 12 • **Step 4: Modeling assumption uncertainties.** PFM analyses may rely on assumptions
13 and approximations that introduce additional uncertainty into the analysis. The impact of
14 uncertain assumptions can be addressed using sensitivity studies.

15 The discussion in this section refers to Section 4, which includes specific technical details about
16 elements of PFM analyses.

17 **3.1 Step 1: Translation of Regulatory Requirements into an Analysis Plan**

18 The first step in a PFM analysis is translating regulatory requirements into a PFM analysis plan.
19 This step involves four key actions:

- 20 (1) Define the regulatory context.
- 21 (2) Define the QoI and how it relates to the PFM model output.
- 22 (3) Determine suitability of the PFM code for the application.
- 23 (4) Identify key elements of the problem that impact analysis choices.

24 **3.1.1 Step 1: Action 1—Define the Regulatory Context**

25 **Purpose:** The purpose of this step is to define how PFM analyses will be used as a technical
26 basis for a regulatory action, including the criteria to be used to support a proposed regulatory
27 action.

28
29 **Description:** When using PFM in support of an application to the NRC, it is important to
30 understand how PFM informs whether regulatory requirements have been met, specifically why
31 a probabilistic approach is appropriate for the problem at hand, and how the probabilistic
32 approach is used to demonstrate compliance with the regulatory criteria. It is particularly
33 important to explain how the probabilistic approach informs the regulatory action when no
34 specific acceptance criteria exist to demonstrate compliance for the problem at hand.

35 **3.1.2 Step 1: Action 2—Define the Quantity of Interest and How it Relates to the Model** 36 **Output and Acceptance Criteria**

37 **Purpose:** The purpose of this step is to directly map regulatory requirements onto specific
38 model outputs, ensuring that the model is predicting appropriate and relevant quantities.
39

1 **Description:** The model output is directly linked to one or more Qols and the acceptance
2 criteria.

3
4 A Qol is a quantity that is directly tied to a regulatory decision. The Qol is related to a model
5 output or to a function of outputs; for a PFM model to be useful, understanding the relationship
6 between the model output and the Qol is critical. For example, suppose the Qol is the
7 probability of rupture by year for a single pipe. For each set of inputs, the PFM model may
8 output the year in which rupture occurs. The Qol is then estimated by calculating the frequency
9 of rupture by year across many realizations of this single pipe's performance.

10
11 In PFM analyses, the Qol will frequently be a probability of an adverse event; however, using a
12 proxy for an adverse event may be necessary when its probability is too small to accurately
13 estimate using computer simulation. For example, probability of rupture could be related to
14 crack length or crack depth, and one or both of these quantities could potentially be used as
15 surrogates for rupture.

16
17 The Qol is typically tied to the acceptance criteria. Often, an acceptance criterion is expressed
18 as a point in the Qol space at which decisions are determined based on whether the Qol
19 exceeds the threshold. An example acceptance criterion is "the 95th percentile of the predicted
20 leak rate must remain below the makeup capacity of the system."

21
22 Both the Qol and acceptance criteria are defined relative to the unit of measurement and the
23 time period over which the Qol is calculated.

24
25 The unit of measurement specifies the target population for inference, defined as the entire set
26 of objects to which the analyst is trying to generalize the results of the analysis. The Qol is
27 interpreted relative to the units of measurement, such as a fleet of power plants, a single plant,
28 a line within the plant, or a single weld within a plant. The units of measurement can also be
29 defined spatially, such as per kilometer of pipe.

30
31 The time period is the interval of time over which the Qol is calculated, such as per year, per
32 decade, or over the life of the plant.

33
34 As an example, consider an analysis intended to show that the likelihood of a single pipe
35 leaking is small over the life of a plant. The Qol is the probability of pipe leakage, the
36 acceptance criterion is the acceptable upper limit on the probability of leakage, the time period
37 is the plant life duration, and the units are the single pipe of interest. All quantities are
38 dependent on the modeling assumptions. For example, no mitigation, 10-year inspection
39 intervals, and the leak detection system all impact the assessments.

40 **3.1.3 Step 1: Action 3—Determine the Suitability of the Probabilistic Fracture** 41 **Mechanics Code for the Specific Application**

42 **Purpose:** The purpose of this step is to determine whether a specific PFM code is suitable for
43 the application of interest and to identify any potential limitations of the code with regard to the
44 application.

45
46 **Description:** The SQA process should follow the graded approach suggested in Section 2.2.2.
47 It is intended to provide assurance that the software was developed in a deliberate and
48 controlled manner, such that every aspect of the software is known and understood.
49 Furthermore, the SQA process ensures source and version control, so as to prevent inadvertent

1 changes to the software that could have unintended consequences on the software predictions.
2 For nuclear regulatory applications, Title 10 of the *Code of Regulations*, Part 50, “Domestic
3 licensing of production and utilization facilities,” (Reference 3-1) Appendix B, “Quality Assurance
4 Criteria for Nuclear Power Plants and Fuel Reprocessing Plants,” requires that the applicants
5 have an approved QA process in place.
6

7 The V&V process is intended to provide the critical evidence for the credibility of a code and a
8 set of analysis tools, and it is composed of two primary activities, known as verification and
9 validation. In general, verification seeks to determine whether a given mathematical model has
10 been solved correctly within the analysis framework. This process has two components, referred
11 to as code verification and solution verification. Code verification specifically focuses on the
12 implementation of software to solve a given set of governing equations (i.e., the mathematical
13 model). Solution verification focuses on approximations to the governing equations that are
14 needed in order to solve them on a computer. These approximations may be made in space,
15 time, or stochastic dimensions. Solution verification has the goal of quantifying the error incurred
16 by these approximations and determining that these effects converge toward zero as resolution
17 is increased (e.g., time steps are reduced or spatial approximations are refined).
18

19 Validation seeks to determine whether a chosen mathematical model is an accurate description
20 of reality. Traditional validation involves comparing outcomes of a simulation to experimental
21 data taken from a representative real-world scenario to determine the accuracy of the overall
22 model representation. An alternative validation approach in the absence of experimental data
23 includes benchmarking the software with comparable software that has been verified and
24 validated previously. The model fidelity has several components, including the physics-based
25 models, the material models, and the geometric description of a system of interest.
26

27 Researchers have detailed these elements in a variety of references (e.g., References 3-2, 3-3,
28 3-4) and a set of standards produced by ASME (References 3-5 and 3-6). While nominally
29 discipline specific, the methods described in these guides and the references therein are very
30 general in nature and provide a good basis for foundational V&V activities in support of model
31 credibility.
32

33 Section 2.2.2 provides information on SQA and V&V documentation for all PFM analysis codes
34 used in analysis. Individual analyses will apply the code in a specific manner; an important
35 aspect of the credibility of the overall analysis is the degree of confidence in the code for the
36 intended application. The intent of this action is to identify and resolve any important gaps in the
37 code capabilities for the intended application.
38

39 *Code capabilities.* Code capabilities refer to all scenarios for which a code has been through an
40 appropriate set of V&V activities. Examples of code capabilities include (1) the range of inputs
41 that were included in verification tests and validation test data, (2) the set of material models or
42 geometries that have an established pedigree, (3) the underlying physics models and the
43 assumptions underlying their range of applicability, and (4) the numerical approximation
44 schemes (e.g., grid size, spatial and temporal resolution) with appropriate solution verification.
45

46 Examples of questions to consider with regard to code capabilities include the following:

- 47 • How well does the chosen model represent the application?
- 48 • Is there a rationale for defining certain model assumptions as conservative?

- 1 • Is the coding for the physics-based models available for review?
- 2 • Are the physics-based models well understood and established?
- 3 • Are code limitations that may impact the regulatory question/issue identified?
- 4 • Is mathematical justification for the model representation of the physics well
5 established?
- 6 • Are limitations of the methodology identified with respect to interpolation or
7 extrapolation?

8 *Analysis features and code capabilities.* An important first step in an analysis is to compare
9 features of the intended application to the code capabilities to determine whether the code is
10 suitable for the application. This process identifies any features that are incompatible with the
11 code capabilities. Further, it identifies any features of the analysis for which the code does not
12 have sufficient V&V evidence. As an example, if a PFM code was validated and calibrated for a
13 specific range of weld residual stresses, then considering the implications of applying the code
14 outside of this input range is critical for interpreting the model credibility.

15
16 The following are some of the key considerations for code capability:

- 17 • Does the range of inputs for which the code has been calibrated and validated include
18 the range of inputs that are required for the specific application? Have the physics
19 models been changed for the specific application? Are the numerical approximations
20 sufficient for the application?
- 21 • If application-specific changes have been made, is the phenomenological behavior of
22 the code expected to be similar for this application relative to the applications for which
23 validation occurred (i.e., are the same physics models still relevant and adequate?)
- 24 • Are there any additional test data to support the applicability of the code for the current
25 application?

26 *Addressing code limitations.* Potential limitations of the code for the application can be
27 addressed in two ways:

- 28 (1) Risk can be mitigated by collecting additional information to improve the vetting of the
29 code in the identified risk areas.
- 30 (2) When it is not possible to collect additional information, justification for the credibility of
31 the code capabilities for the application can often be based on appropriate engineering
32 arguments. When sufficient evidence cannot be collected to address certain gaps,
33 understanding the associated risk to the analysis credibility is critical to interpreting the
34 final results.

35 **3.1.4 Step 1: Action 4—Identify Key Elements of the Problem that Impact Analysis** 36 **Choices**

37 **Purpose:** The purpose of this step is to identify key elements of the PFM application that will
38 determine how to conduct the analysis. Simplifying assumptions and approximations may be

1 necessary based on the complexity of the problem or, conversely, may be justified because the
2 problem at hand is inherently not complex.

3

4 **Description:** Specific aspects of the application drive the methods used in a PFM analysis. In
5 an ideal situation, simple analysis techniques can be applied. More sophisticated analysis
6 methodologies are useful when the following is true:

- 7 • *The model is computationally expensive.* When models are computationally inexpensive
8 to run, sampling uncertainty due to limited model realizations is a secondary issue,
9 because the sample size can often be made arbitrarily large such that sampling
10 uncertainty is negligible. On the other hand, computationally expensive models require
11 more forethought about how to select model realizations and how to design model
12 sampling schemes to achieve converged results.
- 13 • *The QoI is a rare event probability.* Estimating rare event likelihoods typically requires
14 more realizations, more sophisticated sampling schemes, or both. Rare event
15 probabilities (e.g., adverse event or failure probabilities) are defined as probabilities that
16 are close enough to zero that the number of samples needed to estimate the probability
17 is large with respect to computational budget. For example, to estimate a
18 1×10^{-6} probability using simple Monte Carlo sampling (Section 4.3.1), more than
19 1×10^6 model realizations are required.
- 20 • *There are many model inputs.* When the number of model inputs is large, then there are
21 more input uncertainties to characterize. Also, identifying important/sensitive model
22 inputs is more difficult because there are more candidate inputs.
- 23 • *Separation of aleatory and epistemic uncertainty is maintained.* Uncertainty can arise
24 from different causes; the most commonly considered types of uncertainty are aleatory
25 and epistemic uncertainty (Section 4.1.1). For a specific adverse event, the
26 quantification of aleatory uncertainties targets the question, “How likely is the event to
27 happen?” while the quantification of the epistemic uncertainties targets the question,
28 “How confident are we in this estimate of the event likelihood?” PFM analyses can treat
29 aleatory and epistemic uncertainties separately to distinguish the frequency of event
30 occurrence from the confidence in the frequency estimate. Separating uncertainty
31 introduces additional complexity and computational burden into an analysis, because of
32 the double-looping algorithm for separation described in Section 4.1.1. Section 3.2.1 and
33 Section 4.1.1 provide more details about classifying and separating aleatory and
34 epistemic uncertainty.

35 For each of these attributes, it is generally more challenging to conduct SAs to identify important
36 inputs (Section 3.3.3) and design sampling algorithms to achieve statistical model convergence
37 (Sections 3.3.1 and 3.3.2). The points at which these elements can impact the analysis
38 decisions are highlighted throughout the PFM analysis process.

39 **3.2 Step 2: Model Input Uncertainty Characterization**

40 The second step in a PFM analysis is characterizing input uncertainty. This step involves two
41 key actions:

- 42 (1) Identify uncertain model inputs.
- 43 (2) Specify probability distributions on uncertain inputs.

1 The end goal of this step is to determine probability distributions to represent input uncertainty.

2 **3.2.1 Step 2: Action 1—Identify Uncertain Model Inputs**

3 **Purpose:** The purpose of this step is to determine which model inputs are treated with
4 uncertainty and, if relevant, the type of uncertainty (aleatory or epistemic) for each input.

5
6 **Description:** This action includes classifying deterministic versus uncertain inputs and
7 classifying aleatory versus epistemic uncertain inputs (if relevant).

8
9 *Deterministic versus uncertain inputs.* Inputs to a PFM analysis can be represented in two ways:

- 10 • Deterministic inputs take on a single value.
- 11 • Uncertain inputs can take on a range of potential values.

12 Deterministic inputs are fixed to a single value across all model realizations. Such inputs can be
13 fixed for several reasons: (1) they have known physical values (e.g., a known yield strength of a
14 material), (2) the chosen fixed value is determined to be a value of interest (e.g., a conservative
15 value used for a specific reason or a value of relevance for sensitivity studies (see Section 3.4)),
16 or (3) including uncertainty would not affect decisionmaking. Uncertain inputs determine the
17 amount of variability in the model output, conditional on the values of the deterministic inputs.
18 This uncertainty in model inputs is what distinguishes a purely deterministic analysis from a
19 PFM analysis. If the QoI is a failure probability, this probability is determined based on the
20 uncertainty in the model's uncertain inputs, conditional on the values of the deterministic inputs.
21 Data, expert judgment, and SA (Section 3.3.3) inform whether an input is modeled as
22 deterministic or uncertain.

23
24 Understanding the rationale for classifying inputs as deterministic or uncertain is important when
25 interpreting the analysis results. If there is uncertainty as to whether an input is deterministic or
26 uncertain, then modeling the input as uncertain is preferable.

27
28 *Avoiding excessive conservatism in model inputs.* Deterministic fracture mechanics models
29 have historically relied on conservatisms; introducing conservatism into a PFM analysis makes
30 the results difficult to interpret. Conservatisms in inputs may propagate to produce an
31 unrealistically conservative output. For example, the probability that 10 independent variables all
32 take values at or above their respective 90th percentile is 1×10^{-10} , or 1 chance in 10 billion.
33 Hence, taking a conservative approach and setting each of these inputs to their 90th percentile
34 in a deterministic model realization results in a highly unlikely output. Even setting a single input
35 to a conservative value can substantively change the interpretation of the model results; if the
36 model output is highly sensitive to this input, then subsequent modeling results will on average
37 be conservative. Additionally, conservative assumptions in submodels may be anticonservative
38 in full system models. For example, increases in leak rate may be considered conservative at a
39 submodel level. However, when combined with leak rate detection, this conservatism could lead
40 to the suppression of failures due to increased leak rate detection.

41 Understanding when and why conservative inputs are used is important to interpreting the final
42 model results. The influence of conservative choices can be addressed using sensitivity studies.
43 These sensitivity studies are especially important when specifying a best estimate or
44 conservative value is difficult due to limited information.

1 The best estimate is defined as an approximation based on the best available information.
2 Using a best estimate does not imply the chosen deterministic value or input distribution has no
3 uncertainty.

4 *Aleatory versus epistemic uncertain inputs.* If an analysis maintains separation between aleatory
5 and epistemic uncertainty, then uncertain inputs are classified as epistemic or aleatory.
6 Section 4.1.1 provides more details on aleatory versus epistemic uncertainty. This classification
7 is not necessarily straightforward, because the uncertainty type often depends on the context
8 and granularity of the problem. As an example, in a conventional linear elastic fracture
9 mechanics model, the uncertainty in the linear elastic fracture toughness (K_{Ic}) may be regarded
10 as aleatory (irreducible or inherent). Conversely, in a micromechanics model that accounts for
11 features such as grain size, inclusions, and dislocations (i.e., the factors that create the
12 uncertainty in K_{Ic}), this uncertainty may be regarded as epistemic. Mixed situations (part
13 aleatory, part epistemic) are also possible. The categorization of uncertainty is therefore not
14 totally objective and may change depending on the context of the problem.

15
16 To interpret modeling results, it is important to understand how aleatory and epistemic
17 uncertainty are defined in the context of the application and to understand the rationale for
18 classifying inputs as epistemic or aleatory. If it is uncertain whether an important input is
19 aleatory or epistemic, sensitivity studies (Section 3.4) can be conducted to determine the impact
20 of changing the classification.

21 **3.2.2 Step 2: Action 2—Specify Probability Distributions on Uncertain Inputs**

22 **Purpose:** In PFM analyses, uncertainty in model inputs is represented through probability
23 distributions. This uncertainty is propagated forward to the model outputs to estimate and
24 quantify uncertainty in Qols.

25
26 **Description:** This action includes considering attributes of input distribution specification,
27 including the following:

- 28 • iterative nature of input distribution specification
- 29 • importance of analysis context in characterizing input uncertainty
- 30 • nonprobabilistic representations of input uncertainty
- 31 • expert judgment
- 32 • distribution specification methods
- 33 • bounding input distributions
- 34 • accounting for correlation in model inputs

35 *Iterative nature of input distribution specification.* A PFM analysis focuses on those inputs that
36 have the most influence on the model output. These influential inputs are typically identified
37 using SA (Section 3.3.3). If an input's uncertainty has little impact on the output uncertainty, a
38 strong technical basis for the input distribution may not be necessary and a deterministic value
39 could be used. If these results indicate a large impact, additional data, more refined statistical
40 techniques, or further expert elicitation may be needed to further refine the input's probability
41 distribution. In this way, the development of inputs for PFM analysis is an iterative process, and
42 the distributions specified in this step may be iteratively refined in the analysis process.

43
44 *Importance of analysis context in characterizing input uncertainty.* The context of the analysis
45 impacts the input uncertainty. Specific analyses will often have narrower uncertainty ranges
46 than more general analyses. For example, if an analysis is specific to a certain pipe in a specific

1 plant, then the geometry and other characteristics of the system are likely to be defined
2 precisely and the uncertainty range may be relatively small. In contrast, for an analysis meant to
3 represent a series of welds or generic configurations across the U.S. reactor fleet, the variability
4 in geometry, operating conditions, materials, and possible flaw mitigation is likely to be larger.

5
6 *Nonprobabilistic representations of input uncertainty.* In PFM applications, it is common practice
7 to represent input uncertainty by specifying probability distributions on the inputs. In some
8 analyses, it may be appropriate to use other nonprobabilistic representations of uncertainty to
9 characterize an unknown input. Specifically, for epistemic uncertainties, if the lack of knowledge
10 is too great to specify a probability distribution on an input, then nonprobabilistic, interval-based
11 bounding methods can be considered (References 3-2,3-7). Probabilistic representation of
12 uncertainty is often sufficient in PFM applications; understanding the rationale for deviating from
13 a fully probabilistic analysis is important to interpreting the analysis results.

14
15 *Expert judgment.* In PFM applications, relevant data needed to define input distributions are
16 often sparse or unavailable. In these cases, literature and expert opinion can be leveraged. The
17 NRC has provided specific guidance on expert elicitation, with applications to uncertain model
18 inputs (Reference 3-9).

19
20 *Distribution specification methods.* Proper selection of a probability distribution for an uncertain
21 input requires detailed knowledge of the available data as well as qualitative judgments. Expert
22 judgment and the amount and pedigree of the data, as well as the importance of the particular
23 input on the analysis results, are relevant considerations when justifying a distribution.
24 Distribution specification can be highly subjective and uncertain when data are limited.

25
26 Inputs with substantial uncertainty about the probability distribution or uncertainty representation
27 may be candidates for future sensitivity studies to understand the impact of the chosen
28 distribution on analysis results.

29 Section 4.2.1 contains more information about fitting probability distributions to data.

30 *Bounding input distributions.* Input bounds are the upper and lower truncation points defining the
31 physical range of the input. In PFM applications, uncertain inputs are often bounded within a
32 known range. Probability distributions that place nonzero likelihood only within this range can be
33 used to prevent the sampling algorithm from selecting input values that are undesirable,
34 nonphysical, or both. Section 4.2.1 discusses methods for specifying bounded probability
35 distributions.

36
37 Inputs with substantial uncertainty about the ranges may be candidates for future sensitivity
38 studies.

39
40 *Accounting for correlation in model inputs.* In a PFM analysis, some uncertain input variables
41 may be statistically dependent (i.e., correlated). Accounting for the dependence between inputs
42 often ensures a physically possible input set (i.e., ensures that physical laws are preserved).

43 Section 4.2.2 contains more information on dependent inputs.

44 **3.3 Step 3: Estimation of Quantity of Interest and Associated Uncertainty**

45 The third step of a PFM analysis is propagating input uncertainty established in Step 2 through
46 the model to provide a converged estimate of the QoI and characterize its uncertainty. The QoI

1 uncertainty characterized in this step includes uncertainty induced by input uncertainty and
2 sampling uncertainty.

3
4 The goal is to estimate the QoI and its uncertainty with sufficient sampling precision
5 (i.e., achieve converged model results). This step includes four key actions:

- 6 (1) Select a sampling scheme for sampling uncertain model inputs.
- 7 (2) Assess sampling uncertainty.
- 8 (3) Conduct SA to determine input uncertainty importance.
- 9 (4) Conduct output uncertainty analysis.

10 These actions are iterative. First, a sampling scheme is selected and used to estimate the QoI.
11 The second action uses the sampling scheme to estimate the sampling uncertainty in the QoI
12 and determines whether the estimate has converged. The third action uses SAs to identify the
13 input uncertainties that drive the problem. SAs help to better understand the input-output
14 relationship. Results from the second and third actions can be used as a basis to update a
15 sampling scheme to improve convergence. Once a converged solution is found, the fourth
16 action provides a final estimate of the QoI and associated uncertainty.

17 **3.3.1 Step 3: Action 1—Select a Sampling Scheme for Sampling Uncertain Model** 18 **Inputs**

19 **Purpose:** The purpose of this step is to select a method for propagating uncertainty in the
20 model inputs through the model to estimate the QoI and the associated uncertainty.

21
22 **Description:** This action involves selecting a sampling scheme and using it to estimate the QoI
23 and its uncertainty. While many PFM analyses will rely on Monte Carlo sampling methods to
24 estimate a QoI, nonsampling based methods are also available and may be appropriate in some
25 applications.

26
27 *Nonsampling approaches.* Reliability methods, such as the first-order reliability methods
28 (FORM) and second-order reliability methods (SORM), use gradient-based methods to calculate
29 failure probabilities (Section 4.3.4). These methods work best when the model output is
30 sufficiently smooth and differentiable. In such conditions, they can estimate low probabilities
31 (i.e., 1×10^{-4} probability or less) with greater accuracy and fewer realizations than with Monte
32 Carlo sampling methods. These derivative methods are limited by the fact that calculating
33 second-order derivatives can quickly become impracticable as the number of uncertain inputs
34 increases beyond 15 or 20.

35
36 *Sampling approaches.* PFM analyses often use Monte Carlo methods to propagate input
37 uncertainty through the model. Selecting a sampling scheme includes specifying the following:

- 38 • sampling method
- 39 • sample size
- 40 • random seed
- 41 • method for sampling aleatory and epistemic uncertainties, if relevant

42 *Sampling methods.* Inputs can be sampled in different ways. The simplest form of Monte Carlo
43 sampling is simple random sampling (SRS), described in Section 4.3.1. SRS is easy to
44 implement but is often not the most statistically efficient method. Relative to SRS, other
45 sampling schemes can produce more precise estimates of a QoI with the same number of

1 model realizations. When models are computationally expensive or the QoI is a rare probability,
2 or both, more targeted sampling methods can be implemented to decrease the number of
3 realizations required for model convergence. Examples of targeted sampling methods include
4 Latin hypercube sampling (LHS) (Section 4.3.2), importance sampling (Section 4.3.3), and
5 adaptive sampling.
6

7 Importance sampling (Section 4.3.3) is a common sampling method for oversampling important
8 regions of the input space to reduce the sampling uncertainty of QoI estimates. When
9 estimating rare probabilities, the regions of the input space where failures are more likely are
10 oversampled to estimate the probability with less sampling uncertainty (making importance
11 sampling particularly relevant for PFM applications targeting adverse event likelihoods). To
12 implement importance sampling, the analyst selects variables on which the technique is to be
13 applied and their respective importance distributions. One general strategy is to first find the
14 failure regions that contribute to the probability of the rare event and construct the importance
15 distributions based on this information. SAs (Section 3.3.3) and subject matter expertise on
16 important inputs can inform this process. The choice of importance distributions is paramount,
17 since poor choices can lead to higher variance estimates with higher sampling uncertainty.
18 Inefficiency in importance sampling often occurs in high-dimensional problems where many
19 variables are importance sampled (Reference 3-10).
20

21 *Sample size.* The sample size is the number of realizations at different input settings (i.e., the
22 number of sets of inputs that are propagated through the model). There is a natural relationship
23 between the computational burden of the model, the sample size, and the sampling scheme.
24 Specifically, computationally inexpensive models can be run many times, resulting in large
25 sample sizes. In such cases, simple sampling schemes such as SRS are likely sufficient. If the
26 model is computationally expensive, sample sizes will be lower and more efficient sampling
27 schemes are required. Relatedly, if the QoI is a probability, the sample scheme and sample size
28 are related to the magnitude of the probabilities. As the probability gets closer to 0 or 1, more
29 samples or more efficient sampling schemes, or both, are required.
30

31 *Random seed.* Sampling-based approaches rely on random number generators to select the
32 random sample. Random seeds can be selected for a random number generator to ensure that
33 the same random sample is selected each time a set of model realizations is run so that exact
34 results can be reproduced.
35

36 *Separation of aleatory and epistemic uncertainty.* If the analysis maintains separation of aleatory
37 and epistemic uncertainty, then the input uncertainties are typically sampled using a double-loop
38 method, described in Section 4.1.1. This method first samples epistemic inputs. Then, for each
39 set of epistemic inputs, aleatory inputs are sampled numerous times to obtain a distribution of
40 model outputs over aleatory uncertainty. The double-loop structure is computationally
41 expensive, because the QoI is estimated for each set of epistemic samples. Surrogate modeling
42 (Section 4.3.10) is often used to increase the computational efficiency of the double-loop
43 method, relying on a computationally inexpensive statistical model approximation to post hoc
44 separate aleatory and epistemic uncertainty (as described in Section 4.1.1).

45 **3.3.2 Step 3: Action 2—Assess Sampling Uncertainty: Statistical Convergence** 46 **Analysis**

47 **Purpose:** The purpose of this step is to assess the statistical convergence of QoI estimates
48 from model outputs given a sampling scheme.
49

1 **Description:** PFM analyses are based on a finite number of realizations. Since the model
2 cannot be run at all points in the input space, sampling uncertainty is associated with estimating
3 a QoI. Quantifying the sampling uncertainty of QoI estimates is important to determine whether
4 the analysis conclusions might change with an improved sampling scheme. Methods to assess
5 sampling uncertainty convergence include the following:

- 6 • assessing stability of an estimate as the sample size increases
- 7 • calculating statistical sampling uncertainty metrics
- 8 • comparing replicates and assessing variation in the QoI estimates
- 9 • using surrogate modeling to estimate sampling uncertainty
- 10 • updating the sampling scheme

11 Section 4.3.5 discusses these methods in more detail.

12
13 *Assessing stability in an estimate as the sample size increases.* For a given sampling scheme,
14 the sample size can be increased iteratively until QoI estimate is sufficiently stable, suggesting
15 statistical convergence. For example, with an SRS sampling scheme, more input samples can
16 be selected to increase the sample size. Augmented LHS designs can be used to add input
17 samples to an initial LHS design. Stability in the QoI can again be measured as the sample size
18 increases. The major advantages of this approach are that it can be applied to any sampling
19 scheme and that it does not require multiple independent model realizations. However, the
20 approach does not provide a direct measure of sampling uncertainty in the QoI estimate and
21 can be rather computationally expensive.

22
23 *Calculating statistical sampling uncertainty metrics.* Statistical sampling uncertainty metrics
24 quantify the sampling uncertainty in the QoI estimate using statistical sampling theory. Methods
25 for calculating statistical sampling uncertainty metrics are specific to the sampling scheme.
26 Using an SRS sampling scheme, the standard deviation, coefficient of variation (CV), or
27 confidence interval for a QoI can be calculated directly from the sample (Section 4.3.6). If the
28 QoI for a PFM analysis is a rare probability and zero events are observed, then an upper bound
29 on the probability can be calculated using statistical metrics under an SRS scheme. However,
30 SRS is not the most efficient sampling scheme when the QoI is a rare probability.

31
32 Under sampling schemes other than SRS, statistical resampling methods, such as
33 bootstrapping, can be used to calculate statistical sampling uncertainty metrics (Section 4.3.7).
34 Resampling methods are easy to implement but can be more computationally expensive;
35 further, resampling methods can produce inaccurate estimates when the QoI is a rare
36 probability.

37
38 LHS schemes do not offer a simple analytic form for an unbiased estimate of sampling
39 uncertainties from a single sample (References 3-11, 3-12). Furthermore, bootstrapping an LHS
40 is not possible. Since LHS is more efficient (Reference 3-11), SRS uncertainty metrics applied
41 to an LHS scheme will be conservative.

42
43 *Comparing replicates and assessing variation in the QoI estimates.* Another method to assess
44 QoI convergence is to run the sampling scheme several independent times with unique random
45 number seeds. The QoI is estimated for each independent realization, and the variation across
46 the realizations is measured. Example metrics to assess convergence based on these
47 realizations include the standard deviation of the QoI across realizations, the CV (ratio of the
48 standard deviation to the mean), or a confidence interval on the QoI. These metrics can be
49 compared to the desired level of convergence for the application. The major advantages of this

1 approach are that it can be applied to any sampling scheme and that it gives a direct
2 measurement of estimate variability; however, the approach is computationally expensive. In
3 general, the number of replicates is selected to be large enough that the conclusion would not
4 change significantly if more replicates were provided.

5
6 *Using surrogate modeling to estimate sampling uncertainty.* When the model is computationally
7 expensive to run and only a small number of input samples can be propagated through the
8 model, surrogate models (Section 4.3.10) can be used to provide a computationally efficient
9 alternative to the full model. A surrogate model is a statistical approximation to the full,
10 computationally expensive model and is estimated from a set of model realizations. The
11 sampling uncertainty in the surrogate model can be propagated to sampling uncertainty
12 estimates for the QoI.

13
14 *Updating the Sampling Scheme.* If the selected sampling scheme does not provide converged
15 results, then this scheme can be updated by increasing the sample size, changing the sampling
16 method, or both.

17 **3.3.3 Step 3: Action 3—Conduct Sensitivity Analyses to Determine Input Uncertainty** 18 **Importance**

19 **Purpose:** SAs help identify problem drivers, defined as uncertain model inputs that explain
20 substantial uncertainty in the model output. Understanding problem drivers allows the analyst to
21 do the following:

- 22 • Confirm that the model is behaving as expected.
- 23 • Identify inputs whose uncertainty distribution is itself uncertain and that may need
24 refinement before final estimation of the QoI.
- 25 • Identify assumptions that are uncertain and thus may be candidates for sensitivity
26 studies (Step 4).
- 27 • Improve the accuracy of the output uncertainty analysis by reducing the dimension of the
28 input space and identifying important inputs that can be used in more targeted sampling
29 methods such as importance sampling.

30 SA plays a critical role in improving output uncertainty analysis. A common goal of a PFM
31 analysis is to accurately estimate a QoI along with its associated uncertainty. By informing the
32 final sampling scheme, SAs can improve QoI estimation. For example, SA can identify inputs
33 with a large impact on the model output; these inputs may be candidates for importance
34 sampling (Section 4.3.3) to increase the precision of QoI estimates. This action is closely tied to
35 Step 3, Action 1, which provides more detail on selecting an appropriate sampling scheme for
36 the estimation of a QoI and its uncertainty.

37
38 **Description:** In broad terms, SA focuses on identifying how the input uncertainties contribute to
39 the uncertainty in the outputs of interest. References 3-13, 3-14, 3-15, 3-16, 3-17, 3-18, and 3-
40 19 are some of the sources that describe SA techniques and examples. The discussion below
41 addresses the following:

- 42 • the types of SA
- 43 • forward propagation of uncertainty for SA

- 1 • the stages of SA
- 2 • modeling nonlinearities and interactions in SA
- 3 • SA for submodels
- 4 • uncertainties in SA

5 *Types of SA.* There are two general types of SA:

- 6 (1) *Global SA* is the process of decomposing variance in the model output according to the
7 model inputs (see Section 4.3.8).
- 8 (2) *Local SA* is the process of determining how changes to uncertain inputs affect outputs
9 with respect to a reference point in the input domain (see Section 4.3.9).

10 *Forward propagation of uncertainty for SA.* SA is performed after an initial set of uncertain
11 inputs has been propagated through the model, resulting in a distribution of model outputs. SA
12 is often conducted on an initial set of model realizations, with uncertain inputs sampled using a
13 standard Monte Carlo-based sampling scheme with broad coverage of the input space, such
14 that model input-output relationships can be discerned from the sample. The number of model
15 realizations needed depends on the goals of the SA and the computational burden of the model.
16 For example, if the goal of global SA is to understand how inputs vary with the output to select
17 the number of model realizations, the analyst can consider the complexity of the input-output
18 relationship and the number of uncertain model inputs. Local SA typically requires fewer model
19 realizations. After model results are obtained from forward propagation of uncertainty, the
20 analyst can proceed with the two stages of SA described below.

21
22 *Stages of SA.* Typically, SAs have two stages:

- 23 (1) *Exploratory data analysis* involves graphically exploring input-output relationships using
24 scatter plots and calculating local SA metrics, as needed. The SA results can present
25 scatter plots for important inputs. Reference 3-20 describes formal procedures for the
26 analysis of scatterplots.
- 27
28 (2) *Global sensitivity metrics estimation* involves the estimation of the proportion of variance
29 in the model output explained by each model input (first-order sensitivity index) and its
30 interactions with other inputs (total-order sensitivity index).

31 In practice, SA is an iterative process, and these two stages may repeat multiple times. For
32 example, given a large number of inputs and complexities in the input-output relationships,
33 selecting the correct visualizations and interpreting them can be difficult. Estimation of the global
34 sensitivity metrics in the second stage can help to identify the important inputs to visualize.
35 These visualizations can inform results of the global SA.

36
37 *Model nonlinearities and interactions in SA.* PFM applications often involve systems of linked
38 models with complex relationships. SAs allow for the identification and quantification of the
39 input-output relationship, including nonlinearities in the input-output relationship and interactions
40 between model inputs.

41
42 Global SA is a commonly used tool for summarizing input importance in PFM studies and
43 identifying the effects of nonlinearities and interactions. Local SAs identify sensitivities within a
44 small neighborhood around a point of interest and therefore do not identify nonlinearities and
45 interactions; local SA is informative if the goal is to understand local variations (see

1 Section 4.3.9 for more detail). Global and local SA are often used together in the same PFM
2 analysis at different iterations of the SA.

3 *SA for submodels.* Since PFM applications often involve systems of linked models, it may be
4 appropriate to conduct SA on specific submodels (in addition to the full PFM model) as some
5 dominant submodels may hide the results and impacts of other submodels. For example, to
6 investigate the impact of active degradation mechanisms on the probability of leakage or
7 rupture, it may be appropriate for the analysis to exclude fatigue damage. As another example,
8 it may be prudent to identify inputs impacting crack growth before a full SA determining
9 important inputs for rupture.

10

11 *Uncertainties in SA.* Understanding aspects of the model and input uncertainty characterization
12 informs how to conduct SA, as shown in the following examples:

13 • Model approximations for computationally expensive models. Estimating sensitivity
14 metrics is computationally expensive, often requiring many model realizations. As a
15 solution, model approximations or surrogates (Section 4.3.10) are often used in SA as a
16 computationally practical approximation to the full model. Sufficiently flexible model
17 surrogates allow for nonlinearities and interactions between inputs. If the model
18 approximation contains substantial uncertainty, then multiple different model
19 approximation methods can be compared to assess robustness of the SA results to the
20 model approximation method.

21 • High-dimensional inputs. Building an accurate model approximation requires more
22 model realizations when the input space is high dimensional. Without enough
23 realizations, true input-output relationships may not be identified.

24 • Continuous versus binary or discrete outputs for SA. Binary or discrete outputs (such as
25 failure events) inherently contain less statistical information than continuous outputs.
26 More realizations will be needed to identify important model inputs impacting a binary
27 indicator variable than for a continuous model output. An alternative is to identify
28 continuous responses associated with the binary event for SAs, insofar as there is a
29 clear, justifiable connection between the binary event and the continuous variable. For
30 example, instead of conducting SA on the binary indicator for rupture, the analyst could
31 use crack length as the output for SA.

32 • Separation of aleatory and epistemic uncertainty. If the separation of uncertainty types is
33 maintained, SA is conducted for both uncertain aleatory and epistemic inputs. The SA
34 can be run over all uncertainties to determine which inputs have the largest impact on
35 the outputs of interest. Additional SAs can be conducted for aleatory and epistemic
36 inputs separately to identify the impacts of irreducible and reducible uncertainties,
37 respectively.

38 **3.3.4 Step 3: Action 4—Conduct Output Uncertainty Analysis**

39 **Purpose:** The purpose of this step is to provide a final estimate, with associated uncertainty, of
40 the QoI and to visualize results.

41

42 **Description:** A summary of the QoI results may include the following:

43 • a best estimate of the QoI

- 1 • an estimate of uncertainty in the QoI
- 2 • a graphical display of the QoI estimate and uncertainty

3 *Best estimate of a QoI.* The definition of the best estimate of a QoI will depend on the
4 application. When the QoI is uncertain, the best estimate is often quantified using either the
5 mean or median of the QoI distribution. The mean is the arithmetic average over the QoI
6 distribution, and the median is the 50th percentile of this distribution.

7
8 *Estimate of uncertainty in the QoI.* When estimating and visualizing uncertainty in a QoI
9 estimate, it is critical to be clear about the type of uncertainty being summarized. QoI
10 uncertainty can refer to different types of uncertainty, depending on the relationship between the
11 model output and the QoI. Example types of uncertainty include the following:

- 12 • Input uncertainty. If the QoI is a model output, QoI uncertainty may refer to uncertainty in
13 the QoI due to uncertain inputs. A best estimate of the QoI is the mean or median of the
14 QoI over the input space, and the uncertainty in the QoI refers to the distribution of the
15 QoI over uncertain inputs.
- 16 • Sampling uncertainty (also called aleatory uncertainty). QoI uncertainty may also arise
17 due to a limited number of model realizations resulting in uncertain QoI estimates. When
18 convergence analyses (Section 3.3.2) suggest sampling uncertainty is negligible, then
19 visualizing sampling uncertainty will not be necessary. If the sampling uncertainty is not
20 sufficiently small based on convergence analysis results, then this sampling uncertainty
21 can be measured and presented as a source of QoI uncertainty.
- 22 • Epistemic (lack of knowledge) uncertainty. When aleatory and epistemic uncertainty are
23 separated, the QoI is typically calculated for each epistemic sample. Epistemic
24 uncertainty in the QoI measures how the QoI varies due to knowledge uncertainty. A
25 best estimate of the QoI is the mean or median QoI estimate over all epistemic samples.
- 26 • Uncertainty in the QoI results in a distribution of QoI estimates. This uncertainty can be
27 summarized using percentiles of the uncertainty distribution; measures such as variance
28 and standard deviation can also provide useful summaries of QoI uncertainty.

29 If the QoI is a failure probability calculated from uncertain model outputs, then the QoI already
30 incorporates uncertainty in the model inputs. In this case, if aleatory and epistemic uncertainty
31 are not separated and sampling uncertainty is negligible (i.e., a high degree of statistical
32 convergence has been achieved), then there may be no need to present a measure of
33 uncertainty about the failure probability estimate.

34
35 *Graphical display of the QoI estimate and uncertainty.* Graphical displays of the best estimate
36 and uncertainty in the QoI can be used to communicate the results of an uncertainty analysis.
37 The form of the graphical display will depend on the types of uncertainty being visualized and
38 whether the QoI is a function of time or a single scalar. The best approach to visualizing results
39 is application specific. Section 4.3.11 provides more details on output uncertainty analysis.

40 **3.4 Step 4: Sensitivity Studies to Assess the Credibility of Modeling**

41 **Assumptions**

42 The fourth step in a PFM analysis is conducting sensitivity studies, defined as additional
43 analyses conducted under different, yet plausible, assumptions. The purpose of sensitivity

1 studies is to challenge uncertain analysis assumptions that could substantively change the
2 analysis results. Sensitivity studies involve two key actions:

- 3 (1) Determine a set of sensitivity studies.
- 4 (2) Conduct sensitivity studies and present results.

5 **3.4.1 Step 4: Action 1—Determine a Set of Sensitivity Studies**

6 **Purpose:** The purpose of this action is to identify important assumptions that merit further
7 scrutiny to understand what might happen if these assumptions were changed. For example, in
8 the study of a plant, the distribution of a specific input could have been calibrated using
9 information from a global set of similar but different plants. This calibration raises the question of
10 what might be different about the distribution for the individual plant and how that would change
11 the conclusions of the analysis.

12
13 **Description:** Given the complexity of PFM analyses, it is not possible to enumerate all plausible
14 changes in the assumptions. Instead, to evaluate whether a sensitivity study is needed for a
15 specific assumption, two criteria are evaluated:

- 16 (1) Plausible alternate assumptions can be identified.
- 17 (2) Changes to the assumption in question can substantively impact the calculated QoI.

18 The specific number of sensitivity studies will depend on the application, but the goal is to
19 conduct enough studies such that there is a sufficiently low chance that the results of the
20 analysis depend heavily on unverifiable or uncertain assumptions.

21
22 Uncertain analysis assumptions can often be classified as either modeling assumptions or input
23 parameter specification assumptions. Modeling assumptions include any assumptions in the
24 computational modeling framework, while input parameter specification assumptions refer to
25 any assumptions made when specifying the values of the input parameters to the PFM model.
26 Common types of sensitivity studies include considering changes in the results if the following
27 occurs:

- 28 • A plausible alternative model is used.
- 29 • A different probability distribution for an uncertain input (or several uncertain inputs) is
30 used.
- 31 • The categorization of an input as aleatory or epistemic is changed.

32 Reference 3-21 provides guidance on selecting sensitivity studies, which this report reviews in
33 Section 4.4.

34 **3.4.2 Step 4: Action 2—Conduct Sensitivity Studies and Present Results**

35 **Purpose:** The purpose of this action is to perform the sensitivity studies.

36

37 **Description:** Sensitivity studies can take on many different forms, and there is no prescriptive
38 method for conducting sensitivity studies. However, they will all include some common
39 elements:

- 1 • a reference realization (or baseline case) with a documentation of the QoI
- 2 • one or several modified realizations illustrating the concept that needs to be represented
- 3 • a comparison between the reference realization and the modified realization(s)
- 4 • a comparison criterion to decide whether the change is significant
- 5 • a conclusion, including potential consequences

6 **3.5 Step 5: Draw Conclusions from Analysis Results**

7 The fifth step in a PFM analysis is to draw conclusions using the results of Steps 1–4. This step
8 includes two key actions:

- 9 (1) Interpret analysis results.
- 10 (2) Iterate on the analysis process to refine model results.

11 **3.5.1 Step 5: Action 1—Interpret Analysis Results**

12 **Purpose.** The purpose of this action is to synthesize the information gathered in Steps 1–4 and
13 draw conclusions from this information.

14
15 **Description.** In an ideal situation, PFM analysis results can be compared directly to acceptance
16 criteria to make a regulatory decision. In practice, determinations about whether acceptance
17 criteria are met are typically not made based on a single PFM calculation or analysis but rather
18 based on a set of analyses that are compiled into an overall evidence package. Information
19 about the analysis results, scope, and limitations must be considered when drawing final
20 conclusions, considering all elements of the PFM analyses described above in Steps 1–4.
21 Subsequently, drawing final conclusions based on the analysis requires substantial expert
22 judgment to synthesize all information together to make actionable guidelines.

23 **3.5.2 Step 5: Action 2—Iterate on the Analysis Process to Refine Model Results**

24 **Purpose.** The purpose of this action is to determine whether additional analyses are required to
25 draw informative conclusions from the modeling.

26
27 **Description.** If analysis results are inconclusive concerning whether the acceptance criteria are
28 met, then the analyst can consider additional refinements to the analysis to provide the required
29 additional information. For example, the analyst can consider the following:

- 30 • changing or clarifying aspects of the PFM code (Section 3.1.3)
- 31 • refining the input uncertainty distributions (Section 3.2.2)
- 32 • choosing a different sampling scheme or increasing the number of model realizations
33 (Section 3.3.1)
- 34 • adding more sensitivity studies to address existing limitations (Section 3.3.3)

35 PFM analyses are typically iterative in nature, such that initial modeling results inform future
36 analyses. The iterative process continues until the analyst has sufficient information to draw
37 clear conclusions about whether the acceptance criteria are met for the application.

1 **3.6 References**

2 3-1. *U.S. Code of Federal Regulations*, “Domestic licensing of production and utilization
3 facilities,” Part 50, Chapter I, Title 10, “Energy.”

4 3-2. Oberkampf, W.L., Pilch, M., and Trucano, T.G., “Predictive Capability Maturity Model for
5 Computational Modeling and Simulation,” SAND2007-5948, Sandia National
6 Laboratories, 2007.

7 3-3. Oberkampf, W.L., and Roy, C.J., *Verification and Validation in Scientific Computing*,
8 Cambridge University Press, United Kingdom, 2010.

9 3-4. Roache, P.J., *Verification and Validation in Computational Science and Engineering*,
10 Hermosa Publishers, Albuquerque, NM, 1998.

11 3-5. American Society of Mechanical Engineers, “Guide for Verification and Validation in
12 Computational Solid Mechanics,” ASME V&V 10, 2006.

13 3-6. American Society of Mechanical Engineers, “Standard for Verification and Validation in
14 Computational Fluid Dynamics and Heat Transfer,” ASME V&V 20, 2009.

15 3-7. Helton, J.C., Johnson, J.D., and Oberkampf, W.L., “An exploration of alternative
16 approaches to the representation of uncertainty in model predictions,” *Reliability
17 Engineering and System Safety*, 85, 39–71, 2004.

18 3-8. Helton, J.C., Johnson, J.D., Oberkampf, W.L., and Sallaberry, C.J., “Representation of
19 analysis results involving aleatory and epistemic uncertainty,” *International Journal of
20 General Systems*, 39, 605–646, 2010.

21 3-9. Xing, J., and Morrow, S., “White Paper: Practical Insights and Lessons Learned on
22 Implementing Expert Elicitation,” U.S. Nuclear Regulatory Commission, Office of Nuclear
23 Regulatory Research, 2016 (ML16287A734).

24 3-10. Au, S.K., and Beck, J.L., “Important sampling in high dimensions,” *Structural
25 Safety*, 25(2), 139–163, 2003.

26 3-11. Stein, M., “Large Sample Properties using Latin Hypercube Sampling,” *Technometrics*,
27 239–245, 1987.

28 3-12. Helton, J.C., and Davis, F.J., “Latin Hypercube Sampling and the propagation of
29 uncertainty in analyses of computer systems,” *Reliability Engineering and System
30 Safety*, 23–69, 2003.

31 3-13. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., and
32 Tarantola, S., *Global Sensitivity Analysis. The Primer*, John Wiley & Sons, Ltd., 2008.

33 3-14. Helton, J.C., “Uncertainty and sensitivity analysis techniques for use in performance
34 assessment for radioactive waste disposal,” *Reliability Engineering and System Safety*,
35 42, 327–367, 1993.

- 1 3-15. Helton, J.C., Johnson, J.D., Sallaberry, C.J., and Storlie, C.B., "Survey of
2 sampling-based methods for uncertainty and sensitivity analysis," *Reliability Engineering
3 and System Safety*, 91, 1175–1209, 2006.
- 4 3-16. Saltelli, A., Ratto, M., Tarantola, S., and Campolongo, F., "Sensitivity analysis for
5 chemical models," *Chemical Reviews*, 105, 2811–2828, 2005.
- 6 3-17. Storlie, C.B., Swiler, L.P., Helton, J.C., and Sallaberry, C.J., "Implementation and
7 evaluation of nonparametric regression procedures for sensitivity analysis of
8 computationally demanding models," *Reliability Engineering and System Safety*, 94,
9 1735–1763, 2009.
- 10 3-18. Wei, P., Lu, Z., and Song, S., "Variable importance analysis: A comprehensive review,"
11 *Reliability Engineering and System Safety*, 142, 399–432, 2015.
- 12 3-19. Borgonovo, E., and Plischke, E., "Sensitivity analysis: A review of recent advances,"
13 *European Journal of Operational Research*, 248, 869–887, 2016.
- 14 3-20. Kleignen, J.P.C., and Helton, J.C., "Statistical analyses of scatterplots to identify
15 important factors in large-scale simulations," "1: Review and comparison of techniques,"
16 *Reliability Engineering and System Safety*, 65, 147–185, 1999.
- 17 3-21. EricksonKirk, M.T., Dickson, T., Mintz, T., and Simonen, F., "Sensitivity Studies of the
18 Probabilistic Fracture Mechanics Model Used in FAVOR," NUREG-1808, U.S. Nuclear
19 Regulatory Commission, 2010 (ML061580349).

20

4 USEFUL METHODS FOR ESTABLISHING CONFIDENCE IN PROBABILISTIC FRACTURE MECHANICS ANALYSIS

This section details a concise review of analysis methodologies, including notional examples for context that are linked directly to an action introduced in Section 3. While this is not a comprehensive list of acceptable methodologies, this section can be used by applicants who are seeking to further understand the theoretical underpinnings of the processes that are used to establish the credibility of a PFM analysis. For example, Section 4.1 provides technical detail that could be used to develop the technical basis for the action defined in Section 3.1. Each section introduces a concept/method and provides the following information about it:

- **What is it?**—Gives a high-level description of the concept/method.
- **How to use?**—Provides general details on how the concept/method is used, including specific steps or an algorithm where appropriate.
- **When/Why?**—Discusses the PFM context in which the concept/method is used and maps this use to the process described in Section 3.
- **Technical details**—Describes technical details and complexities that are important to the use/implementation/interpretation of the method in the PFM context.
- **References**—Lists references that provide further technical details.

4.1 Useful Methods for Translating Regulatory Requirements into an Analysis Plan

4.1.1 Separation of Aleatory and Epistemic Uncertainty

When constructing an analysis plan, one aspect to consider is the treatment of uncertainty—namely, will uncertainty be treated probabilistically, and, if so, will different types of uncertainty be distinguished? Separating types of uncertainty can be necessary when there is a need to quantify the uncertainty on a statistical QoI (a frequency or probability) or to separate inherent variability from lack-of-knowledge uncertainty. Such separation generally provides additional insights on the magnitude of the uncertainties and on whether they can be reduced. However, separation of uncertainties also comes with increased computational cost and analysis effort, and the decision to maintain the separation influences many steps of the subsequent analysis workflow. As a result, this tradeoff decision needs to be considered at an early stage of the analysis planning.

Specifically, the strategy for handling uncertainty may vary for different types of analysis questions. If the analysis objective is to compute a single best estimate event probability, it is likely sufficient to consider all sources of uncertainty together to arrive at this probability. However, this approach can obscure information. Separating types of uncertainty instead of considering all sources of uncertainty together can lead to more interpretable analysis. For example, rather than computing a single best estimate probability, the analyst may want to understand the confidence in the computed frequency of an event given the current state of knowledge. Some elements of this knowledge uncertainty may be reducible, potentially improving confidence in the frequency estimate and increasing the precision of the analysis

1 results. As described in this section, these reducible sources of uncertainty (referred to as
2 epistemic uncertainty) can be treated separately to maintain this information for the
3 communication of results and decisionmaking about additional activities to conduct.

4 4.1.1.1 *What Is It?*

5 Two primary types of uncertainty sources are often considered in risk analysis (References 4-1,
6 4-2, 4-3, 4-4, 4-5, 4-6, 4-7, 4-8):

- 7 (1) **Aleatory uncertainty** is defined as “uncertainty based on the randomness of the nature
8 of the events or phenomena that cannot be reduced by increasing the analyst’s
9 knowledge of the systems being modeled” (Reference 4-9). Aleatory uncertainty
10 represents the (perceived) randomness in the modeled system that cannot be reduced.
11 Aleatory uncertainties reflect natural, intrinsic, or stochastic variability.
- 12 (2) **Epistemic uncertainty** is defined as “the uncertainty related to the lack of knowledge or
13 confidence about the system or model and is also known as state-of-knowledge
14 uncertainty” (Reference 4-9). Epistemic uncertainty represents the lack-of-knowledge
15 uncertainty in the modeled system that can be reduced.

16 Historical PFM analyses of nuclear power plant structures either (1) do not distinguish between
17 types of uncertainty (References 4-10, 4-11, 4-12, 4-13) or (2) treat the uncertainty as either
18 aleatory or epistemic (References 4-14, 4-15).

19 4.1.1.2 *How to Use?*

20 If an analysis separates aleatory and epistemic uncertainty, then it requires additional effort to
21 separate uncertainty types and iterate over epistemic samples in a “double-loop” sampling
22 algorithm, as described below. This involves the following three steps:

- 23 (1) **Classify types of uncertainty for an application.** The first step in separating types of
24 uncertainty is classifying input variables as aleatory or epistemic. The specific PFM
25 application typically drives classification choices. If it is unclear how to classify an
26 uncertainty, it may be worth considering a sensitivity study (Section 4.4) to understand
27 the impact of the classification.
28
- 29 (2) **Determine how to represent uncertainty.** After classifying uncertainty types, the next
30 step is to determine how to represent the different types of uncertainty. PFM analysis
31 typically represents uncertainties using probability distributions, though other options are
32 possible (see Section 4.1.1.4).
33
- 34 (3) **Propagate uncertainty while maintaining separation of types.** Given a model output,
35 a QoI, and uncertainties related to the model input parameters, the next step is
36 propagating both types of uncertainty. For sampling-based uncertainty propagation,
37 separation of aleatory and epistemic uncertainty is maintained using a double-loop
38 (i.e., nested loop) framework. The following steps can be applied to propagate input
39 uncertainty:
40
- 41 – Epistemic variables are sampled in an outer loop.
 - 42 – For each epistemic sample, aleatory variables are sampled in an inner loop.

1 – Qols are calculated for each epistemic sample (calculated over all aleatory
2 samples), generating an epistemic distribution of Qols.

3 The sections below provide more information about the double-loop procedure.

4 4.1.1.3 *When/Why?*

5 The risk community has made the distinction between the inherent risk (aleatory) and the
6 uncertainty due to lack of knowledge (epistemic). The purpose of this distinction is to
7 acknowledge there will always be a risk given a specific situation, and the consequences can
8 lead to different interpretations in terms of decisionmaking. The analyst can choose whether to
9 separate types of uncertainty. The decision to separate uncertainty types typically depends on
10 several factors:

- 11 • **Computational feasibility.** In practice, maintaining separation during uncertainty
12 propagation (for example, through Monte Carlo sampling) can be computationally
13 challenging, due to the need to construct a “double-loop” sampling scheme.
14 Section 4.1.1.4 contains more information and suggestions for efficiently implementing
15 the double-loop scheme.
- 16
- 17 • **Conceptual interpretation of results.** The interpretation of the results of a PFM
18 analysis changes depending on whether the separation of uncertainty types is
19 maintained (Reference 0). Section 4.1.1.4 contains more information.
- 20
- 21 • **Strength of technical basis.** Ultimately, the separation of uncertainties can help to
22 make a stronger, more comprehensive case and help the analyst understand what
23 needs to be done to improve the accuracy of the answer.

24 4.1.1.4 *Technical Details*

25 **Representing epistemic uncertainty.** In many risk analysis applications, it can be difficult to
26 specify probability distributions on epistemic uncertainties because, by definition, these
27 uncertainties arise due to lack of knowledge. While probabilistic representation of epistemic
28 uncertainty will be sufficient for most PFM applications, nonprobabilistic representations may be
29 appropriate in certain instances (Reference 4-3, 4-16, 4-17). For example, sensitivity studies
30 (Section 4.4) conducted at deterministic (i.e., fixed) values of the epistemic inputs can inform
31 about a “worst case scenario.”

32

33 **Computational burden of separating uncertainty.** The double-loop framework for sampling
34 typically requires a large sample size. For each epistemic sample, the aleatory sample is
35 selected to be sufficiently large for the accurate estimation of the Qol (e.g., failure frequency).
36 More sophisticated sampling schemes (Sections 4.3.2 and 4.3.3) may be needed to make
37 double-looping computationally feasible. If the model is too computationally expensive to directly
38 implement the double-loop sampling, there are two options: (1) do not separate uncertainty
39 types, or (2) build a computationally efficient surrogate model to approximate the full model.
40 Surrogate models are data-driven approximations of the physics model output across the input
41 space, as discussed in Section 4.3.10. Surrogate models introduce additional uncertainty into
42 the problem because the surrogate is itself a model approximation.

43

44 **Interpretation of results.** Maintaining separation of the two types of uncertainty facilitates
45 making statements about confidence in the frequency of an event or the probability of

1 frequency. Specifically, probability of frequency refers to an analysis that models aleatory and
2 epistemic uncertainties probabilistically and separates them in presenting the results
3 (References 4-2, 4-18). As an example, in a PFM analysis aiming to characterize the likelihood
4 of an adverse event, aleatory probabilities represent the frequency of an adverse event
5 (e.g., crack, rupture) given a set of epistemic inputs/assumptions. These frequencies will vary
6 with the set of epistemic inputs/assumptions. This variation represents the epistemic
7 uncertainty/confidence in the frequency.

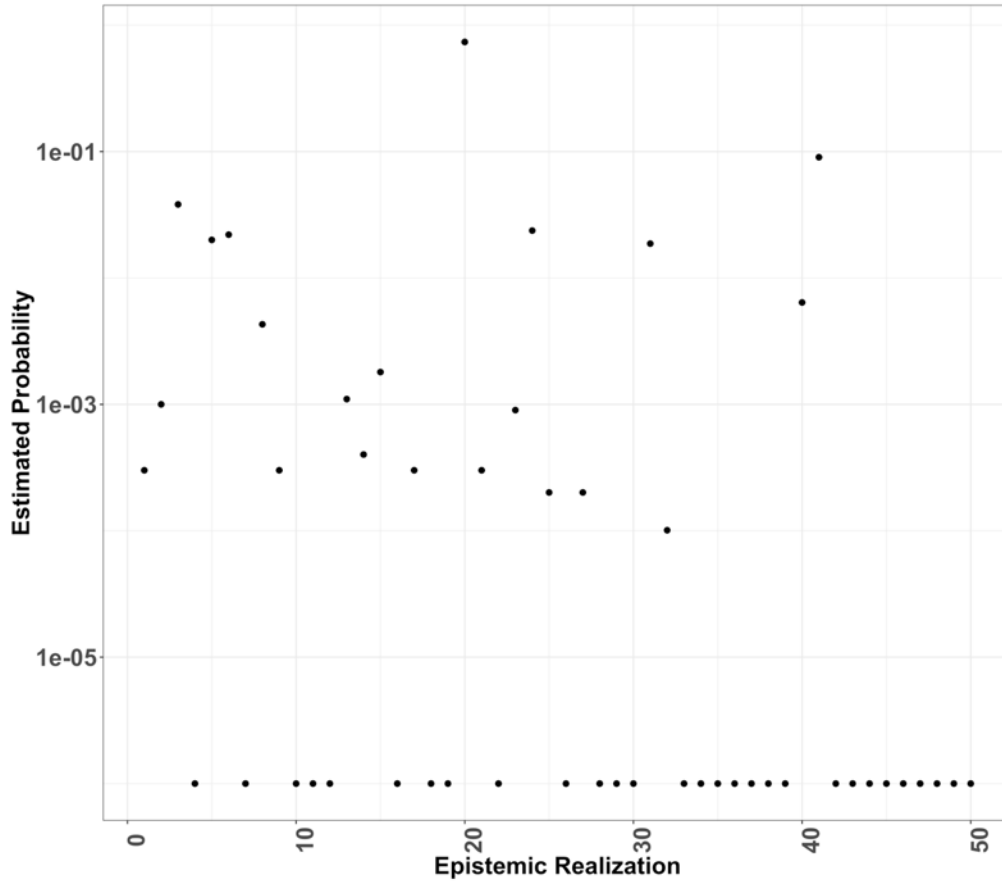
8
9 If an analysis does not distinguish between aleatory and epistemic uncertainties, frequencies of
10 an adverse event are computed over all uncertainties, and the analysis cannot quantify the
11 impact of uncertainties that arise due to lack of knowledge. The implications of such a choice
12 are explained with an example of the double-loop procedure below.

13
14 To illustrate the double-looping procedure, consider estimating the frequency of pipe rupture.
15 The model output is a binary indicator taking the value 1 if the pipe ruptured and 0 otherwise.
16 Each input is categorized as either epistemic or aleatory and is assigned a probability
17 distribution to represent its uncertainty. Then, the model is run, each time with different inputs,
18 using the double-loop algorithm to separate uncertainty:

- 19 (1) A set of epistemic variables is sampled randomly from the variables' probability
20 distributions.
- 21 (2) Fixing this set, many samples (e.g., 1×10^4) of the aleatory variables are sampled
22 randomly and the model is run, collecting the binary output for each realization.
- 23 (3) Steps 1 and 2 are repeated many times (e.g., 1×10^3). The separation of the results by
24 epistemic variable is maintained.

25 Example results appear in Figure 4-1, which shows the proportion of the 1×10^4 aleatory samples
26 that resulted in pipe rupture for the first 50 epistemic samples. For each epistemic realization,
27 this proportion is the estimated frequency of pipe rupture, given the set of epistemic variables.

28



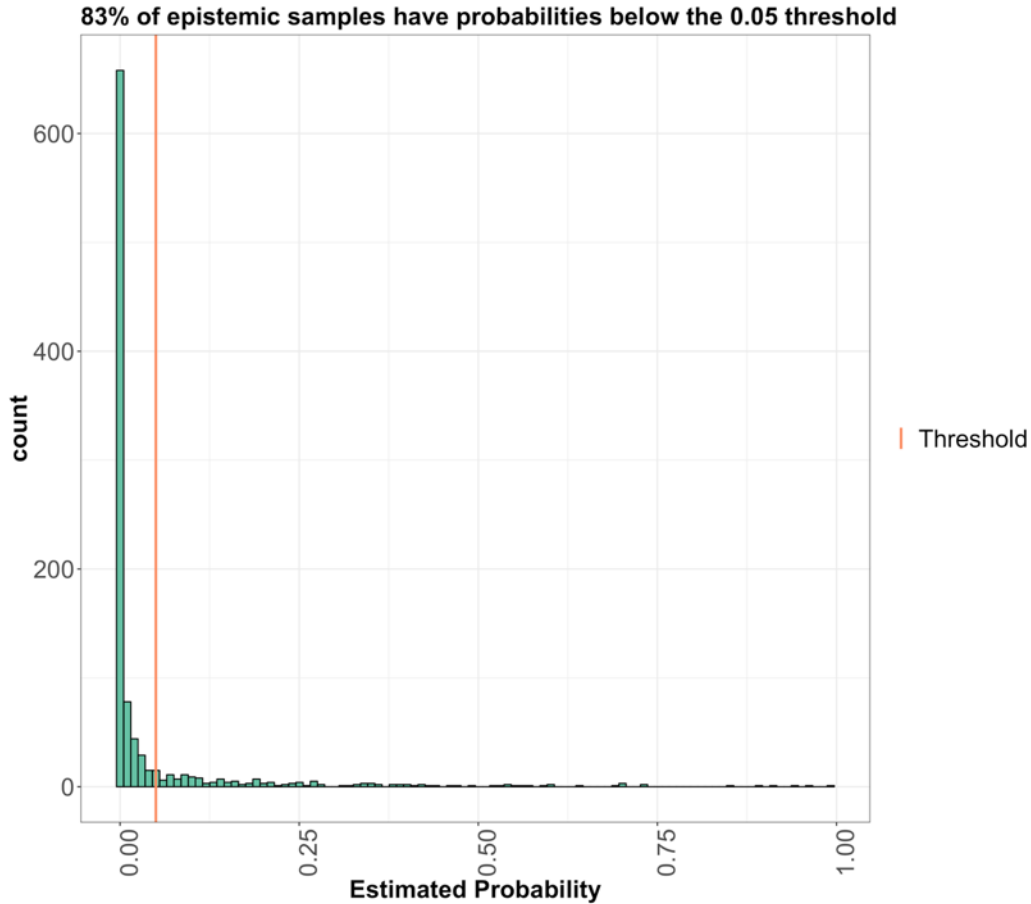
1

2 **Figure 4-1 The Estimated Probability of Pipe Rupture for the First 50 Epistemic Samples**

3

4 For many of the epistemic realizations, the estimated failure frequency is zero, meaning none of
 5 the 1×10^4 random realizations resulted in a pipe rupture. However, several of the estimates are
 6 nonzero. Across the 1×10^3 epistemic realizations, the estimated failure frequency ranges from 0
 7 to 0.99, with roughly 83 percent falling below 0.05 (the vertical line). The histogram of the
 8 estimated failure frequencies in Figure 4-2 shows this.

9



1

2 **Figure 4-2 Histogram of Estimated Probabilities Across 1,000 Epistemic Realizations**

3

4 With maintaining the separation of aleatory and epistemic uncertainties, the results can be
 5 interpreted as follows: there is roughly 83-percent confidence that the rupture probability is
 6 below 0.05. This is likely an optimistic estimate of confidence in the sense that the sampling
 7 uncertainty (i.e., finite sample size uncertainty) for each of the estimated probabilities has not
 8 been considered.

9

10 Without maintaining separation of aleatory and epistemic uncertainties, the estimate of pipe
 11 rupture probability would be the proportion over all samples. This proportion is 0.046, which is
 12 below the 0.05 threshold used above. However, such an approach mixes the likelihood of
 13 rupture (i.e., aleatory uncertainty in rupture) and the confidence associated with rupture
 14 (epistemic uncertainty of rupture). When separating, the conclusion is that roughly 17 percent of
 15 epistemic values result in rupture probabilities above 0.05. When not separating, the conclusion
 16 is the estimated probability of pipe rupture is 0.046. These are two very different conclusions.

17

1 **4.2 Methods for Model Input Uncertainty Characterization**

2 **4.2.1 Statistical Distribution Fitting**

3 *4.2.1.1 What Is It?*

4 Given a set of representative data about the input parameter, statistical distribution fitting is the
5 process of estimating the probability distribution of the input parameter using the available data.

6 *4.2.1.2 How to Use?*

7 Statistical distribution fitting has five steps:

- 8 (1) Determine relevant data.
- 9 (2) Select candidate probability distributions.
- 10 (3) Fit distributions to the data.
- 11 (4) Evaluate the fit of the distributions to the data.
- 12 (5) Select a final input distribution model.

13 Given a candidate distribution and ample data, most statistical software programs can produce
14 estimates of input distribution parameters, uncertainty in these parameters, and evaluations of
15 model fit. Important considerations for statistical distribution fitting include the following:

- 16 • How many data are available and what is the pedigree of those data?
- 17 • How much subject matter knowledge is available about the range and shape of the input
18 parameter distribution?
- 19 • How much accuracy is needed in the input distribution? (More important inputs require
20 more accuracy.)
- 21 • After distribution fitting, how much uncertainty is there in the final estimate of the input
22 distribution?

23 Reference 4-19 provides specific guidance on fitting models to input distributions. The sections
24 below provide more technical details on distribution fitting.

25 *4.2.1.3 When/Why?*

26 Probability distributions are often used to represent uncertainty in model inputs. Statistical
27 distribution fitting is used when data are available to learn about the form of the input
28 distribution.

29
30 The chosen input distribution can impact the PFM results. Expert judgment can inform the
31 distribution, especially when limited or inexact data are available. Additionally, sensitivity studies
32 (Section 4.4) on important input distributions may be needed to assess the impact of
33 assumptions made in the distribution fitting process. When data are not available to estimate
34 probability distributions, expert elicitation can be used (Reference 4-20).

1 4.2.1.4 *Technical Details*

2 This section discusses the five steps in statistical distribution fitting in more detail.

3 **(1) Determine relevant data.**

4 **Data quality.** The amount and pedigree of the source data are important considerations when
5 determining an input distribution. In practice, cost and time limit data quality. Data quality
6 considerations include the following:

- 7 • limited data/small sample size (i.e., the sample is too small to estimate the input
8 distribution with sufficient accuracy)
- 9 • data relevance (i.e., not all data points are direct measures of the outcome of interest)
- 10 • data uncertainty (i.e., individual data points can contain uncertainty due to measurement
11 error)

12 The minimum number of data points needed for a suitable fit is subjective and context specific,
13 but smaller sample sizes lead to larger uncertainty in the best fitting input probability distribution.
14 Additionally, very small sample sizes do not allow for data-driven statistical distribution fitting.
15 Expert judgment about the input and its impact on the results can provide additional insight into
16 the process of choosing the input probability distribution.

17 **(2) Select candidate probability distributions.**

18 **Distribution models.** There is a large set of possible distribution models for input parameters,
19 but, in most cases, simple parametric forms for inputs are used. Common choices include the
20 normal, truncated normal, lognormal, uniform, triangular, and Weibull distributions. Other
21 distributions can be selected, based on their appropriateness for the application at hand.

22

23 Considerations in choosing probability distributions include the following:

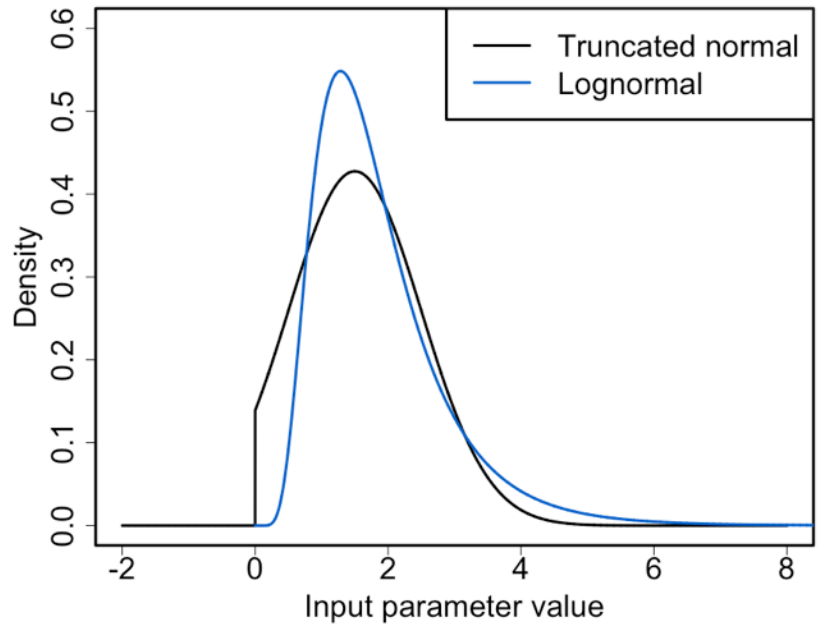
- 24 • range of values the input takes
- 25 • tail behavior and overall shape of the distribution

26 **Input Ranges.** To specify an input distribution, it is important to consider the range of inputs.
27 Specifically, the range of a distribution should be broad enough to include all possibilities but
28 narrow enough to exclude unrealistic or nonphysical values.

29

30 There are two options for bounding the range of an input: (1) select a probability distribution
31 whose range is consistent with the known range of the data, or (2) use a truncated form of a
32 probability distribution. For example, suppose we know an input parameter, such as material
33 strength, is always greater than 0. Then, we can use a distributional model that puts
34 0 probability mass on values less than 0, such as the lognormal, uniform, or Weibull model.
35 Alternatively, we could use a truncated normal model that truncates the normal distribution such
36 that the input is always greater than 0. Figure 4-3 depicts these two options.

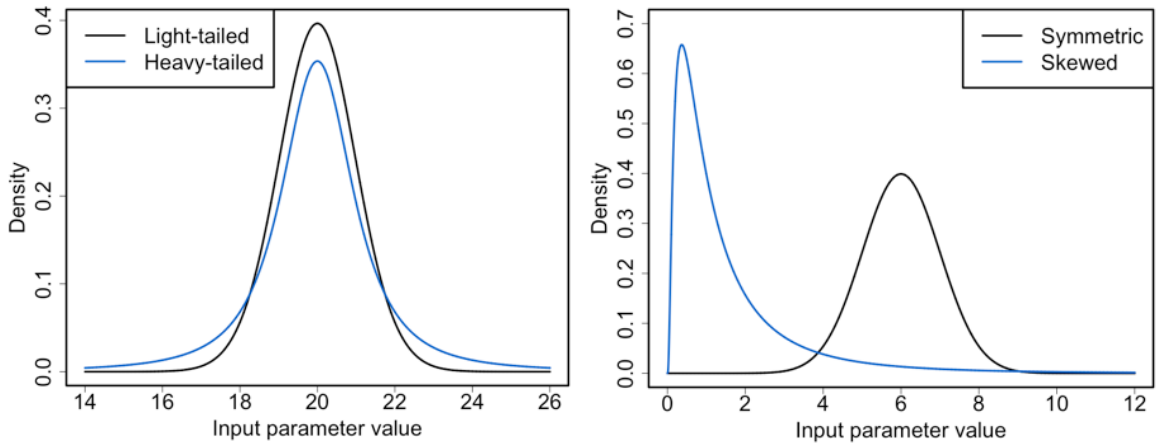
37



1
2
3
4
5
6
7
8
9

Figure 4-3 Distributions with Nonnegative Input Parameters

Tail behavior and shape. Examples of heavy-tailed and skewed distributions appear in Figure 4-4. Determining the tail behavior and shape requires large sample sizes or expert judgment. Given that the tails of distributions often drive structural failures, it is important to investigate the confidence in the underlying probability distributional form and whether the specified distribution fits the underlying data well in the tails.



10
11
12
13
14
15
16

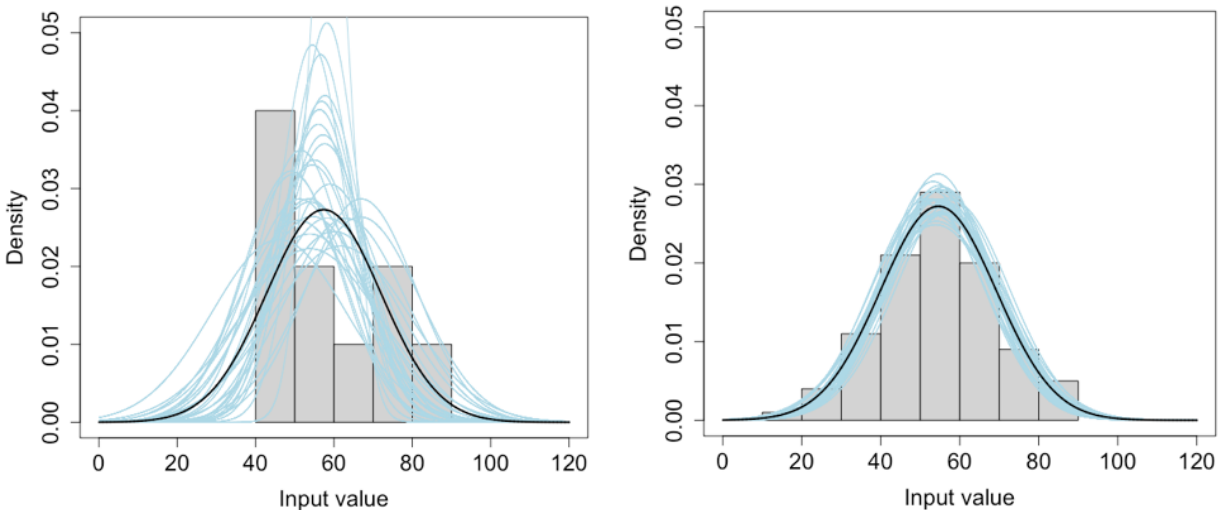
Figure 4-4 Heavy-Tailed and Skewed Distributions

Data transformations. Inputs can be modeled on different scales. A common data transformation is the natural logarithm, where inputs are modeled on the natural log scale, rather than the absolute scale, of the data. This transformation is particularly useful for skewed, positive inputs.

1 **(3) Fit the distributions.**

2 **Parameter estimation.** Given a candidate probability distribution and a set of data, statistical
3 inference can be used to estimate the parameters of that distribution. Most statistical software
4 programs (e.g., R, Python, Minitab, MATLAB with the appropriate toolbox, Easyfit) can estimate
5 distribution parameters, along with uncertainty in those parameters. These parameters are
6 typically estimated using statistical inference techniques such as maximum likelihood
7 estimation, Bayesian inference, or method of moments. If the data cannot be modeled well
8 using a known probability distribution, then nonparametric approaches can be applied.
9

10 **Input uncertainty.** The estimated distribution parameters contain sampling uncertainty,
11 because they were estimated based on a finite sample of data (Figure 4-5). In the figure, the
12 grey bars are a histogram of the data, with best fit normal distribution shown as the black line.
13 The blue lines are sampling uncertainty in the distributional fit due to the limited sample size
14 when $n=10$ (left) and $n=100$ (right).
15



16

17 **Figure 4-5 Sampling Uncertainty in Input Distribution Fits**

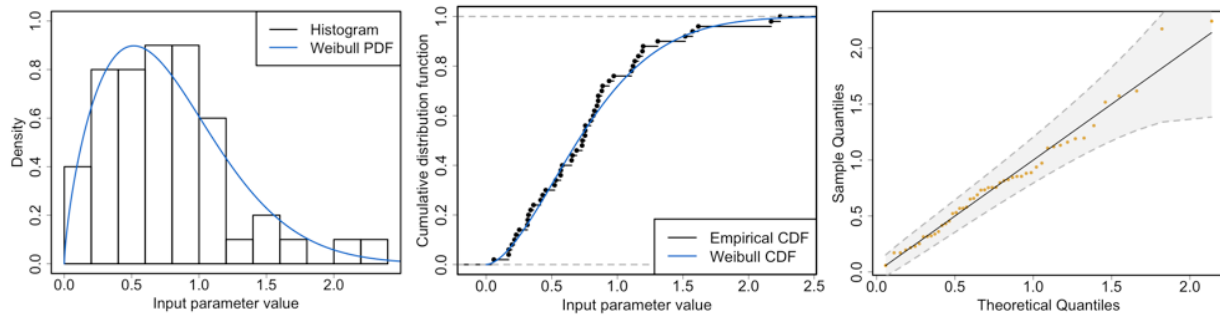
18 **(4) Evaluate distributional fit.**

19 After fitting a distribution to data, it is important to evaluate how well the distribution matches the
20 observed data. From Reference 4-19, the basic principle behind evaluating distributional fits is
21 to compare the parametric estimates from the model fit to nonparametric quantities that are not
22 based on a fitted model. Representations of some of the graphical tools, described below,
23 appear in Figure 4-6:

- 24 • Overlay a parametric fit of the probability density function onto a histogram of the data.
25 The left plot shows $n=50$ data points fit to a Weibull distribution. Large differences
26 between the histogram and parametric probability density function (PDF) estimate would
27 signal poor model fit.

- 1 • Overlay a parametric fit of the cumulative distribution function (CDF) onto the empirical
2 CDF of the data (middle plot). Large differences between the empirical CDF and
3 parametric CDF estimate would signal poor model fit.
- 4 • Construct a probability plot, also called a Quantile-Quantile plot (right plot). Probability
5 plots compare model-estimated versus empirical quantiles of the data. A departure from
6 the reference line indicates a region where the model poorly fits the data. Because there
7 is sampling uncertainty in the quantile estimates, confidence intervals can help assess
8 whether there is statistical evidence of a lack of model fit. Points falling outside the
9 bounds indicate a lack of fit of the probability distribution.

10



11

12 **Figure 4-6 Graphical Diagnostics for Parametric Model Fit**

13

14 Additionally, statistical goodness-of-fit hypothesis tests can also be used to detect evidence of a
15 poor model fit. These tests have strong, known limitations that limit their applicability in practice
16 and must be supplemented with graphical tools and expert judgment to determine whether a
17 model is a reasonable fit to data (Reference 4-19).

18 **(5) Select a final input distribution model.**

19 The final input distribution is an estimate of input uncertainty that reflects both data-driven
20 evidence and expert judgment (particularly in limited-data scenarios). To select an input
21 distribution, the analyst selects the following:

- 22 • **Distribution model.** It is best practice to consider several different candidate probability
23 distributions (e.g., normal, lognormal, Weibull), and select the final distribution based on
24 which is the best fit to the data, a process known as model selection.
- 25 • **Values for the distribution parameters.** Recall that input parameters are uncertain,
26 due to estimation based on a finite sample size.

27 In many instances, there will be uncertainty in the choice of the distribution model and
28 distribution parameter values. Input distribution uncertainty can have a large impact on the final
29 estimate in a PFM analysis. This source of uncertainty is most important when one of the
30 following is true:

- 31 • The output is sensitive to the input.

- 1 • There are insufficient data to accurately estimate an input distribution.
- 2 • The acceptance criterion relates to bounding a probability below a very low threshold
3 (e.g., $p < 1 \times 10^{-6}$).
- 4 This uncertainty can be incorporated into the final PFM analysis in different ways:
- 5 • Treat uncertain probability distributions and their parameters as additional sources of
6 epistemic uncertainty.
- 7 • Choose values of the distribution parameters resulting in conservative values for the
8 inputs with respect to the application at hand.
- 9 • Examine the robustness to changes in the input distribution using a sensitivity study
10 (Section 4.4).

11 **4.2.2 Preserving Physical Relationships between Inputs**

12 *4.2.2.1 What Is It?*

13 In a PFM analysis, most uncertain inputs are assumed to be statistically independent; that is,
14 changing the value of one input does not impact the value of other inputs. However, a subset of
15 input variables is often statistically dependent. For the input set to be physically realistic, these
16 dependencies should be preserved.

17 *4.2.2.2 How to Use?*

18 Before modeling, expert judgment is applied and exploratory data analyses are conducted to
19 understand the relationship between inputs. Relationships can manifest themselves as
20 correlations or as more general dependencies such as nonlinear relationships or ordering
21 relationships (i.e., input 1 must be larger than input 2).

22 Some approaches to specifying statistically dependent inputs include the following:

- 24 • inducing correlation in random samples through the transformation of independent
25 samples
- 26 • constructing a joint probability distribution for the inputs that includes the dependencies
- 27 • specifying a conditional probability distribution for one input as a function of the other
28 input
- 29 • constraining the parameter space

30 Section 4.2.2.4 provides technical details on these approaches.

31 *4.2.2.3 When/Why?*

32 Preserving dependence between variables is often needed to ensure a physically realistic input
33 set or to maintain the physical laws that drive the problem. For example, the inner diameter of a

1 pipe must be smaller than its outer diameter, so a relationship between these variables may be
2 imposed to ensure the physicality of the inputs.

3 4.2.2.4 Technical Details

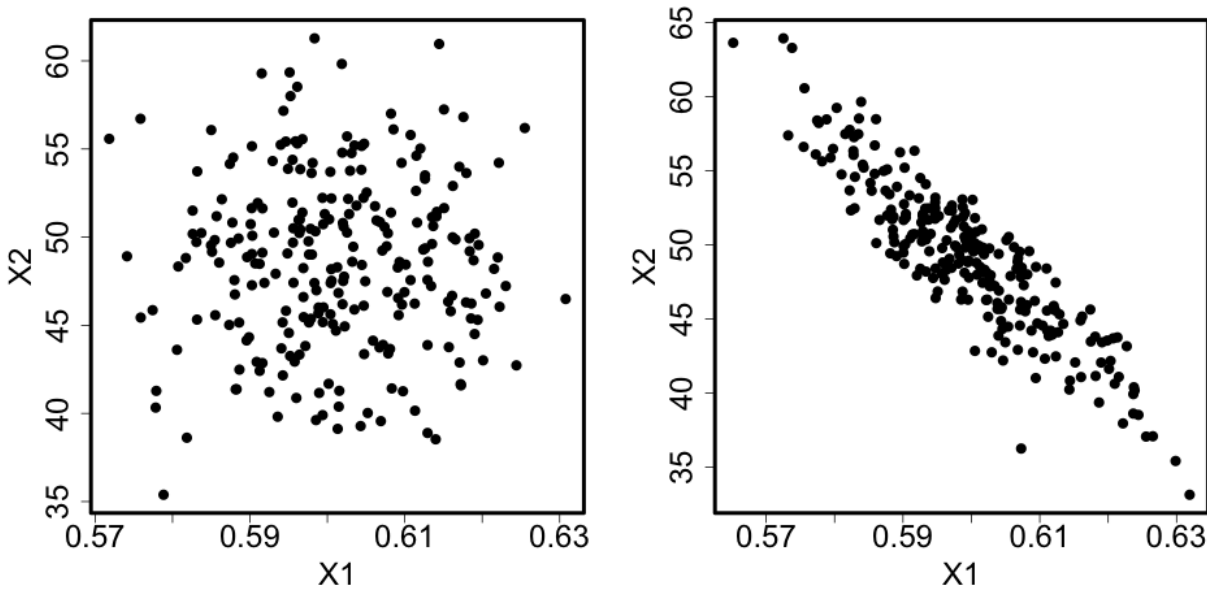
4 The technical details below describe the four approaches to specifying statistically dependent
5 inputs given above.

6
7 **Transforming Independent Samples.** Correlated inputs can be generated by transforming
8 independent random samples. If inputs follow a multivariate normal distribution, then we can
9 directly transform the inputs to induce correlation. Specifically, consider two inputs x and y and a
10 random, independent sample of size n for each. Center and scale the samples of x and y to
11 have mean 0 and standard deviation 1. Let \mathbf{X} be the $n \times 2$ matrix whose columns are formed by
12 the centered and scaled samples of x and y , respectively. Let \mathbf{S} be the 2×2 matrix specifying the
13 correlation:

$$14 \quad \mathbf{S} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

15
16 where ρ is the desired correlation between the inputs. Let $\mathbf{S} = \mathbf{C}^T \mathbf{C}$ be the Cholesky
17 decomposition of \mathbf{S} and set $\mathbf{X}^* = \mathbf{X}\mathbf{C}$. The correlation between the columns of \mathbf{X}^* is $\text{Corr}(\mathbf{X}^*) =$
18 $(\mathbf{X}\mathbf{C})^T(\mathbf{X}\mathbf{C}) = \mathbf{C}^T \mathbf{X}^T \mathbf{X} \mathbf{C} = \mathbf{C}^T \mathbf{I} \mathbf{C} = \mathbf{S}$. The desired standard deviation for each input can be
19 achieved by scaling each column by its desired standard deviation. Next, the desired mean can
20 be added to each column. If inputs are not normally distributed, then this transformation method
21 will not preserve the probability distributions of the individual inputs (i.e., marginal distributions).

22
23 If inputs are not normally distributed, then an alternative approach is to induce correlation on the
24 ranks of the inputs (Reference 4-21). This approach has the advantages of being distribution
25 free and preserving the marginal distribution of the inputs. To implement this approach, the
26 analyst specifies the correlations between the ranks of the inputs, which ideally can be
27 estimated using experimental data or expert judgment. This estimate can be applied with SRS
28 and LHS. Reference 4-21 gives details on implementing the rank correlation method. Figure 4-7
29 provides an example of the rank approach for two input variables. On the left is a scatterplot of a
30 random sample of two variables. Transforming these points using the rank method results in the
31 scatterplot on the right where a strong negative correlation now exists.
32



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

Figure 4-7 Randomly Sampled Inputs (Left) and Transformed Inputs (Right)

Joint distribution modeling. Input parameters can be directly sampled from a joint distribution for the parameters that include a correlation structure. The multivariate normal distribution is a straightforward model for correlated inputs but is only appropriate when a normal distribution can reasonably represent the inputs. The multivariate normal distribution is parameterized by the mean and variance of each variable, along with the statistical (Pearson) correlation between pairs of variables.

If a multivariate normal distribution cannot reasonably represent the joint distribution, more sophisticated statistical models can be applied to specify a joint distribution. Specifically, copula methods are a popular statistical approach for specifying joint distributions of correlated variables (Reference 4-22).

Conditional probability. If one input is dependent on another input, the relationship can be modeled to induce correlation between inputs. Specifically, the joint distribution of inputs X and Y can be factorized as the product between the marginal distribution of X and the conditional distribution of Y given X : $f(X, Y) = f(X)f(Y|X)$. As an example, suppose the marginal distribution of X is Weibull:

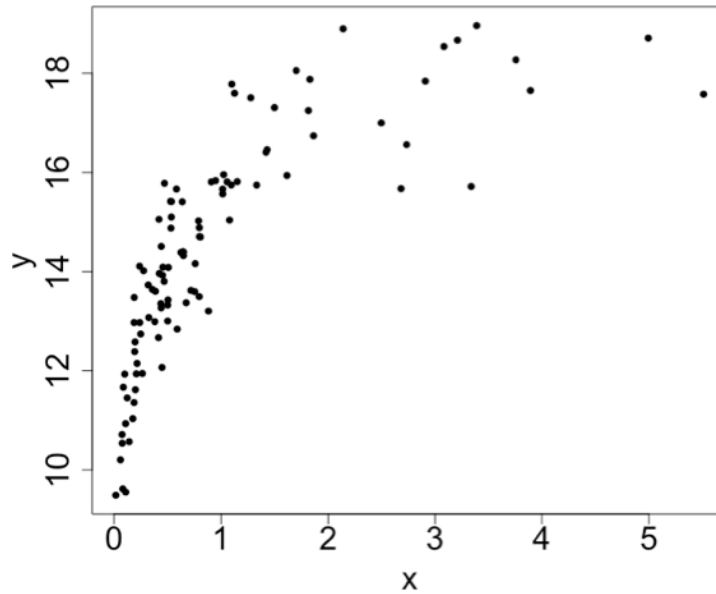
$$X \sim Weibull(1,1), \tag{1}$$

and the conditional distribution of Y given $X = x$ is normal with a mean dependent on x :

$$Y = Normal(15 + 2 \log(x), .3). \tag{2}$$

Figure 4-8 displays samples from the joint distribution of X and Y . To sample from the joint distribution of X and Y , first sample a value of X from the Weibull distribution (Eq. 1), and then sample the value of Y from a normal distribution with mean $15 + 2 \log(x)$ and variance 0.3 (Eq. 2).

1

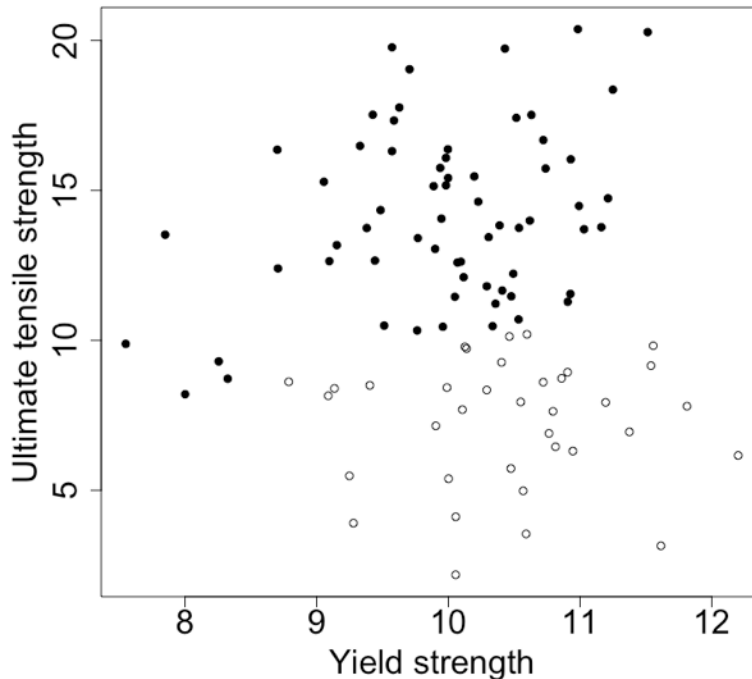


2

3 **Figure 4-8** Sampled Inputs from the Joint Distribution of X and Y

4

5 **Constrain the parameter space.** The input set can be constrained to ensure consistency and
6 physicality. For example, yield strength of a material is lower than its ultimate tensile strength
7 (Figure 4-9), and any samples not satisfying this constraint can simply be discarded. Note that
8 constraining the input space changes the uncertainty distribution on the inputs and can induce
9 correlation between inputs. Therefore, this approach should be used with caution to ensure that
10 imposed constraints accurately represent uncertainty in the inputs.



1

2 **Figure 4-9 Constraining Input Distributions to Ensure that Yield Strength Is Less than**
 3 **Ultimate Tensile Strength (Open Circles Are Not Admissible)**

4

5 All four of the above approaches are viable options for modeling dependencies in inputs. An
 6 advantage of the first approach (transforming independent samples) is that it is distribution free
 7 and requires only knowledge of the correlation between inputs (Reference 4-21). An advantage
 8 of the second approach (joint distribution modeling) is that correlation is directly built into the
 9 input parameter distribution. The third approach (conditional probability) gives flexibility with
 10 respect to the functional form of the dependency between variables. The fourth approach
 11 (constrain the parameter space) offers simplicity in implementation.

12 **4.3 Useful Methods for Forward Propagation of Input Uncertainty**

13 **4.3.1 Simple Random Sampling**

14 *4.3.1.1 What Is It?*

15 Simple random sampling (SRS) is a Monte Carlo sampling technique in which each uncertain
 16 input is sampled randomly from its corresponding probability distribution.

17 *4.3.1.2 How to Use?*

18 The SRS approach follows four steps:

- 19 (1) Specify probability distributions for the uncertain inputs.
- 20 (2) Choose the sample size.
- 21 (3) Implement SRS by randomly sampling the inputs from their probability distributions.

1 (4) Evaluate the computer model at each of the sampled inputs. The sampled outputs
2 represent a random sample of outputs corresponding to the probability distribution
3 implied by the distributions on the inputs.

4 4.3.1.3 *When/Why?*

5 SRS is easy to implement and therefore serves as a good “first pass” sampling scheme for
6 understanding output variability. However, SRS is often less efficient than alternative sampling
7 schemes (see Sections 4.3.2, 4.3.3, and 4.3.4 for alternatives); that is, more realizations are
8 needed using SRS than alternative sampling schemes to estimate a QoI with the same
9 precision. Section 4.3.1.4 gives information on choosing a sample size for SRS.

10 4.3.1.4 *Technical Details*

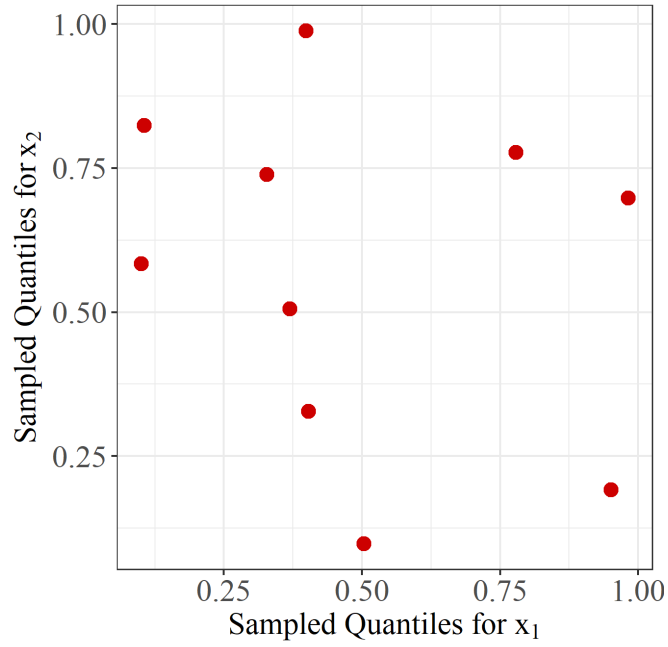
11 **Specifying probability distributions for the uncertain inputs.** Section 4.2 discusses
12 methods to identify and specify probability distributions for uncertain inputs.

13
14 **Choosing a sample size n .** It is usually feasible to directly estimate the amount of sampling
15 uncertainty in a QoI associated with SRS. For example, the law of large numbers indicates that
16 the magnitude of the sampling error associated with many QoIs estimated using SRS will be
17 proportional to $1/\sqrt{n}$ where n represents the sample size.

18
19 To use SRS to estimate a probability, the number of samples should be large relative to the
20 probability of the event occurring. As a rule of thumb, the sample size should be at least 10 to
21 20 times larger than $1/p$, where $p < 0.5$ is the probability of interest, to generate stable results
22 (Reference 4-23).

23
24 **Implementing SRS.** Most software programs can directly implement SRS for many common
25 probability distributions. Alternatively, a simple random sample can be generated for general
26 distributions by transforming uniform random samples on the interval 0 to 1 using the probability
27 integral transform (Reference 4-24). The uniform samples represent the quantiles of the
28 distribution from which the sample is desired. These quantiles are transformed by applying the
29 inverse CDF for the desired distribution. All that is needed to implement the probability integral
30 transform is the ability to randomly sample uniform variables and evaluate the inverse CDF. As
31 an example, a two-dimensional sample of uniform variables appears in Figure 4-10. Each
32 dimension is transformed using the probability integral transform to obtain the simple random
33 sample of variables shown in Figure 4-11, where the first dimension is distributed uniformly on
34 the interval $-1, 1$ ($U(-1,1)$) and the second dimension is normally distributed with mean 0 and
35 variance 1 ($N(0,1)$).

36

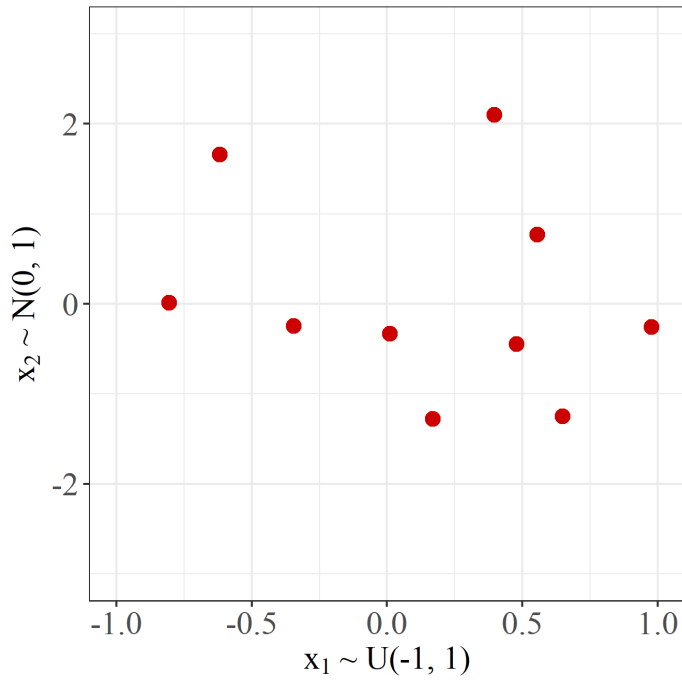


1

2

Figure 4-10 SRS Sample in the Quantile Space for Two Input Variables (n=10)

3



4

5

Figure 4-11 SRS Sample Transformed into the Input Space for Two Input Variables (n=10)

1 4.3.2 Latin Hypercube Sampling

2 4.3.2.1 What Is It?

3 Latin hypercube sampling (LHS) is a Monte Carlo sampling technique. LHS is a method to
4 obtain a sample that is more spread out across the input space than a typical SRS sample,
5 producing estimates with more statistical precision on average.

6 4.3.2.2 How to Use?

7 Many statistical software programs can implement LHS. The following steps describe a method
8 for generating an LHS of size n from independent input distributions associated with p uncertain
9 inputs (Reference 4-25):

- 10 (1) Stratify the input space by dividing the range of each input, x_j , into n disjoint intervals of
11 equal probability.
- 12 (2) For each input, randomly sample a single value from each interval, resulting in n
13 sampled values for each input. For a given input and interval, the sample is taken from
14 the conditional distribution of the input on the interval.
- 15 (3) Randomly combine samples without replacement:
 - 16 a. Randomly pair, without replacement, the n values sampled from the first input,
17 x_1 , with the n values from the second input, x_2 , to produce n pairs.
 - 18 b. Randomly combine these pairs, without replacement, with the n values sampled
19 from the third input, x_3 .
 - 20 c. Continue this process iteratively on x_4, x_5, \dots, x_p resulting in a set of n p -tuples.

21 The correlation between inputs can be incorporated using the Iman-Conover procedure that
22 induces correlation based on the ranks of inputs (Reference 4.3; see Section 4.2.2).

23
24 To summarize, LHS stratifies each input dimension into equally probable strata. In each
25 dimension, each stratum is sampled once (the regions formed by the sampled strata create a
26 pattern akin to a Latin hypercube in experimental design, such as described in Reference 4-26).
27 Within each of the sampled regions, a single sample is randomly sampled according to the
28 distribution within the region (Reference 4-25, 4-27).

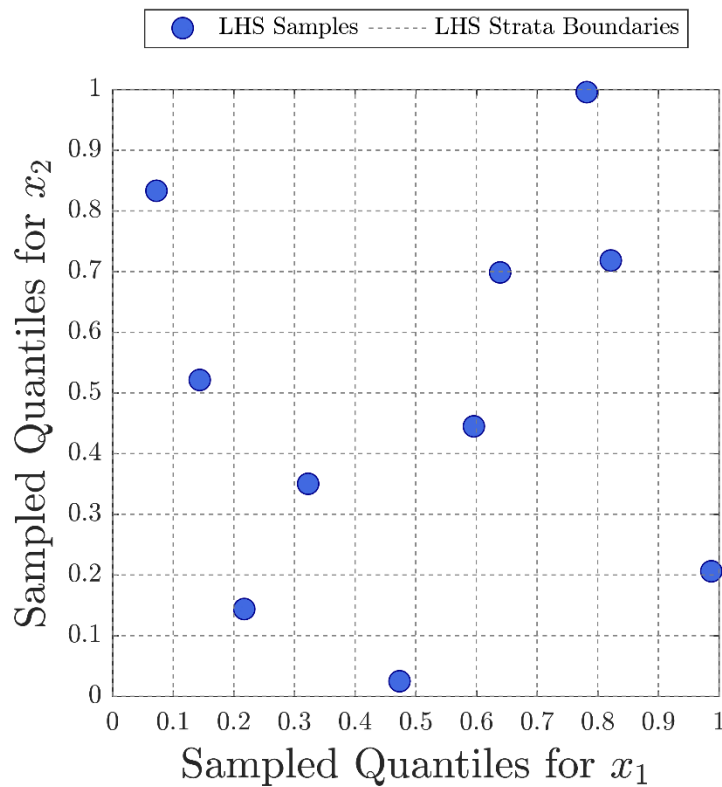
29 4.3.2.3 When/Why?

30 LHS is designed to cover the range of the input space more efficiently than SRS (Section 4.3.1).
31 For this reason, it is a common technique for forward propagation of uncertainty and for building
32 surrogate models (Section 4.3.10). Compared to SRS, LHS will typically result in more
33 statistically precise estimates of a QoI; however, the increase in precision diminishes as the
34 sample size increases. Quantifying statistical uncertainties on QoIs calculated using LHS is
35 more challenging.

1 4.3.2.4 Technical Details

2 The following simple example demonstrates the steps of the LHS algorithm outlined above. It
3 shows an LHS for two input variables, x_1 and x_2 , with $n=10$ samples. In this example, x_1 is
4 uniformly distributed from -1 to 1, and x_2 is normally distributed with a mean of 0 and a standard
5 deviation of 1:

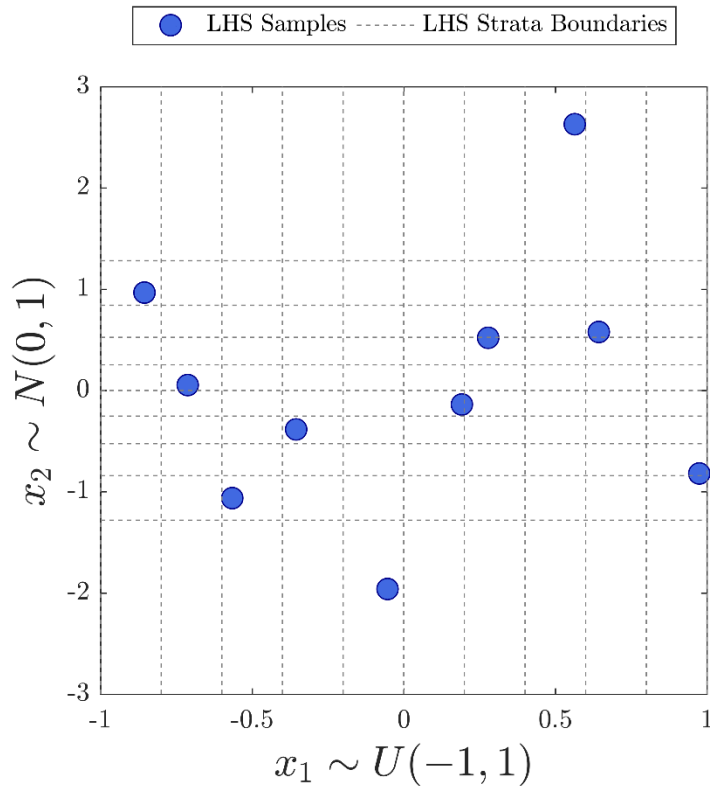
- 6 (1) **Stratify the input space.** First, each input distribution is divided into $n = 10$ intervals
7 (strata) of equal probability according to their respective distributions. This stratification
8 can be done first in the quantile space defined as the two-dimensional hypercube on
9 $(0,1)$. The intervals (strata) in the quantile space are evenly spaced, as displayed in
10 Figure 4-12.



11

12 **Figure 4-12 Example of an LHS in the Quantile Space for Two Input Variables (n=10)**

13



1
2 **Figure 4-13 LHS Transformed into the Input Space for Two Input Variables (n=10)**
3

- 4 (2) **Randomly sample from each interval.** Next, random uniform samples are taken on
5 each interval in each dimension. These samples are transformed to a sample from their
6 specified input distributions using the probability integral transform (Reference 4-28).
7 Figure 4-12 shows uniform samples for each stratum plotted in the two-dimensional
8 quantile space. The transformed samples (and strata) appear in Figure 4-13. The next
9 step describes the displayed pairing of each x_1 sample with an x_2 sample.
10
11 (3) **Randomly combine samples without replacement.** Finally, the values sampled from
12 the first input, x_1 , are randomly paired without replacement with the values sampled from
13 the second input, x_2 . Once a sampled value of x_2 is randomly paired with a sampled
14 value of x_1 , this sampled value of x_2 cannot be paired with a different value of x_1 . In the
15 example shown in Figure 4-12, the sample from the first strata of x_1 is randomly paired
16 with the sample from the ninth strata of x_2 . The sample from the first strata of x_1 is not
17 paired with any other samples of x_2 . Similarly, the sample from the ninth strata of x_2 is
18 not paired with any other samples of x_1 . Next, the sample from the second strata of x_1 is
19 randomly paired with the sample from the sixth strata of x_2 , and so on, until all $n=10$
20 pairs are selected.

21 **Augmenting LHS designs.** The design of an LHS depends on a known sample size. If, after a
22 sample is selected, the analyst wants to decrease sampling uncertainty in a QoI by increasing
23 the sample size, care must be taken to preserve the properties of an LHS. Reference 4-29
24 outlines one method for augmenting an initial LHS while preserving the LHS structure.

1 **Discrete Probability Distributions.** A discrete probability distribution is a related method that
2 produces a discrete approximation to a continuous distribution (Reference 4-30). Like LHS,
3 strata are created in each dimension. Rather than sampling within each stratum, the conditional
4 mean within the stratum is chosen as the sample point.

5 **4.3.3 Importance Sampling**

6 *4.3.3.1 What Is It?*

7 Importance sampling is a Monte Carlo technique that can be used to more efficiently sample
8 from the model input space than equal-probability sampling methods such as SRS and LHS
9 (Reference 4-31). Importance sampling concentrates samples in a specific area of interest in
10 the input space to improve estimation of QoIs (e.g., failure probabilities).

11 *4.3.3.2 How to Use?*

12 The implementation of importance sampling follows three steps:

13 (1) **Choose an importance distribution for the model inputs.** The importance sampling
14 distribution concentrates samples in regions of the input range that have a strong
15 influence on the estimate of the QoI. The importance distribution depends on the
16 relationship between the QoI and the inputs and therefore can be informed by SA. The
17 distribution should be selected such that the estimated QoI has smaller variance than a
18 QoI estimate from an equal-probability sample.

19
20 (2) **Sample inputs from the importance distribution** and run the model at these inputs.

21
22 (3) **Estimate the QoI.** Importance samples are weighted to obtain an unbiased estimate of
23 the QoI. The importance weights are derived from the density functions of the original
24 input distribution and the importance distribution (Reference 4-32).

25 *4.3.3.3 When/Why?*

26 Importance sampling is used to reduce the sampling uncertainty in the estimate of a QoI. If a
27 good importance sampling distribution has been selected, then the estimate of the QoI will be
28 more precise with fewer samples relative to SRS or LHS. Importance sampling reduces
29 sampling uncertainty by concentrating samples in the regions of importance (i.e., those regions
30 that contribute most to the QoI). While importance sampling is theoretically used for variance
31 reduction, poor choice of an importance distribution will increase the variance of a QoI estimate
32 (Reference 4-31).

33
34 Importance sampling can be particularly beneficial in PFM applications when the QoI is a rare
35 probability. That is, to estimate a 1×10^{-6} probability, it is more computationally efficient to
36 concentrate more samples around the area where events are more likely to occur. Without
37 importance sampling, an event will only be observed, on average, every one million samples.
38 Importance sampling algorithms can be designed to dramatically increase the number of
39 observed events and, subsequently, decrease the variance of the probability estimate for a fixed
40 number of samples.

1 4.3.3.4 Technical Details

2 Many QoI estimation problems can be formulated as the estimation of an expectation:

3
4
5

$$E_{\pi}[f(\mathbf{x})] = \int f(\mathbf{x})\pi(\mathbf{x})d\mathbf{x}$$

6 A common PFM example is when $\pi(\mathbf{x})$ represents the probability distribution on a multivariate
7 input \mathbf{x} and $f(\mathbf{x})$ is a model output; that is, the indicator of an adverse event (e.g., pipe rupture)
8 at input \mathbf{x} (i.e., $f(\mathbf{x}) = 1$ if the event occurs and 0 otherwise). In this case, the expectation
9 reduces to the probability of the adverse event.

10

11 By the law of large numbers, the average of f over a random sample $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ from $\pi(\mathbf{x})$
12 will converge to $E_{\pi}[f(\mathbf{x})]$ as n grows. Hence, it is straightforward to estimate the integral with
13 the average:

14

$$E_{\pi}[f(\mathbf{x})] \approx \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}^{(i)})$$

15

16 In the case of an indicator of a rare event (and other cases), this average is inefficient since very
17 few of the random samples will result in $f(\mathbf{x}) = 1$ (i.e., it is difficult to randomly sample an input
18 that results in the adverse event). Instead of sampling from $\pi(\mathbf{x})$, importance sampling takes its
19 sample from an importance distribution $h(\mathbf{x})$. Rewriting the above integral as

20

$$E_{\pi}[f(\mathbf{x})] = \int f(\mathbf{x}) \frac{\pi(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) d\mathbf{x} = E_h \left[f(\mathbf{x}) \frac{\pi(\mathbf{x})}{h(\mathbf{x})} \right],$$

21

22 we notice that it can be estimated from a sample $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}$ from $h(\mathbf{x})$ using a weighted
23 average:

24

$$E_{\pi}[f(\mathbf{x})] \approx \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}^{(i)}) \frac{\pi(\mathbf{x}^{(i)})}{h(\mathbf{x}^{(i)})}.$$

25

26 The values $w^{(i)} = \frac{\pi(\mathbf{x}^{(i)})}{h(\mathbf{x}^{(i)})}$ are importance weights on the $\mathbf{x}^{(i)}$. As we demonstrate below, the
27 careful choice of importance distribution h can dramatically reduce sampling uncertainty of the
28 estimate. Reference 4-32 gives the technical conditions on h needed for the weighted average
29 to converge to $E_{\pi}[f(\mathbf{x})]$ as n grows.

30

31 **Choose an importance distribution for the model inputs.** In practice, choosing the
32 importance distribution h usually involves selecting individual importance distributions for a few
33 of the inputs. An incorrect choice of inputs for importance sampling or a poor selection of the
34 importance distribution may lead to increasing the sampling uncertainty in the estimate of the
35 QoI when compared to an estimate generated without the use of importance sampling. As a
36 result, a careful and thorough analysis is necessary before selecting the importance distribution.
37 Effective implementation of importance sampling requires (1) understanding what regions of the
38 input space are important to the QoI and (2) selecting the importance distribution correctly given
39 the relationship between the important input and the QoI (References 4-31, 4-33, 4-34, 4-35).
40 This aspect of importance sampling is often not straightforward in PFM studies and requires
41 SAs to support the selected inputs, which are often confirmed with expert elicitation.

42

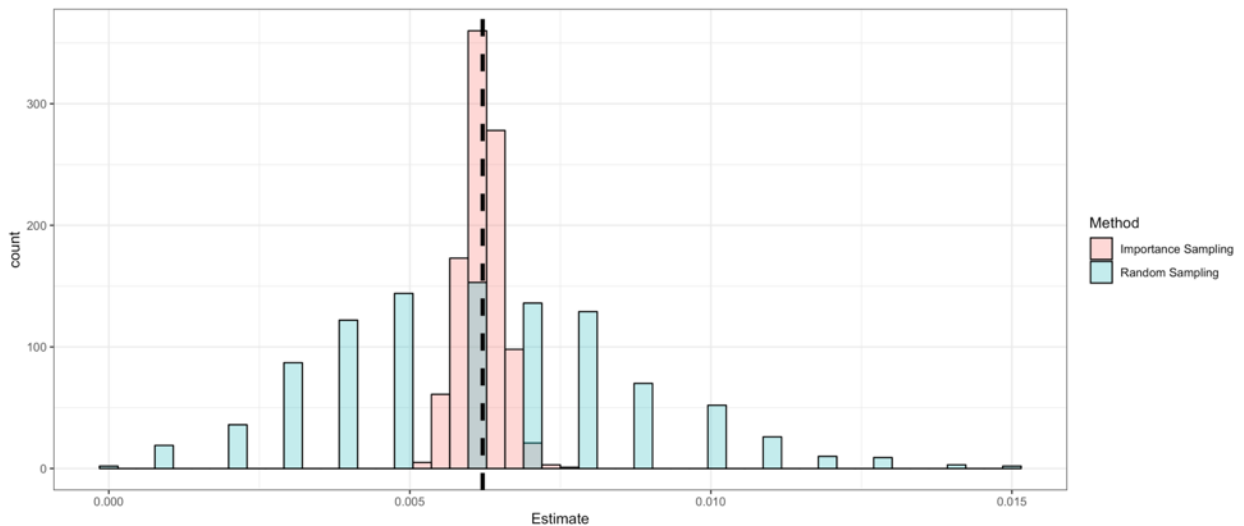
1 Inputs that have a strong relationship with the output are good candidates for importance
2 sampling. SA methods (Section 4.3.8) can be used to quantify the input-output relationship and
3 rank variables in terms of their influence on the QoI. The variables that are found to have the
4 strongest relationship with the QoI are considered for importance sampling.

5
6 Inefficiency in importance sampling often occurs in high-dimensional problems where too many
7 variables are importance sampled (Reference 4-36). To avoid this inefficiency, importance
8 sampling should be limited to only a few important variables.

9
10 **Importance sampling and estimation of the QoI.** After the nontrivial task of choosing a good
11 importance distribution, the implementation of the importance sampling methodology is
12 straightforward. Inputs are randomly sampled from their importance distribution and propagated
13 through the model. The final QoI is then estimated as a weighted average, weighted by the
14 importance weights (References 4-31, 4-33, 4-35).

15
16 **Illustration of importance sampling.** Importance sampling can be demonstrated as follows.
17 Suppose there is one normally distributed input $x \sim N(0,1)$, and the goal is to estimate
18 $E[f(x)] = P(X > 2.5)$. The true probability is known to be 0.0062. Figure 4-14 shows estimates
19 using repeated simple random samples of size $n = 1,000$ from $N(0,1)$ and repeated importance
20 samples from a Student-t distribution centered at 2.5 with 3 degrees of freedom. The
21 importance distribution was chosen to ensure more samples fell inside the failure region (> 2.5),
22 and 3 degrees of freedom were used to produce a heavy-tailed Student-t distribution in that
23 region, as smaller degrees of freedom increase the tail weight of the distribution. The
24 histograms in Figure 4-14 represent sampling uncertainty in the QoI estimate. While both
25 estimates are unbiased around the true probability indicated by the vertical dashed line, the
26 standard deviation of the estimates under importance sampling is 0.0003 compared to 0.003
27 under random sampling. This change represents a reduction in the sampling uncertainty by an
28 order of magnitude. Reference 4-32 describes a similar example.

29



30

31 **Figure 4-14 Example of Estimating a Probability Using Random Sampling and Importance**
32 **Sampling**

33

1 **Adaptive importance sampling.** In reliability analysis (see Section 4.3.4), whose goal is to
2 estimate a failure probability, adaptive importance sampling techniques are often used to help
3 refine the importance distribution. The goal in these applications is to detect the failure
4 boundary, defined as the boundary of the region separating failures and nonfailures
5 (Reference 4-37), in order to improve failure probability estimates. Adaptive methods iteratively
6 update the importance distribution to better estimate this boundary by considering the outputs of
7 previously sampled points in the domain (References 4-38, 4-39).

8
9 To implement adaptive importance sampling, first, an optimization problem is solved to find a
10 particular point on the failure boundary known as the most probable point (MPP). The MPP can
11 be used as a starting point to define an initial importance distribution, which through adaptive
12 sampling is updated as model evaluations are obtained. For example, multimodal sampling
13 (Reference 4-38) and curvature-based sampling (Reference 4-39) begin centered at the MPP
14 and then update the sampling distribution by assigning weights to various candidate density
15 functions. Many software packages such as DAKOTA (Reference 4-31) have the capability to
16 implement adaptive sampling methods.

17
18 Instead of using additional evaluations of the computational model, adaptive sampling can also
19 use information from surrogate models (Section 4.3.10). This feature is most useful when there
20 is a need to reduce the computational expense of model evaluations. For example, efficient
21 global reliability analysis (Reference 4-40) aims to create a Gaussian process (GP) surrogate
22 model for the function of interest and then adaptively select sample points in the domain near
23 the failure region to improve the quality of the surrogate model.

24 **4.3.4 First- and Second-Order Reliability Methods**

25 *4.3.4.1 What Is It?*

26 Reliability methods estimate a failure probability by approximating the probability of violating a
27 certain threshold criterion for a probabilistic analysis of continuous random variables. For
28 example, these methods can estimate the probability that a particular material stress is greater
29 than the yield stress. Often, the estimate requires many fewer samples of the computer model
30 than using Monte Carlo sampling methods. The methods described here are known as the
31 first-order reliability method (FORM) and the second order reliability method (SORM).

32 *4.3.4.2 How to Use?*

33 The use of FORM and SORM follows three steps:

- 34 (1) Define the failure region in terms of a continuous output and a threshold value.
35 Specifically, a failure occurs when the output exceeds the threshold. The failure
36 probability is the integral of the input probability distributions over the failure region.
- 37 (2) Approximate the failure region using FORM or SORM Taylor series approximation
38 around the MPP. The MPP is the point on the failure region boundary with highest input
39 probability density. Determining the location of the MPP requires evaluating the
40 computational model within an optimization algorithm.
- 41 (3) Estimate the failure probability using the integral over the approximate failure region.

1 4.3.4.3 When/Why?

2 Reliability methods are particularly useful because of the computational efficiency of the
3 algorithms. As described in Sections 4.3.1, 4.3.2, and 4.3.3, sampling-based algorithms can
4 also be used to approximate the failure probability of a system. However, these methods often
5 require thousands or tens of thousands of samples to provide good estimates. In contrast,
6 FORM and SORM are more efficient reliability methods because they seek to directly
7 understand the location and probabilistic distance to the limit state (i.e., the boundary of the
8 failure region). Often, for some low-dimensional input spaces, the MPP can be located
9 accurately with a small number of model evaluations (on the order of 10 points). This difference
10 provides substantial computational savings especially when the analysis model is
11 computationally expensive to evaluate. However, the failure probability estimate obtained from
12 FORM or SORM relies on a Taylor series approximation to the shape of the limit state
13 (first-order series in FORM and second-order series in SORM) and can result in a poor estimate
14 if the approximation is not good.

15 4.3.4.4 Technical Details

16 **Defining the failure region:** Following Reference 4-41, consider a model that predicts an
17 output Y as a function g of some set of input random variables $\mathbf{X} = (X_1, X_2, \dots, X_n)$:

$$Y = g(\mathbf{X})$$

18 Suppose a failure event of interest is defined when $y < 0$. Note that any problem can be
19 formulated in this way by considering y to be a margin against failure. For example, X_1 might be
20 the predicted stress in a material, and X_2 might be the yield stress of the material. If we define
21 failure when the material yields, then the margin Y against failure is defined as $Y = X_2 - X_1$, and
22 failure occurs when $y < 0$ for some values (x_1, x_2) .

23
24 In any such scenario, the goal of reliability methods is to compute the failure probability $p_f =$
25 $P(Y < 0) = \int_{FR} \pi(\mathbf{x}) d\mathbf{x}$, where π is the joint probability density function of the inputs \mathbf{X} and
26 $FR = \{\mathbf{X} : Y = g(\mathbf{X}) < 0\}$ is the failure region. The unknown and potentially complex failure
27 region makes computing the integral difficult. To simplify the computation, reliability methods
28 like FORM and SORM make simplifying assumptions on $g(\mathbf{X})$.

29
30 **Approximating the failure region using FORM with normally distributed inputs.** After the
31 failure region has been defined, the failure probability can be estimated with FORM using the
32 following steps, described in detail below:

- 33 (1) Transform each input variable into the standard normal space.
- 34 (2) Find the MPP, \mathbf{x}^{f*} .
- 35 (3) Calculate the distance, β , from 0 to the MPP.
- 36 (4) Use this distance to estimate the failure probability $p_f \approx 1 - \Phi(\beta)$, where Φ is the CDF
37 of the standard normal distribution.

38 Figure 4-15 depicts this process visually (following Reference 4-42).

39

1 **Transformation of inputs.** A problem first addressed by the Hasofer-Lind method
 2 (Reference 4-43) assumes each input is independently normally distributed. The first step in this
 3 method (and others that relax the normality assumption) is to transform each random variable
 4 into the standard normal space (i.e., $N(0,1)$) so that all variables in the input domain have a
 5 common scale:

$$X'_i = \frac{X_i - \mu_{X_i}}{\sigma_{X_i}}, \quad \forall X_i \in i = 1, 2, \dots, n$$

6 Here, X'_i denotes the standard normal transformation of input random variable X_i . Figure 4-15
 7 depicts this transformation.

8
 9 **Find the MPP.** Once this transformation has been performed for all the input random variables,
 10 determining the location of the MPP (x'^*) involves solving the following inverse problem:

$$\begin{aligned} x'^* &= \operatorname{argmin} \sqrt{x'^T x'} \\ &s. t. g(x') = 0 \end{aligned}$$

11 The point x'^* is estimated by finding the values of x' that fall on the failure region boundary
 12 ($g(x') = 0$) and are the closest (minimum) distance to 0 (the mean of the transformed random
 13 variables). The inverse equation above can be solved by optimization methods such as the
 14 Rackwitz algorithm (Reference 4-44) and the Newton-Raphson recursive algorithm
 15 (Reference 4-45).

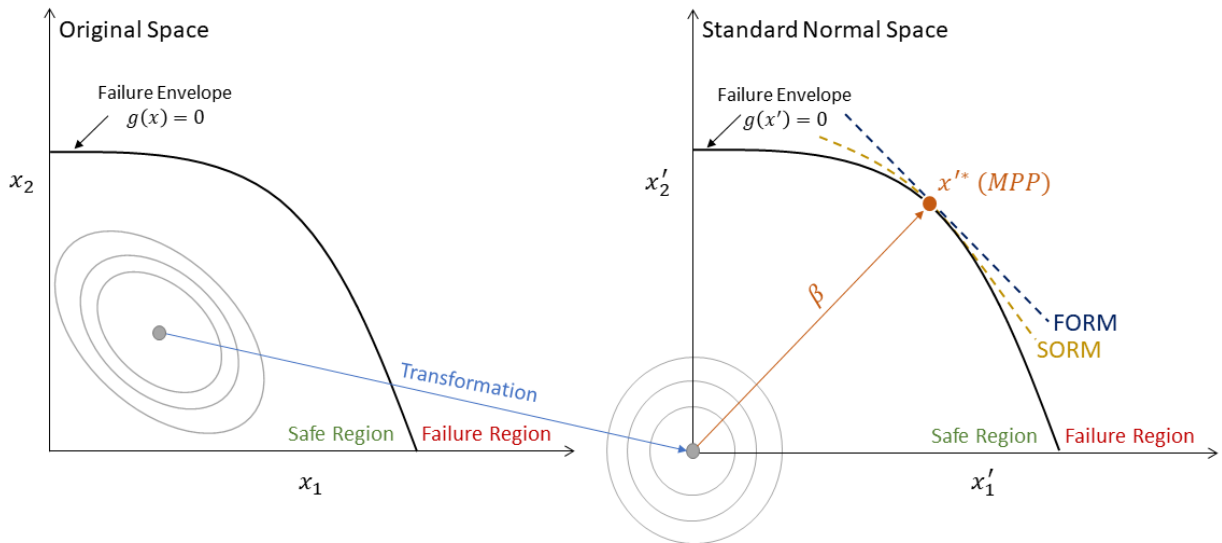
16
 17 **Calculate the distance from the MPP to 0.** Once x'^* is found, its distance from 0 can be
 18 calculated as $\beta = \sqrt{x'^*T x'^*}$. A visualization of x'^* and β appears in Figure 4-15. The parameter
 19 β is the distance from 0 to the MPP x'^* and is known as the safety index for the reliability
 20 problem.

21
 22 **Estimate the failure probability.** The failure probability p_f can then be approximated directly
 23 by making an assumption about the shape of the failure envelope. The simplest assumption,
 24 known as the FORM, is to assume the failure envelope is linear, as shown by the blue dashed
 25 line in Figure 4-15. In this case, the approximation of p_f becomes

$$p_f \approx P(l(x) < 0) = \Phi(-\beta) = 1 - \Phi(\beta),$$

26 where $l(x)$ is the first-order Taylor series approximation of $g(x)$ about x'^* and Φ is the standard
 27 normal CDF. The first equality follows from the linear (and normal) assumption on $l(x)$ with
 28 parameters governed by the Taylor series about x'^* . The second equality follows from the
 29 symmetry of the standard normal distribution. Figure 4-15 shows the linear approximation of the
 30 failure envelope using the FORM method.

31



1
2 **Figure 4-15 Example of the FORM and SORM Methods in the Standard Normal Space**
3 **(Following Reference 4-42)**
4

5 **Nonnormally distributed inputs.** One limitation of the approach above is that all input
6 variables must be independent normal random variables in order for the approach to be valid.
7 When they are instead independent nonnormal random variables, they can be transformed into
8 approximate standard normal random variables by estimating equivalent normal distribution
9 parameters $\mu_{X_i}^N$ and $\sigma_{X_i}^N$ for each nonnormal random variable X_i . The Rackwitz-Fiessler
10 two-parameter equivalent normal transformation (Reference 4-46) achieves this by equating the
11 PDF and CDF of variable X_i to the PDF and CDF of an equivalent standard normal distribution.
12 Once this additional transformation is performed, the same inverse problem can be solved by
13 the previously mentioned optimization methods to arrive at the MPP in equivalent standard
14 normal space.

15
16 **Approximating the failure region using SORM.** Another limitation of the FORM approximation
17 is that it may be overly conservative when the actual failure envelope is highly nonlinear. To
18 improve upon this limitation, curvature of the limit state can be considered by also including the
19 partial derivatives $\frac{\partial g}{\partial X_i}(x^*)$ of the function g with respect to each X_i in the Taylor series
20 expansion of the function g . These improvements to the approximation of the limit state function
21 lead to an improvement to the approximation of p_f . Since the approximation is now a
22 second-order Taylor series, the method is called the second-order reliability method (SORM).
23 The yellow dashed line in Figure 4-15 provides a notional example of the curved approximation
24 to the failure envelope using SORM. References 4-47, 4-48, and 4-49 provide further details.

25 4.3.5 Convergence Analysis

26 4.3.5.1 What Is It?

27 When propagating uncertainty forward through a model, there will be uncertainty in the estimate
28 of the QoI due to the limited number of model realizations. The purpose of a convergence
29 analysis is to assess the magnitude of sampling uncertainty associated with the QoI estimates
30 obtained from Monte Carlo forward propagation of uncertainty (e.g., Sections 4.3.1, 4.3.2, and

1 4.3.3). Ultimately, an estimate has converged if the conclusions of the analysis do not change
2 solely due to sampling uncertainty.

3 4.3.5.2 *How to Use?*

4 To conduct a convergence analysis, the analyst will take the following steps:

- 5 (1) Quantify sampling uncertainty with a metric.
- 6 (2) Compare the metric to a threshold value.

7 The threshold defines the maximum level of uncertainty acceptable for the application.

8
9 When Monte Carlo sampling (e.g., Sections 4.3.1, 4.3.2, and 4.3.3) is used to estimate a QoI,
10 the following are three general methods for quantifying sampling uncertainty:

- 11 • **Calculate sampling uncertainty metrics for an estimate.** Section 4.3.6 covers
12 closed-form sampling uncertainty metrics under SRS for probability estimates.
13 Section 4.3.7 discusses statistical bootstrapping as an alternative to closed-form metrics.
14 The metrics are calculated on a single simulation but require statistical assumptions that
15 must be evaluated in practice.
- 16 • **Assess stability of a QoI estimate as the sample size increases.** The estimate of the
17 QoI is monitored as the sample size grows to determine the appropriate sample size.
- 18 • **Compare QoI estimates over replicate simulations.** Several independent replicates of
19 the model simulations are needed, which may not be feasible to implement in practice.
20 The variation between these replicate simulations is assessed.

21 Section 4.3.5.4 discusses these methods in more depth. In general, the best method for
22 convergence analysis depends on the computational complexity of the model as well as the
23 type of sampling scheme.

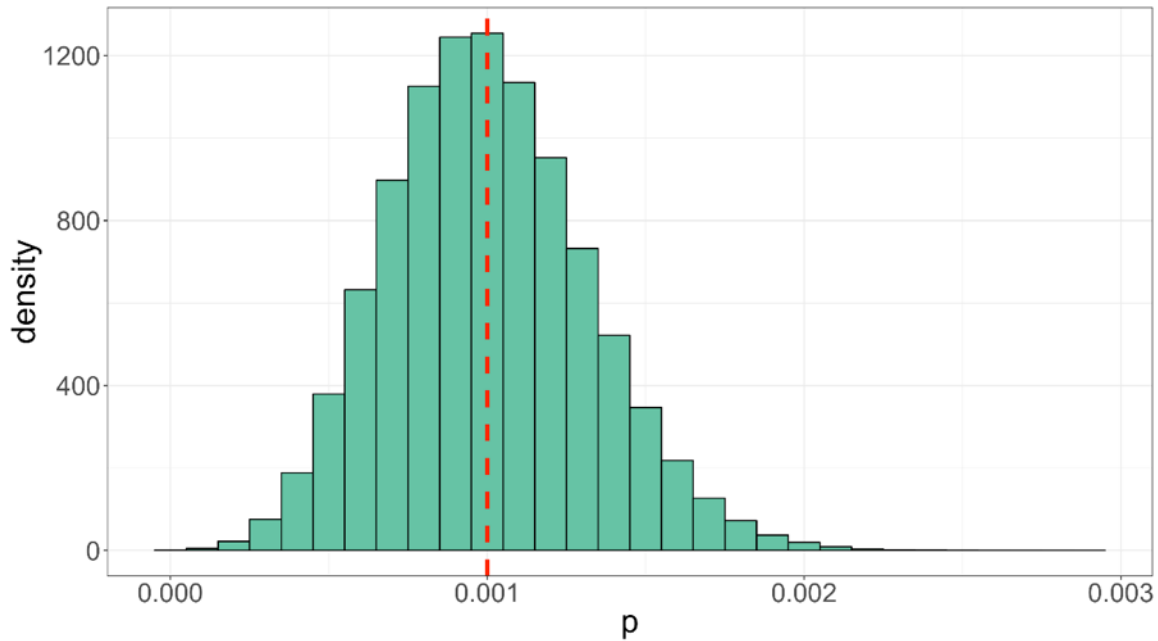
24 4.3.5.3 *When/Why?*

25 In PFM, sampling uncertainty exists in estimates of QoIs. Rigorous assessment of the sampling
26 uncertainty is conducted to ensure that the conclusions of the PFM analysis would not change
27 solely due to random variations of estimates in different simulations.

28 4.3.5.4 *Technical Details*

29 **Sampling uncertainty.** Sampling uncertainty arises because the model can only be run for a
30 finite number of realizations; a set of model realizations used to estimate a QoI such as a failure
31 probability is called a model simulation. Replicate model simulations at different random seeds
32 will produce different results. Consider the problem of estimating a rare probability. The
33 histogram in Figure 4-16 displays estimates of this probability from many independent
34 simulations; each simulation is based on $n=10,000$ model realizations sampled using SRS. This
35 histogram represents the sampling distribution of the probability estimate, defined as the
36 distribution of estimates obtained from repeated simulations. The true probability is 0.001,
37 indicated by the red vertical dashed line, with estimates ranging from 0 to approximately 0.003.
38 This range represents the sampling uncertainty. For a PFM analysis, this range could be
39 acceptable or unacceptable, depending on the requirements of the analysis.

1



2

3 **Figure 4-16 Histogram of Probability Estimates from a Simple Random Sample**

4

5 **Quantifying sampling uncertainty with a statistical metric.** A convergence metric quantifies
6 the sampling uncertainty in the estimate of a QoI, calculated using the output realizations.
7 Convergence metrics can be compared to a prespecified threshold to determine whether the
8 sample size is sufficiently large. Examples of statistical metrics to quantify sampling uncertainty
9 in a convergence analysis include the following:

10

• *Standard error* is the standard deviation of the sampling distribution of the QoI. It is a
11 measure of the variation in the estimate across repeated simulations.

12

• *Coefficient of variation (CV)* is the ratio of the estimated standard error of the QoI to the
13 mean estimate of the QoI. The CV should only be used for a positive QoI, and it is not
14 recommended if the mean estimate of the QoI is close to zero because the estimate of
15 the CV can become very volatile.

16

• *Confidence interval* is an interval estimate of a QoI, providing a range of values for which
17 we have high confidence that the true value of the QoI lies in the interval.

18

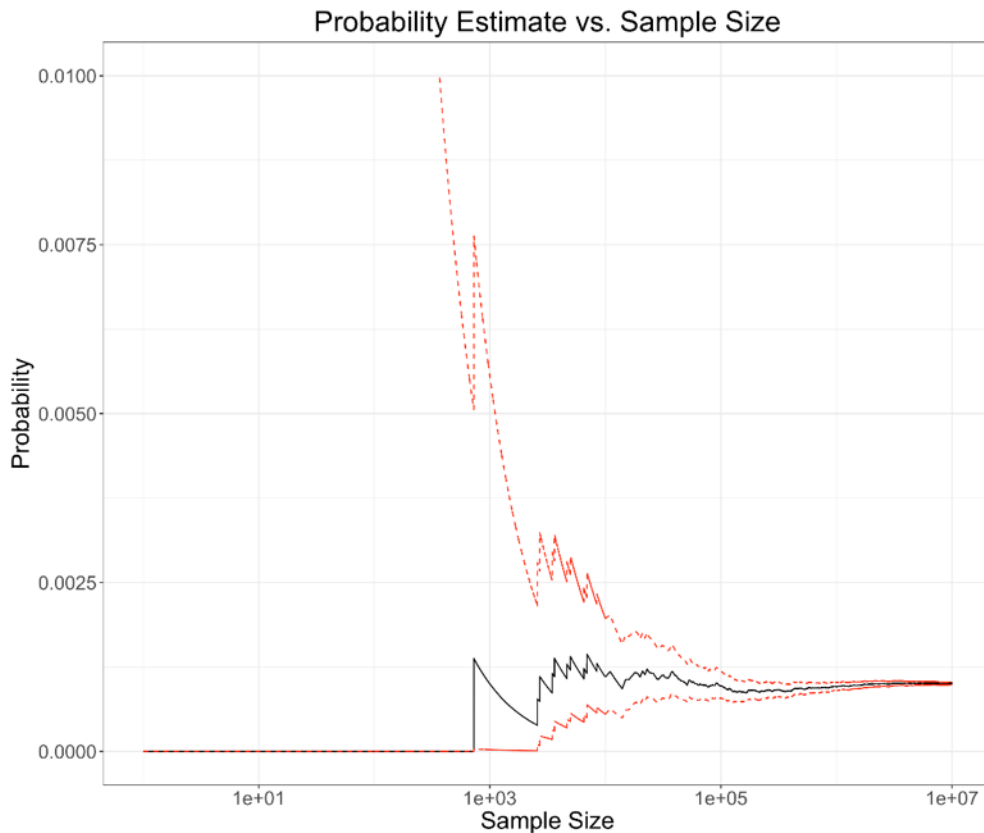
Sections 4.3.6 and 4.3.7 discuss methods for calculating these metrics based on a model
19 simulation. These metrics are compared against predetermined thresholds to determine
20 whether sampling uncertainty is sufficiently low. For example, the standard error or CV can be
21 compared to a threshold defining the maximum acceptable value. The maximum acceptable
22 width of a confidence interval is another possible threshold (Reference 4-50).

23

Assessing the stability of an estimate as the sample size increases. One common method
24 for assessing model convergence is incrementally increasing the sample size and examining
25 the stability of the QoI estimate as a function of sample size. As the sample size increases and
26 sampling uncertainty decreases, the estimate of the QoI will stabilize. Metrics for QoI stability
27

1 include the standard error, CV, and confidence intervals, all of which are calculated from the
2 sample.

3 Figure 4-17 depicts an example demonstrating the convergence of estimating a small probability
4 (1×10^{-3}) using SRS. The simulation was run for 1×10^7 iterations, and the x-axis is plotted on the
5 log scale. The estimated probability is plotted as the black line in the figure. This estimate is 0
6 until a sampling size of about 1×10^3 . Then it is volatile until a sample size of around 1×10^6 ,
7 where it begins to converge to the true value. A two-sided 95-percent confidence interval,
8 represented by the red dashed lines, provides a convergence metric. This bound was
9 constructed using the Clopper-Pearson confidence interval (References 4-50, 4-51). Suppose
10 that the threshold for model convergence is met when the 95-percent confidence interval has
11 width less than 1×10^{-4} . It takes 1,516,000 samples to satisfy the metric in this case.
12



13

14 **Figure 4-17 Confidence Interval used to Assess the Convergence of a Probability Estimate**

15

16 **Comparing estimates over replicate simulations.** A more computationally expensive
17 approach to assessing convergence of a QoI is to conduct many replicate simulations to
18 repeatedly estimate a QoI and then directly estimate variability in the QoI across replicates. The
19 variability in the QoI estimate across simulations provides information for sampling uncertainty.
20 The advantage of this method is that it is easy to apply to any sampling scheme. The
21 disadvantage is that conducting replicate simulations is computationally expensive. To
22 determine whether a sample of size n is sufficient, a total of nr realizations is computed, where
23 r is the number of replications of the simulation. The specific sample size and reasonable
24 number of resamples depend on the application. The sample size of each replicate set should

1 be close to that of the empirical data. Further, these samples can later be combined to produce
2 a more precise final estimate of the Qol.

3
4 This sampling uncertainty can be quantified using different metrics, such as the standard
5 deviation of the Qol estimates, the CV, or a statistical prediction interval for future Qol
6 estimates. A prediction interval is similar to a confidence interval and provides interval bounds
7 such that there is a high level of confidence that a new Qol estimate would lie in this range. An
8 approximate $100(1 - \alpha)\%$ confident prediction interval for a normally distributed random
9 variable is

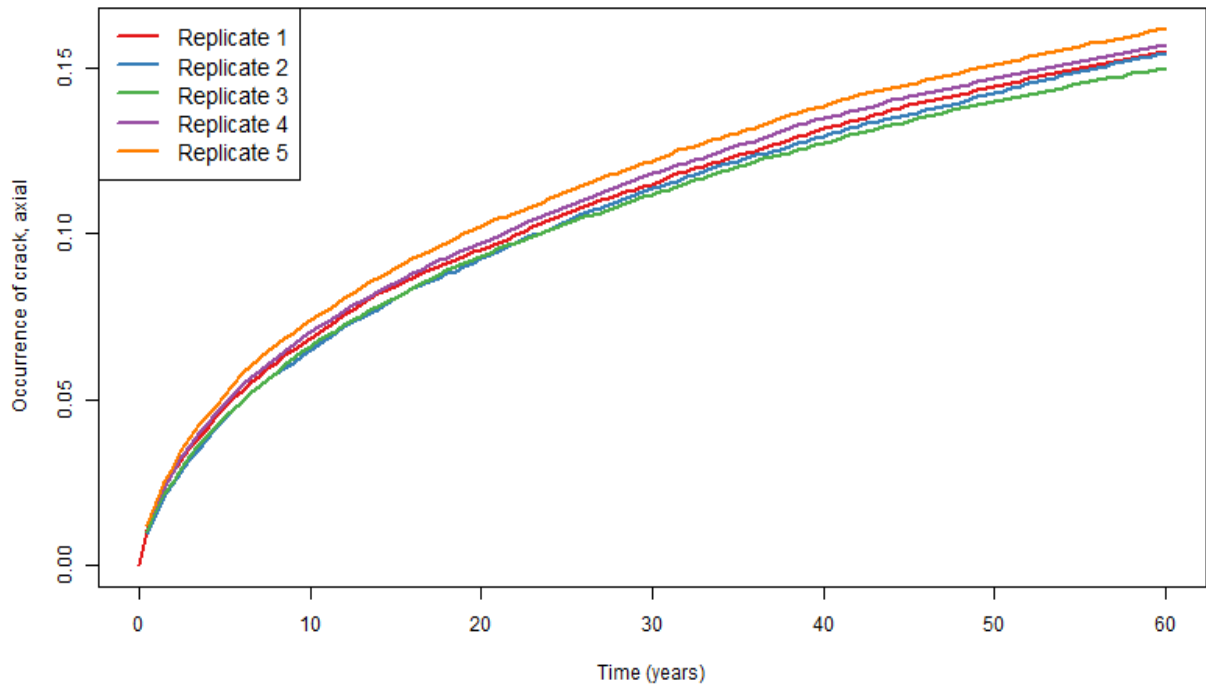
$$\bar{x} \pm t_{\alpha/2, r-1} \sqrt{s^2 \left(1 + \frac{1}{r}\right)},$$

10 where \bar{x} is the average of the Qol estimates, s is the standard deviation of the Qol estimates,
11 and $t_{\frac{\alpha}{2}, r-1}$ is the $(\alpha/2)th$ percentile of the Student-t distribution with $r - 1$ degrees of freedom.

12 Note that, to compute a prediction interval, the distribution of the Qol estimate must be known.
13 Often, the Qol is an average of many model realizations such that Qol estimates will be
14 approximately normally distributed (based on the central limit theorem). When the Qol is a rare
15 probability, this normal approximation can perform poorly, and normal prediction intervals
16 should be interpreted with caution.

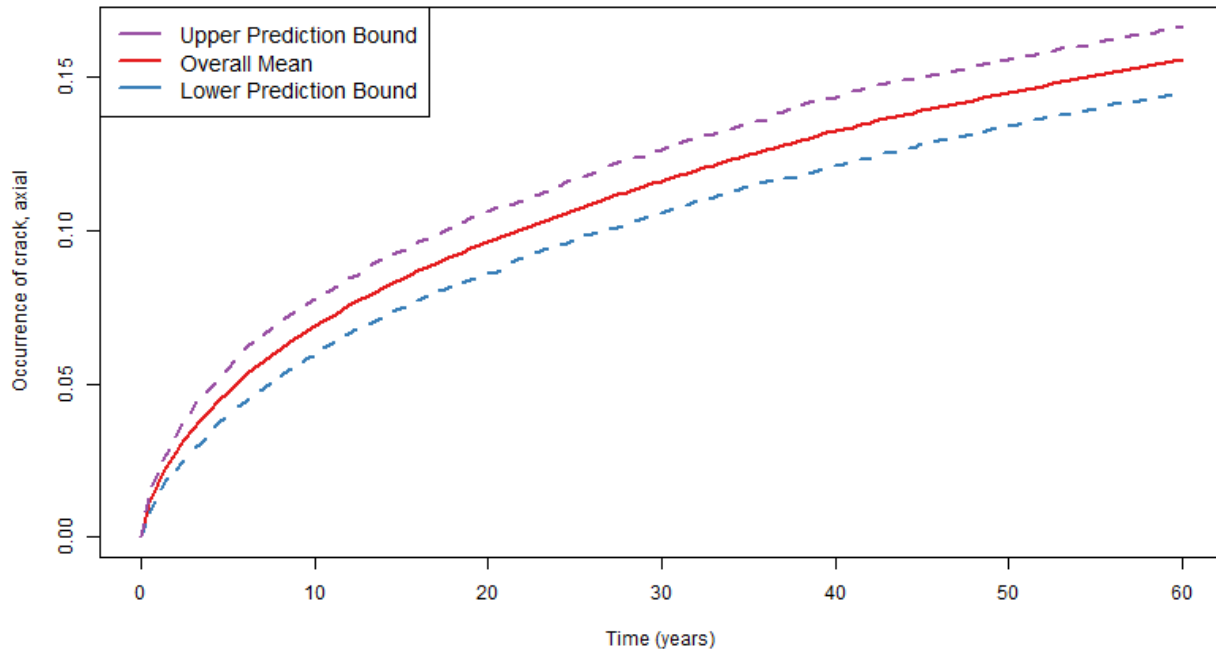
17
18 As an example, suppose we are measuring sampling variability in an estimate of the probability
19 of an axial crack in a pipe over 60 years. Figure 4-18 plots replicate Qol estimates as a function
20 of time based on $r = 5$ replications. Figure 4-19 shows the two-sided 95-percent prediction
21 interval for this example. The width of the prediction interval as a function of time can be
22 compared to a predetermined threshold on the acceptable maximal width to assess
23 convergence. The choice of $r = 5$ should be justified. The more replicates the better, and one
24 can assess the stability of the estimated standard deviation as r increases. Second, the chosen
25 threshold (width of the prediction interval) is important. In more typical PFM analyses where the
26 probability of the event is much lower, the threshold will be much more difficult to achieve than
27 in this simple example.

28



1

2 **Figure 4-18** Estimates of the Probability of Axial Crack for $r=5$ Independent Replications
3 Using the Same Sampling Scheme
4



1

2 **Figure 4-19 Prediction Interval Computed from the Five Independent Simulations**

3 **4.3.6 Closed-Form Metric for Simple Random Sampling Uncertainty in a Probability**
 4 **Estimate**

5 *4.3.6.1 What Is It?*

6 When estimating the probability of an event (e.g., some failure scenario of interest), the
 7 sampling uncertainty in the estimate should be well understood to determine model
 8 convergence (Section 4.3.5). Both the sample size and the rarity of the event under
 9 consideration influence the accuracy of the estimate. This section provides a sampling
 10 uncertainty metric for a probability estimate when SRS is applied in uncertainty propagation.

11 *4.3.6.2 How to Use?*

12 Computing this sampling uncertainty metric involves the following three steps:

- 13 (1) Propagate an SRS of size n from the inputs through the model. Record the number of
 14 events and nonevents.
- 15 (2) Estimate the probability of the event using the total number of recorded events divided
 16 by the sample size.
- 17 (3) Compute a sampling uncertainty metric, such as the standard error, the CV, or
 18 confidence interval (see Section 4.3.5).

19 When LHS or adaptive sampling methods are used, care should be taken when estimating
 20 closed-form metrics for sampling uncertainty since assumptions may be violated. Sampling

1 uncertainty is still present when these sampling algorithms are used and should be assessed
2 using alternative approaches (see Sections 4.3.5 and 4.3.7).

3 4.3.6.3 When/Why?

4 When SRS is used for forward propagation of uncertainty, these metrics can be computed to
5 quantify sampling uncertainty in the probability estimate, providing useful insight about the
6 precision of the estimate. Results may suggest that a larger sample or other variance reduction
7 techniques (see Sections 4.3.2 and 4.3.3) are needed if the precision is insufficient.

8 4.3.6.4 Technical Details

9 **Estimate the event probability.** After propagating inputs sampled using SRS through the
10 model and recording whether the event occurred, the probability of failure p_f can be estimated
11 by the ratio of number of failures (n_f) to the number of trials (n), known as a binomial proportion

$$p_f \approx \hat{p}_f = \frac{n_f}{n}.$$

12
13 **Compute the standard error and CV.** The sampling uncertainty of \hat{p}_f relative to p_f decreases
14 as $n \rightarrow \infty$ and $p_f \rightarrow 1$. That is, increasing n will decrease the sampling uncertainty in the
15 estimate, but the relative decrease depends on the failure probability, with smaller p_f resulting in
16 larger relative sampling uncertainty. References 4-52 and 4-53 explain SRS and its associated
17 uncertainty in estimation in detail.

18
19 Assuming each 0/1 outcome is independent, the number of failures n_f can be assumed to follow
20 a binomial distribution. Based on the binomial distribution, the estimated standard error of \hat{p}_f is

$$\hat{\sigma}_{\hat{p}_f} \approx \sqrt{\frac{(1 - \hat{p}_f)\hat{p}_f}{n}}$$

21 The accuracy of this approximation increases as $n/p(1 - p)$ gets large. Given the estimated
22 standard error, the CV is

$$CV(\hat{p}_f) = \frac{\sigma_{\hat{p}_f}}{\mu_{\hat{p}_f}} = \frac{\sqrt{\frac{(1 - p_f)p_f}{n}}}{p_f}$$

23 where $\mu_{\hat{p}_f}$ and $\sigma_{\hat{p}_f}$ are the mean and standard error of the \hat{p}_f , respectively.

24
25 The CV highlights the fact that the relative uncertainty in a probability estimate \hat{p}_f can be quite
26 large, especially when the target p_f is small. For example, for 10,000 simulations of an event
27 with $p_f = 0.01$, $\hat{\sigma}_{\hat{p}_f}$ is about 0.001 (or 10 percent of the desired estimate), but if $p_f = 0.001$, with
28 10,000 simulations $\hat{\sigma}_{\hat{p}_f}$ is about 0.00032 (32 percent of the desired estimate).

29

1 **Compute a confidence interval.** Statistical confidence intervals provide a plausible range in
2 which a parameter is likely to fall based on the observed data. There are many methods for
3 computing confidence intervals for p_f ; Reference 4-54 discusses several in detail. A commonly
4 used approximate $100(1 - \alpha)\%$ confidence interval for a binomial proportion is

$$\hat{p}_f \pm z_{\alpha/2} \hat{\sigma}_{\hat{p}_f}$$

5 where $z_{\alpha/2}$ is the $(\alpha/2)$ th percentile of the standard normal distribution. This confidence interval
6 relies on approximate normality of \hat{p}_f , which is valid only if \hat{p}_f is not too close to 0 or 1. A rule of
7 thumb is to use this interval only if $n\hat{p}_f > 5$ and $n(1 - \hat{p}_f) > 5$; that is, at least five failures and
8 nonfailures are observed. In PFM applications where the true probability of failure is very small,
9 this confidence interval is unlikely to perform well since the number of observed failures under
10 SRS will often be very small. In addition, when np_f is small, zero failures may be observed, and
11 the interval above is meaningless. Reference 4-54 outlines several alternative confidence
12 intervals for binomial proportions. The next paragraph outlines one method for bounding the
13 probability of failure when no failures are observed.
14

15 **Confidence interval when no failures are observed.** If $n_f = 0$ failures are observed in
16 n realizations, then we can use the fact that n_f follows a binomial distribution to place a
17 one-sided confidence interval on the probability of failure. Specifically, there is $100(1-\alpha)\%$
18 confidence that $p_f < p_u$, where

$$p_u = 1 - \alpha^{\frac{1}{n}}$$

19 For example, if it must be established that $p_f < 10^{-6}$ with 95-percent confidence, a simple
20 random sample of size $n = \log(.05) / \log(1 - 10^{-6}) \approx 3 \times 10^6$ with no observed failures is
21 needed. References 4-55 and 4-56 provide more details.

22 **4.3.7 Statistical Bootstrapping**

23 *4.3.7.1 What Is It?*

24 Statistical bootstrapping is a flexible statistical method for calculating sampling uncertainty in a
25 QoI estimate. Bootstrapping relies on resampling from the observed data to calculate QoI
26 uncertainty and is particularly useful when closed-form metrics (as described in Section 4.3.6)
27 are difficult or impossible to derive.
28

29 While there are many versions of bootstrapping, the general idea is to repeatedly resample from
30 the observed data, each time estimating the QoI. The variability in the QoI estimates across
31 bootstrap resamples provides a measure of sampling uncertainty.

32 *4.3.7.2 How to Use?*

33 The most common bootstrap method is to resample directly from the observed data. This form
34 of bootstrapping has three steps:

- 35 (1) Take a sample from the observed data. The sample size is the same size as the
36 observed data. The sample is taken with replacement, where single observations in the

1 data can be included multiple times in a single bootstrap resample. The sampling at this
2 step should be consistent with the way the data were generated.

3 (2) Calculate the QoI from the sampled data.

4 (3) Repeat steps 1 and 2 many times. Use the collection of calculated QoIs to approximate
5 the sampling uncertainty in the QoI. For example, the standard deviation of the collection
6 of QoIs is an estimate of the standard error.

7 4.3.7.3 *When/Why?*

8 Bootstrapping offers a flexible method for estimating sampling uncertainty. The following are the
9 main reasons to use bootstrapping:

- 10 • The algorithm is generic, so it can be applied to most QoIs and many sampling
11 schemes.
- 12 • The algorithm is simple to implement, requiring only the ability to resample from the data
13 and repeatedly calculate the QoI.
- 14 • Closed-form metrics for sampling uncertainty (e.g., Section 4.3.6 for a probability
15 estimate) are difficult to estimate without violating assumptions in many cases.

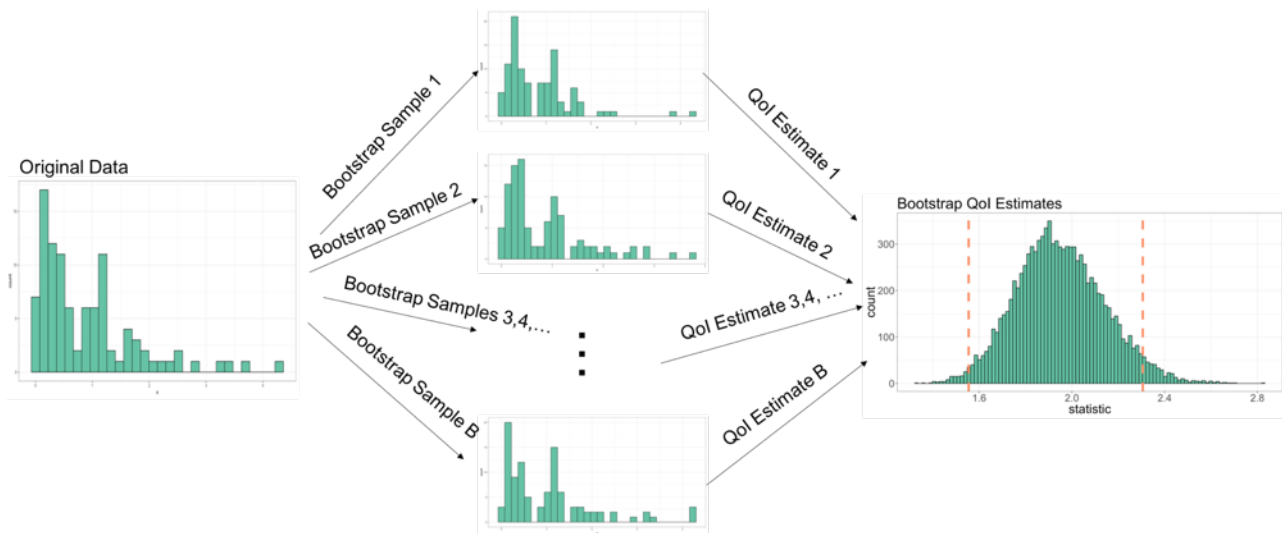
16 However, it is important to understand when not to use bootstrapping. The bootstrap will result
17 in inaccurate measures of sampling uncertainty when either of the following is true:

- 18 • The sample size is small (i.e., sparse data).
- 19 • The original sample was drawn using a complex sampling scheme that cannot be
20 resampled (e.g., LHS).

21 Section 4.3.7.4 contains more information about conditions for bootstrap failure.

22 4.3.7.4 *Technical Details*

23 Figure 4-20 depicts the steps of the bootstrap. The left plot is a histogram of the original data.
24 The middle histograms displayed vertically represent B different bootstrap samples of the
25 original data. Each of these has the same sample size as the original data. However, these
26 samples are taken with replacement, meaning that some values may be observed more than
27 once. The variation among these histograms is an estimate of the sampling variation of the
28 observed sample. For each of the bootstrap samples, an estimate of the QoI is computed and is
29 aggregated in the histogram on the right. This right-most histogram is an estimate of the
30 sampling uncertainty in the estimate of the QoI. The vertical dashed lines are the 0.025 and
31 0.975 quantiles of the bootstrap QoI estimates and correspond to an estimate of a 95-percent
32 confidence interval.
33



1
2 **Figure 4-20 Visualization of the Steps Taken for the Standard Statistical Bootstrap**

3
4 The following are the most common ways the bootstrap fails:

- 5 • Data are too sparse.
6 • The resampling does not reflect how the data were generated.

7 The bootstrap can underestimate uncertainty when data are sparse. Specifically, standard
8 errors will be too small, and confidence intervals will be too narrow. In PFM applications, sparse
9 data are likely to occur when the number of computer realizations is small relative to the QoI.
10 More samples are needed for estimating rare event probabilities and extreme percentiles,
11 because the sparsity of the data is not judged based on the overall number of model realizations
12 but on the overall number of events of interest that occur. Therefore, we may need many more
13 than 1×10^6 simulated realizations to accurately quantify uncertainty in a 1×10^{-6} probability.
14

15 Further, the bootstrap will not accurately estimate sampling uncertainty unless, in the
16 resampling step, the resampling reflects how the data were generated. In PFM applications, the
17 bootstrap can be used with both simple random samples and importance sampling. The
18 bootstrap cannot provide accurate uncertainty quantification in complex sampling schemes such
19 as LHS because there is no way to resample from the observed data in a way that
20 approximates the original LHS scheme.
21

22 More technically, the major assumption of the bootstrap is that, by resampling from the data, we
23 are constructing samples that approximate the empirical distribution of the data. When data are
24 sparse, we cannot approximate this distribution well. When data are generated from a complex
25 sampling scheme such as LHS, we cannot resample from the data in a way that approximates
26 the empirical distribution of the original sample obtained using LHS.
27

28 **Bootstrap confidence intervals.** Confidence intervals are often desired to provide a plausible
29 range in which a parameter is likely to fall based on the sampled data (see Section 4.3.5).
30 Commonly, a probability distribution for the observed data is assumed (either through fitting to
31 data or by expert judgment), and confidence intervals can be derived directly from this
32 assumption. If the choice of probability distribution does not have a strong basis, the bootstrap

1 is an alternative approach as it bypasses the need to analytically derive confidence intervals
 2 using an assumed probability distribution.

3
 4 For example, suppose the QoI is the mean μ from a population from which a sample of data of
 5 size n is collected: $x^{(1)}, \dots, x^{(n)}$. If it is assumed the population is normally distributed, then the
 6 analytically derived $100(1 - \alpha)\%$ confidence interval is

$$\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}},$$

7
8
9
10

11 where \bar{x} and s are the sample mean and standard deviation and $t_{n-1, \alpha/2}$ is the $\alpha/2$ quantile of a
 12 t-distribution with $n - 1$ degrees of freedom.

13
 14 The nonparametric bootstrap approach to the above problem takes a sample of the data of size
 15 n with replacement B times (commonly 1,000 or more), each time computing the sample mean.
 16 This procedure results in a collection of sample means from which confidence intervals can be
 17 constructed. The simplest, but often least accurate, approach to constructing bootstrap
 18 confidence interval for a QoI θ is using empirical quantiles of the bootstrap distribution

$$(\theta^*_{\alpha/2}, \theta^*_{1-\alpha/2}),$$

19
 20
 21 where $\theta^*_{\alpha/2}$ is the $\alpha/2$ th percentile of the bootstrap distribution of θ . Another approach is the
 22 basic method, defined as

$$(2\hat{\theta} - \theta^*_{1-\alpha/2}, 2\hat{\theta} - \theta^*_{\alpha/2}),$$

23
 24
 25 where $\hat{\theta}$ is the estimate of θ from the original sample and $\theta^*_{1-\alpha/2}$ is the $1 - (\alpha/2)$ th percentile of
 26 the bootstrap distribution for θ . Alternatively, the Studentized method for estimating a
 27 confidence interval can be calculated as

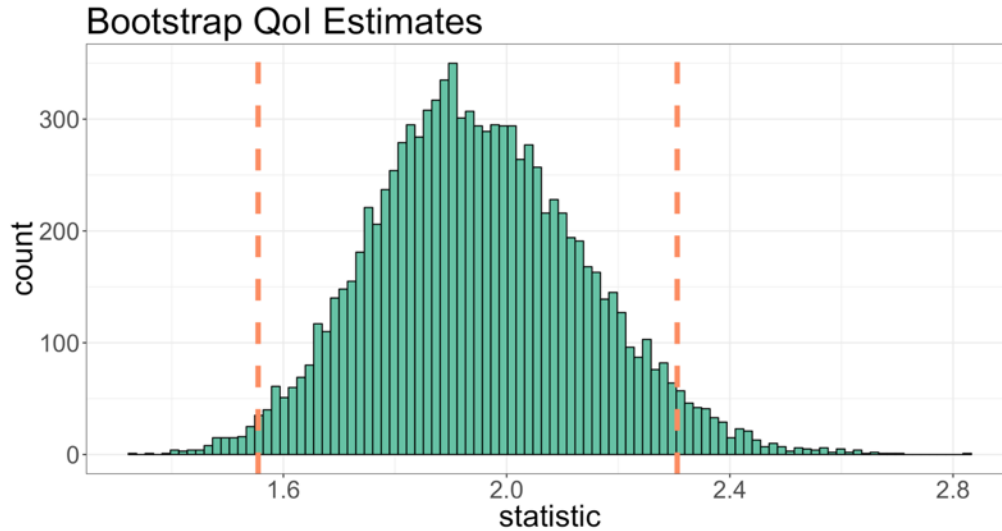
$$(\theta - t^*_{1-\alpha/2} s\hat{e}_\theta, \theta - t^*_{\alpha/2} s\hat{e}_\theta),$$

28
 29
 30
 31 where $t^*_{1-\alpha/2}$ is the $1 - (\alpha/2)$ th percentile of the bootstrapped Student's t-test
 32 $t^* = (\hat{\theta}^* - \hat{\theta})/s\hat{e}_\theta$. Here, $\hat{\theta}^*$ and $s\hat{e}_\theta$ are the estimate and the standard error, respectively, of the
 33 bootstrap distribution of θ . There are many other ways to construct bootstrap intervals, each
 34 with their own advantages and disadvantages. References 4-57 and 4-58 provide more
 35 information.

36
 37 When no closed-form expression is available for a confidence interval, bootstrapping is often a
 38 simple solution for obtaining a confidence interval. As an example, suppose the QoI was some
 39 function of the population mean and standard deviation, such as $(\mu^2 + \mu)/\sigma$. For each bootstrap
 40 sample, the estimated QoI $(\bar{x}^2 + \bar{x})/s$ is computed.

41
 42 As an example, Figure 4-21 shows the bootstrap distribution of this statistic from a sample of
 43 size 100 of data from a normal distribution with mean and variance 1. Using the bootstrap
 44 distribution, a confidence interval for $(\mu^2 + \mu)/\sigma$ can be easily calculated (see References 4-57
 45 and 4-58). The vertical dashed lines in the figure show a 95-percent bootstrap interval. In this
 46 example, the true value of $(\mu^2 + \mu)/\sigma$ is known to be 2, which is clearly within the bootstrap
 47 confidence interval.

48



1
 2 **Figure 4-21 Bootstrap Sampling Distribution along with a 95-Percent Confidence Interval**
 3 **for the Complex Estimator Example**

4 **4.3.8 Global Sensitivity Analysis**

5 *4.3.8.1 What Is It?*

6 Sensitivity analysis (SA) seeks to answer a fundamental question: how sensitive is a model to
 7 its input parameters and which inputs are most important (Reference 4-59)? SA can be used to
 8 identify the inputs that have the strongest impact on the outputs (i.e., most sensitive or important
 9 inputs). Further, SA can help understand the nature of the input-output relationship. Global SA
 10 is used to quantify the amount of output uncertainty that can be attributed to uncertainty in the
 11 input variables (Reference 4-59).

12 *4.3.8.2 How to Use?*

13 Before performing a SA, it is important to choose a relevant output to analyze. The output
 14 should be closely related to the QoI. Further, binary/categorical outputs inherently contain less
 15 statistical information than continuous outputs. Frequently, the binary output is a function of
 16 continuous outputs, and these continuous outputs can often provide better information on input
 17 sensitivity with fewer samples. Because of this, it is generally beneficial to use continuous
 18 outputs for SA when possible.

19
 20 After an output has been chosen, an SA can be performed using exploratory data analysis and
 21 global sensitivity metrics estimation:

- 22 • **Exploratory data analysis.** Exploratory data analysis summarizes characteristics of the
 23 input-output relationships using summary statistics and visualizations (Reference 4-60).
 24 Perhaps the most useful visualization to understand the relationship between PFM
 25 inputs and outputs is a scatterplot. If the number of input and output variables is small,
 26 scatterplots can be produced for each output with each input. With many inputs and
 27 outputs, relevant visualizations may be chosen based on subject matter knowledge.
 28 Alternatively, one can estimate the global sensitivity metrics first and use these to
 29 choose the visualizations.

1 • **Global sensitivity metrics estimation.** Sensitivity metrics provide a quantitative value
2 that characterizes the relationship between inputs and outputs. The following two metrics
3 can be used to quantify the input/output relationship (References 4-59, 4-61, 4-62):

4 – First-order sensitivity indices refer to the proportion of the variance in the output
5 that is explained by the variance in a single input.

6 – Total-order sensitivity indices refer to the proportion of the variance in the output
7 that is explained by the variance in an input and its interactions with other inputs.

8 Section 4.3.8.4 includes details about estimating these sensitivity metrics.

9 4.3.8.3 *When/Why?*

10 SA can be performed to achieve the following:

11 • Understand the problem drivers and rank inputs based on the magnitude of their effect
12 on the output(s).

13 • Improve the precision and accuracy of uncertainty propagation by doing the following:

14 – identifying important inputs whose uncertainty distributions may need further
15 refinement

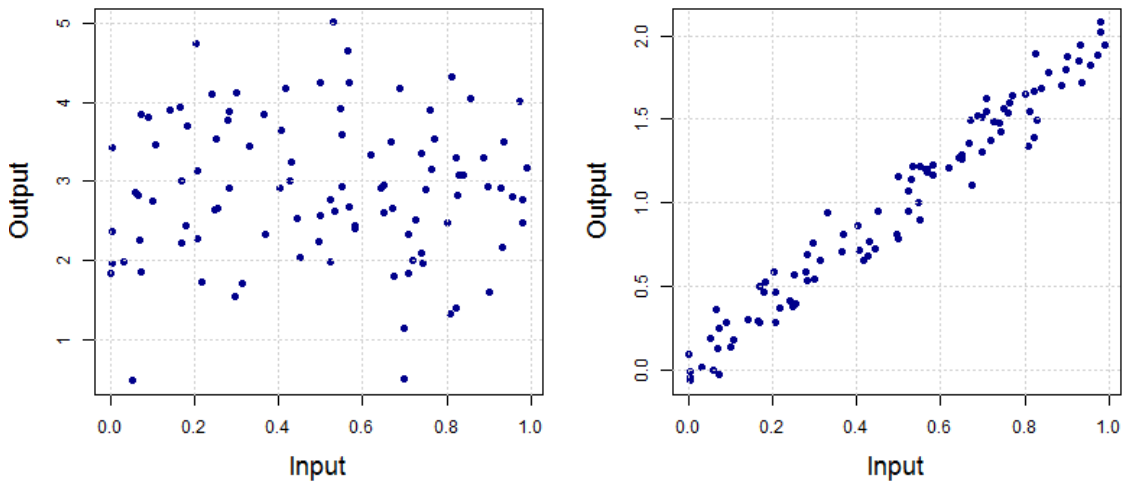
16 – determining candidate inputs for importance sampling

17 4.3.8.4 *Technical Details*

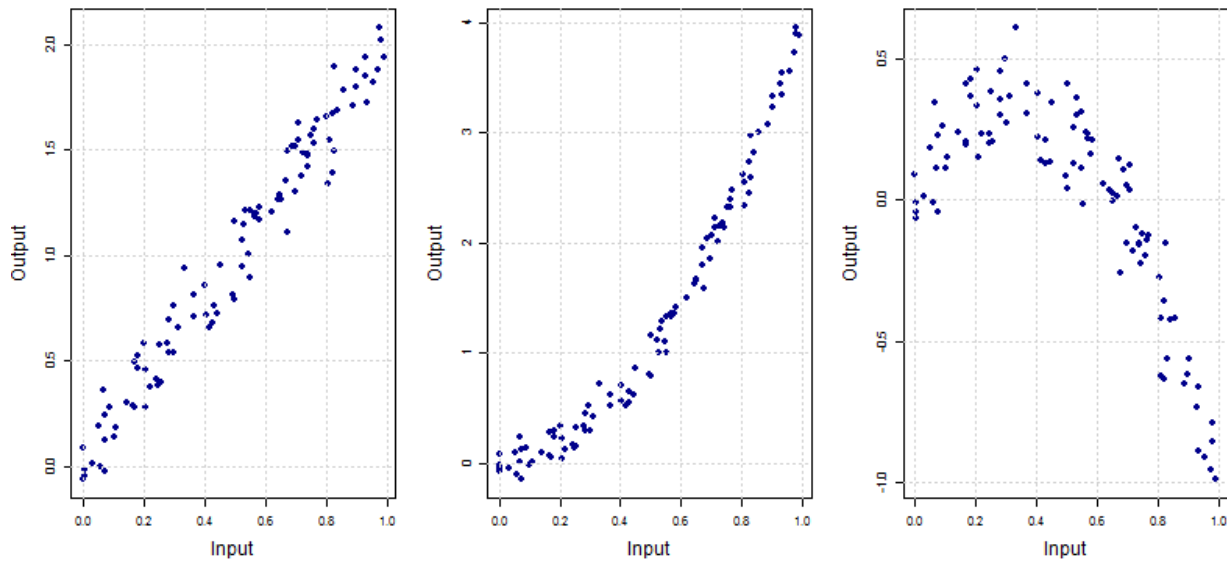
18 **Exploratory data analysis.** Scatterplots can be used to visually assess the nature and
19 magnitude of the relationship between an input and output.

20
21 Figure 4-22 shows an example of an input without (left) and with (right) a strong relationship
22 with the response. For important inputs, scatterplots can also be used to determine whether the
23 relationship is roughly linear, monotonic (i.e., entirely increasing or decreasing), or more
24 complex.

25
26 Figure 4-23 shows examples of linear (left), nonlinear/monotonic (middle), and
27 nonlinear/nonmonotonic (right) input/output relationships. Reference 4-63 gives formal
28 procedures for analyzing scatter plots. In practice, it can be difficult to visually inspect a large
29 number of scatter plots, and more complex relationships involving interactions can often be
30 missed. Estimating sensitivity metrics can help identify the most important relationships to
31 visualize.
32



1
 2 **Figure 4-22 Scatterplots Showing an Input Without (Left) and with (Right) a Significant**
 3 **Relationship with the Output Variable**
 4



5
 6 **Figure 4-23 Scatterplots Showing Linear (Left), Nonlinear/Monotonic (Middle), and**
 7 **Nonlinear/Nonmonotonic (Right) Relationships between the Input and Output**
 8 **Variables**
 9

10 **Global sensitivity metrics estimation.** Input sensitivity can be measured in a variety of
 11 different ways (References 4-64 and 4-65). Variance-based indices are common sensitivity
 12 metrics that decompose the output variance and attribute this variance to certain inputs.
 13 Heuristically, the first-order sensitivity index reflects the proportion of the total output uncertainty
 14 that is explained by the uncertainty in the input x_j alone. The total effects sensitivity index

1 reflects the fraction of the output uncertainty that is explained by x_j by itself and together with its
 2 interaction with other variables.
 3
 4 Mathematically, the first-order and total effects sensitivity indices can be described as follows.
 5 Suppose the output of the computer model is

$$y = f(\mathbf{x}), \quad (3)$$

6 where y is the model output, $\mathbf{x} = [x_1, \dots, x_p]$ is a vector of p input variables, and f is the model.
 7 The first- and total-order sensitivity indices for an input x_j (denoted S_j and T_j , respectively) are
 8 defined as

$$S_j = \frac{\text{Var}(E[f(\mathbf{x})|x_j])}{\text{Var}(f(\mathbf{x}))}, \quad (4)$$

$$T_j = \frac{E(\text{Var}[f(\mathbf{x})|x_{(-j)}])}{\text{Var}(f(\mathbf{x}))} = 1 - \frac{\text{Var}(E[f(\mathbf{x})|x_{(-j)}])}{\text{Var}(f(\mathbf{x}))}, \quad (5)$$

9 where $x_{(-j)}$ is a vector of all p input variables, excluding the j^{th} input (Reference 4-59). The
 10 numerator of the first-order sensitivity metric in Eq. 4 is the variance of $E[f(\mathbf{x})|x_j]$, the average
 11 value of the output $f(\mathbf{x})$, conditional on the input of interest x_j . This variance is taken with
 12 respect to the distribution on x_j ; therefore, the numerator measures how much the average
 13 output varies as x_j varies. A large variance indicates x_j affects the output $f(\mathbf{x})$ and a small
 14 variance indicates it does not. The meaning of “large” and “small” is relative to the total variation
 15 of the output, the denominator of Eq. 4.
 16

17 S_j reflects the proportion of the total output uncertainty that is explained by the uncertainty in x_j
 18 alone, though a similar metric, T_j , is used to assess the proportion of uncertainty in the output
 19 explained by x_j and its interactions with other variables. The numerator in the total-order
 20 sensitivity metric in Eq. 5 is the expectation (average) of $\text{Var}[f(\mathbf{x})|x_{(-j)}]$, the variance of the
 21 output $f(\mathbf{x})$ given all but the j^{th} input. If there is high variation for a wide range of $x_{(-j)}$, then the
 22 outer expectation will be large, resulting in a large value for T_j . If T_j is much larger than S_j , it
 23 implies that there are significant interactions between x_j and the other inputs.
 24

25 **Estimating sensitivity metrics using surrogate models.** The calculation of first- and
 26 total-order sensitivity indices involves the estimation of high-dimensional integrals representing
 27 the expectations and variances in Eq. 4 and Eq. 5. The Monte Carlo integration approaches
 28 detailed in References 4-59 and 4-66 can be used to estimate these indices. However, this can
 29 be computationally prohibitive for many applications because it requires a large number of
 30 realizations when there are a large number of inputs. To address this problem, surrogate
 31 models can be used to estimate the indices (References 4-67 and 4-68). Section 4.3.10
 32 contains more information on surrogate models.
 33

34 **Active subspaces.** At times, the dimensionality of the input space may be prohibitively large for
 35 performing SAs. If this is the case, dimensionality reduction methods such as active subspaces
 36 may be useful. Active subspaces are a way to identify important directions of the input space

1 that affect the QOI. Directions that are not important can be ignored, resulting in reduced
2 dimensionality. References 4-69 and 4-70 provide more information on active subspaces.

3 **4.3.9 Local Sensitivity Analysis**

4 *4.3.9.1 What Is It?*

5 Local SA specifically focuses on how changes to each input at or near a specific reference point
6 in the input domain, like a mean or median, affect outputs of interest (Reference 4-71).
7 Alternatively, global SA attempts to quantify the effects of the uncertain inputs on the output
8 relative to the entire input space (Section 4.3.8).

9 *4.3.9.2 How to Use?*

10 Local SA determines the rate of change of a specified output with respect to a given model
11 input. The aim is to compute the partial derivative with respect to the input at a specific point in
12 the domain (input space). One method for computing this partial derivative for a single input
13 involves the following steps:

- 14 (1) Run the model at the specified value of the input.
- 15 (2) Perturb the input and run the model again.
- 16 (3) Measure the change in the output by estimating the partial derivative.

17 Typically, the other inputs remain fixed during this process, and the measured change in the
18 output is attributed to a single input, conditional on the values of the other inputs.

19 *4.3.9.3 When/Why?*

20 Local SA is a relatively efficient first step toward learning about the important parameters in a
21 model. With only two evaluations of a model, the linear effect on an output of changing a single
22 parameter can be estimated. This step can help down-select to a smaller set of parameters to
23 study in a full uncertainty analysis. It also provides some physical intuition for how certain
24 parameters affect the output. However, the local nature of this analysis should always be kept in
25 mind because (1) a parameter with low local sensitivity can still have a major effect on an output
26 of interest if its associated uncertainty is large, and (2) the local sensitivity of a parameter can
27 sometimes change significantly over the domain of interest. Global SA informs the effect of the
28 parameter over its full uncertainty range and across the entire domain.

29 *4.3.9.4 Technical Details*

30 **Calculating local sensitivity metrics.** Local SA only requires a small number of model
31 evaluations. First, a nominal input value is chosen and that value is perturbed by some amount
32 in one direction (i.e., perturb one dimension of the input space). The amount chosen should be
33 large enough so that a significant change in the output can be observed, but it should be small
34 enough to stay within the region of the input space of concern.

35
36 The sensitivity is measured by the partial derivative, which is estimated by a finite difference.
37 For example, for some output of interest h , the sensitivity of a single input x_i at
38 $\mathbf{x} = (x_1, x_2, \dots, x_p)$ is approximated by perturbing x_i by an amount δx_i and approximating the
39 partial derivative with the finite difference:

$$\frac{\partial h(\mathbf{x})}{\partial x_i} \approx \frac{h(x_1, \dots, x_i + \delta x_i, \dots, x_p) - h(\mathbf{x})}{\delta x_i} \quad (6)$$

1 Partial derivatives can be compared across a set of inputs by repeating Steps 1–3 in
 2 Section 4.3.9.2, perturbing a single input each time. Approximating the partial derivative with a
 3 finite difference is effectively a polynomial approximation using a Taylor series expansion
 4 around the reference point (Reference 4-71, 4-72). Note that the reference point (the model
 5 realization at the nominal input values) can be reused for the computation for each input. From
 6 the results of the local SA, the inputs can be ranked in terms of their contribution to an output of
 7 interest.

8 **4.3.10 Surrogate Models**

9 *4.3.10.1 What Is It?*

10 Surrogate models (also known as emulators, metamodels, and response surfaces) are relatively
 11 fast statistical models that approximate more complex computer models. Surrogates are less
 12 computationally expensive to evaluate than the computer model and can be useful for SA and
 13 uncertainty propagation when conducting a sufficient number of computer model realizations is
 14 computationally prohibitive.

15 *4.3.10.2 How to Use?*

16 Surrogate models are constructed using the following steps (with more information in
 17 Section 4.3.10.4):

- 18 (1) Generate training data by running the computer model using several sets of input values
 19 and obtaining the corresponding output values.
- 20 (2) Use the training data to construct the surrogate model.
- 21 (3) Validate the surrogate on a new set of computer model realizations (testing data) to
 22 check its quality.

23 A surrogate can be used to approximate the full computer model for SA and uncertainty
 24 propagation. The choice of surrogate is dependent on the assumptions the user is willing to
 25 make, which relate to the type of output and the complexity of the input/output relationship. Two
 26 different output types are commonly seen in PFM models:

- 27 (1) **Continuous data** can take on an infinite number of possible (physical) values
 28 (e.g., crack length). Common surrogates for continuous data include linear regression,
 29 multivariate adaptive regression splines, and GP regression.
- 30 (2) **Binary data** take on only two levels for the output. Typically, the binary variable is an
 31 indicator for an event, taking on 0 if the event did not occur and 1 if it did (e.g., rupture or
 32 no rupture). Surrogates for binary data model the probability of the event occurring. An
 33 example surrogate for binary data is a generalized linear model.

1 4.3.10.3 When/Why?

2 Surrogate models can be used to decrease computation time through building a computationally
3 efficient computer model approximation. In PFM applications, surrogates can be used in SA and
4 uncertainty propagation:

- 5 • In SA, surrogates can be used to determine how uncertainty in the inputs affects
6 uncertainty in the outputs (Sections 4.3.8 and 4.3.9).
- 7 • In uncertainty propagation, surrogates can be used for propagating uncertainty in the
8 inputs through the computer model (Section 4.3). Input samples are propagated through
9 the surrogate model rather than the full computer model to allow for many more
10 evaluations.

11 Surrogate models approximate the computer model, and this approximation adds uncertainty in
12 the PFM analysis. Surrogate uncertainty should be considered in the interpretation of the results
13 under the following conditions:

- 14 • If surrogates are used for SA, several different surrogates can be tried to explore the
15 sensitivity of the SA results to the selected surrogate model.
- 16 • If surrogates are used for uncertainty propagation, the magnitude of error associated
17 with the surrogate model approximation can be quantified and included as additional
18 uncertainty in the estimation of the QoI.

19 4.3.10.4 Technical Details

20 **Generate training data.** An output y of a physical process or computer model can be
21 represented as a function of the input x :

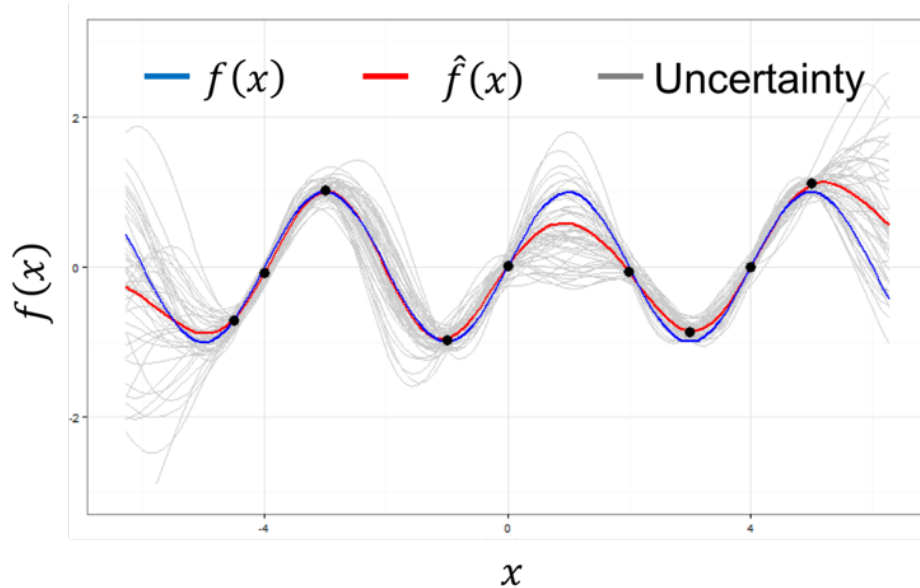
$$22 \quad y = f(x).$$

23
24 The representation here is deterministic; given the same value of x , the same value of the
25 output y will result. A surrogate estimates the true process function f statistically using a set of
26 training data $(y^{(i)}, x^{(i)})$, $i = 1, 2, \dots, n$ where $x^{(i)}$ is the i^{th} set of inputs on which a computer
27 model of the process is evaluated resulting in the output $y^{(i)}$.

28
29 The accuracy of the surrogate increases with the size of the training set. For continuous
30 outputs, a general rule of thumb for the number of data points n is approximately $10p$, where p
31 is the number of input variables (Reference 4-73). Reference 4-74 gives an overview of options
32 for choosing the input combinations that will be used in constructing the surrogate model. A
33 useful and common choice is an LHS (Section 4.3.2). This section outlines several options for
34 constructing a surrogate.

35
36 **Construct the surrogate model.** The training data are used to fit a statistical model
37 approximating $f(x)$ for any x . The choice of surrogate model to use will depend on several
38 aspects of the problem, such as the type of output variable (e.g., continuous or binary),
39 continuity or discontinuity of f , the size of the training data set, and the domain on which a
40 surrogate is required. This section discusses several types of surrogate models. Ideally,
41 uncertainty in the surrogate model predictions are measured. An example surrogate appears in
42 Figure 4-24 where the black points represent the training data, the blue curve represents the
43 true computer model $f(x)$ across the entire input space, the red curve is the surrogate estimate

1 $\hat{f}(x)$, and the gray curves represent statistical uncertainty in the surrogate estimate. This
 2 particular surrogate, a GP, interpolates the training data and has the intuitive property that
 3 statistical uncertainty is larger for locations farther away from the training data
 4 (e.g., Reference 4-75).
 5



6
 7 **Figure 4-24 GP Surrogate Fit to Training Data (Black Points) from the True but Unknown**
 8 **Function $f(x)$**
 9

10 **Surrogate validation.** Validation of the surrogate can be done using the following steps:

- 11 (1) Use the surrogate to predict the response at a set of new input values x not used in
 12 construction of the surrogate.
- 13 (2) Run the full computer model at x .
- 14 (3) Compare the predicted response using the surrogate to the response using the
 15 computer model.
- 16 (4) Determine whether the surrogate is sufficiently accurate. If not, then more realizations
 17 from the computer model are needed to improve the surrogate, or a different surrogate
 18 model is needed.
- 19 (5) If the surrogate is used for uncertainty propagation, the error associated with the
 20 surrogate's approximation should be considered when quantifying uncertainty in the QoI.

21 Reference 4-74 provides more information on surrogate validation.

22
 23 It is important to ensure that the surrogate model is properly approximating the computer model
 24 by checking for potential over- or under-fitting of the surrogate, and multicollinearity. Overfitting
 25 refers to a surrogate representing the training data set so well that the surrogate does not
 26 generalize to new datasets and has low prediction capabilities. Underfitting refers to a surrogate

1 that does not represent the training data well and therefore also does not generalize well.
2 Multicollinearity may arise when two or more independent input variables in a surrogate model
3 are correlated. This is potentially concerning because multicollinearity can result in unstable and
4 unreliable output results.

5
6 Iterating between the surrogate model construction and validation steps is necessary to develop
7 the most appropriate surrogate model for approximating the computer model.

8
9 **Surrogate models for continuous data.** There are many types of surrogates for continuous
10 data. Examples of surrogate models include linear regression (Reference 4-76), multivariate
11 adaptive regression splines (MARS) (Reference 4-77), and GPs (Reference 4-75).
12 References 4-78 and 4-79 provide detailed overviews of these and other techniques, as well as
13 details on how to use these surrogates for SA. The following gives a brief description of them:

- 14 • **Linear regression** is a statistical surrogate that models the output as a linear function of
15 the inputs and tends to be one of the more interpretable models. It includes uncertainty
16 in the coefficients and allows for uncertainty estimates in the outputs. Linear regression
17 is often used as an initial screening tool in SA to identify the most important variables
18 and can be used as a surrogate for the computer model.

19
20 Despite the name, it is possible to model interactions and nonlinearities within a linear
21 model. To sort through the many potential model candidates, fit criteria can be used to
22 find the best model. For example, the Akaike Information Criterion (AIC) and the
23 Bayesian Information Criterion (BIC) can be used to quantify model fit to the data, as
24 well as automated methods to find the optimal AIC/BIC, such as stepwise selection.

- 25 • **Multivariate adaptive regression splines (MARS)** is a machine learning (ML) method
26 that is used for flexible nonparametric regression modeling of high-dimensional data.
27 Separate splines are fit to different intervals of the predictor variables. Variables, knots
28 and interactions are evaluated simultaneously to produce an optimal fit. MARS allows for
29 automatic variable selection and transformations and for variable interactions.

- 30 • **Gaussian process (GP) regression** is an ML method that assumes the input-output
31 relationship can be modeled as a GP, which is a specific type of multivariate normal
32 distribution. Specifically, correlation between the outputs is induced using a correlation
33 structure that is a function of the inputs. The correlation structure is constructed such
34 that inputs close together produce more similar outputs. GP is a flexible tool for
35 interpolating outputs throughout the parameter input space. A primary disadvantage of
36 GP is that it can become computationally expensive and unstable with large training sets
37 or many inputs. Dimension reduction approximation techniques can be applied to make
38 GPs more computationally feasible (Reference 4-78).

39 **Surrogates for binary data.** Binary data can arise in PFM applications when the model output
40 is the occurrence of an adverse event (such as crack or rupture). As with continuous data, there
41 are many different options for fitting surrogates to binary data. Because binary data contain less
42 information than continuous data do, more initial computer model realizations (i.e., a larger
43 training sample) are required to accurately model the relationship between inputs and outputs.
44 In particular, to create a surrogate for rare events, more initial computer model realizations are
45 commonly required, along with a strategic sampling plan, such as importance sampling
46 (Section 4.3.3).

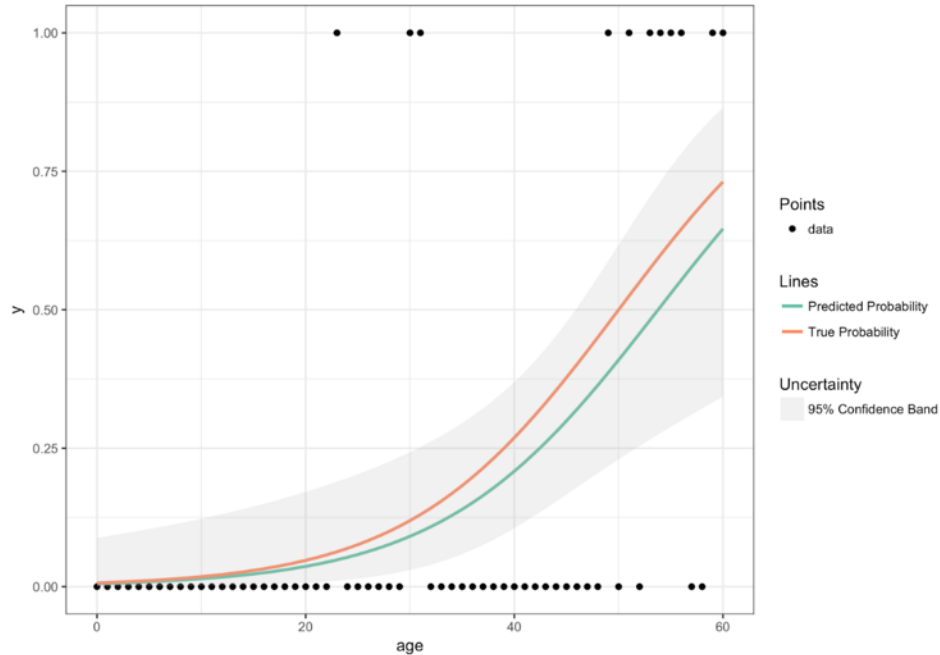
47

1 For example, a generalized linear model is a flexible extension of linear regression that can be
2 used when the response does not satisfy the assumption of having a normal error distribution
3 (e.g., when the response variable is binary). Common examples of generalized linear models for
4 binary data are logistic regression and probit regression (References 4-80, 4-81).

5
6 **Additional methods.** Additional methods (for both continuous and binary data) include the
7 following:

- 8 • **Machine learning** (ML) covers a broad group of flexible techniques that fit complex
9 relationships in the data, with the goal of predicting an unobserved output as accurately
10 as possible. ML methods can be used for both continuous and binary outputs. Examples
11 of ML techniques include the already mentioned MARS and GP models, as well as
12 neural networks (deep learning), regression trees, and support vector machines. Many
13 texts (e.g., Reference 4-82, 4-83, 4-84, 4-85) provide technical details for a wide range
14 of ML and statistical learning methods. Note that whether a model is considered to be an
15 ML model varies from group to group. Some texts also consider select Bayesian models
16 to be ML models. While some ML models may be limited in interpretability and
17 uncertainty quantification, research is underway to improve interpretability and
18 uncertainty estimates for ML models.
19
- 20 • **Bayesian models** integrate information based on probability theory. These models take
21 into consideration prior knowledge with data to produce an output using the Bayes
22 Theorem. The posterior distribution (output) is proportional to the product of the
23 likelihood distribution (probability distribution that represents the observed test data from
24 the computer experiment) and the prior distribution (probability distribution that
25 represents the knowledge before observing the test data). All statistical inferences on
26 the QoI are done on the posterior distribution.

27 Figure 4-25 shows an example of a surrogate prediction for binary data, where the probability of
28 failure is predicted as a function of a single input variable—age. The surrogate models the
29 probability of failure based on observed pass/fail ($y=0$ or $y=1$) outputs; the surrogate model is
30 then compared to predictions of the failure probability from the computer model as validation of
31 the surrogate. The points represent responses for components of varying ages, with a 0
32 meaning the component did not fail and a 1 meaning the component did fail. The orange line
33 shows the true probability of failure (estimated from the computer model). The teal line shows
34 the estimated probability of failure using logistic regression. The gray band represents
35 95-percent confidence bands on the probability.
36



1
 2 **Figure 4-25 Example of a Generalized Linear Model for Binary Data—Component Failure**
 3 **as a Function of Age**

4 **4.3.11 Visualizing Output Uncertainty Due to Input Uncertainty**

5 *4.3.11.1 What Is It?*

6 PFM analyses are conducted to estimate a specific QoI, though QoIs are never estimated as a
 7 single exact value due to uncertainty. Uncertainty analysis is the process of understanding and
 8 documenting uncertainty in a QoI estimate across model realizations. The uncertainty analysis
 9 approach depends on the QoI and the sampling design for the model realizations.
 10 Communication of the uncertainty analysis involves visualization. This section outlines common
 11 techniques for visualizing QoI estimates along with the quantifiable uncertainty in those
 12 estimates.

13 *4.3.11.2 How to Use?*

14 The appropriate visualization technique depends on three considerations:

- 15 (1) whether the analysis separates aleatory and epistemic uncertainties
 16 (2) the type of the QoI (i.e., whether the probability QoI is represented as a function of time
 17 (continuous performance measure) or for a single point in time)
 18 (3) whether the model realizations have equal weight

19 Section 4.3.11.4 describes the appropriate visualization techniques based on these three
 20 dependencies.

1 4.3.11.3 When/Why

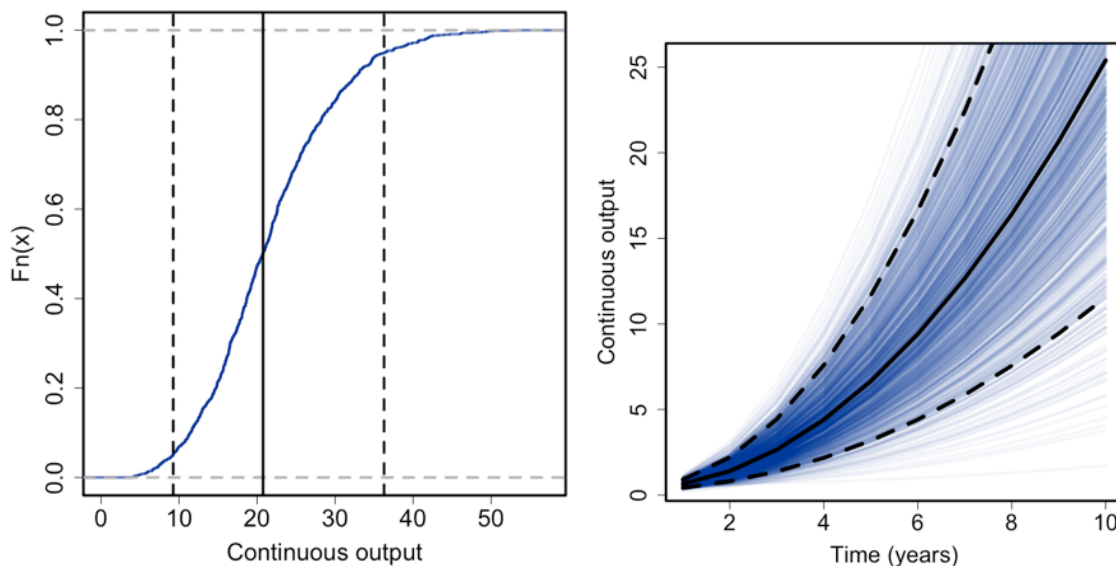
2 Analysts visualize uncertainty and variability throughout a study. Thoughtful visualizations
3 enhance the final results and help communicate the results effectively.

4 4.3.11.4 Technical Details

5 This section distinguishes the cases with and without separation of uncertainty types and
6 discusses the type of the QoI. It also gives an overview of visualizing results with unequal
7 weighting.

8
9 **No separation of aleatory and epistemic uncertainty.** Without separation of uncertainty, the
10 model results will consist of a set of n outputs, where the outputs may be measured over time.

11
12 *QoI is a continuous performance measure.* Often, the QoI is continuous, such as crack length or
13 leak rate. When the output is continuous and is not measured over time, the empirical CDF of
14 the output samples should be plotted. When the output is measured over time, uncertainty in a
15 continuous output can be visualized by plotting the output over time for each of the
16 n realizations and overlaying the best estimate (e.g., a mean or median) and measure of
17 uncertainty (e.g., quantiles of the output) at each time point. Figure 4-26 provides an example
18 for two scenarios: (1) the output is not measured over time, and (2) the output is measured over
19 time. The left plot shows an empirical CDF of a continuous output at a single time point over
20 1,000 realizations, with a solid vertical line at the median and dashed vertical lines at the 5th
21 and 95th percentiles of the output. The right plot shows a continuous output over time over
22 1,000 realizations, with a solid line at the median and dashed lines at the 5th and 95th
23 percentiles of the output.



24

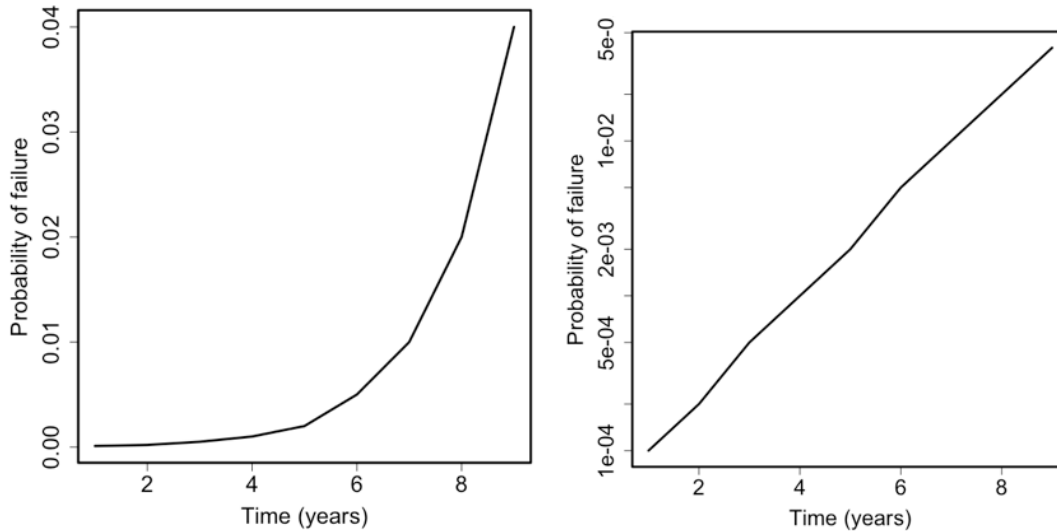
25 **Figure 4-26 Continuous Output at a Single Time Point (Left) and Over Time (Right)**

26

27 *QoI is a probability.* If the QoI is a probability (e.g., probability of failure), the model often outputs
28 a binary 0/1 variable indicating that an event did or did not occur for that realization. The number
29 of the 0/1 outputs divided by n number of samples is then used to estimate the probability of the

1 event. Sampling uncertainty in the probability estimate can be computed using confidence
2 intervals or other methods for computing sampling uncertainty.

3
4 When the Qol is measured over time, a graph of the estimate over time provides insight into
5 how the estimate changes with time (Figure 4-27). When estimating rare probabilities, it is often
6 more informative to plot the estimates on the log scale, so that the order of magnitude of the
7 probability can be easily ascertained from the plot.
8



9
10 **Figure 4-27 Failure Probability Over Time when Aleatory and Epistemic Uncertainty are**
11 **not Separated; Linear Scale (Left) and Log Scale (Right)**

12
13 **Separation of aleatory and epistemic uncertainty.** Separating uncertainties allows for the
14 direct quantification of the impact of epistemic uncertainty (Section 4.1.1). Specifically, when
15 uncertainties are separated using a double-loop algorithm, the set of model realizations will
16 consist of n_e unique epistemic samples and n_a aleatory samples within each epistemic sample.
17 The final sample size is then $n = n_e * n_a$.

18
19 *Qol is a continuous performance measure.* An estimate of the Qol, say Q_i , is computed from the
20 output across the aleatory samples for each unique epistemic sample $i = 1, 2, \dots, n_e$. A best
21 estimate of the Qol is a measure of centrality (e.g., the mean or median) of the set of Q_i s. The
22 epistemic uncertainty in the Qol can be represented using percentiles of the Q_i s. Specifically,
23 the median, 5th, and 95th percentile of the Q_i s can be presented as a best estimate and
24 uncertainty for the Qol.

25
26 *Qol is a probability.* When the Qol is a probability, the average of the 0/1 output across the
27 aleatory samples for each unique epistemic sample is computed to estimate the probability of
28 the event, conditioned on the value of the epistemic input. A best estimate of the probability is
29 the mean or median of these estimates. The epistemic uncertainty in the probability can be
30 represented using percentiles of the estimates. Specifically, the median, 5th and 95th
31 percentiles can be presented as a best estimate with uncertainty.

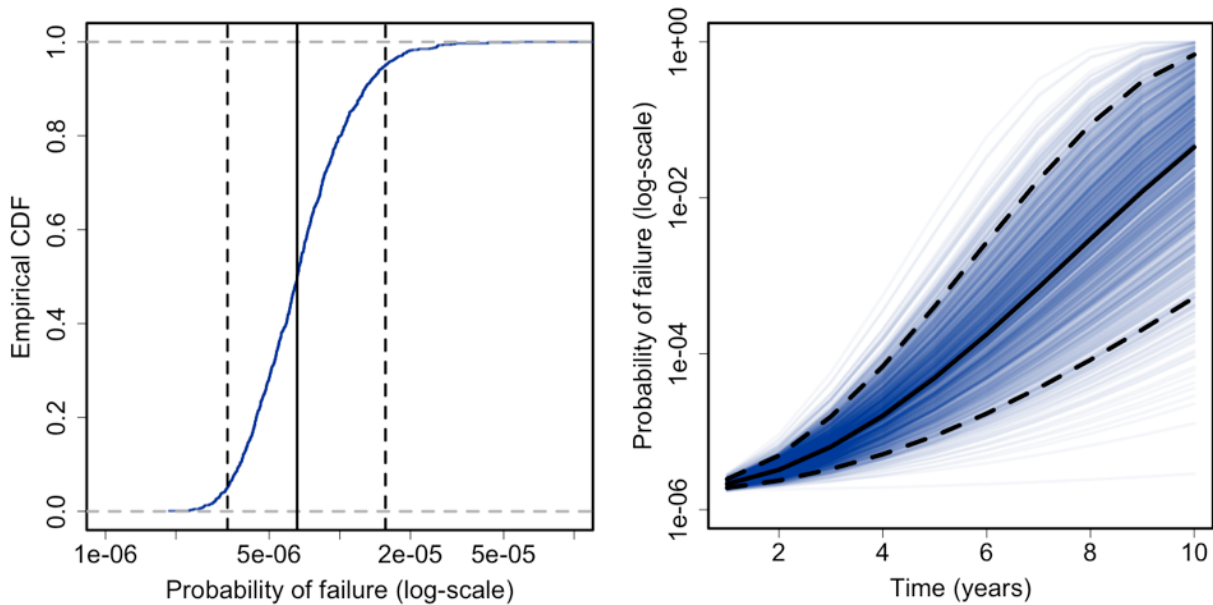
32
33 In general, an estimate Q_i is provided for each epistemic input. These estimates contain
34 sampling uncertainty due to a finite aleatory sample size. The precision of the individual Q_i s

1 should be considered. The number of samples n_a and sampling scheme determine how
 2 accurately each Q_i can be estimated. For example, if the probability of failure is on the order of
 3 1×10^{-3} , then more than 1×10^3 samples ($n_a > 10^3$) will be required to accurately estimate each Q_i
 4 using equal-probability weighted samples of 0/1 outputs (see Section 4.1.1).

5
 6 The left plot in Figure 4-28 displays visualizations for the case when a probability is estimated at
 7 a single time point (or not a function of time) from 0/1 output. The figure plots the CDF of the
 8 estimated probabilities for each epistemic sample out of 1,000 samples. The solid vertical line is
 9 the median, and the dashed vertical lines are the 5th and 95th percentiles of the output. The
 10 estimated probabilities represent frequencies of the event over the aleatory samples. While it
 11 looks similar to the plot in Figure 4-26, its interpretation is different. If the aleatory sample size is
 12 large enough to make the sampling uncertainty in each estimate negligible, the spread in this
 13 CDF represents the spread due to epistemic uncertainty. Likewise, the plot on the right
 14 visualizes probability estimates as a function of time. Each blue curve represents an estimate of
 15 the probability given a fixed epistemic parameter. The solid line is the median, and the dashed
 16 lines are the 5th and 95th percentiles of the epistemic output. If the aleatory sample size is large
 17 enough, the spread in these blue curves represents uncertainty due to epistemic uncertainty in
 18 the inputs.

19
 20 It is important to understand that if the aleatory uncertainty is a significant contributor to the
 21 uncertainty in each Q_i , the variability observed in these plots is due to both aleatory and
 22 epistemic uncertainty (Reference 4-86).

23



24

25 **Figure 4-28 Frequency over Aleatory Samples at a Single Time Point (Left) and as a**
 26 **Function of Time (Right)**

27
 28 **Weighting model realizations.** When inputs are sampled using SRS or LHS, the generated
 29 inputs all have an equal probability of selection and the outputs are weighted equally when
 30 calculating a QoI. When inputs are sampled using importance sampling or another weighted
 31 sampling method, the model outputs are weighted differently. When calculating the best
 32 estimate and uncertainty in a QoI, the relevant weights should be applied. When visualizing

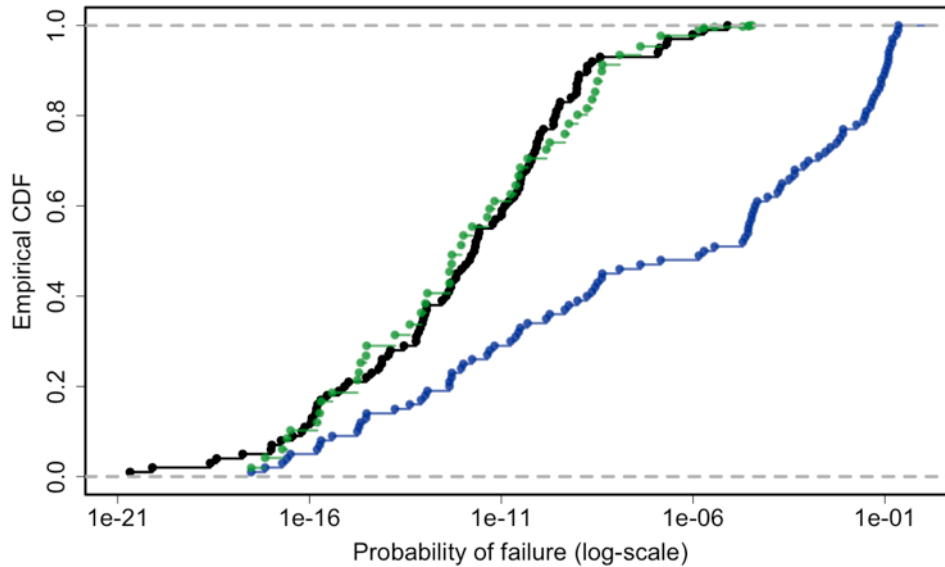
1 uncertainty, realizations from the true output distribution should be plotted, rather than the
2 observed output distribution with unequally weighted outputs. Sampling from the true output
3 distribution can be achieved by sampling with replacement from the observed, weighted data
4 using the following algorithm:

5 (1) Each output i has a corresponding weight w_i based on the selected sampling method
6 (under SRS or LHS with no importance sampling, $w_i = 1$ for each i).

7 (2) The analyst should resample with replacement from the n outputs, where each output
8 has a probability of being sampled proportional to its weight.

9 (3) The resampled data can be considered an unweighted, simple random sample of
10 outputs.

11 Figure 4-29 shows an example of reweighting an importance-sampled distribution. The figure
12 displays empirical CDFs of a probability of failure. The blue CDF is that of the output from the
13 importance sampling. Since the importance sampling oversamples regions where failures are
14 likely to occur, the CDF calculated directly from these outputs results in much larger estimated
15 probabilities than those computed under SRS (black CDF). The green CDF is created by
16 resampling the importance-sampled distribution using the importance sampling weights as
17 described in the algorithm above. As desired, this reconstructed CDF is much closer to the one
18 observed under SRS and actually estimates the distribution of failure probability.
19



20

21 **Figure 4-29** Importance-Sampled Distribution (Blue), Simple Random Sample (Black)
22 Distribution, and Reconstructed Unweighted Distribution (Green) for a
23 Probability of Failure

1 **4.4 Useful Methods for Sensitivity Studies**

2 **4.4.1 Sensitivity Studies**

3 *4.4.1.1 What Is It?*

4 Sensitivity studies are case studies that exercise the PFM computational framework under
5 different assumptions. The goal of sensitivity studies is to determine whether uncertain
6 assumptions impact the conclusions of the PFM analysis.

7 *4.4.1.2 How to Use?*

8 Key aspects of conducting sensitivity studies include the following:

- 9 • determining the set of uncertain assumptions that will be evaluated using sensitivity
10 studies
- 11 • designing and running sensitivity studies

12 **Determining the set of uncertain assumptions.** The complexity of PFM computational
13 frameworks results in a large set of assumptions, some of which may be uncertain and thus
14 candidates for sensitivity studies. Assumption uncertainties can often be categorized as model
15 uncertainty or input uncertainty. Further, uncertain assumptions can often be categorized by the
16 degree of uncertainty in the assumption. Section 4.4.1.4 contains more information on model
17 versus input uncertainty and classifying the degree of uncertainty to determine whether a
18 sensitivity study is needed.

19
20 **Designing and running sensitivity studies.** When setting up a sensitivity study, the settings in
21 the model and inputs are changed to reflect the plausible alternative assumption(s) under study.
22 The analyst has a choice to conduct deterministic model realizations at a single value of the
23 model inputs or to conduct probabilistic analyses over the range of the model inputs. The
24 analyst will select either probabilistic or deterministic analysis for the sensitivity studies based
25 on the change in the assumptions and the specific question being asked.

26
27 The sensitivity studies are designed to evaluate how changing an assumption impacts the
28 results of the analysis. This requires knowledge gained throughout the PFM process as well as
29 subject matter expertise. For example, the inputs whose assumptions are natural candidates for
30 sensitivity studies are those considered important in SAs. Subject matter experts can help
31 determine the credibility of the models and input parameters and identify plausible alternatives.

32 *4.4.1.3 When/Why?*

33 A typical PFM analysis relies on a complex model that consists of many submodels with many
34 inputs and outputs joined together in an overall model framework. The complexity of these
35 models results in a large set of assumptions, some of which may be uncertain. Each submodel
36 and parameter input relies on assumptions that represents a decision to set up the problem and
37 model in a specific way. Since the results of the PFM analysis depend on these uncertain
38 assumptions, the effect of the assumptions should be studied with the goal of understanding
39 whether plausible alternative assumptions will significantly change the results.

1 4.4.1.4 *Technical Details*

2 **Determining a set of uncertain assumptions to study.** PFM analyses contain many uncertain
3 assumptions, but sensitivity studies should not be conducted for all such assumptions. Two
4 primary factors should drive whether sensitivity studies are conducted (Reference 4-87):

- 5 (1) plausibility of assumption violation
- 6 (2) impact on analysis results

7 In general, sensitivity studies should be considered for more plausible assumptions that can
8 impact the QoI. If subject matter experts or SAs cannot determine the plausibility or impact of a
9 particular assumption a priori, a sensitivity study should generally be considered.

10

11 **Types of uncertain assumptions.** To determine a set of plausible alternative assumptions,
12 Reference 4-87 distinguishes between two types of assumptions:

13 (1) **Modeling assumptions** refer to the types of submodels used in the PFM code, the
14 assumptions made to develop each of the submodels, and any approximations made
15 during calculations performed within each of the submodels. Modeling assumptions also
16 include context assumptions that pertain to the context of the PFM analysis. Changes in
17 analysis context are related to completeness uncertainty, defined as “uncertainty caused
18 by the limitations in the scope of the model, such as whether all applicable physical
19 phenomena have been adequately represented, and all accident scenarios that could
20 significantly affect the determination of risk have been identified” (Reference 4-87).
21 Examples of context assumptions for PFM applications include alternate scenarios, such
22 as worst case scenarios and different intervention scenarios (discussed more below).

23 (2) **Input parameter specification assumptions** refer to any assumptions made when
24 specifying the values of the input parameters to propagate through the PFM code. These
25 include the choice to explicitly separate aleatory and epistemic uncertainties and the
26 classification of each variable into these categories, the choice of fixed probability
27 distributions for the inputs, the choice of correlation structure between the inputs, and
28 the choice to treat certain inputs as deterministic (i.e., fixed).

29 **Degree of assumption uncertainty.** Reference 4-87 provides a useful set of categories for
30 models and input parameter assumptions that helps to identify and rank sensitivity studies in
31 terms of their plausibility and impact, as summarized below. The summary describes a list of
32 categories for both models and input parameters, with a corresponding suggestion for whether a
33 sensitivity study is needed:

- 34 • **Model categories:**¹ Models can be categorized according to the uncertainty in the
35 modeling assumptions. Potential categorizations include the following:
36
 - 37 – *The model/submodel is a correct and credible representation of the underlying*
38 *physical process.* Sensitivity studies are typically not needed. There is little
39 benefit in subjecting a correct model to a sensitivity study. This category implies

¹ “Model” and “submodel” are used interchangeably here and should not cause confusion. Typically, it is the individual submodels that are categorized before the categorization of the overall model.

- 1 that there are either no other plausible models or any other plausible model is
2 similar to the current model and would have low impact on the QoI.
- 3 – *The applicability of this model to all conditions of interest cannot be assessed*
4 *reliably with the current state of knowledge.* Sensitivity studies should be
5 considered. The correctness of the model is unknown. It is possible that there is
6 no known plausible alternative model on which to develop a sensitivity study. In
7 such cases, sensitivity studies scrutinizing the engineering decisions made in
8 developing the model can help determine whether these decisions have
9 unforeseen significant effects on the results. That is, there are potentially other
10 plausible engineering decisions that could have been made and that would
11 impact the QoI.
- 12 – *Plausible alternatives to the model adopted exist for a given physical process,*
13 *and these alternatives have roughly equal justification to the model adopted.*
14 Sensitivity studies should be considered. The alternative plausible models with
15 roughly equal justification are usually candidates for sensitivity studies, especially
16 if the model affects the QoI. The alternatives may include context assumptions
17 such as worst case scenarios and intervention scenarios.
- 18 – *A model provides a conservative representation of the underlying physical*
19 *process.* Sensitivity studies might be conducted. Conservative models are often
20 adopted because of a lack of information. It may be necessary to set up studies
21 to quantify the impact of the conservative choices.
- 22 • **Input parameter categories:** Inputs can also be categorized according to uncertainty in
23 their assumptions. Potential categorizations include the following:
- 24 – *The uncertainty distribution for the input parameter accurately represents the*
25 *input for the conditions of interest. Additionally, the choice to classify the input as*
26 *aleatory or epistemic is unambiguous.* A sensitivity study is not needed. This
27 category implies there are no alternatives worth considering for the input
28 parameter specification.
- 29 – *The value or the uncertainty distribution was developed using limited prior*
30 *information or data. Alternatively (or in addition), the choice to classify the*
31 *parameter as aleatory or epistemic is ambiguous.* A sensitivity study should be
32 considered. Given the limited information used to specify the input, plausible
33 alternatives likely exist and are candidates for sensitivity studies. The analyst
34 considers the impact the input has on the QoI when determining whether a study
35 is needed. If the distribution/value is highly uncertain, but SA results and expert
36 judgment agree that the input does not drive variability in the QoI, then a
37 sensitivity study is typically not necessary. Alternatively, if the variable does drive
38 variability in the QoI, a sensitivity study should be conducted.
- 39 – *The distribution for the input parameter is considered a conservative*
40 *representation of the parameter for the conditions of interest.* A sensitivity study
41 might be conducted. Conservative input parameters are often used out of
42 necessity due to a lack of information. As plausible alternative and potentially
43 less conservative input specifications exist, quantifying the impact of these
44 conservatisms could be helpful in building credibility.

1 **Designing sensitivity studies.** Sensitivity studies are designed based on the question of
2 interest. Typical questions asked in sensitivity studies include the following:

- 3 • Do the results change significantly if a plausible alternative model is used? (*Step 1:*
4 *Action 3*)

5 If assumptions about the underlying code or physics model may be violated for the
6 specific application, then sensitivity studies can address how the QoI changes under
7 different model form assumptions (e.g., geometric fidelity, material model selection, new
8 submodels). Sensitivity studies can demonstrate that the overall behavior of the PFM
9 code is consistent with expert understanding of the expected system behavior, including
10 demonstrating expected trends and correlations between inputs and outputs of interest.
11 Benchmarking against other comparable codes may be used to increase confidence in
12 the PFM code by demonstrating that the results produced by the PFM code are
13 reasonable and can be predicted by similar codes (Reference 4-88).

- 14 • Do the results change significantly if a different distribution is used for an important input
15 variable? (*Step 2: Action 2*)

16 Sensitivity studies that vary the type of input distribution or distribution parameters can
17 be conducted to determine the impact of the chosen distribution. The analyst should
18 consider changing the characteristics of input distributions (e.g., shifting the mean,
19 variance, or other distribution moments, such as skewness and kurtosis) as well as
20 changing the distribution itself to highlight the uncertainty in specifying the distribution
21 correctly.

- 22 • Do the results change significantly if a variable is considered aleatory rather than
23 epistemic or vice versa? (*Step 2: Action 1*)

24 If the analysis maintains the separation of aleatory and epistemic uncertainty and the
25 uncertainty of an input cannot be clearly defined as aleatory or epistemic, then sensitivity
26 studies can address how the analysis results change depending on the classification of
27 this uncertainty type.

- 28 • Do the results change significantly under different context assumptions used to set up
29 the problem (*Step 1: Action 1*)

30 Examples of alternate scenarios could include the following:

- 31 – *Worst case scenarios* or any adverse condition (such as accidents) are often
32 considered for sensitivity studies in support of defense in depth. They are either
33 designed by experts or found in benchmarking studies.
- 34 – *Intervention scenarios* study the impact of some (usually positive) changes in the
35 system, such as inspection or mitigation.
- 36 – *Defense-in-depth scenarios* involve changes to nominal assumptions to
37 represent adverse conditions or beyond-design-basis conditions. Such studies
38 can be combined with different intervention scenarios to assess the benefit of the
39 interventions under extreme conditions.

1 Sensitivity studies are important for both deterministic and probabilistic fracture analyses. For
2 example, Reference 4-89 describes sensitivity studies conducted to understand the effects of
3 potential changes to selected inputs and mechanisms in calculating failure probabilities under
4 different inservice inspections programs. Sensitivity studies have also been used in support of
5 defense in depth by considering beyond-design-basis accidents (see Reference 4-9 for some
6 examples). Sensitivity studies seek to assess the credibility of the PFM model and analysis
7 within the domain of the application. This is different from credibility in the context of V&V or
8 statistical and numerical stability, where the goal is to build trust that the results are computed
9 correctly and accurately enough to represent the phenomenon under study. Rather, sensitivity
10 studies seek to inform what may happen under alternative assumptions made when defining the
11 problem under consideration by quantifying the effects of the alternative assumptions.

12 **4.5 References**

- 13 4-1. Kaplan, S. and Garrick, B.J., "On the quantitative definition of risk," *Risk Analysis*, 1(1),
14 11–27, 1981.
- 15 4-2. Kaplan, S., "Formalisms for handling phenomenological uncertainties: the concepts of
16 probability, frequency, variability, and probability of frequency," *Nuclear*
17 *Technology*, 102(1), 137–142, 1993.
- 18 4-3. Hoffman, F.O., and Hammonds, J.S., "Propagation of uncertainty in risk assessments:
19 the need to distinguish between uncertainty due to lack of knowledge and uncertainty
20 due to variability," *Risk Analysis*, 14(5), 707–712, 1994.
- 21 4-4. Helton, J.C., "Treatment of uncertainty in performance assessments for complex
22 systems," *Risk Analysis*, 14(4), 483–511, 1994.
- 23 4-5. Der Kiureghian, A., and Ditlevsen, O., "Aleatory or epistemic? Does it matter?" *Structural*
24 *Safety*, 31(2), 105–112, 2009.
- 25 4-6. Apostolakis, G., "The distinction between aleatory and epistemic uncertainties is
26 important: an example from the inclusion of aging effects into PSA," in *Proceedings of*
27 *PSA '99, International Topical Meeting on Probabilistic Safety Assessment*, 135–142,
28 1999.
- 29 4-7. Helton, J.C., Johnson, J., and Oberkampf, W., "An exploration of alternative approaches
30 to the representation of uncertainty in model predictions," *Reliability Engineering and*
31 *System Safety*, 89, 39–71, 2004.
- 32 4-8. Helton, J.C., and Burmaster, D.E., "Guest editorial: treatment of aleatory and epistemic
33 uncertainty in performance assessments for complex systems," *Reliability Engineering*
34 *and System Safety*, 54(2–3), 91–94, 1996.
- 35 4-9. Drouin, M., Gonzalez, M., Herrick, S., Hyslop, J.S., Stroup, D., Lehner, J., Pratt, T.,
36 Dennis, M., LaChance, J., and Wheeler, T., "Glossary of Risk-Related Terms in Support
37 of Risk-Informed Decisionmaking," NUREG-2122, U.S. Nuclear Regulatory Commission,
38 2013 (ML13311A353).

- 1 4-10. Chapman, O.J., "RR-PRODIGAL—A Model for Estimating the Probabilities of Defects in
2 Reactor Pressure Vessel Welds," NUREG/CR-5505, U.S. Nuclear Regulatory
3 Commission, 1998.
- 4 4-11. Duan, X., Wang, M., and Kozluk, M.J., "Benchmarking PRAISE-CANDU 1.0 With
5 NURBIM Project Fatigue Cases," PVP2013-97785, *ASME 2013 Pressure Vessels and
6 Piping Conference Volume 2: Computer Technology and Bolted Joints*, Paris, France,
7 American Society of Mechanical Engineers, 2013.
- 8 4-12. Harris, D., and Lim, E., Dedhia, D., "Probability of Pipe Fracture in the Primary Coolant
9 Loop of a PWR Plant," Volume 5, "Probabilistic Fracture Mechanics Analysis,"
10 NUREG/CR-2189 Vol. 5, UCID-18967 Vol.5, U.S. Nuclear Regulatory Commission, 1981
11 (ML15300A304).
- 12 4-13. Harris, D.O., "A Probabilistic Fracture Mechanics Code for Piping Reliability Analysis
13 (pcPRAISE code)," NUREG/CR-5864, U.S. Nuclear Regulatory Commission, 1992
14 (ML012490206).
- 15 4-14. Rudland, D., Harrington, C., and Dingreville, R., "Development of the Extremely Low
16 Probability of Rupture (xLPR) Version 2.0 Code," PVP2015-45134, *Proceedings of the
17 ASME 2015 Pressure Vessels and Piping Conference*, American Society of Mechanical
18 Engineers, 2015.
- 19 4-15. Kurth, R., Sallaberry, C., Young, B., Scott, P., Brust, F., and Kurth, E., "Benchmarking
20 Probabilistic Codes for LBB Analysis for Circumferential Cracks," *Proceedings of the
21 ASME 2017 Pressure Vessels and Piping Conference*, Waikola, HI, American Society of
22 Mechanical Engineers, 2017.
- 23 4-16. Helton, J.C., and Johnson, D.J., "Quantification of margins and uncertainties: Alternative
24 representations of epistemic uncertainty," *Reliability Engineering and System
25 Safety*, 96(9), 1034–1052, 2011.
- 26 4-17. Helton, J.C., Johnson, J.D., Oberkampf, W.L., and Sallaberry, C.J., "Representation of
27 analysis results involving aleatory and epistemic uncertainty," *International Journal of
28 General Systems*, 39(6), 605–646, 2010.
- 29 4-18. Helton, J.C., "Quantification of margins and uncertainties: Conceptual and computational
30 basis," *Reliability Engineering and System Safety*, 96(9), 976–1013, 2011.
- 31 4-19. Atwood, C.L., LaChance, J.L., Martz, H.F., Anderson, D.J., Englehardt, M.,
32 Whitehead, D., and Wheeler, T., "Handbook of Parameter Estimation for Probabilistic
33 Risk Assessment," NUREG/CR-6823, U.S. Nuclear Regulatory Commission, 2003
34 (ML032900131).
- 35 4-20. Xing, J., and Morrow, S., "White Paper: Practical Insights and Lessons Learned on
36 Implementing Expert Elicitation, U.S. Nuclear Regulatory Commission, Office of Nuclear
37 Regulatory Research, 2016.
- 38 4-21. Iman, R.L., and Conover, W.J., "A distribution-free approach to inducing rank correlation
39 among input variables," *Communication in Statistics—Simulation and Computation*,
40 11(3), 311–334, 1982.

- 1 4-22. Jaworski, P., Durante, F., Härdle, W.K., and Rychlik, T. (Eds.), *Copula theory and its*
2 *applications: proceedings of the workshop held in Warsaw, 25–26 September 2009*,
3 Volume 198, Springer Science and Business Media, 2010.
- 4 4-23. Haldar, A., and Mahadevan, S., *Probability, Reliability, and Statistical Methods in*
5 *Engineering Design*, John Wiley and Sons, Inc., 2000.
- 6 4-24. Robert, C.P., *Monte Carlo Statistical Methods*, John Wiley and Sons, Ltd., 2004.
- 7 4-25. Helton, J.C., and Davis, F.J., “Latin hypercube sampling and the propagation of
8 uncertainty in analyses of complex systems,” *Reliability Engineering and Systems*
9 *Safety*, 81(1), 23–69, 2003.
- 10 4-26. Mead, R., *The Design of Experiments*, Cambridge University Press, United Kingdom,
11 1988.
- 12 4-27. McKay, M., Beckman, R., and Conover, W., “A Comparison of Three Methods for
13 Selecting Values of Input Variables in the Analysis of Output from a Computer Code,”
14 *Technometrics*, 21(2), 239–245, 1979.
- 15 4-28. Robert, C.P., *Monte Carlo Statistical Methods*, John Wiley and Sons, Ltd., 2004.
- 16 4-29. Sallaberry, C.J., Helton, J.C., and Hora, S.C., “Extension of Latin hypercube samples
17 with correlated variables,” *Reliability Engineering and System Safety*, 93(7), 1047–1059,
18 2008.
- 19 4-30. Kaplan, S., “On the method of discrete probability distributions in risk and reliability
20 calculations—application to seismic risk assessment,” *Risk Analysis*, 1: 189–196, 1981.
- 21 4-31. Adams, B.M., Bauman, L.E., Bohnhoff, W.J., Dalbey, K.R., Eddy, J.P., Ebeida, M.S.,
22 Eldred, M.S., Hough, P.D., Hu, K.T., Jakeman, J.D., Swiler, L.P., Stephens, J.A., Vigil,
23 D.M., and Wildey, T.M., “DAKOTA, A Multilevel Parallel Object-Oriented Framework for
24 Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity
25 Analysis: Version 6.0 User’s Manual,” Sandia Technical Report SAND2014-4633,
26 July 2014, updated May 2017.
- 27 4-32. Robert, C.P. *Monte Carlo Statistical Methods*, John Wiley and Sons, Ltd., 2004.
- 28 4-33. Helton, J.C., and Davis, F.J., “Latin hypercube sampling and the propagation of
29 uncertainty in analyses of complex systems,” *Reliability Engineering and Systems*
30 *Safety*, 81(1), 23–69, 2003.
- 31 4-34. Swiler, L. and West, N., “Importance Sampling: Promises and Limitations,”
32 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials
33 Conference, 2010.
- 34 4-35. Glynn, P.W., and Iglehart, D.L., “Importance Sampling for Stochastic Simulations,”
35 *Management Science*, 35(11), 1367–1392, 1989.
- 36 4-36. Au, S.K., and Beck, J.L., “Important sampling in high dimensions,” *Structural*
37 *Safety*, 25(2), 139–163, 2003.

- 1 4-37. Haldar, A., and Mahadevan, S., Probability, Reliability, and Statistical Methods in
2 Engineering Design, John Wiley and Sons, Inc., 2000.
- 3 4-38. Karamchandani, A., Bjerager, P., and Cornell, A.C., "Adaptive Importance Sampling,"
4 *Proceedings, International Conference on Structural Safety and Reliability (ICOSSAR)*,
5 San Francisco, CA, 855–862, 1989.
- 6 4-39. Wu, Y.T., "An Adaptive Importance Sampling Method for Structural System Reliability
7 Analysis, Reliability Technology 1992," in T.A. Cruse (Editor), *ASME Winter Annual
8 Meeting*, Volume AD-28, Anaheim, CA, 217–231, 1992.
- 9 4-40. Bichon, B.J., Eldred, M.S., Swiler, L.P., Mahadevan, S., and McFarland, J.M., "Efficient
10 Global Reliability Analysis for Nonlinear Implicit Performance Functions," *AIAA Journal*,
11 46(10), 2459–2468, 2008.
- 12 4-41. Haldar, A., and Mahadevan, S., Probability, Reliability, and Statistical Methods in
13 Engineering Design, John Wiley and Sons, Inc., 2000.
- 14 4-42. Lopez, R., and Beck, A., "Reliability-Based Design Optimization Strategies Based on
15 FORM: A Review," *Journal of the Brazilian Society of Mechanical Sciences and
16 Engineering*, 34, 506–514, 2012.
- 17 4-43. Hasofer, A.M., and Lind, N.C., "Exact and Invariant Second Moment Code Format,"
18 *Journal of the Engineering Mechanics Division, ASCE*, 100(EM1), 111–121, 1974.
- 19 4-44. Rackwitz, R., "Practical Probabilistic Approach to Design," *Bulletin No. 112*, Comité
20 European du Béton, Paris, France, 1976.
- 21 4-45. Rackwitz, R., and Fiessler, B., "Structural Reliability Under Combined Random Load
22 Sequences," *Computers and Structures*, 9(5), 484–494, 1978.
- 23 4-46. Rackwitz, R., and Fiessler, B., "Note on Discrete Safety Checking When Using Non-
24 Normal Stochastic Models for Basic Variables," *Load Project Working Session*,
25 Massachusetts Institute of Technology, 1976.
- 26 4-47. Fiessler, B., Neumann, H.J., and Rackwitz, R., "Quadratic Limit States in Structural
27 Reliability," *Journal of Engineering Mechanics*, American Society of Civil Engineers,
28 1095(4), 661–676, 1979.
- 29 4-48. Breitung, K., "Asymptotic Approximations for Multinormal Integrals," *Journal of
30 Engineering Mechanics*, American Society of Civil Engineers, 110(3), 357–366, 1984.
- 31 4-49. Hohenbichler, M., Gollwitzer, S., Kruse, W., and Rackwitz, R., "New Light on First- and
32 Second-Order Reliability Methods," *Structural Safety*, 4, 267–284, 1987.
- 33 4-50. Brown, L.D., Cai, T.T., and DasGupta, A., "Interval estimation for a binomial proportion,"
34 *Statistical science*, 101–117, 2001.
- 35 4-51. Clopper, C.J., and Pearson, E.S., "The use of confidence or fiducial limits illustrated in
36 the case of the binomial," *Biometrika*, 26, 404–413, 1934.

- 1 4-52. Lohr, S., *Sampling: design and analysis*, Nelson Education, 2009.
- 2 4-53. Haldar, A., and Mahadevan, S., *Probability, Reliability, and Statistical Methods in*
3 *Engineering Design*, John Wiley and Sons, Inc., 2000.
- 4 4-54. Brown, L.D., Cai, T.T., and DasGupta, A., "Interval estimation for a binomial
5 proportion," *Statistical science*, 101–117, 2001.
- 6 4-55. Louis, T., "Confidence Intervals for a Binomial Parameter after Observing No
7 Successes," *The American Statistician*, 35(3), 154, 1981.
- 8 4-56. Bickel, P.J., and Doksum, K.A., *Mathematical Statistics: Basic Ideas and Selected*
9 *Topics*, Holden-Day, 1977.
- 10 4-57. Davison, A.C., and Hinkley, D.V., *Bootstrap Methods and their Application*, Volume 1,
11 Cambridge University Press, United Kingdom, 1997.
- 12 4-58. Davison, A.C., Hinkley, D.V., and Young, G.A., "Recent developments in bootstrap
13 methodology," *Statistical Science*, 141–157, 2003.
- 14 4-59. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., and
15 Tarantola, S., *Global Sensitivity Analysis. The Primer*, John Wiley and Sons, Ltd., 2008.
- 16 4-60. Tukey, J.W., *Exploratory Data Analysis*, Volume 2, 1977.
- 17 4-61. Cukier, R.I., Fortuin, C.M., and Shuler, K.E., "Study of the sensitivity of coupled reaction
18 systems to uncertainties in rate coefficients," "I: Theory," *The Journal of Chemical*
19 *Physics*, 59, 2003.
- 20 4-62. Sobol', I.M., "On sensitivity estimation for nonlinear mathematical models,"
21 *Matematicheskoe Modelirovanie*, 2(1), 112–118, 1990. (In Russian)
- 22 4-63. Kleignen, J.P.C., and Helton, J.C., "Statistical analyses of scatterplots to identify
23 important factors in large-scale simulations," "1: Review and comparison of techniques,"
24 *Reliability Engineering and System Safety*, 65, 147–185, 1999.
- 25 4-64. Helton, J.C., Johnson, J.D., Sallaberry, C.J., and Storlie, C.B., "Survey of
26 sampling-based methods for uncertainty and sensitivity analysis," *Reliability Engineering*
27 *and System Safety*, 91(10–11), 1175–1209, 2006.
- 28 4-65. Borgonovo, E., and Plischke, E., "Sensitivity analysis: a review of recent
29 advances," *European Journal of Operational Research*, 248(3), 869–887, 2016.
- 30 4-66. Robert, C.P., *Monte Carlo Statistical Methods*, John Wiley and Sons, Ltd., 2004
- 31 4-67. Storlie, C.B., and Helton, J.C., "Multiple predictor smoothing methods for sensitivity
32 analysis: Description of techniques," *Reliability Engineering and System Safety*, 93,
33 28–54, 2008.
- 34 4-68. Storlie, C.B., Swiler, L.P., Helton, J.C., and Sallaberry, C.J., "Implementation and
35 evaluation of nonparametric regression procedures for sensitivity analysis of

- 1 computationally demanding models,” *Reliability Engineering and System Safety*, 94,
2 1735–1763, 2009.
- 3 4-69. Sullivan, T.J., *Introduction to Uncertainty Quantification*, Volume 63, Springer, 2015.
- 4 4-70. Constantine, P.G., “Active subspaces: Emerging ideas for dimension reduction in
5 parameter studies,” Volume 2, *SIAM Spotlights*, Society for Industrial and Applied
6 Mathematics, 2015.
- 7 4-71. Smith, R.C., *Uncertainty Quantification Theory, Implementation, and Applications*,
8 Society for Industrial and Applied Mathematics, 2014.
- 9 4-72. Helton, J.C., “Uncertainty and sensitivity analysis techniques for use in performance
10 assessment for radioactive waste disposal,” *Reliability Engineering and System
11 Safety*, 42(2–3), 327–367, 1993.
- 12 4-73. Loeppky, J.L., Sacks, J., and Welch, W.J., “Choosing the sample size of a computer
13 experiment: A practical guide,” *Technometrics*, 51(4), 366–376, 2009.
- 14 4-74. Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K.,
15 “Surrogate-based analysis and optimization,” *Progress in Aerospace Sciences*, 41, 1–
16 28, 2005.
- 17 4-75. Rasmussen, C.E., and Williams, C., *Gaussian Processes for Machine Learning*, MIT
18 Press, 2006.
- 19 4-76. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., and
20 Tarantola, S., *Global Sensitivity Analysis. The Primer*, John Wiley and Sons, Ltd., 2008.
- 21 4-77. Friedman, J., “Multivariate adaptive regression splines” (with discussion), *Annals of
22 Statistics*, 1–141, 1991.
- 23 4-78. Storlie, C.B., and Helton, J.C., “Multiple predictor smoothing methods for sensitivity
24 analysis: Description of techniques,” *Reliability Engineering and System Safety*, 93,
25 28–54, 2008.
- 26 4-79. Storlie, C.B., Swiler, L.P., Helton, J.C., and Sallaberry, C.J., “Implementation and
27 evaluation of nonparametric regression procedures for sensitivity analysis of
28 computationally demanding models,” *Reliability Engineering and System Safety*, 94,
29 1735–1763, 2009.
- 30 4-80. Casella, G., and Berger, R.L., *Statistical Inference*, Volume 2, Duxbury, 2002.
- 31 4-81. McCullagh, P., and Nelder, J., *Generalized Linear Models, Second Edition*, Chapman
32 and Hall/CRC, 1989.
- 33 4-82. Breiman, L., Friedman, J., Olshen, R., and Stone, C., *Classification and Regression
34 Trees*, Wadsworth Intl., 1984.
- 35 4-83. Hastie, T., Tibshirani, T., and Friedman, J., *The Elements of Statistical Learning*,
36 Springer-Verlag, 2009.

- 1 4-84. Casarez, C., "Global Sensitivity Analysis of xLPR using Metamodeling," *CNSC-IAEA 2nd*
2 *International Seminar on Probabilistic Methodologies for Nuclear Applications*, Ottawa,
3 Ontario, Canada, 2017.
- 4 4-85. Sutton, C.D., "11—Classification and Regression Trees, Bagging, and
5 Boosting," *Handbook of Statistics*, 24, 303–329, 2005.
- 6 4-86. Jyrkama, M.I., and Pandey, M.D., "On the separation of aleatory and epistemic
7 uncertainties in probabilistic assessments," *Nuclear Engineering and Design*, 303,
8 68–74, 2016.
- 9 4-87. EricksonKirk, M.T., Dickson, T., Mintz, T., and Simonen, F., "Sensitivity Studies of the
10 Probabilistic Fracture Mechanics Model Used in FAVOR," NUREG-1808, U.S. Nuclear
11 Regulatory Commission, 2010 (ML061580349).
- 12 4-88. Brickstad, B., Schimpfke, T., Chapman, O.J.V., Dillstrom, P., Cueto-Felgueroso, C., and
13 Bell, C.D., "Project NURBIM (Nuclear RI-ISI Methodology for Passive Components):
14 Benchmarking of Structural Reliability Models and Associated Software," in *ASME/JSME*
15 *2004 Pressure Vessels and Piping Conference*, American Society of Mechanical
16 Engineers, 109–119, 2004.
- 17 4-89. Khaleel, M.A., and Simonen, F.A., "Evaluation of Structural Failure Probabilities and
18 Candidate Inservice Inspection Programs," NUREG/CR-6986, U.S. Nuclear Regulatory
19 Commission, Office of Nuclear Regulatory Research, 2009 (ML091620307).

5 SUMMARY AND CONCLUSIONS

This document outlines a framework for conducting PFM analyses that recognizes the fact that each regulatory application may have unique characteristics. To address the unique characteristics of diverse PFM applications, this NUREG presents the elements of a graded approach for PFM analyses and the corresponding recommendations for supporting documentation. These elements are aligned with the documentation elements previously given in the NRC's technical letter report, "Important Aspects of Probabilistic Fracture Mechanics Analyses," and the outcomes from the NRC public meeting to discuss a graded approach for PFM codes and analyses for regulatory applications.

The three technical sections (Sections 2, 3, and 4) develop the concept of PFM analysis methodology and outline important considerations for a high-quality and high-confidence PFM analysis. The sections are linked together and progressively dive into more detailed elements of PFM applications, but each may have different audiences:

- Section 2 is intended for applicants of all experience levels. Each subsection introduces an element of content that would be expected in a PFM submittal. It identifies representative circumstances for a submittal and describes a graded approach for the specific information to provide to the NRC.
- Section 3 could be used by applicants who are familiar with PFM submittals but are seeking additional information on the development of an analysis structure or on formalism. Each subsection presents an analytical step that may exist in a PFM submittal.
- Section 4 could be used by applicants who are seeking to further understand the theoretical underpinnings of the processes used to establish the credibility of a PFM analysis. Each subsection presents the fundamental background for the concepts and methods used in a PFM analysis. Examples give details for analysts on (nonprescriptive) approaches for PFM analyses. The concepts and methods are linked directly to the analysis steps presented in Section 3.

The NRC does not require PFM submittals to follow the process outlined in this NUREG. Applicants should refer to DG-1382 / RG-1.245 for specific guidance on using PFM for regulatory submittals.

GLOSSARY

- 1
- 2 Acceptance Criteria
- 3 Set of conditions that must be met to achieve success for the desired application.
- 4
- 5 Accuracy and Precision
- 6 “Accuracy” is the degree to which the result of a measurement, calculation, or specification
- 7 conforms to the correct value (i.e., reality or a standard accepted value). “Precision” is a
- 8 description of random errors and a measure of statistical variability for a given quantity. In other
- 9 words, “accuracy” is the proximity of measurement results to the true value; “precision” is the
- 10 repeatability or reproducibility of the measurement.
- 11
- 12 Aleatory Uncertainty
- 13 Uncertainty based on the randomness of the nature of the events or phenomena that cannot be
- 14 reduced by increasing the analyst’s knowledge of the systems being modeled (Reference 0-1).
- 15
- 16 Assumption
- 17 A decision or judgment made in the development of a model or analysis (Reference 0-1).
- 18
- 19 Bayesian Inference
- 20 Type of data analysis in which an initial estimate about a parameter value is combined with
- 21 evidence to arrive at a more informed estimate (Reference 0-1).
- 22
- 23 Benchmark (in the context of PFM computational analyses)
- 24 An established point of reference against which computers or programs can be measured in
- 25 tests comparing their performance, reliability, output, etc. A standard against which similar
- 26 analyses must be measured or judged. Benchmarks are often a part of validation for scientific
- 27 analysis software.
- 28
- 29 Best Estimate
- 30 Approximation of a quantity based on the best available information (Reference 0-1). Models
- 31 that attempt to fit data or phenomena as best as possible. That is, models that do not
- 32 intentionally bound data for a given phenomenon or are not intentionally conservative or
- 33 optimistic.
- 34
- 35 Calibration
- 36 The process of adjusting physical modeling parameters in the computational model to improve
- 37 agreement with experimental data (Reference 0-2).
- 38
- 39 Code
- 40 The computer implementation of algorithms developed to facilitate the formulation and
- 41 approximation solution of a class of problems. (Reference 0-2).
- 42
- 43 Code Verification
- 44 The process of determining and documenting the extent to which a computer program (“code”)
- 45 correctly solves the equations of the mathematical model (Reference 0-3).

- 1 Completeness Uncertainty
2 Caused by the limitations in the scope of the model, such as whether all applicable physical
3 phenomena have been adequately represented and all accident scenarios that could
4 significantly affect the determination of risk have been identified (Reference 0-1).
5
- 6 Component
7 A part of a system in a nuclear power plant (Reference 0-1).
8
- 9 Conditional Probability
10 Probability of occurrence of an event, given that a prior event has occurred (Reference 0-1).
11
- 12 Confidence Interval
13 A range of values that has a specified likelihood of including the true value of a random variable
14 (Reference 0-1).
15
- 16 Consequence
17 In the context of nuclear regulatory submittals, the health effects or the economic costs resulting
18 from a nuclear power plant accident (Reference 0-1).
19
- 20 Conservative Analysis
21 An analysis that uses assumptions such that the assessed outcome is meant or found to be less
22 favorable than the expected outcome (Reference 0-1).
23
- 24 Convergence Analysis
25 An analysis with the purpose of assessing the approximation error in the quantity of interest
26 estimates to establish that conclusions of the analysis would not change solely due to sampling
27 uncertainty.
28
- 29 Correlation
30 A general term for interdependence between pairs of variables (Reference 0-4).
31
- 32 Credibility
33 The quality to elicit belief or trust in modeling and simulation results (Reference 0-5).
34
- 35 Cumulative Distribution Function
36 A function that provides the probability that a parameter is less than or equal to a given value
37 (Reference 0-1).
38
- 39 Continuous variable
40 See "Discrete versus Continuous Variables."
41
- 42 Dependent
43 Not independent.
44

- 1 Deterministic
- 2 A characteristic of decisionmaking in which results from engineering analyses not involving
3 probabilistic considerations are used to support a decision (Reference 0-1). Consistent with the
4 principles of determinism, which hold that specific causes completely and certainly determine
5 effects of all sorts (Reference 0-6). Also refers to fixed model inputs.
- 6
- 7 Deterministic Fracture Mechanics
- 8 An analysis that uses fixed values of input parameters to a fracture mechanics model to
9 estimate a fixed model output or quantity of interest computed from the output.
- 10
- 11 Discrete versus Continuous Variables
- 12 A discrete random variable is a variable that has a nonzero probability for only a finite, or
13 countably infinite, set of values. A continuous random variable is a variable that has an
14 absolutely continuous cumulative distribution function (Reference 0-3).
- 15
- 16 Distribution
- 17 A function specifying the values that the random variable can take and the likelihood they will
18 occur.
- 19
- 20 Engineering Judgment
- 21 The scientific process by which a design, installation, operation/maintenance, or safety problem
22 is systematically evaluated. The decision made by an engineer based on the available data to
23 propose a design or a line of action.
- 24
- 25 Epistemic Uncertainty
- 26 The uncertainty related to the lack of knowledge or confidence about the system or model; also
27 known as “state-of-knowledge uncertainty.” As defined by the American Society of Mechanical
28 Engineers (ASME)/American Nuclear Society (ANS) probabilistic risk assessment (PRA)
29 standard (Reference 0-1), “the uncertainty attributable to incomplete knowledge about a
30 phenomenon that affects our ability to model it. Epistemic uncertainty is reflected in ranges of
31 values for parameters, a range of viable models, the level of model detail, multiple expert
32 interpretations, and statistical confidence. In principle, epistemic uncertainty can be reduced by
33 the accumulation of additional information. (Epistemic uncertainty is sometimes also called
34 ‘modeling uncertainty.’)” (Reference 0-1)
- 35
- 36 Expert Elicitation
- 37 A formal, structured, and documented process to obtain judgments from expert(s). May be used
38 to obtain information from technical experts on topics that are uncertain. A process in which
39 experts are assembled and their judgment is sought and aggregated in a formal way.
40 (Reference 0-1)
- 41
- 42 Expert Judgment
- 43 Information (or opinion) provided by one or more technical experts based on their experience
44 and knowledge. Used when there is a lack of information, for example, if certain parameter
45 values are unknown, or there are questions about phenomenology in accident progression. May
46 be part of a structured approach, such as expert elicitation, but is not necessarily as formal. May
47 be the opinion of one or more experts, whereas expert elicitation is a highly structured process

1 in which the opinions of several experts are sought, collected, and aggregated in a very formal
2 way. (Reference 0-1)

3 Failure Probability

4 As defined in the ASME/ANS PRA standard (Reference 0-1), “the likelihood that a system or
5 component will fail to operate upon demand or fail to operate for a specific mission time”
6 (Reference 0-1). For components, can also be the likelihood of a component being in a
7 defective, unacceptable condition (adverse condition or event) (e.g., leakage from reactor
8 coolant pressure boundary).

9 10 Frequency

11 The expected number of occurrences of an event or accident condition expressed per unit of
12 time. Normally expressed in events per plant (or reactor) operating year or events per plant (or
13 reactor) calendar year (Reference 0-1).

14 15 Global Sensitivity Analysis

16 The study of how the uncertainty in the output or quantity of interest of a model (numerical or
17 otherwise) can be apportioned to different sources of uncertainty in the model input. The term
18 “global” ensures that the analysis considers more than just local or one-factor-at-a-time effects.
19 Hence, interactions and nonlinearities are important components of a global statistical sensitivity
20 analysis (Reference 0-3).

21 22 Important Variable

23 A variable whose uncertainty contributes substantially to the uncertainty in the response
24 (Reference 0-7).

25 26 Independent

27 Two events are said to be independent if knowing the outcome of one tells us nothing about the
28 other (Reference 0-4).

29 30 Input

31 Data or parameters that users can specify for a model; the output of the model varies as a
32 function of the inputs, which can consist of physical values (e.g., material properties, tolerances)
33 and model specifications (e.g., spatial resolution).

34 35 Input Uncertainty

36 The uncertainty in the values of the inputs to the model represented by probabilistic distributions
37 (Reference 0-1).

38 39 Interaction Effect

40 A term applied when two (or more) explanatory variables do not act independently on a
41 response variable.

42 43 Level of Detail

44 The degree of resolution or specificity in the analyses performed. Generally refers to the level to
45 which a system is modeled; dictated by (1) the level of detail to which information is available,
46 (2) the level of detail required so that dependencies are included, (3) the level of detail so that

1 the risk contributors are included, and (4) the level of detail sufficient to support the application
2 (Reference 0-1).

3 Local Sensitivity Analysis

4 A sensitivity analysis that is relative to the location in the input space chosen and not for the
5 entire input space (Reference 0-7).

6 7 Margin

8 The distance between the quantity of interest and the acceptance criteria.

9 10 Mean

11 The average of a set of numerical values; more technically, the expected value of a random
12 variable (Reference 0-1).

13 14 Median

15 The value that a random variable is equally likely to be above and below. Also known as the
16 50th percentile of the distribution of a random variable (Reference 0-1).

17 18 Model

19 A representation of a physical process that allows for prediction of the process' behavior
20 (Reference 0-1).

21 22 Model Uncertainty

23 Related to an issue for which no consensus approach or model exists and where the choice of
24 approach or model is known to have an effect on the decision made (Reference 0-1).

25 26 Output

27 A value calculated by the model given a set of inputs.

28 29 Parameter

30 A numerical characteristic of a population or probability distribution. More technically, the
31 variables used to calculate and describe frequencies and probabilities (Reference 0-1).

32 33 Percentile

34 The set of divisions that produce exactly 100 equal parts in a series of continuous values
35 (Reference 0-4).

36 37 Point Estimate

38 An estimate of a parameter in the form of a single value (Reference 0-1).

39 40 Precision

41 See "Accuracy and Precision."
42

1 Prediction

2 The use of a model to make statements about quantities of interest in settings (initial conditions,
3 physical regimes, parameter values, etc.) that are inside (interpolative) or outside (extrapolative)
4 the conditions for which the model validation effort occurred (Reference 0-3).

5
6 Probabilistic

7 A characteristic of an evaluation that considers the likelihood of events (Reference 0-1).

8
9 Probabilistic Fracture Mechanics

10 An analysis that uses probabilistic representations of uncertain input parameters to a fracture
11 mechanics model to estimate uncertainty in the model outputs or quantities of interest computed
12 from the outputs (Reference 0-8).

13
14 Probabilistic Risk Assessment

15 A systematic^[SEP] method for^[SEP] assessing the likelihood of accidents and their potential
16 consequences (Reference 0-1).

17
18 Probability

19 A number between 0 and 1 describing the likelihood or chance of an event occurring. There are
20 two main interpretations of probability:

21 (1) *Frequency interpretation.* The probability of an event is the relative frequency of the
22 occurrence of the event in a long sequence of trials in which the event does or does not
23 occur. In other words, the likelihood that an event will occur is expressed by the ratio of
24 the number of actual occurrences to the total number of possible occurrences
25 (Reference 0-1).

26 (2) *Subjective interpretation.* The probability of an event comes from expert judgment about
27 uncertain events or quantities, in the form of probability statements about future events.
28 It is not based on any precise computation but is often a reasonable assessment by a
29 knowledgeable person (Reference 0-3).

30 Probability Density Function

31 A function of a continuous random variable whose integral over an interval gives the probability
32 that its value will fall within the interval (Reference 0-1). Analogous to probability distribution for
33 continuous random variables.

34
35 Probability Distribution

36 A function specifying the values that the random variable can take and the likelihood they will
37 occur (Reference 0-1).

38
39 Quantiles

40 Divisions of a probability distribution or frequency distribution into equal, ordered subgroups
41 (Reference 0-4).

42

- 1 Quantity of Interest
- 2 A numerical characteristic of the system being modeled, the value of which is of interest to
3 stakeholders, typically because it informs a decision (Reference 0-3). Can refer to either a
4 physical quantity that is an output from a model or a given feature of the probability distribution
5 function of the output of a deterministic model with uncertain inputs. (Reference 0-9)
- 6
- 7 Random Uncertainty
- 8 See “Aleatory Uncertainty.”
- 9
- 10 Random Variable
- 11 A variable, the values of which occur according to some specified probability distribution
12 (Reference 0-4).
- 13
- 14 Rank
- 15 The relative position of the members of a sample with respect to some characteristic
16 (Reference 0-4).
- 17
- 18 Rare
- 19 Events that are unlikely to occur. Rare event probabilities are defined as probabilities that are
20 close enough to 0 that the number of samples needed to estimate the probability is large
21 relative to the computational budget.
- 22
- 23 Realization
- 24 The execution of a model for a single set of input parameter values (Reference 0-8).
- 25
- 26 Regression
- 27 A form of statistical analysis in which observational data are used to statistically fit a
28 mathematical function that presents the data (i.e., dependent variables) as a function of a set of
29 parameters and one or more independent variables (Reference 0-3).
- 30 Reliability
- 31 The likelihood that a system, structure, or component performs its required function(s) for a
32 specific period of time (Reference 0-1).
- 33
- 34 Risk
- 35 The combined answer to the three questions that consider (1) what can go wrong, (2) how likely
36 it is, and (3) what its consequences might be (Reference 0-1).
- 37
- 38 Risk-Informed
- 39 A characteristic of decisionmaking in which risk results or insights are used together with other
40 factors to support a decision (Reference 0-1).
- 41
- 42 Robustness
- 43 The degree to which deviations from a “best” decision provide suboptimal values of the desired
44 criterion. These deviations can be due to uncertainty in model formulation, assumed parameter
45 values, etc. (Reference 0-3).

1 Sampling
2 The process of selecting some part of a population to observe, so as to estimate something of
3 interest about the whole population (Reference 0-4).
4
5 Sampling Uncertainty
6 The uncertainty in an estimate of a quantity of interest that arises due to finite sampling.
7 Different sets of model realizations will result in different estimates. This type of uncertainty
8 contributes to uncertainty in the true value of the quantity of interest and is often summarized
9 using the sampling variance.
10
11 Sampling Variance
12 The variance of an estimate of a quantity of interest that arises due to sampling uncertainty
13 (i.e., finite sampling). An estimate of this variance is often used to summarize sampling
14 uncertainty.
15
16 Sensitive Variable
17 A variable that has a significant influence on the response (Reference 0-10).
18
19 Sensitivity Analysis
19 The study of how uncertainty in the output of a model can be apportioned to different sources of
20 uncertainty in the model input (Reference 0-10).
21
22 Sensitivity Metrics
23 Quantitative values that characterize the relationship between input and output variables. The
24 following two metrics can be used:
25 (1) First-order sensitivity indices measure the proportion of the uncertainty in the output that
26 is explained by the uncertainty in a single input.
27 (2) Total-order sensitivity indices measure the proportion of the uncertainty in the output that
28 is explained by the uncertainty in an input and its interactions with other inputs
29 (Reference 0-10).
30
31 Sensitivity Studies
31 Probabilistic fracture mechanics analyses that are conducted under credible alternative
32 assumptions (Reference 0-11).
33
34 Significant
35 A factor that can have a major or notable influence on the results of a risk analysis
36 (Reference 0-1).
37
38 Simulation
39 The execution of a computer code to mimic an actual system (Reference 0-3). Typically
40 comprises a set of model realizations.
41

1 Software Quality Assurance
2 A planned and systematic pattern of all actions necessary to provide adequate confidence that a
3 software item or product conforms to established technical requirements; a set of activities
4 designed to evaluate the process by which the software products are developed or
5 manufactured (Reference 0-12).
6
7 Solution Verification
8 The process of determining as completely as possible the accuracy with which the algorithms
9 solve the mathematical-model equations for a specified quantity of interest (Reference 0-3).
10
11 State-of-Knowledge Uncertainty
12 See “Epistemic Uncertainty” (Reference 0-1).
13
14 Statistic
15 A numerical characteristic of a sample, such as the sample mean and sample variance
16 (Reference 0-4).
17
18 Statistical Model
19 A description of the assumed structure of a set of observations that can range from a fairly
20 imprecise verbal account to, more usually, a formalized mathematical expression of the process
21 assumed to have generated the data (Reference 0-4).
22
23 Stochastic Uncertainty
24 See “Aleatory Uncertainty” (Reference 0-1).
25
26 Subjective Probability
27 Expert judgment about uncertain events or quantities, in the form of probability statements
28 about future events. Not based on any precise computation but often a reasonable assessment
29 by a knowledgeable person (Reference 0-3).
30
31 Surrogate
32 A function that predicts outputs from a model as a function of the model inputs (Reference 0-3).
33 Also known as response surface, metamodel, or emulator.
34
35 Uncertainty
36 Variability in an estimate because of the randomness of the data or the lack of knowledge
37 (Reference 0-1).
38
39 Uncertainty Analysis
40 A process for determining the level of imprecision in the results of the probabilistic analysis and
41 its parameters (Reference 0-1).
42
43 Uncertainty Distribution
44 See “Probability Distribution” (Reference 0-1).
45

1 Uncertainty Interval/Range

2 A range that bounds the uncertainty value(s) of a parameter or analysis result by establishing
3 upper and lower limits (see “Confidence Interval,” “Probability Distribution”) (Reference 0-1).

4
5 Uncertainty Propagation

6 Characterizing the uncertainty of a model’s responses that results from the propagation through
7 the model of the uncertainty in the model’s inputs (Reference 0-3).

8
9 Uncertainty Quantification

10 The process of characterizing all relevant uncertainties in a model and quantifying their effect on
11 a quantity of interest (Reference 0-3).

12
13 Validation

14 The process of determining the degree to which a model is an accurate representation of the
15 real world from the perspective of the intended uses of the model (Reference 0-3).

16
17 Variable

18 Some characteristic that differs from subject to subject or from time to time (Reference 0-4).

19
20 Variance

21 The second moment of a probability distribution, defined as $E(X - \mu)^2$, where μ is the first
22 moment of the random variable X . A common measure of variability around the mean of a
23 distribution (Reference 0-3).

24
25 Verification

26 The process of determining whether a computer program (“code”) correctly solves the
27 mathematical-model equations. This includes code verification (determining whether the code
28 correctly implements the intended algorithms) and solution verification (determining the
29 accuracy with which the algorithms solve the mathematical-model equations for specified
30 quantities of interest) (Reference 0-3).

31
32 **Glossary References**

33
34 0-1. Drouin, M., Gonzalez, M., Herrick, S., Hyslop, J.S., Stroup, D., Lehner, J., Pratt, T.,
35 Dennis, M., LaChance, J., and Wheeler, T., “Glossary of Risk-Related Terms in Support
36 of Risk-Informed Decisionmaking,” NUREG-2122, U.S. Nuclear Regulatory Commission,
37 2013 (ML13311A353).

38 0-2. Kumar, M., “Guide for Verification and Validation in Computational Solid Mechanics,”
39 ASME V&V 10, American Society of Mechanical Engineers, 2006.

40 0-3. National Research Council, *Assessing the Reliability of Complex Models: Mathematical
41 and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*,
42 The National Academies Press, 2012.

43 0-4. Everitt, B.S., and Skrondal, A., *The Cambridge Dictionary of Statistics, 4th Edition*,
44 Cambridge University Press, United Kingdom, 2010.

- 1 0-5. National Aeronautics and Space Administration, "Standard for Models and Simulations,"
2 NASA-STD-7009A, 2016.
- 3 0-6. U.S. Nuclear Regulatory Commission, "Glossary," Web site last updated
4 March 19, 2020, <http://www.nrc.gov/reading-rm/basic-ref/glossary.html>.
- 5 0-7. Hamby, D.M., "A review of techniques for parameter sensitivity analysis of environmental
6 models," *Environmental Monitoring and Assessment*, 32.2: 135–154, 1994.
- 7 0-8. Raynaud, P., Kirk, M., Benson, M., and Homiack, M., "Important Aspects of Probabilistic
8 Fracture Mechanics Analyses," U.S. Nuclear Regulatory Commission, 2018
9 (ML18178A431).
- 10 0-9. Pasanisi, A., Keller, M., and Parent, E., "Estimation of a quantity of interest in uncertainty
11 analysis: Some help from Bayesian decision theory," *Reliability Engineering & System
12 Safety*, 100, 93–101, 2012.
- 13 0-10. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., and
14 Tarantola, S., *Global Sensitivity Analysis. The Primer*, John Wiley and Sons, Ltd., 2008.
- 15 0-11. Erickson Kirk, M., Dickson, T., Mintz, T., and Simonen, F., "Sensitivity Studies of the
16 Probabilistic Fracture Mechanics Model Used in FAVOR," NUREG-1808, U.S. Nuclear
17 Regulatory Commission, 2010 (ML061580349).
- 18 0-12. IEEE Standards Board, *IEEE Standard Glossary of Software Engineering Terminology*,
19 Institute of Electrical and Electronics Engineers, 1990.

20

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

2. TITLE AND SUBTITLE

Technical Basis for the use of Probabilistic Fracture Mechanics in Regulatory Applications

Draft Report for Comment

3. DATE REPORT PUBLISHED

MONTH

YEAR

September

2021

4. FIN OR GRANT NUMBER

N/A

5. AUTHOR(S)

Lauren Hund, John Lewis, Nevin Martin, Michael Starr, Dusty Brooks, Adah Zhang, Remi Dingreville, Aubrey Eckert, Josh Mullins (Sandia National Laboratory)

Patrick Raynaud, David Rudland, David Dijamco, Stephen Cimblidge (U.S. NRC)

6. TYPE OF REPORT

Technical

7. PERIOD COVERED (Inclusive Dates)

N/A

8. PERFORMING ORGANIZATION - NAME AND ADDRESS (If NRC, provide Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Sandia National Laboratories, P.O. Box 8500, Albuquerque, NM 87185

U.S. NRC, Division of Engineering, Office of Nuclear Regulatory Research

U.S. NRC, Division of New and Renewed Licenses, Office of Nuclear Reactor Regulation

9. SPONSORING ORGANIZATION - NAME AND ADDRESS (If NRC, type "Same as above", if contractor, provide NRC Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address.)

Division of Engineering

Office of Nuclear Regulatory Research

U.S. Nuclear Regulatory Commission

Washington, DC 20555-0001

10. SUPPLEMENTARY NOTES

None.

11. ABSTRACT (200 words or less)

This NUREG, on probabilistic fracture mechanics (PFM), is a companion document to Draft Regulatory Guide 1382 (DG-1382) / Regulatory Guide 1.245 (RG-1.245), "Preparing Probabilistic Fracture Mechanics (PFM) Submittals." This NUREG describes a graded approach to developing PFM submittal documentation, provides a generalized technical basis for conducting PFM analyses, and constitutes the technical basis for DG-1382 / RG-1.245.

The graded approach that is outlined below represents a balance between the benefits of clear, consistent, and comprehensive submittals and the need to maintain flexibility for PFM analyses that, by their nature, include many situation-specific aspects. The resulting guidance, provided in DG-1382/RG1.245, outlines a procedure that includes this suggested graded approach for PFM analyses and submittals. The unique characteristics of the underlying regulatory application dictate the breadth and depth of content included in the submission.

This document also describes a hypothetical process for conducting a PFM analysis. This process is aligned with the position on documentation elements given previously in the U.S. Nuclear Regulatory Commission's (NRC's) technical letter report, "Important Aspects of Probabilistic Fracture Mechanics Analyses," issued in 2018. The NUREG provides fundamental background for the concepts and methods introduced in the analysis process. The examples give details for analysts on (nonprescriptive) approaches for PFM analyses

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

Probabilistic Fracture Mechanics, PFM, RG-1.245, Graded Approach, PFM analysis, PFM submittal.

13. AVAILABILITY STATEMENT

unlimited

14. SECURITY CLASSIFICATION

(This Page)

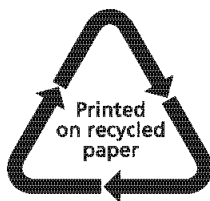
unclassified

(This Report)

unclassified

15. NUMBER OF PAGES

16. PRICE



Federal Recycling Program



UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, DC 20555-0001

OFFICIAL BUSINESS



@NRCgov

**NUREG/CR-7278
Draft**

**Technical Basis for the use of Probabilistic Fracture
Mechanics in Regulatory Applications**

September 2021