Important Aspects of Probabilistic Fracture Mechanics Analyses

Technical Letter Report prepared by Patrick Raynaud, Mark Kirk, Michael Benson, and Matthew Homiack, RES/DE/CIB, in response to Task 1 in UNR NRR-2016-004

Contents

Executive Summary

In response to increasing interest in probabilistic fracture mechanics (PFM) analyses to support regulatory decisionmaking, this report develops the concept of a PFM analysis methodology and outlines important considerations for a high-quality and high-confidence PFM analysis. Fracture mechanics analyses have historically been performed deterministically; however, in response to the need to quantify risk in regulatory decisionmaking, an increase in the use of probabilistic approaches has occurred. While both deterministic and probabilistic are only as good as the models we develop to represent reality, they differ in the way that they treat uncertainty. Each approach has its advantages, however increased emphasis on cost-benefit analyses and risk based metrics may provide opportunities for use of probabilistic methods.

A general framework to describe, perform, and evaluate a PFM analysis is presented in this report. 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 model 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 between inputs.
- Uncertainty characterization and treatment is at the core of a PFM analysis. In many PFM analyses, it may be useful to separate epistemic and aleatory uncertainty. 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. Concepts and methods to verify and validate a probabilistic framework are discussed.
- Ways to demonstrate PFM convergence include varying sample size and sampling strategy, as well as performing stability analysis. Output uncertainty analysis can take various forms depending on the problem being analyzed. Sensitivity analyses can help to identify the drivers of uncertainty for a given problem or output. Sensitivity studies are useful to understand which parameters drive the issue being investigated, and to show that some expected trends are indeed reflected in the analysis results. Methods to perform such studies are presented.

Following this report, the next step towards the development of guidance to perform and evaluate PFM analyses is the development of a draft regulatory guide, accompanied by a more detailed technical basis NUREG document. In parallel, a pilot study will be defined and executed to test the draft guidance produced and to identify deficiencies and areas that need improvement. The lessons-learned from this process will be incorporated in a final guidance document, and the technical basis NUREG will be revised accordingly.

It is important to keep in mind that creating a proper PFM analysis by following the regulatory guidance that will be developed in this project may be only one part of the acceptance of analyses in a risk-informed submittal. The NRC's process of risk informed decision making is rooted in several tenants that are aimed at ensuring that known and unknown uncertainties are accounted for. Therefore, PFM (or PRA) analyses should allow for the direct consideration of the uncertainties that one can quantify.

The other aspects, such as defense in depth or condition monitoring, all ensure that one can address phenomena and occurrences that are not accounted for. Regulatory Guide 1.174 describes the requirements for risk-informed design basis changes based on assessing their impacts on core damage frequency. Assessing the impacts of relief requests based on changes in core damage frequency alone may not be appropriate in the case of very low probability events

Definitions

This section lists and defines important concepts and terms mentioned throughout this Technical Letter Report, so as to promote a clear common understanding of the ideas and discussions to follow.

Accuracy and Precision:
Accuracy is the degree to which the result of a measurement, calculation, or specification conforms to the correct value (i.e. reality or a standard accepted value). Precision is a description of random errors, a measure of statistical variability for a given quantity. In other words, ACCURACY is the proximity of measurement results to the true value; precision, the repeatability, or reproducibility of the measurement.

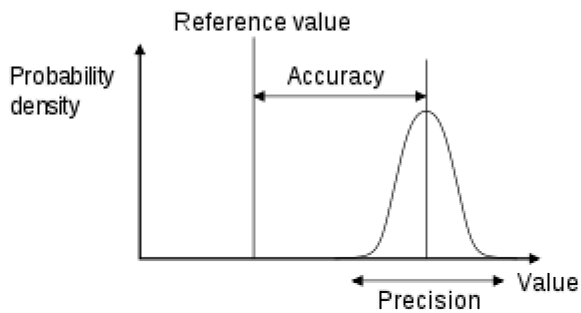

Figure 1: Illustration of the concepts of accuracy and precision

Aleatory Uncertainty:
Also defined as statistical uncertainty, aleatory uncertainty is one that is presumed to be the intrinsic randomness of a phenomenon.

Epistemic Uncertainty:
Also defined as systematic uncertainty, epistemic uncertainty is one that is presumed as being caused by lack of knowledge (or data).

Probabilistic Fracture Mechanics
PFM is any application of fracture mechanics that considers some of the inputs to be random. A prime example of a random input is initial crack size, which is seldom accurately known and usually has a strong influence on lifetime. In nuclear power plants, PFM may be used to calculate probabilities or frequencies of failure for components of interest. These probabilities or frequencies of failure may then be used as inputs in PRA calculations.

Probabilistic Risk Assessment:
PRA, also sometimes called “Probabilistic Safety Assessment” (PSA), is an engineering analysis process that utilizes the current state of knowledge to provide answers to three fundamental questions [1]:

- What can go wrong?
- How likely is it?
- What are the consequences?

A fourth question related to PRA is often “how confident are we in the answer to the first three questions?”

Within the regulatory context, PRA can be used as an engineering analysis process that models a nuclear power plant (NPP) as a system. By answering the three fundamental questions, it provides both qualitative and quantitative information. It builds on and represents the current state of knowledge regarding (1) what design, operation, and environmental (physical and regulatory) aspects of the plant are important to risk, what elements are not important, and why, and (2) what is the magnitude of the risk associated with the overall plant.

A NPP PRA model will, for some plant operational mode (e.g., at-power operation, low power conditions, shutdown conditions), typically delineate and analyze a large number of accident scenarios, starting with an "initiating event," (i.e., a disturbance to plant normal operations such as the loss of coolant), an internal hazard (e.g., a fire or flood starting within the plant boundaries), or an external hazard (e.g., an earthquake or external flood), and ending with some predicted level of consequence. Regarding consequences, the definition of PRA is sufficiently broad to encompass a wide variety of analysis endpoints. Thus, depending on the needs of the problem at hand, a PRA can assess consequences in terms of, as examples, core condition, radioactive material release from the plant, or offsite effects.

Realization:
In probability and statistics, a realization, observation, or observed value, of a random variable is the value that is actually observed (what actually happened) in a given single simulation. Within the context of probabilistic analyses, a realization often refers to a single run (or simulation) of the model.

Sensitivity Analysis:
Sensitivity analysis is the process of characterizing the impact of variability in the model inputs on variability in the model output. That is, it is the study of uncertainty in an output of interest (e.g. probability of rupture) due to the uncertainty in the inputs (and model) of the system under consideration. Therefore, sensitivity analysis should be used throughout a PFM analysis. During stability analysis, sensitivity analysis is used to help determine what sampling options must be adjusted to achieve a converged solution. Once a converged solution is found, sensitivity analysis techniques can be used to determine the input variables that have the most impact on the uncertainty in the results. Further discussion of sensitivity analysis is provided in Section 6.3.

Sensitivity Study:
Sensitivity studies are used to determine the effect of different models and assumptions on the quantities of interest (QoI). In this sense, they are "what if" analyses performed for individual sub- or full-model assumptions to determine how results might change under different, yet relevant, assumptions. Ideally, results would be insensitive to a collection of assumptions in the sense that regulatory decisions would not change across this collection. However, a well-planned sensitivity study can elucidate the assumptions that could change regulatory decisions. The plausibility of these assumptions can then be further evaluated by subject matter experts. Additional discussion of Sensitivity Study is provided in Section 6.4.

Stability Analysis:
Stability analysis compares results from several probabilistic simulations. The goal is to find the simulation options that provide solutions which provide converged, stable solutions for the specific application. Further discussion of Stability Analysis is provided in Section 6.1.

Uncertainty Analysis

Uncertainty analysis is the process of characterizing the uncertainty of the analysis output. This may include calculation of means and percentiles for quantities of interest, as well as determining the confidence intervals on quantities of interest for outputs of interest.

Validation:

The process of determining the degree to which a model is an accurate representation of the physical phenomenon considered from the perspective of the intended uses of the model.

Verification:

The process of determining that a model implementation (1) accurately represents the conceptual description of the model, and (2) correctly implements the chosen coding solution for the model. In particular for software, this process implies ensuring that the code being verified meets the software requirements.

Uncertainty propagation:

The uncertainty in the input is propagated through the model using sampling-based methods. The available means to propagate uncertainty are ultimately dependent on the computational constraints as well as the nature of the inputs and outputs under consideration.

xLPR:

The Extremely Low Probability of Rupture code, a fully probabilistic fracture mechanics code to simulate pressurized water reactor primary piping cracking, leakage, and failure.

1 Introduction

1.1 Motivation for Increasing Confidence in Probabilistic Fracture Analyses

In the past, the US Nuclear Regulatory Commission (NRC) has typically regulated the use of nuclear structural materials on a deterministic basis. Safety factors, margins, and conservatisms were used to account for model and input uncertainties. However, in the mid-1990s, the NRC issued a policy statement [2] that encouraged the use of Probabilistic Risk Assessments (PRA) to improve safety decision making and improve regulatory efficiency. Since that time, the NRC has made progress in its efforts to implement risk-informed and performance-based approaches into its regulation and continues to revisit and update the approaches on a regular basis.

A major component to the overall safety of nuclear structures is the fracture behavior of the materials. Consensus Codes and Standards responsible for the design and analysis of such structures, such as ASME Boiler and Pressure Vessel Code, typically rely on conservative fracture models with applied safety factors and conservative bounding inputs to account for the numerous uncertainties that may be present. Improving the efficiency of such models by truly understanding the impacts of the assumptions and uncertainties becomes difficult because of the conservative nature of the models and inputs and the inadequate documentation of the basis for safety factors. PFM is an analysis technique that can be used to complement deterministic fracture mechanics analyses to gain greater insight into the structural integrity of components. Two notable NRC efforts in the area of PFM include the FAVOR (Fracture Analysis of Vessels - Oak Ridge) [3, 4] and xLPR (eXtremely Low Probability of Rupture) projects [5].

PFM allows the direct representation of uncertainties through the use of best-estimate models and distributed inputs. (i.e., inputs for which the likelihood of being equal to a given value is represented by a probabilistic distribution). The term ‘best-estimate’ refers here to models that attempt to fit data or phenomena as well as possible. That is, models that do not intentionally bound or bias data for a given phenomenon. The PFM analysis methodology permits determination of the direct impact of uncertainties on the results, which gives the user the ability to determine and possibly refine the specific drivers to the problem. However, PFM analyses can be more complicated and difficult to conduct than deterministic analyses. Determining validated best-estimate models, developing input distributions with limited experimental data, characterizing and propagating input and model uncertainties, and understanding the impacts of problem assumptions on the adequacy of the results, can complicate the development and approval of PFM analyses in a regulatory application. Moreover, the underlying deterministic models that drive PFM are sometimes simplistic to permit reasonable solution times. The uncertainty introduced in a PFM by overly-simplified deterministic models can be difficult to quantify.

The Electric Power Research Institute (EPRI) has submitted several Materials Reliability Program (MRP) reports containing PFM analyses to the NRC for review and approval. Such efforts include MRP-105, MRP-113, MRP-116, MRP-206, MRP-335, and MRP-395. In general, NRC staff reviews of PFM submittals have been difficult and lengthy, although there have been successful applications of PFM in the cases of risk-informed inservice inspection, and the elimination of vessel weld inspections, for example [6, 7]. The reasons for these difficulties are complex and include, but are not limited to: lack of proper input choice justifications; choice of inputs that does not reflect observations from operational experience; insufficient description of the PFM tools used in the analyses; insufficient verification and validation of the tools used for PFM analyses; and insufficient analysis of results’ sensitivity to choice of inputs. The overall impression from the staff’s perspective has been that defining quality assurance requirements and acceptable methodologies to perform a PFM analysis would be very beneficial for the development and use of PFM analyses in a regulatory applications.

This Technical Letter Report provides some thoughts on how to improve confidence in structural analyses performed using PFM, by focusing on topics such as problem definition, PFM model development, input definition, uncertainty analyses, probabilistic framework development, and output analysis including uncertainty analysis, sensitivity analyses (to determine impact of uncertainties on result) and sensitivity studies (to determine impact of mean values on the results). By determining the main drivers to the probabilistic results and investigating the impacts of the assumption made to develop those drivers, the confidence in the overall results can be greatly improved. Some preliminary thoughts on regulatory review of PFM analyses are also discussed.

1.2 Objective

The aim of this work is to develop a methodology that is sufficiently general to be suitable to guide the development and/or critique of any PFM analysis, and that also effectively and logically manages the level of detail and breadth of scope that PFM analyses can take on. It is anticipated that the project will produce not only reports documenting such a methodology but also Regulatory Guidance describing its implementation.

1.3 Report Structure

Chapter 2 proposes a definition of a fracture mechanics analysis of component structural integrity. The Chapter discusses the commonalities and differences between deterministic and probabilistic approaches to these analyses, and reflects on differences in situations that may favor each methodology. Chapters 3, 4, and 5 discuss (respectively) some of the major building blocks of a PFM analysis:

- Chapter 3 discusses questions and issues associated with models (e.g., a model of fracture toughness and how it varies with temperature).
- Chapter 4 discusses questions and issues concerning inputs representation within a PFM analysis (e.g., a value of reactor coolant temperature might be an input).
- Chapter 5 discusses questions and issues associated with how uncertainties in both models and inputs are represented, and discusses questions and issues associated with the so-called "framework" that binds the models, the inputs, and their uncertainty characterization into the overall mathematical representation of the system. Key here is the critical need to evaluate uncertainty propagation within not only the individual pieces of the model but also within the framework itself.
- Chapter 6 discusses questions and issues associated with the PFM outputs, which includes such topics as convergence and sensitivity studies, and code validation/verification.

2 Definition of a PFM Analysis

2.1 The Framework for Structural Integrity Analysis Provided by Fracture Mechanics

Any analyses involving fracture mechanics, be they deterministic (DFM) or probabilistic (PFM), involve the same key elements, as explained by way of example using the following equations motivated by the classical work of George Irwin and colleagues [8]:

$$\text{Is } K_{APPLIED} \begin{bmatrix} \leq \\ \text{or} \\ > \end{bmatrix} K_{RESISTANCE}\ ? \tag{1}$$

$$K = \sigma \times \sqrt{\pi a} \times F \tag{2}$$

Eq. (1) represents, an example of a simple fracture mechanics-based failure criterion. Within Irwin's so-called linear-elastic fracture mechanics[1] (LEFM) K_{APPLIED} represents the applied driving force to fracture produced by the structural loading while $K_{\text{RESISTANCE}}$ represents the resistance to this driving force provided by the materials of construction. These two quantities are compared to assess the likelihood of structural failure: if K_{APPLIED} remains below $K_{\text{RESISTANCE}}$ then failure is unlikely, conversely conditions for which K_{APPLIED} exceeds $K_{\text{RESISTANCE}}$ have a greater probability of failure.

Either of these K values can be calculated using eq. (2), where σ represents the applied stress, a represents the crack size, and F is a non-dimensional factor that accounts for the geometry of a structure or of a fracture toughness test specimen

To determine $K_{\text{RESISTANCE}}$ any one of several standard test methods developed by the American Society of Testing and Materials (ASTM) can be employed [9, 10, 11]; these will give the form of eq. (2) that is appropriate for the specimen geometry being tested and the conditions of the test.

To determine K_{APPLIED} for a structure (e.g., a pipe, a pressure vessel, a baffle bolt, &c.):

- The stress (σ) is typically determined by either finite-element analysis of the actual structure or by a closed-form analysis of a simplified representation of the structure.
- The crack size (a) may be determined by non-destructive evaluation (NDE) of an indication found to exist in the structure, might represent the size of a flaw NDE could miss, or might represent a nominal flaw size agreed to as appropriate for certain types of assessments.
- The non-dimensional geometry factor (F) is typically determined by either finite-element analysis of the actual structure or by a closed-form analysis of a simplified representation of the structure.

1 LEFM applies to situations where the effects of plasticity are highly localized to the crack tip and do not interact with any structural boundaries. Elastic Plastic Fracture Mechanics (EPFM) extend LEFM concepts to higher loading levels. That being said, the key elements of both LEFM and EPFM are identical. High Temperature Fracture Mechanics (HTFM), which may involve the active degradation mechanism of time dependent creep crack growth, is not important for current US NPP but will be important for future advanced plants being considered.

While the techniques and technologies needed to determine these variables ($K_{\text{RESISTANCE}}$, K_{APPLIED}, σ, a, F) can be complex, their ultimate use in a fracture-mechanics assessment of structural integrity involves only simple algebra.

It can be said that fracture toughness, applied stress, flaw size, and component geometry are common to any fracture mechanics analysis. Only two additional variables, environment and time, need be added to be able to completely understand, characterize, and evaluate virtually any fracture mechanics analysis. All of these variables may or may not evolve with time and spatial location within a component or structure. Furthermore, as a whole, the structure is evolving toward some end state, an end state that is characterized by PFM using a quantitative metrics (e.g., failure occurrence).

Regardless of the details of a particular problem, these six categories of variables (i.e., material properties, stress, flaw size, environment, spatial extent, and time) comprise the most important aspects of any fracture mechanics analysis, whether it be DFM or PFM. It should be noted that human interaction such as inspection is considered as a measure to confirm the validity of the fracture mechanics model, and repair or mitigation is considered as a measure to change the key, if not all, variables.

There are two more aspects of real-world problems that fracture mechanics analyses need to reflect, as follows:

- **Dependencies between variable categories:** Interdependences may exist between these six variable categories; their accurate representation in the DFM or PFM model is key to the development of a model that reasonably approximates reality. As one example, the environment may change with time, a change that produces changes in other variables such as the material properties (e.g., material degradation or time-dependent aging) and/or the flaw size (environmentally assisted crack growth).
- **Multiple time and spatial scales:** There may be multiple time scales that need to be modeled, for example:
    - o *In a reactor pressure vessel* loading that may give rise to fatigue damage occurs over the months and years the plant operates. Superimposed over this are possible accident loadings from unanticipated/faulted conditions; this would occur over the time scale of minutes to hours. Superimposed over this is the possibility, albeit remote, that cleavage fracture could occur because of one of these accidents (e.g., a pressurized thermal shock transient). Were this to happen the fracture would make its way through the vessel wall in fractional seconds.
    - o *In piping systems* inspections are performed that can result in the subsequent weld repair (including weld overlay or inlay) or mechanical stress improvement (MSIP) of found-cracks. Both actions (repair or MSIP) could be viewed as altering the material properties, stress pattern, or flaw geometry, and resetting position on the time axis for the repaired component.

2.2 Similarities and Differences between Probabilistic and Deterministic Fracture Mechanics Analysis Methodologies

Historically, most assessments of structural integrity using fracture mechanics are performed deterministically. That is, a single-value estimate of the magnitude of the driving force for structural failure (K_{APPLIED}) produced by loading events is compared to a single-value estimate of the resistance of the materials from which the structure is fabricated to that driving force ($K_{RESISTANCE}$). The use of single values is an approximation. As an example, the $K_{APPLIED}$ produced by loading cannot always be the same because applied pressures and temperatures, even when controlled, vary somewhat. Likewise, the size of cracks (that is, the variable *a* in Eq. (2)) is uncertain. Similarly, $K_{RESISTANCE}$ is a distributed quantity because real engineering materials are not "*homogeneous and isotropic,*" which is but one reason why fracture toughness data shows a range of values even under tightly controlled conditions. ***Thus, $K_{APPLIED}$ and $K_{RESISTANCE}$ are both <u>inherently</u> distributed quantities*** (that is, they are best represented by a range of values, and not by a single value).

Figure 2-1 depicts the actual situation just described. The bell curves located on the left of each subfigure represent the distribution of the driving force in the structure ($K_{APPLIED}$), while the curves on the right of each sub-figure are the toughness distribution ($K_{RESISTANCE}$). If these distributions are widely separated (top row, left-most column) then there is either no probability of structural failure, or it is exceedingly small. Other degrees of distribution overlap are of course possible; these correspond to different probabilities of structural failure. The actual situation depicted by the left-most column is ***approximated differently*** by the DFM and PFM modeling approaches, as follows:

- <u>DFM analysis (the middle column).</u> Deterministic models represent the uncertainties in $K_{APPLIED}$ and $K_{RESISTANCE}$ by a number of techniques (e.g., conservative bounding values, conservative sub-models, margins) all of which permit bounding of the distributed $K_{APPLIED}$ and $K_{RESISTANCE}$ quantities using single values, usually representative of some high or low percentile bound. To ensure a safe, or "conservative," outcome (i.e., failures are over-estimated), features that contribute to driving force are systematically over-estimated (e.g., higher loads, bigger flaw sizes) while features that contribute to material resistance (i.e., fracture toughness) are systematically under-estimated. Furthermore, as shown in the first row, a very unlikely occurrence of events would be considered as a "zero probability of occurrence" without actual estimate of what the real probability of occurrence is. Without characterization of the conservatism involved, a zero deterministic probability of occurrence is not informative. Similar comment can be made on the second row, with the separation between "failure" and "no-failure", which as serious consequence, is decided by the level of conservatism and needs to be documented. These issues associated with DFM can be mitigated by imposing a safety factor (e.g., 3 as specified in ASME Code, Section XI) on top of the conservatisms in overestimating driving force and underestimating material resistance. The former represents explicit conservatism, and the latter represents implicit conservatism.
- <u>PFM analysis</u> <u>(the right most column).</u> Probabilistic models represent the uncertainties in $K_{APPLIED}$ and $K_{RESISTANCE}$ by one of two ways:
    - As distributed quantities that seek to represent the distributions in the component parts as accurately as possible. The estimates of the distributed quantities are shown as darker bell curves superimposed over the actual distributions shown by the lighter bell curves. In principle the estimated variability is always less than the actual variability, if only by a little, because <u>all</u> factors affecting variability can never be fully identified. This approach is typically used for the model components thought to be most important or significant to predicting the outcome
- In some situations, it is necessary to represent a probability distribution with a discrete input or model which produces a discrete output. This approach is typically used for model components that are either thought to not contribute greatly to the predicted outcome, and/or for those model components about which the state of knowledge is low. In these instances, best-estimate values or models should be employed. As in the case in purely

probabilistic assessments, conservatisms should be applied in the final step of the analysis (e.g. in the acceptance criteria) and not at intermediate steps of the process. This embrace of certain DFM aspects within a PFM should be the exception rather than the rule, and should be undertaken only after consideration, documentation, and elimination of approaches that represent the uncertainties as probability distributions.

- As illustrated previously, the overlap in the driving force and resistance distributions provides a means to estimate of the probability of structural failure.

Both DFM and PFM approaches are thus fundamentally similar: both are mathematical abstractions used to approximate reality. Moreover, both share the common goal of representing the uncertainties in a mathematical form that enables problem solution, although this is achieved by different means.

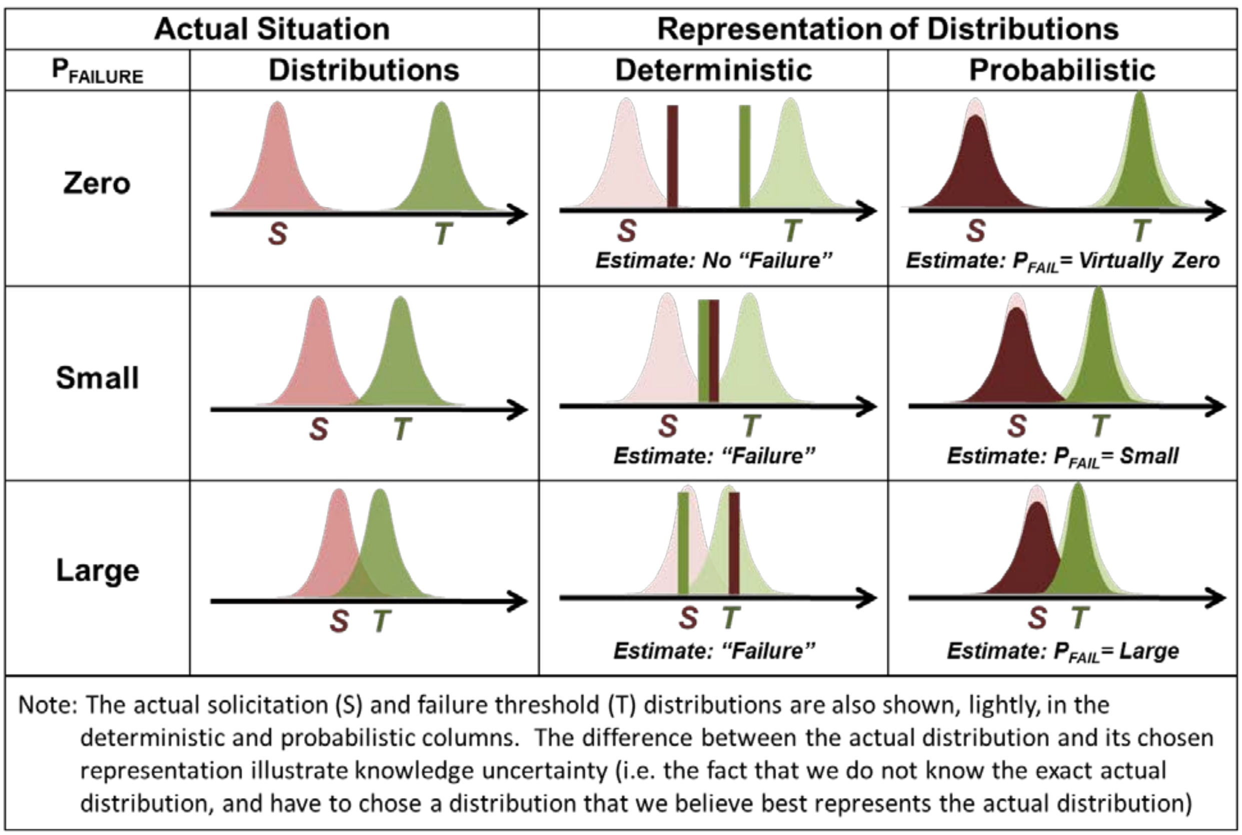

Figure 2-1. Illustration of how deterministic and probabilistic modeling approaches represent the actual distributions of driving force ($K_{APPLIED}$) and structural resistance ($K_{RESISTANCE}$), and how these representations influence the estimate of structural performance that is the outcome of the analysis.

The very fact that a failure *probability* is one possible numeric outcome of a PFM analysis has been viewed by some as an impediment to the acceptance of PFM techniques. The following are possible reasons for this impediment and arguments to alleviate them:

- Statements regarding a probability of failure, however remote, explicitly acknowledge this possible outcome. In contrast, DFM obscures this possibility via its lexicon; DFM results are often expressed in terms of factors of safety or margins against failure. Nevertheless, it must be acknowledged that the language engineers use to describe DFM analysis results do not alter outcomes: these structures can, and sometimes do, fail. In this respect, PFM language is seen as more accurate and transparent than that used to express DFM results.

- Because engineers have been trained to think in terms of factors of safety or margins against failure, some of them may find the idea of a failure probability too abstract to grasp. However, the terms "factor of safety", "margin against failure", and "probability of failure" are *all* numbers reflecting levels of conservatism. Recognizing this similarity promotes an easier acceptance of the PFM as an alternative to DFM.
- Procedures to review PFM assessments, and probabilistic acceptable criteria are, in general, less well established and less formalized than for DFM assessments. Indeed the point of this project is to directly address this issue.

To summarize, the key points of this section are as follows:

- Both DFM and PFM analyses are models used to represent reality, and although use of DFM has met its objectives, as reflected in various engineering support for different purposes throughout the nuclear industry, there is more room for using PFM as an alternative to better quantify implicit and explicit conservatism in nuclear components
- Both DFM and PFM analyses treat uncertainties mathematically, but in different ways. Furthermore, it may not be possible to properly account for every uncertainty.
- PFM analyses will represent most, but not necessarily all, uncertainties identified so far as important distributed quantities.
- Both PFM and DFM use specific metrics to quantify results.
- Use of DFM or PFM methods does not represent an either-or choice. Indeed, in many circumstances DFM and PFM methods can, and have, been used as complimentary parts of a safety case.

2.3 Considerations in the Use of PFM

While not exclusively so, the great majority of engineering structures operating today were originally designed on a deterministic basis. This is also true of many NRC regulations, which were developed using prescriptive, deterministic requirements based on a combination of experience, test results, and expert judgment. In developing those requirements, the NRC staff considered engineering margins (both implicit as associated with the use of conservative models, and explicit as associated with safety factors) as well as the principle of defense-in-depth (also called "redundancy"). It was assumed that undesirable events, however improbable, might occur, and so the plant designs included safety systems capable of preventing and/or mitigating the consequences of those events.

Over time, the reliance on a deterministic basis for engineering designs and regulations has given way to an increased use of probabilistic techniques for such purposes. Many factors support and motivate this evolution, including the following:

- **<u>Public Policy Decisions</u>:** In the early 1990s Congress passed a law called the "Government Performance and Results Act" (GPRA) [12]. One objective of that law is to *"improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction."* In response to this law, the NRC committed to move toward risk-informed, performance-based regulation and issued a policy statement that encouraged the use of Probabilistic Risk Assessments (PRA) to improve both safety decisionmaking and regulatory efficiency. This policy statement formalized the Commission's commitment to the expanded use of PRA, and stated in part that, "the use of PRA technology should be increased in all regulatory matters to the extent supported by the sate-of-the-art in PRA methods and data and in a manner that compliments the NRC's deterministic approach and supports the NRC's traditional defense-in-depth philosophy." Therefore, use of PFM for passive pressure boundary components, which is the topic of this report and this research effort, is aligned with the Commissions' statement.

- **Factors unanticipated in the design phase, and/or not addressed by Codes and Standards:** There is a fundamental difference between how deficiencies, or potential deficiencies, discovered during the design and construction of a structure are addressed versus those discovered later, often after many years or decades of safe service. During design and construction, deficiencies that do not meet specification can be addressed by repair, replacement, or reconstruction without consideration of whether the deficiency poses an actual challenge to the safe operation and functionality of the structure. However, once operation begins, repairs that were considered feasible during construction can become cost prohibitive (here "cost" can be expressed in terms of dollars, time, or dose). While the NRC's prime mission is safety, it is nevertheless obligated [see 10 CFR 50.109(c)(5) and 10 CFR 50.109(c)(7)] to assess if safety benefits justify the attendant cost.

  PFM assessments are ideally suited to such situations, because PFM metrics relate directly and clearly to systems that can challenge safety (i.e., probability of structural failure). Indeed an assessment of risk is explicitly required by the backfit rule [see 10 CFR 50.109(c)(3)]. Conversely, DFM analyses are implicit and therefore the safety impact is quantified differently. Typically, in DFM, safety is assessed based on the concept of safety margin: a larger margin implies added safety. In contrast, PFM is deemed safe if acceptance criteria are met: being farther from the acceptance criteria (as long as they are met with a high degree of confidence), implies a greater degree of safety.

  PFM also provides methods to account for factors that occur during service (e.g., new damage mechanisms, unanticipated loadings, and aging) that were not considered during design. Especially when such factors are encountered for the first time, the performance of analyses following the guidelines of codes, standards, and regulations can be difficult because these established procedures may not account for the new factors. Historically unanticipated material degradation mechanisms have arisen approximately every seven years [13] in NPPs, the latest of these being the PWSCC aging issue in PWRs in Alloy 600 and 182/82 welds which led in part to the development of the xLPR code. Attempts to force-fit the new phenomena into an existing codified framework is complicated by the many compromises and judgments needed to fit the new phenomena into an existing framework. Rarely is the result transparent or quantifiable. In contrast the adaptability inherent to PFM fosters the clear treatment of such situations. However, unanticipated degradation mechanisms may require the implementation of new deterministic models into the PFM code to capture the new degradation mechanisms. Because of this, both DFM and PFM codes must continually be updated.

- **Use of structures beyond their design, or licensed, lifetimes:** One currently important case of a factor unanticipated during the initial construction of NPPs is operation for more than the original licensed life. As plants consider additional 20-year license extensions, and imposing a safety factor of three in the DFM method becomes impractical, PFM methods could be used to better quantify the risks associated with operation of NPPs beyond the initial licensed life.

Beyond these specific benefits, PFM also offers some intangible benefits beyond the reach of DFM. In particular, PFM may force analysts to take a deeper look at the data used to develop models or inputs, and to actually be more critical of this data. This stems from the fact that it is often easier to conservatively bound data than to attempt a best-fit model with quantified error bounds. As a result of this process, analysts are likely to achieve a better understanding of the data they use in their inputs and models.

In the two recent cases where the NRC has pursued a PFM assessment (i.e., for reactor pressure vessels in the Alternate PTS rule and for piping systems in the xLPR project) the rigor of the process needed to build the models revealed deficiencies in them that may have been otherwise overlooked. As some examples, in developing the FAVOR PFM model that supported the Alternate PTS rule it was discovered that the initial sampling scheme implemented for the embrittlement trend curve double-counted uncertainties, and that crack-face pressure for internal surface breaking cracks had not been modeled. Similar improvements are expected as the NRC staff adopts a uniform process for its review and assessment of PFM analyses submitted by licensees.

3 Models for PFM Analyses

While the probabilistic framework manages input sampling and compilation of results, the individual models utilize the sampled inputs to approximate a physical system and generate the outputs. Most individual models, as part of the overall PFM analysis, are deterministic by nature. Multiple models may be interdependent within the same analysis, with the probabilistic framework passing variables from model to model before the final result is generated. This section discusses various aspects of selecting and substantiating models for a PFM analysis.

3.1 Model Selection

3.1.1 General Concepts

The goal of any engineering assessment methodology is to determine the response of a system to a variety of inputs. The assessment of the system in question should be designed and developed using a set of models that best represents the physical behavior of the system of interest[2]. At times there may be more than one model available for predicting the response of the system to a particular phenomenon. Expert engineering judgment may be used in selecting which model to employ, but including all available models allows for sensitivity studies and determining the impacts of model predictions on the response of interest.

Model selection may involve balancing the accuracy and practicality of various mathematical approaches. In weighing these concerns, the analyst must consider the desired level of convergence of the PFM analysis (i.e., the number of realizations required to achieve confidence in the result). The other factor here is computational resources. While a particular approach may be considered "best estimate," it may not be practical for a PFM analysis given the time and resource constraints imposed on the analyst. The occasional need for choosing a model that has less fidelity, but is easy to solve for due to solution speed requirements of PFM, may affect results. In such cases, biases or uncertainties relative to the "best-estimate" model should be quantified and accounted for by propagating the associated uncertainty through the probabilistic model. The model choice can further be complicated by the fact that PFM requires the use of the most accurate deterministic models rather than conservative models, but such models are not always a very good representation of the data, resulting in known model biases and uncertainties. A simple example of this is with regard to the surface crack stability models in the xLPR code. Currently in xLPR, surface crack stability is based on limit load solutions. This is because surface crack J-estimation schemes are not fully vetted and currently the only available J-Resistance curves are for compact tension specimens, which produce lower bound resistance curves. If a surface crack J-estimation scheme were used in conjunction with compact tension J-Resistance curves, the predictions would be overly conservative. As it is now, limit load surface crack stability methods will be non-conservative for lower toughness materials.

In all cases, the rationale behind model selection should be thoroughly documented. The documentation should include:

- A discussion of alternative approaches that were rejected

[2] Note that the chosen model must be simple, fast, and practical for the many realizations required by PFM.

- A discussion of how the selected model compares to potential other more complex models that better represent reality, or to reality itself
- A qualitative or quantitative discussion of the modeling uncertainties

3.1.2 Theoretical Models

The mathematical models used in PFM codes to represent physical processes or phenomena are based on evidence/understanding from a variety of information sources. Models chosen should be as realistic a representation of the actual physical system as is possible. Some models have a greater degree of theoretical underpinning, but even theoretical models usually have simplifying assumptions (e.g. isotropy, homogeneity). For example, models used to describe the stress in a pipe wall or the stress intensity factor for a crack rely on a solution of the equilibrium and compatibility equations of basic solid -mechanics [14]. As such, these models generally differ from the physical reality only inasmuch as the boundary conditions and/or initial conditions differ from that reality. Even so, approximations are sometimes introduced in these models. In the case of the stress intensity factor calculation, a complicated loading condition on the crack faces may be approximated by a polynomial fit [15]. But the use of a general weight function method for calculating K for complex stresses may yield more accurate results [16]. The analyst should identify and account for the impacts of this type of approximation.

Theoretical models may present themselves in terms of closed-form equations, which are ideal computationally. Numerical discretization approaches (e.g., finite difference methods and nonlinear equation solving) in an individual model may have several drawbacks in a PFM framework, including high computational cost and possible spatial and temporal non-convergence. Interpolation of tabulated coefficients may be necessary in many physical models. Interpolation itself may be considered computationally burdensome. Fitting closed-form equations to tabulated data may lead to greater efficiency, but uncertainties inherent to uncertainty modeling (i.e. confidence bounds) or due to fitting errors should be accounted for.

3.1.3 Empirical Models

Other models that were developed using test data are sometimes called empirical models. While the data underlying these models often exhibit considerable scatter (uncertainty), it should nevertheless be recognized that this data reflects the underlying physical processes under the specific conditions where the data was gathered, and within the range in which it was gathered. As a result, if the data source in actual operational experience from components in service, the data may better represent the real situation being modeled than data from a controlled laboratory environment that simulates in-service conditions, though this may not always be the case. When incorporating empirical models into a PFM analysis, several considerations should be taken into account.

First, the analyst should have knowledge of and confidence in the dataset that gave rise to the empirical fit. This dataset should be well documented in terms of experimental procedures and data processing. Relevant information on the dataset may include:

- Experimental design, statistical design of experiments (Were the experiments completely random or was a certain range targeted for optimization? What was the degree of statistical replication?), or data generating mechanism
- Consensus standards referenced while conducting experiments and processing data
- Location or locations within a test specimen where the data were gathered

- Data screening procedures that may have led to exclusion of some data points from the dataset
- Ranges of parameters that were varied in the experimental work
- Measurement uncertainty information

The analyst should also establish an understanding of the statistical model used to fit the data, including any potential theoretical underpinnings of the statistical model. The model may have applicability limitations that originate from the mathematical form of the equation and/or the range of conditions covered by the underlying dataset. These limitations should be understood and incorporated in the computer coding of the model.

Finally, uncertainties related to goodness-of-fit and scatter about the best estimate curve should be assessed. For example, it should be assessed what estimator constitutes the best-estimate, and whether the model incorporated into the PFM framework uses the best estimate or a conservative bound. Using conservative bounds without scrutiny may lead to biased PFM analysis results. In both cases (conservative bound or best estimate), the analyst should assess the sensitivities of the approach on the final results of the PFM calculation.

3.1.4 Computational Models

Finite element or finite difference models are also directly part of some PFM models. As computational resources and solution speeds improve in the future, these models could become part of practical PFM models. However, presently limited by computational resources, these self-contained PFM models are for the most part not practical and are not further discussed here.

3.1.5 Model Validity

Model applicability limits should be considered for both theoretical and empirical models. Applicability limits for empirical models will most likely come from the underlying dataset (e.g., the model is valid only for the range of temperatures considered in the testing program). The applicability limits of a theoretical model may not be as straight forward to determine. The analyst may have to consider the physical system, the mathematical derivation of equations, relevant test data, and expert judgment in order to arrive at appropriate limits. Extrapolation (the use of models beyond their strict applicability limits) should be avoided when possible, regardless of the type of model (theoretical or empirical). Note that for example, an alternative option to extrapolation could be to define the input probability distributions so that such values are not sampled, by truncating the distributions to avoid regions where no information exists for the model validation. However, sometimes the joint collection of sampled inputs may produce a result outside the range of applicability and these distributions are not easy to specify a-priori due to the complex input-output relationship.

Since inputs are sampled randomly from distributions it is important to define the bounds of applicability/validity of any model (including the input distribution) used within PFM analyses. It is equally important to ensure that the data used to construct distributions for model parameters are indeed relevant for the problem being modeled, which may imply using only the subset of a large database that is most relevant to the problem at hand. The bounds of model applicability can be:

1. Determined from the limits of the data from which the model was developed,
2. Established through validation testing or through expert judgment, or

3. Established by physical limits of the system, and so on.

In any case, the bounds of model applicability should be explicitly documented, along with how the code handles those situations where the sampled inputs may be outside these bounds. An extreme example of these bounds might be to ensure that yield stress sampled is always positive.

When an applicability limit is violated, the coded model will have to take appropriate action. Exiting the PFM analysis completely upon violation may not be an ideal approach, especially when long runtimes are required. Warnings explaining to the user which input was in violation and the action the code took may be an acceptable alternative. However, the user should be able to make an informed judgement about the uncertainties introduced by violated limits.

3.2 Gaining Confidence in Models

Establishing confidence in the individual models is an important step in substantiating the PFM analysis as a whole. Confidence in models relies heavily on careful documentation of the limitations of the models and of the verification and validation activities that support the technical basis of the software. After identifying limitations of the models, the impact of these limitations on results should be understood. These aspects of PFM analysis are expounded upon in this section.

An important part of establishing confidence may involve running the code deterministically and initially focusing primarily on the physics. As any complex system with a large number of inputs, the confidence in the code and its models is not obtained only via mathematical verification, but also by gaining experience in exercising the code by analyzing the trends of the code outputs.

3.2.1 Model Definition, Verification and Validation

As introduced in Section 3.1, models will likely have a limited validity input domain specified by the PFM code developer. The analyst should document the rationale behind applications outside this validity domain. The actions taken by the coded model when the domain is violated should also be documented. This documentation should include guidance to the user on how to interpret results when the domain is violated and the action that the code took to remedy the situation. In addition, the likelihood of being outside of the model bounds should be mathematically assessed, including the physical reality of such states and the downstream consequence of such situations. For example, NRC analysts have encountered a case where K solutions were limited to a radius to thickness ratio smaller than 2, and thus could not be used on a thick-walled pressurizer surge nozzle that needed to be analyzed. In this case the NRC analysts were forced to artificially increase the pipe diameter while keeping the thickness constant, to obtain a result from the code.

Within the bounds of the validity domain, the model will have uncertainties. Potential sources of uncertainties include data fitting errors, scatter in the dataset used to define an empirical model, and simplifying assumptions in deriving equations. These sources of uncertainties should be documented for each model in a PFM analysis. The documentation should identify the relevant sources of uncertainty for each model and assess their impacts on final and intermediate results. An intermediate result may be fed to another model in the PFM analysis. In this case, the logic tying each model together to the final result should be documented.

Two important analyses used to gain confidence in models are often referred to as model verification and model validation. Only brief descriptions are given here, but good starting points for the vast amount of literature on the subject include [17] and [18]. Model verification is a quality assurance process by which the coded model is verified to meet the defined software requirements, i.e. the mathematics of the model are correctly coded. Someone independent of the model development should review all software requirements, define and execute relevant software testing, and report on any anomalies. Potential testing activities include comparing coded equations to documented equations, comparing code results to independent spreadsheet calculations, and any other activities to verify that the code requirements are met.

Model validation scrutinizes the degree to which the chosen model represents the physical system. This validation may take different forms. For example, in the case of a closed-form solution for arriving at mechanical stress, a finite element solution could be used to validate the range of applicability of the closed-form solution. Conversely, both theoretical and empirical models can be validated against experimental data not used in their original development. The owner of the PFM code should document the steps used to validate all models.

3.2.2 Documenting Model Uncertainty

As alluded to in Section 3.1, there are various sources of model uncertainty, such as:

- Simplifying assumptions in model derivation
- Scatter in the dataset used to create an empirical model
- Lack of experimental data justifying theoretical models
- The need to use a simple model instead of a more accurate but computationally intense model

All sources of uncertainty should be identified and documented for each model in the PFM analysis. The documentation should explain the impact of the uncertainty on intermediate and final results. Given this known uncertainty the user should take care in interpreting the results, in particular when assessing whether the final answer is conservative or non-conservative for the application. Caution must be used when determining the impacts of a given input range on the final output of a code. In fact, some inputs may affect different models in different ways: an input that may lead to conservatism in one sub-model may have the opposite effect in a different sub-model. Any relevant bases that justify the positions described on modelling uncertainty should also be documented. Quantifying uncertainties is useful but not always practical. If a qualitative description of model uncertainties and their impacts on results is chosen, then the analyst should explicitly state the reasons a quantitative approach is not feasible. Expert engineering judgment may be adequate in some cases to account for qualitative uncertainty in the results.

Careful documentation of sources of model uncertainty has several benefits. First, it provides assurance that the owner of the PFM code recognizes and accounts for the limitations of the software. This can be useful for users and decision-makers in appropriately interpreting the results. As an additional benefit, this process can identify potential areas of research for improving the PFM analysis. Section 7 of this report provides a more detailed explanation of uncertainty characterization.

3.2.3 Assessment of Alternate Models

Consideration of alternative models, when available, can provide an additional layer of confidence in the quality of the PFM analysis. If the model is considered by the technical community to be best estimate, then results from this model should fall between those of a known conservative model and a known non-conservative model. It should nonetheless be noted that most models are not consistently conservative or non-conservative, and many models will have a different bias depending on the input used. This type of analysis can bolster the technical community's consensus opinion. It may also be useful at the model selection stage.

Diversity of modeling approaches is an important concept here, as well as in other ideas expressed in this report. A numerical approach can be compared to a closed-form solution, or vice-versa. A theoretical model can be compared to an empirical model and the associated dataset. Such comparisons are part of an overall sensitivity study discussed above. Careful use of a high fidelity DFM model may help in this assessment.

4 Inputs for PFM Analyses

In PFM analyses, the common practice is to treat the inputs and associated uncertainties by representing them through probability distributions. Estimating the functional forms of these distributions and applying potential simplifications or corrections, such as for example bounding the distribution to avoid unphysical inputs, are essential to building confidence in the end results. This section addresses these issues and related considerations.

4.1 Input Types

Inputs to a PFM analysis can be represented in two ways: constant and random. The approach to choose the type of representation for each input may focus on available data and may rely on expert judgment. Distribution Fitting remains an important and ever growing subject of interest in statistics (see chapter 1 of [19] for instance). The two most classical techniques are moments matching and maximum likelihood (an example of such techniques can be found in [20]). Distribution fitting tools can be found both in dedicated software and generic mathematical and statistical software (Matlab©, R, JMP…). When the available data is relatively small or qualitative only, a broader approach relaxing the probabilistic framework can be considered, as presented in [21].

Constant inputs characterize an input variable by a single, point value. No matter the number of realizations used for the analysis, the input value is always the same. In contrast, random inputs can take on a range of potential values. Uncertainty in these inputs is typically represented by a probability distribution on the input space, and for this reason, uncertain inputs are sometimes also referred to as "distributed" inputs. Uncertain inputs are the core of a PFM analysis.

Many different probability distributions exist that can be used to represent data [22]. These distributions can be classified as either continuous or discrete. Some common parametric forms of continuous distributions include uniform, normal, log-normal, and Weibull. The binomial distribution is a common discrete distribution. To the analyst, the parametric probability distributions are characterized using several inputs. The analyst would first specify the type of distribution, then he or she should enter the necessary parameters to define its shape. For example, values for the mean and standard deviation would be entered to define a normal distribution. Joint parametric distributions, incorporating dependencies within a group of input variables, can also be specified. After the distributions are selected, a sampling algorithm is usually used to randomly select values from the corresponding probability distribution for use in the PFM analysis. In some less common cases, methods that are not sampling-based may be used, such as First/Second Order Reliability Methods (FROM/SORM), factorial design, or discretization of input values (for smaller problems). There are also other ways besides applying common parametric forms to specify probability distributions, such as nonparametric/empirical density estimation. Regardless, the framework is similar in that choices regarding the type of input distribution must be made. The next section discusses this decision process.

4.2 Construction of Probability Distributions

Proper development of a representative probability distribution for an uncertain input requires a detailed knowledge of the available data as well as qualitative judgments. In reality, however, collection of available data is always limited by cost and time, and developing a robust uncertainty distribution for every input may be impossible due to these limitations. The data itself also contains biases through measurement uncertainties, omissions, etc. Ultimately, the analyst should develop the uncertainty distributions with relevant data and be able to defend the chosen distribution. The amount and pedigree of the source data, as well as the influence of the particular input parameter on the results of the overall PFM analysis are all important considerations when justifying a distribution.

The typical approach is to gather a set of empirical data and then choose one of the parametric probability distribution forms as the best representation of that data. Many goodness of fit statistical tests exist and may also be used to assess how well a given distribution is suited to a dataset. Such tests include Kolmogorov-Smirnov [23, 24], Anderson-Darling [25], Shapiro-Wilk [26], and many others. In practice, the selection of the probability distribution form is frequently driven by engineering judgment. Large data sample sizes are typically needed to accurately determine the form of the underlying probability distribution, though the minimum number of datum needed for a suitable fit may be subjective and context-specific. However, the sample size is directly linked to the resolution of the desired probability distribution. Thus, fewer data can be acceptable when a coarse resolution is sufficient. Uncertainty analyses can be used to gauge the level of resolution that is appropriate for the analysis.

When choosing a probability distribution, several important parameters must be considered. First, it is important to consider the minimum and maximum of a distribution, to ensure that the correct support is selected. Next, a list of all that is known about the input being considered should be created. Last but not least, the shape of the distribution should also be considered. Common summaries of distribution shape include skewness and kurtosis. Skewness is a measure of the lack of symmetry of a distribution. Kurtosis is a measure of how heavy the tails of the distribution are and is typically quantified relative to the normal distribution. Estimating higher order moments of statistical distributions, such as skewness and kurtosis, typically requires large sample sizes; subsequently, expert judgment, coupled with data, is often used to inform distributional shape. Given that the tails of distributions often drive low probability events such as structural failures, it is important to investigate how much confidence there is in the underlying probability distributional form and whether the specified distribution fits the underlying data well in the tails. It should be pointed out here that similar but not identical distributions may sometimes lead to the same answer for a given problem, for example in the case of a beta distribution matching a truncated lognormal distribution or in the case of a triangular distribution that is close to the truncated normal distribution.

Once the form of the probability distribution has been specified, statistical estimation techniques can be applied to estimate the parameters of the distribution. Examples of such statistical methods include maximum likelihood estimation, moment-matching, and Bayesian inference. The analyst may need to assess how confident he or she is with the estimates of the distribution's parameters. Even small changes to the standard deviation can produce markedly different behavior in the tails of a distribution, so understanding and accounting for uncertainty in distributional parameter estimates is important. This can be achieved by performing sensitivity studies on input distribution parameters or by studying the impact if changing distributions for important inputs.

In PFM, it's sometimes necessary to estimate a probability distribution when sparse or no data are available. Literature and expert opinion are often sought in these cases. Key considerations are how the literature and expert information was developed and the extent to which it applies to the system under study. Different opinions from the literature and experts should also be sought and considered as applicable. Formal procedures for use of expert judgement have been developed at NRC in other technical areas (see https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1563/sr1563.pdf); these, and the more general literature on this subject, can be employed effectively in such situations. When compared to probability distributions developed from data, it can be challenging to justify distributions developed from limited data, however use of formalized procedures for treatment of expert opinion coupled with physical constraints inherent to the distributed quantity can add robustness to the justification. Sensitivity studies aimed at quantifying the impact of the choice of distribution type and distribution parameters may be used as a means to increase confidence in the ultimate outcome of the analysis.

4.3 Input Bounds

Many standard probability distributions have 'infinite support,' i.e. they can take on an infinite range of possible values. In practical engineering applications, inputs are typically bounded, and probability distributions need to be 'truncated' to accommodate the finite range of the inputs. Input bounds are the upper and lower truncation points applied to input probability distributions. These bounds are typically used to prevent the sampling algorithm from selecting input values that are undesirable and/or non-physical. For example, an analyst might select a certain bound to constrain the sampled inputs so that they remain within a model's specified range of applicability. Such an approach is often used on the front end of PFM analyses to prevent calculation errors, but limiting the input range because of model deficiencies should be properly documented and justified. The analyst may also wish to impose bounds to prevent a non-physical response from the system. Negative material strength is such an example. Selecting appropriate input bounds can be just as important as selecting the distribution itself because the bounds determine where the tails of the distribution are cut-off and thereby can have a large impact on uncertainty in the outputs. If bounds are to be applied to an input, their effect could be assessed in the sensitivity studies discussed in the previous section.

The selection of input bounds can range from straightforward to complex. Choosing bounds to constrain a model within its range of applicability can be straightforward if the bounds are derived directly from the technical basis for the model, but should nonetheless be accounted for. Importantly, care should be exercised if that same input is also used in another model with a different range of applicability. For example, one can easily picture a PFM code that can model crack growth as a function of applied stress intensity factor K for both fatigue and primary water stress corrosion cracking (PWSCC. In such a case, if one of the crack growth models is defined over a wider range for K values than the other, restricting K values to the smaller of the two ranges of applicability may not be an adequate choice. In fact, choosing to limit input values within a model's domain of applicability may artificially suppress uncertainty as well as some outcomes in the model predictions, particularly if the domain of applicability of the model only covers a small fraction of what is physically possible. Accordingly, it is important for the analyst to understand how the various inputs may interact within a model of a complex system. The selection of physical bounds can be more complex. Sometimes, the true physical limit may not be known, even if the distribution is a based on a large number of high quality data. Expert opinion is thus often necessary to establish physical limits. The technical basis supporting these opinions needs to be documented along with explanations for alternate opinions that were excluded. Additionally, although it is critical to set an appropriate bound a distribution tail that is known to drive failures, it may also be just as important to set a bound on the non-critical part. Such is the case because unrealistic values sampled from this region could still alter the overall results when a number of realizations are analyzed in aggregate.

4.4 Quantification of Input Assumptions

In a PFM analysis, focus is placed on those inputs that have the most influence on the output of interest. Uncertainty analyses are used to quantify the contributions of individual uncertain inputs to the uncertainty in the results. Once the input sensitivities are known, the analyst can focus on those inputs that have the most impact. If the input uncertainty has very little impact on the output uncertainty, a strong technical basis for the input distribution may not be necessary. If low probability events are the desired outcome, it may be appropriate to scrutinize the appropriateness of the tails of the input distributions that are driving the output uncertainty. In this case, sensitivity analyses and sensitivity studies that vary the input distribution type or parameters can be conducted to determine the impact of the chosen distribution type and its representation of the tails. If these results indicate a large impact, additional data, more refined statistical techniques, or further expert elicitation may be needed to further refine the input's probability distribution. In this way, the development of inputs for PFM analysis can be an iterative process.

Note that there are many different methods for determining input sensitivity [27]. These are further discussed in Section 6.

4.5 Conservatisms

Conservative treatments are the norm in deterministic analyses; however, they may not be appropriate in a PFM analysis where the objective is often to generate a best estimate distribution for the probability of occurrence of an event, and use a specific percentile (which will likely depend on the application) of that distribution as the final answer to the PFM analysis. Of principle concern is the bias that conservative inputs can impart on the PFM analysis results. In fact, in complex models, use of 'conservative assumptions' may not always lead to 'conservative results'. For example, choosing a high through-thickness crack growth rate may increase likelihood of a leak instead of failure, because less time will elapse before a leak, leaving less time for a long part-wall circumferential crack to develop and become unstable. On the other hand, some conservative assumptions can mitigate complex and uncertain inputs, such as for example assuming that a crack has initiated as opposed to using a problematic crack initiation model. Therefore, it is important to document all inputs that were considered to be conservative rather than best estimate. In some cases, it may be acceptable to have multiple inputs that are either conservative or non-conservative, if they are self-correcting in nature. The analyst should quantify the influence the conservative inputs have on the results so that the end impact of all conservatisms (and potential non-conservatisms) is understood. If the influence is minor, then the conservatisms may be sufficient. However, if there is a strong influence, additional work might be needed to refine the conservatisms.

Two other risks are associated with the use of conservatism. First is the risk of having too many variables set to conservative values, leading to a case that is extremely unlikely (for instance 10 variables set at their 90th percentile limits the analysis to an area in the input space with a probability of 10^{-10} or 1 chance in 10 billion). Second is the risk of erroneously considering a parameter as conservative due to its role in a sub-model, while its overall role in the whole system becomes non-conservative (e.g. high crack growth rate assumption leading to an increased leak rate may be considered conservative, but may lead to events being suppressed due to leak rate detection). In general, for a PFM analysis, it is appropriate to have all random variables set at their 50th percentile, or mean values. However, it may be appropriate to have constant variables set at a higher percentile, depending on the level of uncertainty pm these variables and the specific context of the analysis.

4.6 Dependencies

In a PFM analysis, most distributed inputs are assumed to be statistically independent from each other, and are sampled as such. However, a subset of input variables are often statistically dependent. The association between variables is often needed to ensure a physically possible input set, or to preserve the physical laws that drive the problem being modeled. For example, the yield stress of a material is usually smaller than its ultimate tensile stress, so a relationship between these variables may be imposed to ensure physicality of the inputs. Dependence between variables may or may not be causal. If two or more variables should be dependent according to, for example, the laws of physics, the input set should be checked for consistency and physicality. For example, this if the case for the yield strength of a material, which is always lower than its ultimate tensile strength.

Known variable dependencies can be represented by correlating two or more variables via mathematical relationships. This is desirable underestimating the influence of correlated variables. For example, if two variables are highly correlated with an output variable, it is likely that they are also correlated with each other, possibly even more strongly than they are correlated with the output of interest. If so, putting both of them in the model will result in some multicollinearity, risking both of them being declared insignificant, and skewing the correlation coefficient estimates.

5 Uncertainty Characterization and Framework for PFM Analyses

In this Section, epistemic and aleatory uncertainty types are defined, and the process for identification and classification of uncertain parameters between epistemic and aleatory are described. A probabilistic framework for uncertainty propagation in a PFM analysis is discussed, and thoughts on how to verify and validate such a probabilistic framework are presented.

5.1 Epistemic and Aleatory Uncertainty

The purpose of safety analyses is usually to assess the risk of some unwanted outcome, which can be interpreted discretely as a pair (c_i, p_i), where c_i represents the consequence (or severity) of a particular outcome *i*, and p_i its likelihood expressed as probability (which can also be expressed as a frequency). PFM analyses can be used to estimate this probability p_i. To do so, Monte Carlo techniques essentially regroup all uncertainties into a single integral, reducing the dimensionality of the problem to a more manageable size. The tradeoff is that it is not possible to directly reconstruct the multi-dimensional solution. In other words, the full extent of how the output is affected by the uncertainty is only approximated via statistical techniques or the use of response surfaces.

While it would be impractical to consider the multidimensional problem with all its aspects (not only with respect to the unrealistic number of runs that would be required but also because the interpretation would rapidly become too complex to comprehend), the risk community has for decades made the distinction between the inherent risk of each problem and the uncertainty due to lack of knowledge.

The purpose of this distinction is to acknowledge that there will always be a risk given a specific situation. The risk can be due to unwanted (and sometimes unforeseen) events without control (for instance natural disasters such as seismic events or other), human interaction (risk of error), or physical limitation (strength of the material with respect to the condition of use, undetected flaws due to their size or nature…). While this risk can somewhat be mitigated (thus changing the problem considered), it is inherent and present in each problem studied, and ultimately is the output that needs to be quantified and estimated for decision making. In addition to the risk itself, there is the uncertainty due to lack of knowledge. This may be due to simplified models, lack of data to express a poorly known input, indecision with respect to which model or input is the best representation for the problem at hand, or any combination of the above.

While both risk and lack of knowledge need to be acknowledged to have a good understanding of what may happen and what would be the consequences, their interpretation in terms of decision making differ:

- One (risk) would require a change in the problem considered to be affected (such as mitigation or change in inspection schedule). This uncertainty is grouped in what is called **aleatory uncertainty.**
- The other (lack of knowledge) may (in theory) be reduced by more research on a particular model or input and does not necessarily require a physical change in the problem in consideration. It is called **epistemic uncertainty.**

Daneshkhah [28] provides the following definitions for aleatory and epistemic uncertainty:

- Aleatory Uncertainty: This uncertainty arises because of (perceived) natural, unpredictable variation in the performance of the system under study over which there is little control (the variability in fracture toughness), or the inherent randomness in the future (for example adverse events such as natural disasters or possible human error). The knowledge of experts cannot be expected to reduce aleatory uncertainty although their knowledge may be useful in quantifying the uncertainty. Thus, this type of uncertainty is sometimes referred to as irreducible uncertainty. This implies cases where there is no way to know for sure what will happen or what the value is/will be, but based on past experiment/observation, there is a sense of what may happen. It can also represent error in experiments used to gather data (accuracy of the material used, inherent randomness from the experiment).

- Epistemic Uncertainty: This type of uncertainty is due to a lack of knowledge about the behavior of the system that is conceptually resolvable. That is, it could potentially be reduced thanks to further research, model development, model choice, or testing. It may also result from the choice of a model that is necessary for PFM due to solution time constraints. This uncertainty reflects the level of confidence that people have with regards to a possible value or outcome (considered as unique but poorly known); this confidence may be increased by gathering more information.

It is important to point out that most inputs contain both aleatory and epistemic uncertainties, but are nonetheless usually categorized as either one or the other, but not both. However, these two uncertainties are often combined, since there often exists some information that may reduce uncertainty up to a certain point. This is why the expression “perceived randomness” is used to define aleatory uncertainty.

The approach sometimes used by the risk community to separate aleatory and epistemic uncertainties is to split the unique Monte Carlo integral into two: one to represent aleatory uncertainty and the other epistemic uncertainty. The usual numerical implementation consists in a nested sample (estimated with a double loop). The outer sample is traditionally reserved for the epistemic uncertainty while the inner sample is used for aleatory uncertainty. For each outer loop, a single sample of the epistemic parameters is selected and held constant while within the inner loop, aleatory parameters are sampled the desired number of times (NA). This is repeated for the total number of epistemic realizations (NE). Thus each epistemic outer loop has NA number of possible outcomes and (NE x NA) represents the total number of possible outcomes generated in the model simulation. The group of aleatory samples under a single epistemic set of parameters is a sample from the conditional distribution of the aleatory variables *given* the epistemic variables (supposing thus that there is no lack of knowledge). One possible way to summarize this conditional distribution is to estimate a quantity of interest from each conditional distribution, e.g. an average, a percentile, or in the specific case of PFM the probability of failure for a given realization of the outer epistemic loop. Results stemming from the epistemic loop represent the level of knowledge we have with respect to the probability of an event (in which the probability is represented by the aleatory distribution function, usually represented as a complementary cumulative distribution function).

The cost of separating between aleatory and epistemic uncertainty is evident in the ‘double loop’ design. As an example, a 100 x 100 sample size would require 10,000 calls from the code while giving only 100 different values for each epistemic parameter. If 10,000 epistemic realizations are required in this example, separating aleatory and epistemic uncertainty would imply performing 1,000,000 calls to the code (10,000 x 100). Therefore, the cost is increased tremendously. However, the gain is, as described below, more insight with respect to the problem. Therefore the choice of separating between those two categories in a PFM depends on the nature of the problem and the outcome. It is important to point out that the uncertainty classification does not affect the mean for a given quantity. That is, using the previous example, the mean over all 10,000 aleatory realizations of the simulation is the same as the mean over the 100 epistemic realizations, since each epistemic output value is actually the mean of the 100 aleatory realizations that comprise each epistemic realization. If an analyst can answer "yes" to the following questions, then a separation would be appropriate:

1. Can the probability be estimated accurately enough with the size of each sample? (if the probability of the event is driven by epistemic uncertainty and one cannot afford an epistemic sample large enough to generate enough events, then the answer is **no**)
2. Can one distinguish between aleatory and epistemic uncertainty for the most influential parameters? (if epistemic and aleatory uncertainty are mixed together without possibility of distinction then the answer is **no**)
3. Does the separation help to make a stronger, more comprehensive case, and/or help to understand what needs to be done to improve the accuracy of the answer? (if no mitigation or no improvement on the model can be made or if the separation confuses the decision maker, then the answer is **no**)

The appropriate separation of aleatory and epistemic uncertainty can be an important component of the design and computational implementation of an analysis of a complex system as well as the decisions that are made on the basis of that analysis. For instance, 'knowledge uncertainty', which is akin to epistemic uncertainty, must be differentiated from spatial and temporal variability, which may be more aleatory in nature. Quantification of the aleatory uncertainties may answer the question, "How likely is the event to happen," while quantification of the epistemic uncertainties may answer the question, "How confident are we in the first answer." Understanding and properly defining these uncertainty types allows for a quantitative categorization of the uncertainty of PFM results.

Characterization of the uncertainty as being aleatory or epistemic is not absolute; it depends on the context and granularity of the problem. The most common approach to categorize uncertainty is via expert elicitation. This does not mean that the process is arbitrary, but that a careful effort should be placed on the description of each uncertain parameter, including a rationale for its characterization. The interpretation of results depends on this characterization. The more important part is thus to clearly explain the rationale behind each decision. As an example, in a conventional LEFM (linear elastic fracture mechanics) model the uncertainty in the linear elastic fracture toughness (K_{Ic}) may be regarded as aleatory (irreducible). Conversely in a micro-mechanics model that accounts for features such as grain size, inclusions, dislocations, and so on, (that is the factors that create the uncertainty in K_{Ic}), this uncertainty may be regarded as epistemic. Mixed situations (part aleatory, part epistemic) are also possible.

Usually, if the uncertainty of a particular parameter, variable, or model cannot be clearly defined as aleatory or epistemic, one may want to treat that uncertainty as epistemic within a sensitivity study. This allows the user to rank the epistemic uncertainty (per uncertain parameter) according to its contribution to the output uncertainty, which may give the user an understanding of the importance of that epistemic uncertainty to the output uncertainty. For example, one may categorize the weld residual stress as epistemic. One can then look at the output uncertainty with that input uncertainty separated and see, for example, that the output 95th and 5th quantiles are 5 orders of magnitude apart, suggesting this uncertainty is overwhelming the results. Once the major contributors to the uncertainty in the output are known, it is customary (and recommended) to revisit the uncertainty associated with these inputs, both in term of representation (distribution type, parameters) and characterization (epistemic, aleatory).

5.2 PFM Analysis Uncertainty Identification and Context

A key aspect in the development of PFM models is the determination of random variables and their treatment throughout the analyses. When random variables have been selected, their influence on the performance metrics needs to be evaluated through quantitative techniques such as analysis of variance (ANOVA), as well as qualitative techniques such as scatterplots (see for instance [29, 30]). Constant (deterministic) variables may also need to be evaluated, usually via one-at-a-time sensitivity studies, as the choice of a constant value may affect the results. As most PFM codes include a large number of inputs, it is often the responsibility of experts and analysts to select which constant inputs will be studied with respect to its impact on the performance metric (based on the equations implemented). Once the important variables are identified, technical justification is needed to (1) confirm that the variable representations appropriately reflect major sources of uncertainty, (2) ensure that model parameter representations (i.e., probability distributions) are reasonable and have a defensible technical basis, and (3) classify the random variables as epistemic or aleatory for the purpose of the PFM analysis.

The context of the analysis may have a direct impact on the uncertainty representation and type classification. The more specific an analysis, the narrower the uncertainty ranges may be. For example, if an analysis is specific to a given pipe in a given plant, the geometry and other characteristics of the system are likely to be precisely defined. As a result the uncertainty range may be relatively small, and the associated distributions relatively narrow, potentially warranting the classification of geometric variables as constants. Furthermore, the geometry and other characteristics of the system actually correspond to a physical quantity that could be measured, thus any uncertainty with regards to dimensions could be categorized as epistemic if the geometric variables are considered random instead of constant. In contrast, for an analysis meant to represent a series of welds or generic configurations across the US reactor fleet, the variability in geometry, operating conditions, materials, and possible flaw mitigation is likely to be larger, and thus result in wider distributions for uncertain variables. Furthermore, since the exact geometry and other characteristics of the system vary from plant to plant and do not correspond to a single value that could physically be measured, the uncertainty may in this case be classified as aleatory instead of epistemic. In turn, this is likely to have an impact on the variability of the results of the PFM analysis.

5.3 Probabilistic Framework, Uncertainty Propagation, and Sampling Techniques

In order to determine the probabilistic distribution of a predicted response, one propagates the input uncertainties through models (which may be themselves uncertain) to estimate uncertainties in the results. To achieve this goal in the context of PFM, it is useful to create a probabilistic framework that can adequately represent input uncertainties, propagate them through a fracture mechanics model, and adequately process the output uncertainty. The most effective way to propagate uncertainty is ultimately dependent on the computational constraints as well as the nature of the inputs and outputs under consideration.

A variety of methods exist for propagating uncertainty using a probabilistic framework. For those situations where the response surface is regular and continuous, methods such as First or Second Order Reliability Method (FORM, SORM) [31] reliable computational methods for structural reliability based on the joint probability density of all the factors influencing failure or non-failure are available. However, for many complex problems where the response surface is difficult to analytically describe or highly discontinuous, probabilistic sampling-based methods may be used. In this case, the probabilistic framework shall be capable of sampling the input space using a pre-defined sampling technique and simulate the response of the system for each sample generated. Each simulation is referred to as a realization of the system, and uncertainty propagation is achieved through the use of the probabilistic framework as follows:

1. For each realization, all of the uncertain parameters and variables are sampled by the probabilistic framework. If aleatory and epistemic uncertainties have been identified and separated, two nested sampling loops may exist in the framework: an outer loop for the epistemic variables, and an inner loop for the aleatory variables. Potential reasons for separating aleatory and epistemic uncertainty are discussed in section 5.1.
2. The system is then simulated (given the particular set of input parameters) such that the response of the system is estimated. This results in a large number of separate and independent results, each representing one possible “future” state of the system.
3. The results of the independent system realizations are assembled into probability distributions of possible outcomes, and may be analyzed by the framework or by using other post-processing tools to study the probabilistic response of the system.

An important role of the probabilistic framework used to propagate uncertainty through a system is input space sampling. As a result, a short discussion of sampling techniques commonly used in PFM analyses to propagate uncertainty is warranted here. Monte-Carlo sampling techniques [32] are perhaps the most commonly used, and in this case the uncertain parameters and variables should be sampled from their distributions. There are many techniques for sampling, all of which intend to represent the distribution for the particular analyses desired. For instance, simple random sampling (SRS) is capable of providing adequate cover of the input space, but accurately modeling low-likelihood events requires a large number of samples which can be prohibitive. Latin Hypercube Sampling (LHS) is one method that may be used to get more uniform coverage of the sample space [33]. LHS divides the input space into equal-probability strata based on input distribution quantiles and a fixed number of samples (usually one) is obtained from each stratum. LHS covers the input space more uniformly than random sampling for the same sample size and thus typically increases precision in characterizing the output distribution. Discrete Probability Distribution (DPD) is another method increasing the stratification by discretizing each probability and reducing the mono-dimensional coverage of the input space in the benefit of multidimensional aspect [34]. If low probability events are being assessed, techniques such as importance sampling, adaptive importance sampling, or stratified sampling [35], may be used. These techniques allow the analyst to specify increased weighting for the part of the distribution that has the most effect on the response (based on the results of sensitivity studies on an initial analysis) so that low probability events can be found with limited realizations. However, the weighting should be consistent and properly propagated throughout the analyses. Furthermore, different sampling techniques may require different levels of convergence testing to ensure converged solutions. For instance, LHS and importance sampling have been built into the xLPR Code.

5.4 Verification and Validation

High confidence in a PFM analysis requires that any PFM software developed for the analysis undergo verification and validation. Verification establishes the correspondence between the PFM computer code and its specifications, while validation establishes whether the PFM code is fit for its intended purpose. The primary purpose of the PFM software is usually to estimate the failure probability of a component in a fracture-mechanics driven problem to the best extent possible based on the best available input distributions. However, examining other outputs of the PFM model can be as important in meeting the performance requirements for the system being analyzed. Furthermore, as seen in the development of the xLPR code for example, small probabilities are hard to predict with high accuracy and difficult to validate, and thus other outputs may be more informative of risk.

All plant-specific activities affecting the safety-related function of nuclear power plant structures, systems, and components must comply with the Quality Assurance (QA) requirement of 10CFR50 Appendix B [36]. For software applied to plant-specific issues by the licensee, the implication of this requirement is that a documented quality assurance program should exist that contains the requirements for the PFM software, including development, procurement, maintenance, testing and configuration management. The purpose of the QA program is to ensure software verification. Furthermore, the QA plan should follow well defined criteria, such as, for example, those specified in ASME NQA-1 [37] or NUREG/BR-0167 [38].

In addition to following a strict QA process for the development of a PFM code, it is important to validate both the individual deterministic models that compose the PFM code and the integrated PFM code. Ideally, all outputs from individual deterministic fracture mechanics models can be validated against analytical solutions that have been shown to accurately represent reality, field (or relevant experimental) data, advanced finite element solutions, or all three, if applicable. Once the individual models have been validated, the relevant outputs from the integrated PFM code can also be validated for each specific application of the PFM code. For a given application, it is desirable to demonstrate validation of all important outputs (those that represent the Quantities of Interest for the problem being analyzed) over the entire range of possible inputs. One should also recognize that validation of a given output of interest may not ensure validation of all outputs, nor does it ensure validation of that specific output for a range of inputs outside of what was used during the validation process. For example, if one considers primary coolant system piping in a nuclear reactor, the validation of crack growth rate and through-wall crack opening displacement predictions does not ensure the validation of a leak rate prediction. Furthermore, still using the same example, the predictions of crack growth and crack opening displacement may only be validated for the temperature, pressure, and chemical environment applicable to steady-state operation in a light-water reactor, and may thus not be applicable to other reactor types that use different operating conditions or a different type of coolant[3]. Consequently, the validation of an integrated PFM code is application-specific, and should be tailored to the application. It is possible that validation for one application is sufficient to demonstrate validation for some other applications, but this should be verified on a case-by-case basis.

Establishing and meeting the acceptance criteria for the validation of an integrated PFM code over a broad range of conditions can be complicated by the fact that limited field data may be available, and because of the probabilistic nature of the code. As a result, validation of the software could be accomplished using a graded approach, where the extent of testing applied to different individual models of the software is commensurate with the importance of the individual model (for the problem being analyzed) and the risk that the software element does not function properly at the integral level. For instance, if a piping PFM code is used to calculate leakage and rupture, and the purpose of the

[3] Note that extensive deterministic model changes would be necessary for a PFM code that was developed for light water reactors compared to a high temperature air or sodium cooled reactor since crack growth mechanisms, damage development, and stability will be different.

analysis is to demonstrate the impact of inspection on leakage times, the rupture portions of the code (i.e. through-wall flaw stability) may not need the same verification and validation (V&V) rigor as the leak rate calculations. Nonetheless, when possible, the software results should be validated to the extent practical against available field data and hypothetical test cases (e.g., impact of piping rupture probability on varying inspection intervals or probability of detection) that have been developed based on engineering judgment. For instance, the xLPR PFM code was partially validated by comparing predictions of crack depth versus time to crack occurrences that occurred in operating plants. The level of testing should be chosen such that reasonable assurance is provided that the PFM code is reasonable in the context of the intended use of the software.

The extent of the V&V and other aspects of quality assurance such as documentation of the PFM code technical basis could vary depending on a number of factors, such as:

- The application of the code, for example whether the code is used to support an inspection interval, license extension, or used to support a permanent change to the plant configuration (e.g., remove pipe whip restraints). The more permanent a change, the higher the requirements should be for V&V.
- Scale and complexity of the code and models: a more complex code would require more extensive V&V.
- Whether the code is pre-verified software or one-time-use software (in which case the calculation results are verified and validated on an individual run basis, thus requiring less V&V up front), though many ‘single uses’ of a one-time-use software may result in a high degree of verification.
- Whether submodels are applied that have consensus support as standard modeling approaches (similar to standard deterministic approaches that are accepted without specific validation documentation).

In many cases, the events that a PFM code is designed to model and predict are very low probability events, for example, the rupture of a pipe in the primary reactor coolant system. As a result, little or no field data may be available to validate such low probability events. In some cases, precursor events may be used to validate the PFM code, such as incidences of cracking in RCS piping, which could lead to piping rupture if not detected and repaired. In other cases, experimental data applicable to the phenomenon being studied might be available. Nonetheless, validation based on field or experimental data can be challenging and limited. In such cases, other studies can be performed, such as sensitivity or model behavior studies, and benchmarking against other comparable codes. Sensitivity studies can be used to demonstrate that the overall behavior of the PFM code is consistent with expert understanding of the expected system behavior, including demonstrating expected trends and correlations between inputs and outputs of interest. Benchmarking against other comparable codes may be used to increase confidence in the PFM code by demonstrating that the results produced by the PFM code are reasonable and can be predicted by similar codes [39]. Although benchmarking alone is not sufficient to validate a PFM code, it can provide a significant contribution to ensuring the accuracy of a PFM code.

In general, peer review by one or more independent parties, detailed documentation of the theory and structure, and independent effort to confirm or reproduce results (either with the same PFM code or another one with similar capabilities), would also increase confidence in the PFM code or analysis.

6 Analysis of PFM Outputs and Results

Producing reliable, realistic results from probabilistic fracture mechanics analyses requires more than a single probabilistic analysis with a given PFM code (hereafter referred to as a 'run' of the code). The user should conduct multiple runs to demonstrate that the solution is converged, the uncertainties are properly represented, and the parameters driving the problem have been sufficiently characterized. This chapter discusses how convergence of a PFM analysis can be demonstrated, how to study output uncertainty distributions, how to extract information from a PFM analysis via uncertainty analyses, and finally how to identify problem drivers (i.e. important contributors to the output) and gain confidence in the PFM analysis via sensitivity studies.

6.1 Convergence of Results

The concept of convergence is intrinsically linked to the decision that needs to be made after reviewing the PFM results. Convergence is often a matter of perspective and really comes back to the consideration of whether one has enough confidence in the results to make a decision one way or another. As such, before addressing the convergence of an output, it is necessary to consider a threshold value (or range) for each output considered in final decisionmaking. This threshold constitutes the "pass/fail" criterion for the problem considered. The second point of consideration is how much confidence is required in the final answer or decision, often linked to the seriousness or magnitude of the consequence. With relatively reasonable consequences, having a (theoretical) higher 95^{th} or 99^{th} percentile close or at the threshold may be considered acceptable. With more serious consequences, it may be required to have the 95^{th} or 99^{th} percentile at least at one order of magnitude from the threshold for increased safety.

Once these two criteria are defined, a solution is considered converged enough if the uncertainty is low enough that it does not change the conclusion that could be drawn from the analysis. As a result, output statistics far from the threshold can afford large uncertainties while output statistics close to the threshold will require tighter confidence bounds.

To demonstrate convergence, one should consider both temporal (if the problem is time-based), spatial (if the problem is location-based), and statistical solution convergence. Temporal or spatial convergence demonstration can be accomplished by simply running the code with different (either increasing or decreasing) time steps or spatial discretization until the solution diverges. It is important to note that the solution convergence for any given output does not guarantee convergence for all outputs. As a hypothetical example, convergence could be achieved for crack length for a coarse azimuthal spatial discretization, while convergence on crack azimuthal position may require a much finer discretization. Similar convergence tests may need to be conducted for each unique desired output.

Demonstrating solution, or statistical, convergence can be a bit more complicated, and may be achieved via stability analysis. Simulation options to explore include (but are not limited to)

1. Sample Size
2. Sampling strategy (Simple Random Sampling [SRS], Latin Hypercube Sampling [LHS], Discrete Probability Distributions [DPD] [40], Importance Sampling [IS])
3. Number of strata to use for DPD/LHS
4. Degree of spatial or temporal discretization for the problem being analyzed, as applicable.

The sampling options chosen to consider are application specific but without further guidance should include at a minimum the sample size and sampling strategy. Plots of the quantities of interest as driven by acceptance criteria (e.g. mean probability of crack depth, leakage, and rupture) under the different simulation options should be provided and interpreted. These plots will help with decisions regarding the convergence of the solutions. When considering number of strata to use for DPD, the number of segments for spatially varying parameters, or time step size, results from more granular simulations should be sufficiently close to results with finer resolution. Note that results concerning convergence depend on the output of interest and the purpose of the application. For example, acceptable convergence for a mean probability will likely be achieved sooner than convergence for a large quantile of the distribution of probabilities.

Sampling techniques are numerical procedures and, as such, may result in varying degrees of accuracy. The accuracy is dependent on the method and on the sample size used. As discussed in Section 5.3, several types of sampling techniques can be used in the development of a probabilistic output. For Monte Carlo-type analyses, conducting multiple runs replicated samples using a different random seed is the most straightforward solution. Another approach is to increase sample size until the desired result converges. The advantage of the first approach is that if SRS or LHS is used, it is possible to increase the sample size while testing replicates. However, with computational time limitations and/or advanced sampling techniques (importance sampling, etc.), these convergence runs may not be practical. Statistical uncertainty quantification (UQ) methods can often be applied to characterize uncertainty in estimates of a quantity of interest due to sampling uncertainty, but these methods must account for the sampling scheme (e.g. importance sampling, LHS, etc.). By either repeated convergence runs or statistical UQ, demonstrating solution convergence adds greatly to the confidence in the probabilistic output.

6.2 Output Uncertainty Quantification (Uncertainty Analysis)

Once convergence of a PFM analysis has been shown, one can perform uncertainty analyses on the outputs of interest, for example to compare the likelihood of an event to a safety metric. Uncertainty analyses are the calculations required to develop the uncertainty distribution of output from distributed inputs. Output uncertainty analysis can take different forms depending on the nature of the output or answer being sought. Temporal outputs may consist of creating distributions for the probability of an event or the value of a parameter of interest over time, while spatial outputs may consist of distributions of the location of a given event if it happens. In order to represent the uncertainty of the output, one may choose to show the temporal or spatial output for each and every realization of the problem, or to represent the result via one or more statistical quantities.

The relevant quantities of interest (QoI) that can be defined as a result of uncertainty analysis are dependent on the acceptance criteria for a given problem, and should be justified. Whatever the choice, the analyst should explain why the statistical quantities being chosen are a relevant representation of the output uncertainty. For example, for the overall probability of leakage as a function of time, the mean of the probability of leakage over all epistemic runs is probably the most representative statistical quantity, and could be termed the 'expected' leakage probability. For conservatism, a large quantile for the probability of leakage, such as the 95^{th} percentile, may be preferred as the primary QoI in a safety analysis. In this particular example, the probability of leakage is known for each individual realization and based on an indicator function: it is set to 0 if the system is not leaking, and 1 if the system is leaking. As another example, a relevant way to represent the uncertainty of a spatial output may be via a probability density function (PDF) or cumulative distribution function

(CDF) as a function of location (i.e. a CDF and PDF for each location, resulting in a set of PDF and CDF at different locations), or a PDF or CDF representing the location variation. However, it should be recognized that such risk numbers are not absolute but depend on the deterministic models chosen to characterize the PFM code. Introduction of different deterministic models will change these numbers.

6.3 Sensitivity Analysis (Input Importance Quantification)

An important aspect of PFM analyses consists of understanding the relationship between problem input and output uncertainties. In particular, it is useful and important to identify the drivers of uncertainty for any given problem, and this goal can be achieved by performing sensitivity analyses. Sensitivity analyses are also important in helping to develop sampling options which provide converged results.

Sensitivity analyses can be performed using a variety of methods that can be grouped into two main categories: graphical qualitative methods and quantitative 'variance-explained' methods. A common graphical qualitative means to look at the relationship between uncertain inputs and outputs is through the use of scatterplots of the output variables against individual input variables. The advantage of this approach is that it can deal with any set of arbitrarily-distributed input and output variables, and gives a direct qualitative visual indication of sensitivity. Scatterplots can also be used to help guide decisions about which inputs may be good candidates for importance sampling. For large problems with many uncertain inputs and outputs, it is not practical to create scatterplots for all variables, and data reduction may be needed.

Quantitative metrics should also be used to quantify and rank the importance of the relationship between inputs and outputs. The proportion of variance in the output driven by each input can be calculated using a large number of model runs; when this is not possible, regression analysis methods are often applied as a computationally efficient means of quantifying and ranking the importance of the inputs. Many methods exist for regression analysis, and in general the NRC's experience with running sensitivity analyses [41, 42, 43] dictates that three or more methods should be combined to ensure that false positives for correlation are eliminated, as well as to increase confidence that the correlations found are significant. The specific number of regression methods is not as important as showing that the final results are comprehensive, well supported, and well documented for the problem under analysis. Some common regression analysis methods include linear regression, rank regression, quadratic regression, Gaussian regression, regression trees, and multivariate adaptive regression splines (MARS), among others.

6.4 Sensitivity Studies

In addition to uncertainty analysis and sensitivity analyses, it is expected that any PFM analysis submitted to NRC for review include sensitivity studies to explore the specific influence of important parameters and analyze specific scenarios of interest. Sensitivity studies are useful to better understand the physics of a problem, to show that some expected trends are indeed reflected in the analysis results. Sensitivity studies are also used to perform "what if" analyses, to revisit some assumptions (constant values or distributions changed), and to investigate alternative scenarios in support of the defense in depth. [44]

Deterministic sensitivity studies consider one (or a set of) input(s) of interest and change it (them) in the deterministic reference to estimate how much it impacts a specific output. Experts change the parameters that could influence the output (based on the previous sensitivity analysis or their own expert judgment) one at a time. Such analyses are also performed to assess the impact of some alternative scenarios (such as extreme condition and/or mitigation). Deterministic sensitivity studies usually focus on the physics of the system and build confidence on the theoretical aspect of the analysis.

Probabilistic sensitivity studies are the probabilistic equivalent of the deterministic sensitivity studies. They allow the comparison of the effect of changing one input, not only from a median reference but from different points of reference in the input space and globally. The change in input can be as for the deterministic case (with a shift) but also as a change in uncertainty (spread) or point of interest (skewness). If epistemic and aleatory uncertainties have been separated as part of a PFM analysis, one may also change the uncertainty classification for some variables to investigate the effects of such changes. These studies include the probabilistic aspect which focuses more on the changes in probabilities of occurrence of an event of interest associated with each change.

6.5 Analysis of Problem Drivers and Confidence Demonstration

Once the main uncertainty drivers have been identified for all outputs of interest, the uncertainty of the inputs that drives the problem should be re-visited to gain confidence in the PFM analysis conclusions. It is important to ensure that the distributions that characterize the uncertainty of the problem drivers are representative, and it is even more important to study the consequences if these distributions are not representative. For example, if the uncertainty of the problem drivers is so large that the 95th percentile of the PFM results does not meet acceptance criteria for the problem being analyzed, one may conclude that more research is needed to reduce this uncertainty. Such a reduction may require additional analyses or additional data to be collected. Similarly, if the distributions of the problem drivers are found to not be good representations of actual data, the analyst should investigate the consequences of changing the distributions of the problem drivers. If the change in output is minimal, it may still be possible to reach the same conclusion. It should be noted here that varying the distribution of an input of interest does not necessarily require rerunning the entire analysis as long as the support of the original distribution (i.e. the domain of the distribution containing those elements which are not mapped to zero) contains the support of the new distribution. Instead, the change in distribution may be achieved by changing the weights of each realization according to the relative change in probability of occurrence when changing the distribution for the input of interest.

To further increase confidence in the conclusions, the areas in the distribution that could change the conclusion should be identified and sufficiently sampled for each significant driver to the problem. If these areas are in the tails of the distributions, which is often the case, importance sampling may be used to get more results for those areas of the distributions. When confidence cannot be gained in the chosen distribution for a given parameter that drives the problem, several alternatives can be explored. One such alternative relies on expert judgement to narrow the input distribution or to qualitatively assess the confidence in the mean values of the input that drive the problem. When such an alternative is chosen, sensitivity studies should be performed to determine whether the input mean selection is driving the solution to the problem, and if so, steps should be taken to increase the confidence in the mean values of key inputs. This typically requires the collection of additional data on those inputs, and can be resource intensive.

7 Conclusions

7.1 Summary

In response to increasing interest in PFM analyses to support safety demonstrations and regulatory relief requests from NRC licensees, this report defined the concept of a PFM analysis and outlined important concepts that should be considered when producing a high-quality, high-confidence PFM analysis. Fracture mechanics analyses have historically been performed deterministically, though an increase in the use of probabilistic approaches has occurred in response to the need to quantify risk in safety decisionmaking. Deterministic and probabilistic approaches are only as good as the models we develop to represent reality, and primarily differ in the way that they treat uncertainty. Each approach has its advantages, though policy has resulted in an increased emphasis on cost benefit analyses and risk based metrics, which favor an increased use of probabilistic methods.

A general framework to describe, perform, and evaluate a PFM analysis was presented in this report. The important pieces of a PFM analysis that should be considered with great care 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 model uncertainty quantification are key steps to gain confidence in the adequacy of the models used. Often models used in PFM are simplified representation of the reality due to solution time constraints. The possible errors introduced by this must be understood. It is also important to V&V the system as a whole and not only each individual model, since a complex system is more than a simple assembly of independent models.
- Treatment of uncertain inputs may consist of constructing probability distributions, determining input bounds if applicable, and quantifying any assumptions, conservatisms, or dependencies between inputs, which insure that the input space is limited to and include all of the domain of physical realm.
- Uncertainty characterization and treatment is at the core of a PFM analysis. In many PFM analyses, it may be useful to separate epistemic and aleatory uncertainty to be able to assess the confidence levels for probabilities of interest. Uncertainty identification, quantification, and propagation are essential elements in describing a PFM methodology or analysis. The proper choice of sampling techniques (when sampling-based methods are considered) is also an important step that needs justification. Concepts and methods to verify and validate a probabilistic framework were discussed.
- Ways to demonstrate PFM convergence when sampling-based methods are used include varying sample size, use of replicated samples, and sampling strategy, as well as performing stability analysis. Output uncertainty analysis can take various forms depending on the problem being analyzed. Sensitivity analyses can help to identify the drivers of uncertainty for a given problem. Sensitivity studies are useful to better understand the physics of a problem, to show that some expected trends are indeed reflected in the analysis results, and to perform “what if” analyses and revisit some assumptions (constant values or distributions changed) as well as investigate alternative scenarios in support of the defense in depth. Methods to perform such studies were presented.

7.2 Future Work

Following this report, the next step towards the development of guidance to perform and evaluate PFM analyses is the development of a more detailed technical basis NUREG document. The NUREG will dive deeper into the concepts described in this report, and will attempt to more precisely define PFM quality metrics and criteria, or a process to define such metrics and criteria. From this NUREG document, a more succinct draft Regulatory Guide will be produced highlighting the steps required to produce a high confidence PFM analysis, and the ways to evaluate such a PFM. In parallel, a pilot study will be defined and executed to test the draft guidance produced and to identify deficiencies and areas that need improvement. The lessons-learned from this process will be incorporated in a final guidance document, and the technical basis NUREG will be revised accordingly.

8 References

[1] S. Kaplan and B. J. Garrick, "On The Quantitative Definition of Risk," *Risk Analysis,* vol. 1, no. 1, pp. 11-27, 1981.

[2] "60 FR 42622 - Use of Probabilistic Risk Assessment Methods in Nuclear Regulatory Activities; Final Policy Statement," *Federal Register,* vol. 60, no. 158, August 16, 1995.

[3] P. T. Williams, T. L. Dickson, B. R. Bass and H. B. Klasky, "ORNL/LTR-2016/309: Fracture Analysis of Vessels – Oak Ridge - FAVOR v16.1, Computer Code: Theory and Implementation of Algorithms, Methods, and Correlations," Oak Ridge National Laboratory, 2016.

[4] T. L. Dickson, P. T. Williams, B. R. Bass and H. B. Klasky, "ORNL/LTR-2016/310: Fracture Analysis of Vessels – Oak Ridge - FAVOR, v16.1, Computer Code:User's Guide," Oak Ridge National Laboratory, 2016.

[5] D. L. Rudland and C. Harrington, "NUREG 2110: xLPR Pilot Study Report," U.S. Nuclear Regulatory Commission, Washington, DC, 2012.

[6] EPRI, "BWR Reactor Pressure Vessel Shell Weld Inspection Requirements," EPRI-TR-105697, BWRVIP-05, September 1995.

[7] Westinghouse Electric Company, "Risk-Informed Extension of the Reactor Vessel In-Service Inspection Interval," WCAP-16168-NP, Revison 2, 2006.

[8] G. R. Irwin, "Linear fracture mechanics, fracture transition, and fracture control," *Engineering Fracture Mechanics,* vol. 1, no. 2, pp. 241-257, 1968.

[9] ASTM E399-12e3, Standard Test Method for Linear-Elastic Plane-Strain Fracture Toughness KIc of Metallic Materials, West Conshohocken, PA: ASTM International, www.astm.org, 2012.

[10] ASTM E1820-15ae1, Standard Test Method for Measurement of Fracture Toughness, West Conshohocken, PA: ASTM International, www.astm.org, 2015.

[11] ASTM E1921-16, Standard Test Method for Determination of Reference Temperature, To, for Ferritic Steels in the Transition Range, West Conshohocken, PA: ASTM International, www.astm.org, 2016.

[12] U.S. Congress, *Government Performance and Results Act of 1993,* Washington, DC: 103rd Congress, Congressional Record, 1993.

[13] G. Wilkowski, R. Tregoning, P. Scott and D. Rudland, "Status of Efforts to Evaluate LOCA Frequency Estimates Using Combined PRA and PFM Approaches," in *Proceedings of 28th MPA Seminar*, Stuttgart, Germany, October 2002.

[14] T. L. Anderson, Fracture Mechanics: Fundamentals and Applications, Third Edition, CRC Press, 2005.

[15] API 579-1/ASME FFS-1: Fitness-for-Service, Washington, DC: ASME, 2007.

[16] S. X. Xu, D. A. Scarth and R. C. Cipolla, "Technical Basis for Explicit Inclusion of Weight Function Method for Calculation of Stress Intensity Factor for Surface Flaws in ASME Section XI Appendix A," in *Proceedings of the 2011 ASME Pressure Vessels and Piping Conference*, Baltimore, MD, USA, July 17-21, 2011.

[17] W. L. Oberkampf, T. G. Trucano and C. Hirsch, "Verification, validation, and predictive capability in computational engineering and physics," *Applied Mechanics Reviews,* vol. 57, no. 5, pp. 345-384, 2004.

[18] T. G. Trucano, L. P. Swiler, T. Igusa, W. L. Oberkampf and M. Pilch, "Calibration, validation, and sensitivity analysis: What's what," *Reliability Engineering & System Safety,* vol. 91, no. 10, pp. 1331-1357, 2006.

[19] Z. A. Karian and E. J. Dudewicz, Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods, CRC Press, 2000.

[20] M. Kirk, C. Sallaberry, M. Erickson and R. Kurth, "Draft report "A User's guide to the xLPR Uncertainty Characterization Spreadsheet"," 2012.

[21] S. Ferson, J. O'Rawe and M. Balch, "Computing with confidence: imprecise posteriors and predictive distributions," in *International Conference on Vulnerability and Risk Analysis and Management and International Symposium on Uncertainty Modeling and Analyis*, Liverpool, UK, 2014.

[22] [Online]. Available: https://en.wikipedia.org/wiki/List_of_probability_distributions.

[23] A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," *G. Ist. Ital. Attuari,* vol. 4, pp. 83-91, 1933.

[24] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Annals of Mathematical Statistics,* vol. 19, pp. 279-281, 1948.

[25] T. W. Anderson and D. A. Darling, "A Test of Goodness-of-Fit," *Journal of the American Statistical Association,* vol. 49, pp. 765-769, 1954.

[26] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika,* vol. 52, no. 3-4, pp. 591-611, 1965.

[27] M. Ionescu-Bujor and D. G. Cacuci, "A Comparative Review of Sensitivity and Uncertainty Analysis of Large-Scale Sytems: Part I and II," *Nucliear Science and Engineering,* vol. 147, no. 3, pp. 189-217, 2004.

[28] A. R. Daneshkhah, "Uncertainty in Probabilistic Risk Assessment: A Review," http://www.shef.ac.uk/content/1/c6/03/09/33/risk.pdf, August 2004.

[29] A. Saltelli, K. Chan and E. M. Scott, Sensitivity Analysis, Wiley series in probability and statistics, 2000.

[30] C. B. Storlie and J. C. Helton, "Multiple predictor smoothing methods for sensitivity analysis: Description of Techniques," *Reliability Engineering and System Safety,* vol. 93, no. 1, pp. 28-54, 2008.

[31] B. Fiessler B, H. J. Neumann and R. Rackwitz, "Quadratic limit states in structural reliability," *Journal of Engineering Mechanics, ASCE,* vol. 105, no. 4, pp. 661-676, 1979.

[32] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association,* vol. 44, p. 335, 1949.

[33] R. Iman, J. Helton and J. Campbell, "An approach to sensitivity analysis of computer models, Part 1. Introduction, input variable selection and preliminary variable assessment," *Journal of Quality Technology,* vol. 13, no. 3, pp. 174-183, 1891.

[34] R. E. Kurth and D. C. Cox, "Discrete Probability Distributions for Probabilistic Fracture Mechanics," *Risk Analysis,* vol. 5, no. 3, pp. 235-240, 1985.

[35] C.-E. Särndal and e. al., "Stratified Sampling," in *Model Assisted Survey Sampling*, New York, Springer, 2003, pp. 100-109.

[36] "Appendix B to Part 50—Quality Assurance Criteria for Nuclear Power Plants and Fuel Reprocessing Plants," in *Title 10 - Code of Federal Regulations, Chapter I*, Washington, DC, U.S. Government Publishing Office, 2015.

[37] ASME NQA-1-2015: Quality Assurance Requirements for Nuclear Facility Applications, ASME International, 2015.

[38] NUREG/BR-0167: Software Quality Assurance Program and Guidelines, Washington, DC: U.S. Nuclear Regulatory Commission, 1993.

[39] B. Brickstad, T. Schimpfke, O. J. Chapman, P. Dillstrom, C. Cueto-Felgueroso and C. D. Bell, "Project NURBIM (Nuclear RI-ISI Methodology for Passive Components): Benchmarking of Structural Reliability Models and Assocaited Software," in *ASME/JSME 2004 Pressure Vessel and Piping Conference*, San Diego, 2004.

[40] "Discrete distribution," Encyclopedia of Mathematics, 7 February 2011. [Online]. Available: http://www.encyclopediaofmath.org/index.php?title=Discrete_distribution&oldid=16276. [Accessed 24 February 2017].

[41] P. Mattie, R. Gauntt, K. Ross, N. Bixler, D. Osborn, C. Sallaberry, J. Jones and T. Ghosh, "NUREG/CR-7155 / SAND2012-10702P: State-of-the-Art Reactor Consequence Analyses Project: Uncertainty Analysis of the Unmitigated Long-Term Station Blackout of the Peach Bottom Atomic Power Station," U.S. Nuclear Regulatory Commission (ML16133A461), Washington, DC, USA, May 2016.

[42] Sandia National Laboratories, Severe Accident Analysis Department, "Draft Report: State-of-the-Art Reactor Consequence Analysis (SOARCA) Project: Sequoyah Integrated Deterministic and Uncertainty Analyses," U.S. Nuclear Regulatory Commission (ML17156A270), Washington, DC, USA.

[43] K. Ross, N. Bixler, S. Weber, C. Sallaberry and J. Jones, "Draft Report: State-of-the-Art Reactor Consequence Analysis Project: Uncertainty Analysis of the Unmitigated Short-Term Station Blackout of the Surry Power Station," U.S. Nuclear Regulatory Commission (ML15224A001), Washington, DC, USA.

[44] C. J. Sallaberry, R. E. Kurth, F. W. Brust and E. A. Kurth, "Proposed Approach of Scenario Analysis Using a Probabilistic Code," in *Proceedings of the ASME 2017 Pressure Vessels and Piping Conference*, Waikoloa Village, Hawaii, USA, July 16-20, 2017.