Technical Basis of a Probabilistic Method for PT-CT Contact Assessment

A Review

Mahesh Pandey
Waterloo, ON
January 24, 2016

Table of Contents

1. Summary ... 7

1.1 Background ... 7

1.2 Review of benchmarking analysis ... 8

1.3 Review of reliability analysis ... 9

1.4 Closing remarks ... 11

2. Introduction ... 13

2.1 Background ... 13

2.1.1 Summary of PT-CT Assessment Approach ... 14

2.2 Objectives of the Review ... 15

2.3 Scope of the Review ... 15

2.4 Organization ... 16

3. Probabilistic Analysis and Estimation: Basic Background ... 18

3.1 Introduction ... 18

3.2 A Function of Random Variables ... 19

3.2.1 Statistical Moments: Nonlinear Function ... 19

3.2.2 Statistical Moments: Linear Function ... 20

3.2.3 Probability Distribution: Nonlinear Function ... 20

3.2.4 Probability Distribution: Linear Function ... 21

3.2.5 Special Case: A Linear Function of Gaussian Random Variables ... 21

3.3 Statistical Estimation ... 22

3.3.1 Data for basic variables are available ... 22

3.3.2 Data for the function are available ... 22

3.3.3 Data for function available but distributions of basic variables are unknown ... 22

3.4 Closing Remarks ... 23

4. Review of the Model Calibration (Benchmarking) 24
4.1 Basic Problem 24
4.2 Summary of Probabilistic Analysis Problem 26
4.3 Review of the Simplex Method 27
4.4 Evaluation of the Probabilistic Model of the CE-PM Report 28
4.5 Remarks: Fuel Channels with Loose Spacers 29
4.6 Closing Remarks 30
4.7 Points of Clarification 30
5. Prediction of the PT Failure Frequency 35
5.1 Background 35
5.2 Reliability of a non-repairable component 37
5.2.1 Basic Concepts 37
5.2.2 Lifetime Distribution 38
5.2.3 Failure Frequency 38
5.2.4 Mission Reliability 39
5.2.5 Illustration 40
5.3 Reliability of a repairable component 42
5.3.1 Failure Rate 43
5.3.2 Homogeneous Poisson Process (HPP) Model 43
5.3.3 An Example 44
5.4 Reliability Analysis and Limit States of the Pressure Tube: A discussion 48
5.4.1 Background 48
5.4.2 Repairable versus Non-Repairable Reliability Problems 50
5.4.3 Limit States of Pressure Tubes 51
5.5 Closing Remarks 53

5.6 Points of Clarifications 53
6. Other Issues and Points of Clarification 55
6.1 Introduction 55
6.2 Points of Clarification 55
7. References 58

Table of Figures

Figure 1.1: Steps in the probabilistic assessment method developed by Candu Energy [1] 7
Figure 2.1: Flowchart of the probabilistic assessment of CE-PM report [1] 15
Figure 5.1: Concept of reliability of non-repairable components 37
Figure 5.2(a): Lifetime distribution, and (b) the hazard function: Non-repairable component 40
Figure 5.3: The fraction of components having lifetime (a) between 10 and 20 years, and (b) between 60 and 70 years. 40
Figure 5.4: The service life of a repairable component with recurring failures 42
Figure 5.5: (a) Lifetime distribution and the hazard rate, and (b) reliability function of the gamma distribution 45
Figure 5.6: Difference between annual probability of failure and failure frequency: Non-repairable problem 46
Figure 5.7: Failure rate with gamma distributed time between failures: Repairable problem 46
Figure 5.8: Comparison of failure rate of repairable equipment with the hazard rate of a non-repairable equipment.. 47
Figure 5.9: Comparison of the failure rate of repairable and non-repairable equipment: Example of the gamma distribution 48

Notations

CDF	Cumulative distribution function
CE	Candu Energy
CT	Calandria Tube
CE-PM	Candu Energy's Probabilistic Method
EFPH	Effective Full Power Hour
H_{eq}	Equivalent hydrogen concentration
MCS	Monte Carlo Simulation
PDF	Probability Density Function
PSA	Probabilistic Safety Assessment
PT	Pressure Tube
RCRA	Reactor Core Risk Assessment
RV	Random Variable
SLAR	Spacer Location and Repositioning

1. Summary

1.1 Background

This report presents a detailed review of a probabilistic method developed by Candu Energy (CE) for assessing the risk of pressure tube – Calandria tube (PT-CT) contact in the reactor core.

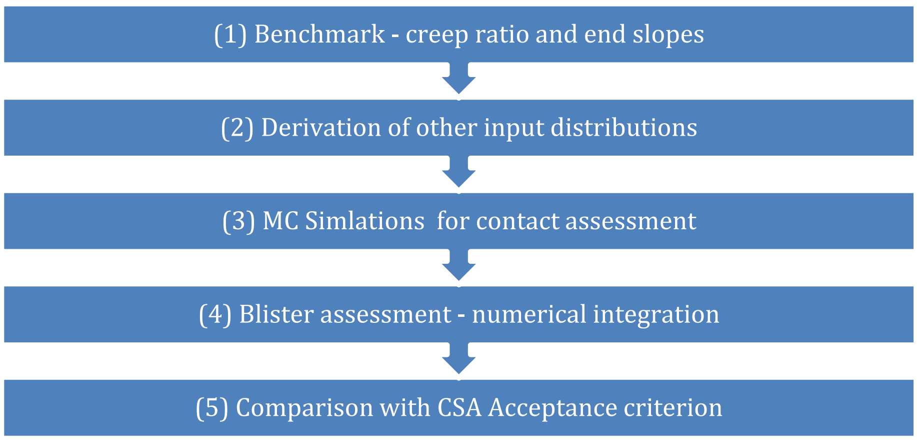


Figure 1.1: Steps in the probabilistic assessment method developed by Candu Energy [1]

The probabilistic method developed by Candu Energy, referred to as CE-PM [1] in this review report, consists of 5 major parts, as shown in Figure 1.1. The first step of benchmarking involves the calibration of creep factor and end slope distributions using the gap measurement data obtained from past inspection campaigns. The estimation of distributions of other variables, such as the geometry of fuel channels, spacer locations and measurement error are subsequently discussed. A simulation method is presented for estimating the probability distribution of the time and location of PT-CT contacts, which is based on the CDEPTH code. The blister assessment model predicts the probability of blister cracking in contact locations. The last part compares the calculated failure frequency with the acceptance criterion specified in the CSA Standard [2]. Many important details of the benchmarking method are revealed in [3].

It must be stressed that the CNSC staff considers the “contact leading to blister formation” as the ultimate failure state.

A key premise of this review is that since a probabilistic assessment criterion can be represented by a function of random variables, the proposed solution approach must be consistent with basic principles of probabilistic functional analysis. With this goal in mind, this review carries out an in-depth evaluation of: (1) the conceptual accuracy of probabilistic formulation, (2) the relevance and impact of the modeling assumptions, (3) mathematical accuracy and consistency of the proposed estimation method, and (4) validity of the interpretation of final results. All together this review poses 33 points of clarification in Sections 4 to 6 of this report. A gist of these points is summarized in the following.

1.2 Review of benchmarking analysis

A probabilistic assessment criterion can be technically defined by a function, $g(\mathbf{X})$, that can represent the gap prediction model based on a computer code like CDEPT. Here $\mathbf{X}$ denotes a vector of input random variables, such as the creep factor and end slopes. Therefore, the review begins with a concise overview of basic concepts of the probabilistic functional analysis and discusses possible approaches to statistical estimation of the distribution parameters (see Section 3). This discussion is essential to provide a context to the technical analysis of the benchmarking method presented in the CE-PM report. This discussion also provides necessary background to questions and points of clarification raised throughout this report.

The benchmarking step of the CE-PM report is rephrased in a clear and precise mathematical language in Section 4, which allows one to evaluate the implications of modeling assumptions made in the development of CE-PM. For example, an assumption that the PT-CT gap and the primary variables influencing the gap (namely, the creep factor and end slopes) follow the normal distribution is a cornerstone of CE-PM. But, the empirical evidence provided to substantiate this assumption is rather weak and questionable.

Given that a large amount of gap measurement data is available from past inspections of several fuel channels, it is anticipated that the assumption of normality of the gap can be verified, at least in an empirical manner.

An important assumption of CE-PM is that the creep ratio has negligible influence on the gap in the end spans of the fuel channel, and the end slope on the gap in internal spans. Results presented in Figures 27 and 28 do not seem to support this assumption.

For instance, Figure 27 shows the sensitivity of the creep ratio to the PT-CT gap in the channel P08. As seen by the width of the bounds, the gap in the end spans seems to be fairly sensitive to the creep ratio. In fact, the order of this sensitivity is similar to that of the sensitivity of end slopes as shown by Figure 28. Similarly, comparing the results for the two internal spans (excluding the central span), the following observation can be made. The sensitivity of the end slope to gap is higher than that of the creep ratio, which appear to contradict the assumption of the CE-PM report..

At the outset of a probabilistic assessment, the form of the performance function, $g(\mathbf{X})$, such as being a linear versus nonlinear function, must be carefully examined, since this will govern the subsequent analysis. It is rather strange that the CE-PM report does not discuss at all the dependence of the PT-CT gap on the basic variables in a deterministic sense. For example, in the absence of any randomness, how does the gap varies with end slopes and the creep factor? If the dependence indeed is linear or approximately linear, as implied by the modeling assumptions of the CE-PM report [1] and a supporting reference paper [3], then the statistical estimation would become a much more manageable task. On the other hand, if the PT-CT gap turns out to be a highly non-linear function of the creep factor and end slopes, the basic tenet of CE-PM would become questionable.

This review also shows that if the assumptions of the CE-PM report are indeed valid, then the gap model can be calibrated in a direct manner using a standard estimation method like the maximum likelihood method. Apparently, there is no need for an iterative Monte Carlo method for model calibration, as proposed by the CE-PM report. If, however, the main assumptions of the CE-PM report, namely, PT-CT gap follows a normal distribution, is not valid, the model calibration problem must be revisited.

1.3 Review of reliability analysis

In the context of PT-CT contact assessment, the next important issue tackled in the review is about the modeling of the risk/reliability analysis problem and the use of a correct reliability measure. The overall purpose of a probabilistic assessment is to quantify the risk associated with the operation of a component or system and ensure that mitigating actions are in place in case that the risk exceeds the acceptance standard. A fundamental question then arises as to what is the relevant reliability metric that should be adopted in the PT-CT contact and blister formation problem. Is the failure frequency as defined and calculated in CE-PM a relevant measure of reliability? The answer to this question depends on the type of reliability analysis to which the PT failure problem belongs.

In the scientific literature, the reliability analysis problem is classified into two broad categories: non-repairable and repairable equipment (or system or component) reliability problems [4]. A detailed explanation of the "repairable" vs. "non-repairable" component reliability problem is given in Chapter 5 of this report.

In the non-repairable or "first-failure" problem, the *mission reliability or probability of failure* over an evaluation period (or inspection interval) is a relevant metric of reliability, whereas the *failure rate* (*or failure frequency*) is a relevant risk metric in the repairable system problem [4]. Every probability assessment must begin with identifying the type of problem, repairable or non-repairable, that it intends to address in the probabilistic assessment. In the CE-PM report, the failure frequency is adopted as a reliability measure without any justification and discussion of this sort.

Since the conceptual distinction between the repairable and non-repairable reliability problems cannot be described in a few paragraphs, Section 5 of this report presents an exposition of key concepts of the time dependent reliability theory with simple illustrative examples. Based on this discussion, the importance of using a correct measure can be better understood.

The CSA Standard [2] describes several limit states of the performance of the fuel channel, such as critical flaw sizes, minimum fracture toughness and PT-CT contact. However, the Standard does not provide any guidance about the classification of such problems as repairable or non-repairable failures. The Standard also omits the relation and relevance of the chosen risk measure, i.e., failure frequency, to the reliability with respect to postulated failure modes.

How to classify a risk/reliability assessment problem related to the fuel channel performance into a repairable or non-repairable category is discussed in detail in Section 5.4. This discussion is somewhat generic and applicable to probabilistic assessment of any reactor system.

The CE-PM report estimates the PT lifetime distribution, i.e., the distribution of the time of blister formation for each pressure tube in the reactor core. Since the simulation stops after the "first" failure of a PT, and it does not consider any repair or replacement followed by a failure in the simulation, it is concluded that the PT failure has NOT been analyzed as a repairable reliability problem. Since CE-PM implicitly analyzed the PT-CT contact as a non-repairable reliability problem, the probability of failure over an appropriate evaluation period is a meaningful reliability measure.

A fundamental problem with the definition of the failure frequency in a non-repairable problem is that its long term value approaches zero as the operating life increases to a large value as shown in Section 5.2. Even for an aging system where an increase in failure frequency is normally expected, the failure frequency in a non-repairable problem will exhibit a non-monotonic trend, which is an incorrect solution.

In essence, the reliability measure adopted in a probabilistic assessment must be commensurate with the type of performance limit state (serviceability or ultimate), nature of failure mode (self-announced or latent), rate of progression of limit states and the maintainability of the system. For the sake of conceptual and procedural accuracy of the probabilistic assessment, it is recommended that the CE-PM report include a section that describes the nature of limit states analyzed in the contact assessment and accordingly justify the chosen reliability measure.

1.4 Closing remarks

In addition to the review of statistical estimation and reliability analysis methods, this report also reviews (see In Section 6,) several issues related to sample size used in simulations, convergence, error bounds, sensitivity analysis and the spacer movement model used in the CE-PM report.

An overall conclusion of this review is that the proposed PT - CT contact assessment method is based on several modeling assumptions which are rather weakly supported by the evidence presented in the report. The questions and points of clarification raised in the review can be adequately addressed by CE, since the Candu industry has developed robust models and computer codes and gathered a substantial amount of data from fuel channel inspections. The use of existing information and database can strengthen the technical foundation of the PT-CT contact assessment by better justifying the modelling assumptions used in the CE-PM report.

In closing, a considerable emphasis is placed in this review on the conceptual accuracy of the problem formulation and the use of a correct reliability metric (i.e., failure rate versus failure probability) in a probabilistic assessment with a definite purpose. The purpose of this in-depth discussion is to pave the way for correct applications of the reliability theory to the risk-informed decision making in the nuclear industry.

2. Introduction

2.1 Background

The Candu Energy (CE) has developed a probabilistic method for assessing the impact of PT-CT contact in a CANDU reactor core [1]. This methodology, referred to as **CE-PM** here onwards, is basically a probabilistic version of a deterministic computer code C-DEPTH. The computer code C-DEPTH uses the finite element method (FEM) to model the sag of fuel channel assembly and calculate the gap between the pressure tube and calandaria tube at any given time in effective full power hours (EFPH).

Followed by a PT-CT contact, blister formation and blister cracking are considered as “non-performance” conditions or limit states of the pressure tube.

It must be stressed that the CNSC staff considers the “contact leading to blister formation (i.e., Heq > BFT)” as the ultimate failure state.

With respect to the limit states of blister formation and cracking, the CE-PM report predicts the expected number of pressure tube failures over an evaluation period. If the predicted failure frequency turns out to be in compliance with the acceptance criterion given in CSA Standard N285.8 [2], it is concluded that the reactor core is fit for service in the evaluation period. Obviously, the prediction of time of PT-CT contact is an important first step of the analysis.

The probabilistic version of PT-CT contact assessment is based on randomizing several input variables of the C-DEPTH code. Some input parameters can be directly measured through the gauging of fuel channels, such as dimensions of pressure tube and the location of spacers. Using the inspection data from several fuel channels, suitable probability distributions can be assigned to these variables. However, there are a few latent variables in the problem, such as the creep factor, which cannot be measured through inspection. Note that a latent variable means a variable that is not amenable to a direct measurement. Therefore, an indirect approach is adopted in the CE-PM report to estimate the distribution parameters of latent variables using a sample of measured PT-CT gap (i.e., model output itself).

The process of estimating distribution parameters of random variables contained in a model (or a computer code) using a random sample of model output is known as “model calibration” in the statistical literature. Some of the commonly used calibration methods in the literature are based on the Gaussian process model and the Bayesian updating methods [5].

The CE-PM report does not utilize a standard calibration method; rather it develops its own simulation-based method and refers to it as the "Benchmarking". The benchmarking is based on several key assumptions. It is therefore important to evaluate whether or not the rationale and assumptions of the CE-PM approach – (1) capture the essence of the problem, and (2) follow them consistently throughout the analysis. If the assumptions of the proposed approach are contradictory, they would challenge the foundation of the entire calibration method and the validity of numerical results of the subsequent probabilistic assessment.

2.1.1 Summary of PT-CT Assessment Approach

The CE-PM report aims to predict the distribution of PT-CT gap at a given axial location and a time in the service life, which is done in two parts:

1. Model calibration or Benchmarking: Use the available gap measurement data to calibrate some unknown parameters of the creep and sag model used in C-DEPTH.
2. Model prediction: Using values derived in step 1, predict the PT failure frequency over the evaluation period of the risk assessment.

In the CE-PM report, the prediction of PT-CT contacts is based on the Monte Carlo simulation method, which requires information about the probability distributions of all input random variables. Here one difficulty that arises is that some input variables cannot be measured directly during an inspection.

Probability distributions of variables like the PT wall thickness and diameter can be easily estimated from abundantly available fuel channel gauging data. In contrast, the reactor specific data for the end slopes of fuel channels and the creep rate cannot be obtained through inspection of fuel channels. To circumvent this limitation, the CE-PM developed a procedure to estimate the distributions of end slopes and the creep factor using the PT-CT gap data available from past inspection campaigns. This step is referred to as "statistical calibration of the model".

There is another parallel module that consists of the deuterium ingress and blister cracking models, which predicts the probability of blister formation and blister cracking at contact locations.

In the model prediction stage, simulations are carried out to predict contact locations, and blister formation/cracking following the contact. The probability distribution of the time to failure is finally converted into the frequency of pressure tube failures.

2.2 Objectives of the Review

The overall objective of this report is to examine the technical basis and assumptions of the PT-CT contact assessment presented in the CE-PM report [1]. An abbreviated form of this methodology was published in [3].

Specific objectives of the review are to:

1. Evaluate the conceptual accuracy of the simulation-based statistical estimation method adopted by CE-PM.
2. Assess the rationale of the model calibration (or benchmarking) method proposed by CE-PM in comparison to well established principles of probabilistic functional analysis.
3. Evaluate the conceptual accuracy of time dependent reliability predictions of CE-PM report.
4. Evaluate the methods used for analyzing sensitivity, convergence, and error bounds.

2.3 Scope of the Review

The scope of this technical review includes the evaluation of the assumptions and technical details of the PT-CT assessment methodology presented in the CE-PM report. The assessment method is divided into 5 major parts, as shown in Figure 1 of the CE-PM report, and reproduced below.

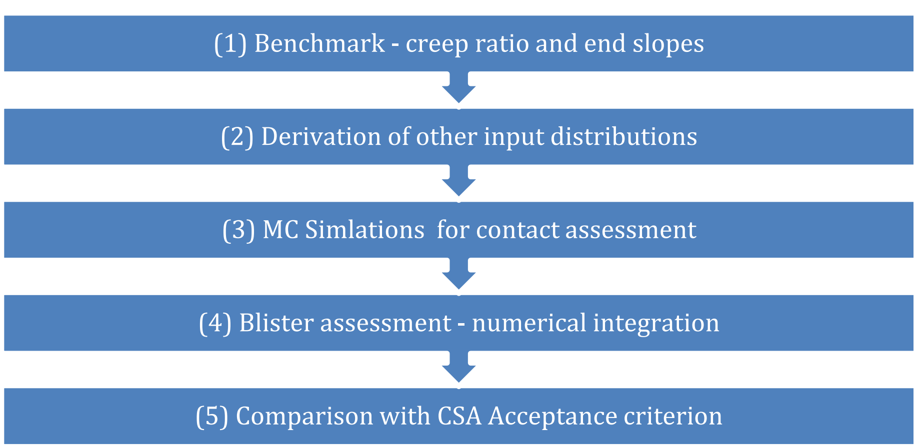


Figure 2.1: Flowchart of the probabilistic assessment of CE-PM report [1]

The PT-CT contact assessment utilizes a large body of knowledge that already exists in the CANDU industry about the fuel channel deformation, creep equation, reactor operation and inspection methods. Naturally, this review cannot cover all those broad aspects and established models that are already in use. Therefore, the attention is focussed evaluating the probabilistic formulation of the PT-CT contact problem as well as details of statistical estimation and computational methods.

This review presents an in-depth evaluation of technical details of Parts 1-3, the benchmarking to the Monte Carlo simulation method. The blister cracking assessment given in Part 4 is cursorily examined, since the blister cracking is not yet accepted by CNSC as a permissible performance criterion. The last Part 5 simply states the acceptable failure frequency criteria given in the CSA Standard [2], which are considered beyond the scope of this review.

The items that are not addressed in this review are the deuterium ingress model, creep equation, CDEPTH code, and the estimation of inspection measurement error.

The deuterium ingress model is a regression equation that correlates deuterium ingress to temperature and fluence. This equation is derived using the available inspection data in a straightforward manner. Furthermore, the use of a regression-based ingress model is an established practice in the industry. The CE-PM report does not provide any data that would allow one to review the calculations.

The creep equation and CDEPTH code are not the part of this review for an obvious reason that they are well established models with a long history of use by the industry. The review of statistical estimation of the inspection measurement error is not possible, because the CE-PM report does not provide any relevant data and description of the estimation method.

In a philosophical sense, this technical review is not about checking the "numerical values" of the input parameters and calculations, rather it is about evaluating: (1) the conceptual accuracy of probabilistic formulation, (2) mathematical accuracy and consistency of the estimation method, and (3) interpretation and relevance of the final results.

2.4 Organization

The review of the technical basis for PT-CT contact assessment is organized in four major sections. Since the model calibration method is the backbone of CE-PM, this review presents a detailed examination of the fundamental basis of the calibration method. Section 3 presents essential features of the probabilistic analysis of a function of random variables and basic approaches for statistical estimation. Section 4 analyzes the statistical model calibration problem or “Benchmarking” presented in the CE-PM report. Section 5 presents an overview of the concepts of time-dependent reliability analysis and their relation to PT-CT contact assessment problem. The PT failure problem in CE-PM appears to be treated as a non-repairable reliability problem. The appropriate reliability metric for this class of problems is the probability of failure over an evaluation period (e.g., an inspection interval), not the failure frequency as defined in the CE-PM report. Section 6 evaluates the issues related to selection of the sample size in simulations, convergence and sensitivity analyses, and spacer movement model. Each of Sections 4 – 6 concludes with a list of questions and points of clarification which can be followed through by CNSC experts.

3. Probabilistic Analysis and Estimation: Basic Background

3.1 Introduction

This section presents basic concepts of the probabilistic analysis of a function of random variables, and discusses possible approaches to statistical estimation of the distribution parameters. This discussion is essential to provide a context to the technical analysis of benchmarking model presented in the CE-PM report. This discussion will also provide necessary background to questions and points of clarification raised in the next Section of this report (see Section 4.7).

The basic idea discussed in this Section is that a probabilistic assessment criterion can be technically defined by a function of random variables. For example, a function, $g(\mathbf{X})$, can be used to define the performance requirement in terms of a vector of basic random variables (or inputs) $\mathbf{X}$. The function, $g(\mathbf{X})$, is also referred to as the "performance function" or the "limit state function".

The basic aim of the probabilistic analysis is to determine the distribution of the performance function, $F_G(u;\ \alpha_G, \beta_G)$, which is subsequently used in the reliability prediction. Note that $F_G(\dots)$ is the cumulative distribution function (CDF) of $g(\mathbf{X})$, and α_G and β_G are distribution parameters.

If the assessment involves a time dependent degradation phenomenon, then the probabilistic analysis should also provide a conceptual discussion about the modeling of the degradation process. For example, should the irradiation creep be modelled as a stochastic process (e.g. a cumulative process) or as a growth curve with covariates. The final model should of course be justified based on the mechanistic knowledge of the degradation process.

The statistical estimation means the estimation of the distribution parameters, α_G and β_G using the available data. The maximum likelihood method is a standard method of the estimation. Other useful approaches are the probability paper plot method and the method of moments. If there is interest in updating the parameters as new data become available, the Bayesian framework has to be adopted for the estimation purposes.

3.2 A Function of Random Variables

For the sake of brevity of discussion, consider a scalar function of two random variables, $g(X_1, X_2)$. In other words, the vector **X** has two variables, X_1 and X_2. The cumulative distribution function of a random variable, X_i, is denoted as $F_{X_i}(x)$, and the probability density function is denoted as $f_{X_i}(x)$. The distribution parameters of X_i are denoted as (α_i, β_i). The mean and standard deviations of X_i are denoted as μ_i and σ_i. X_1 and X_2 are referred to as basic variables.

The distribution of the performance function depends on distributions of the underlying basic random variables X_1 and X_2. The derivation of statistical moments and the distribution of the performance function are studied in the following sub-Sections.

3.2.1 Statistical Moments: Nonlinear Function

The expected value of the function can be calculated in general using the following bivariate integral:

$$E[g(X_1, X_2)] = \int \int g(x_1, x_2) f(x_1, x_2) dx_1 dx_2 \tag{3.1}$$

Note that $f(x_1, x_2)$ is the bivariate PDF of basic variables. If the function involves n basic variables, evaluation of the expected value would require solution of an n-dimensional multivariate integral, which is a computationally challenging task especially when $n > 4$. To overcome this difficulty, the Taylor series approximation of the function is often used.

A Taylor series expansion of the function, $g(X_1, X_2)$, can be written about the mean values of X_1 and X_2 in the following manner

$$\begin{aligned} g(X_1, X_2) = g(\mu_1, \mu_2) + \sum_{i=1}^{2} (X_i - \mu_i) \left(\frac{\partial g(X_i, X_j)}{\partial X_i} \right)_{\substack{X_1 = \mu_1, \\ X_2 = \mu_2}} + \\ \frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} (X_i - \mu_i)(X_j - \mu_j) \left(\frac{\partial^2 g(X_i, X_j)}{\partial X_i \partial X_j} \right)_{\substack{X_1 = \mu_1 \\ X_2 = \mu_2}} + \cdots \end{aligned} \tag{3.2}$$

The first order approximation of the mean of the function can be given as

$$\mu_G \approx g(\mu_1, \mu_2) \tag{3.3}$$

This approximation is useful for mildly nonlinear functions.

The variance can be approximated as

$$\sigma_G^2 \approx \sum_{i=1}^{2} c_i^2 \sigma_i^2 + \sum_{i \neq j} \sum c_i c_j E[(X_i - \mu_i)(X_j - \mu_j)] \tag{3.4}$$

Note that $E[(X_i - \mu_i)(X_j - \mu_j)]$ is the covariance between X_1 and X_2 and

$$c_i = \left(\frac{\partial g(X_1, X_2)}{\partial X_i} \right)_{X_1=\mu_1, X_2=\mu_2} \quad \text{(for } i = 1, 2\text{)} \tag{3.5}$$

It should be noted that the first order approximation of the variance requires the derivatives of the function.

3.2.2 Statistical Moments: Linear Function

A linear bivariate function can be written in a standard form as

$$g(X_1, X_2) = b_0 + b_1 X_1 + b_2 X_2 \tag{3.6}$$

Considering X_1 and X_2 as independent random variables, the mean and variance of the function can be evaluated exactly as

$$\mu_G = b_0 + b_1 \mu_1 + b_2 \mu_2 \quad \text{and} \quad \sigma_G^2 = b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2 \tag{3.7}$$

In summary, statistical moments of a linear function of *n* random variables can be exactly evaluated.

3.2.3 Probability Distribution: Nonlinear Function

In case that $g(X_1, X_2)$ is a nonlinear function of general form (an explicit or implicit function), then its distribution is difficult to derive analytically. The Monte Carlo simulation method is the easiest way of estimating the distribution of a general function.

In the simulation method, a random sample of $g(X_1, X_2)$ is simulated from known distributions of X_1 and X_2. The simulated sample is then fitted by a suitable distribution based on a standard method like maximum likelihood method or probability paper plot method.

The simulation method can be numerically intensive, if a single evaluation of the function requires a complex computer program with significant CPU time. In such a situation, simulation of 1000 to 10,000 simulations would take an inordinate amount of time.

If the function is simple and exxplicit, then an analytical form of the distribution can be derived.

3.2.4 Probability Distribution: Linear Function

In case of a linear function of independent basic variables given by Eq.(3.7), the probability distribution can be computed from the following convolution integral

$$F_G(u) = \int F_{X_1}\left(\frac{u - b_0 - b_2 x_2}{b_1}\right) f_{X_2}(x_2) dx_2 \tag{3.8}$$

Note that the integration limits depend on the range of validity of basic variables.

In case of *n* random variables, this approach will become quite tedious, as one will have to evaluate nested multivariate integrals.

3.2.5 Special Case: A Linear Function of Gaussian Random Variables

A linear function of Gaussian (or normal) random variables can be analyzed exactly. It means that statistical moments as well as the probability distribution can be exactly derived.

The mean and standard deviation of the performance function can be obtained from Eq.(3.7). The probability distribution of the function will also be a Gaussian distribution with the mean μ_G and standard deviation σ_G, and the density function given as

$$f_G(x) = \frac{1}{\sigma_G \sqrt{2\pi}} \exp(-\frac{x - \mu_G^2}{2\sigma_G^2}) \tag{3.9}$$

<u>A linear function of normally distributed random variables will always follow the normal distribution. This implies that if a function consisting of normally distributed random variables follows the normal distribution, then the performance function must be linear.</u>

3.3 Statistical Estimation

Here we discuss methods for estimating the distribution type and parameters of the distribution of $g(X_1, X_2)$. This problem can be approached in different ways, depending on the availability of data.

3.3.1 Data for basic variables are available

If random samples of X_1 and X_2 are available, then their distributions can be directly estimated using any standard methods, such as, probability paper plot, method of moments, and maximum likelihood method.

After the estimation of the distributions $F_{X_1}(x_1;\ \alpha_1, \beta_1)$ and $F_{X_2}(x_2;\ \alpha_2, \beta_2)$, the Monte Carlo simulation method can be used to simulate a random sample of the performance function. Simulation is the most direct method of estimating the distribution of the function of any general nature, such as a nonlinear, time dependent and implicit function. The simulated sample of the performance function can be subsequently fitted by any suitable method, such as the probability paper plot or the maximum likelihood method.

In case of a linear function of non-normal random variables, the distribution can be obtained by calculating the convolution integral given in Eq. (3.8). This method would work for a bivariate case.

In the special case of a linear function of Gaussian random variables (Section 3.2.5), the problem becomes quite simple. In this case, the performance function will follow the normal distribution with the mean and variance calculated using Eq.(3.7).

3.3.2 Data for the function are available

In this case, a random sample of values of the performance function, $g_1, g_2, \cdots, g_n$, is available through a direct inspection method. The estimation of the distribution, $F_G(u)$, is quite straight forward using any of the standard estimation methods (probability paper plot, method of moments or maximum likelihood).

3.3.3 Data for function available but distributions of basic variables are unknown

This is a rather interesting problem in which a random sample of the function, $g_1, g_2, \cdots, g_n$, is available, but distributions of underlying basic variables (X_1 and X_2) are not known since they are not directly observable, i.e., X_1 and X_2 are latent variables.

If a random variable is not amenable to direct measurement, it is referred to as "latent" variable.

In this case, it is obvious that the estimation of distribution parameters of the basic variables has to be done in some indirect manner, except in a few special cases, such as a linear function of Gaussian variables.

Technically, the distribution, $F_G(u)$, has to be written as a function of parameters $(\alpha_i, \beta_i), i = 1, 2$, of the basic variables. Then the maximum likelihood method can be used for estimation. The other alternative is to use the method of moments to calculate the distribution parameters.

In the literature, the estimation of parameters of the basic variables using a random sample of the function is known as statistical model calibration. Since the statistical calibration problem is not easy to solve in a general setting, various approximations have been developed in the literature [5].

The Bayesian method also provides an elegant approach to the calibration problem. In this approach, we can start with suitable prior distributions for the parameters (α_i, β_i), and update them using the random sample of performance function.

3.4 Closing Remarks

An assessment criterion can be mathematically represented as a performance function of random variables. The nature (linear versus nonlinear) of the performance function must be carefully examined, since this will govern the derivation of the distribution as well as estimation of the distribution parameter.

The probabilistic analysis and estimation of a linear function of normally distributed variables is the simplest problem, and therefore widely used in engineering applications.

A linear performance function consisting of normally distributed random variables will always follow the normal distribution. This implies that if a function consisting of normally distributed random variables follows the normal distribution, then the performance function must be linear.

In closing, a probabilistic model adopted for the risk assessment must be consistent with basic principles of probabilistic analysis of a function of random variables.

4. Review of the Model Calibration (Benchmarking)

4.1 Basic Problem

The CE-PM report [1] describes the statistical estimation of parameters of the PT-CT gap prediction model in such colloquial terms that it becomes difficult to evaluate the precise mathematical basis of this approach. An additional reference, [3], is quite useful in understanding the assumptions and details of the benchmarking method. In this Section, the CE-PM approach is rephrased in a more precise mathematical language, and implications of the assumptions and approximations are investigated in a systematic manner. Finally, points of clarification related to CE-PM are summarized at the end of this Chapter.

The PT-CT gap in an i^{th} fuel channel at a k^{th} axial location, a_{ik}, and time, t, in the service life can be written as a general function of random variables:

$$Y_i(a_{ik}, t) = g(X_1, X_2, \cdots, X_n, a_{ik}, t) \tag{4.1}$$

where $X_1, X_2, \cdots, X_n$ are basic random variables that affect the PT-CT gap, as summarized in Table 6.1 of CE-PM. For example, PT and CT dimensions, flux fluctuations, axial load, spacer location accuracy, creep factor and end slopes are some of the variables that influence the PT-CT gap. This function in essence represents the C-DEPTH code used in CE-PM.

If computation of the function (i.e. running of C-DEPTH code) is a time consuming task, then computer codes are typically approximated by an appropriate response surface, fitted by polynomial or spline functions. Another approach is to replace the computer code by a surrogate function. These topics are rich areas of research in uncertainty analysis in engineering and science. The dimensional reduction method [6], and the Gaussian process model [5] provide efficient means to solve this class of problems. The CE-PM report did not seem to explore any such formal methods.

In case of a fuel channel with tight fitting spacers, the CE-PM report concludes that only two variables are most influential, namely, X_1= creep factor, and X_2 = end slopes. Furthermore, CE-PM assumes that X_1 and X_2 are independent random variables, which are also invariant with time and axial location.

The probability distribution of gap in an i^{th} fuel channel at an axial location, a_{ik}, depends on the distributions of X_1 and X_2 and the form of the function, $Y_i(a_{ik}, t)$, (nonlinear or linear).

Based on available inspection data, the end slope (X_2) is modelled as a normal random variable. However, the distribution parameters for a specific reactor have to be determined from the reactor specific data.

The creep factor is basically an adjustment factor applied to the creep equation to account for a departure in the creep of a channel from the nominal estimate provided by the standard creep model. The creep factor in the CE-PM report is assumed to be invariant with respect to time and axial location within a pressure tube. Being an adjustment factor, the creep factor distribution has to be derived from the reactor specific information. The creep factor is also assumed to follow the normal (or Gaussian) distribution.

In summary, the PT-CT gap prediction problem is modelled as a function of time and location invariant random variables, which results in a considerable simplification of the analysis.

The location invariance means that the probability distribution of a random variable does not change with axial location along the length of the pressure. For example, a constant creep factor is applied to every axial location. Similarly, the time invariance means that the distribution does not change with the time, i.e., service life of the pressure tube. An example is a constant creep factor used over the entire service life of the pressure tube.

The CE-PM report assumes that the PT-CT gap follows the Gaussian (or normal) distribution. This assumption is the corner stone of the statistical estimation method presented in CE-PM. This issue will be examined later in this Chapter.

The basic problem of calibration tackled in the CE-PM report is that given a set of gap measurements, $y_i(a_{ij}, t_k)$, how one should estimate parameters of X_1 and X_2. Note that the data are collected from a sample of inspected fuel channels, $i = 1, n$, at different locations a_{ij}, j = 1, *m*, and possibly at different points in time (t_k), *k* = 1, *l*, in the service life of the reactor. This step in the CE-PM report is called a "Benchmarking" step.

In essence, the basic variables X_1 (creep factor) and X_2 (end slope) are used to model variability at the reactor level. It is conceptualized that each pressure tube takes a set of random values (x_1, x_2) from the population distribution X_1 and X_2, and then proceeds along the creep curve as calculated by the C-DEPTH code for the rest of its service life (or evaluation period). In other words, there is no temporal uncertainty considered in the evolution of fuel channel deformation process over time.

The CE-PM report describes an iterative Monte Carlo simulation method for benchmarking purposes. Since the description is devoid of accurate mathematical functional equations and relationships, it is hard to interpret and evaluate various steps of this method.

After the model is calibrated, CE-PM predicts the probability distribution of PT-CT gap at any future time using the creep evolution law built in the C-DEPTH computer code. The creep law as such is taken as a deterministic law, modified to a channel specific condition by the creep factor and end slopes. This step is referred to as the prediction step. The prediction is based on the standard Monte Carlo simulation method.

4.2 Summary of Probabilistic Analysis Problem

In the case of fuel channels with tight fitting spacers, the PT-CT gap in an i^{th} pressure tube at a k^{th} location, a_{ik}, and time, t, in the service life is modelled as a function of only two random variables

$$Y_i(a_{ik}, t) = g(X_1, X_2, a_{ik}, t) \tag{4.2}$$

where X_1 and X_2 are the creep factor and the end slope, respectively.

The gap measurement tool introduces some noise (or random error W_M), such that the measured gap, $Y_{Mi}(a_{ik}, t)$, randomly deviates from the actual gap. The measurement error is assumed to be additive as shown below:

$$Y_{Mi}(a_{ik}, t) = Y_i(a_{ik}, t) + W_M = g(X_1, X_2, a_{ik}, t) + W_M \tag{4.3}$$

The measurement error has a mean zero and a standard deviation of σ_M.

The statistical estimation problem in the CE-PM report is formulated based on some assumptions, which are collectively presented in the following list:

1. The creep factor (X_1) and end slope (X_2) are modelled as normally distributed and independent random variables [3].
2. The PT-CT gap is also assumed to follow the normal distribution.

3. The PT-CT gap in internal spans is largely affected by the creep factor (X_1), but the end slope (X_2) has negligible influence.
4. The PT-CT gap in end spans is largely controlled by the end slope (X_2), but the effect of creep factor is negligible.
5. The creep factor is treated as a time invariant random variable and it is also constant for all axial locations in a pressure tube.
6. The end slopes are also modelled as the time and location invariant random variables.

The model calibration means the estimation of the four unknown distribution parameters, namely, $X_1(\mu_1, \sigma_1)$ and $X_2(\mu_2, \sigma_2)$ given a sample of measured gap values ($y_{M1}, y_{M2}, \cdots, y_{Mn}$).

The CE-PM report presents its own "non-conventional" method for model calibration without explaining why a standard estimation method, such as maximum likelihood, is not applicable to this problem.

4.3 Review of the Simplex Method

The Simplex method, as described in Section 5.1 of [1], is proposed as a method of estimating variability in the creep rate and end slope as a part of model calibration.

Given gap measurement data, the simplex method calculates values of the creep factor ($X_1 = x_1$) and end slope ($X_2 = x_2$) by minimizing the square of the error between measured and calculated gap profiles. In this way, given a sample of *n* measured gap profiles, corresponding samples of creep rates, $x_{11}, x_{12}, \cdots, x_{1n}$, and end slopes, $x_{21},\ x_{22}, \cdots, x_{2n}$, can be calculated. These calculated samples can then be fitted by any suitable method like probability paper plot.

Although the Simplex method is conceptually attractive, it is fraught with practical difficulties as discussed below.

If the model given by Eq.(4.2) is not a perfect representation of the real deformation of a fuel channel, which is expected to be the case, then it will not be possible to find values of the creep factor and end slope that would match precisely the calculated and observed gap profiles. The CE-PM report also states that it is hard to achieve a close match between the measured and calculated gap profiles.

The measured gap value is already contaminated by the measurement error of a random magnitude, as shown by Eq.(4.3). Thus, the creep factor and end slope calculated using this gap value will also be contaminated by the measurement error. The magnitude of measurement error in the creep factor and end slope cannot be directly quantified.

The consideration of the movement of loose spacers in simplex method is not discussed in clear terms.

Because of the above two factors, the Simplex method cannot provide reliable estimates of distributions of creep factor and end slope. The CE-PM report does acknowledge limitations of the Simplex method and does not use this method in the final calibration of the PT-CT gap model.

4.4 Evaluation of the Probabilistic Model of the CE-PM Report

Based on available gap measurement data, CE-PM assumes that the PT-CT gap follows the Gaussian (or normal) distribution (Assumption 2 on page 26). This assumption has a profound implication in the statistical estimation method developed in the CE-PM report.

Since the PT-CT gap with tight fitting spacers is modelled as a function of two independent normal random variables (creep factor and end slope), and the function (i.e., PT-CT gap) itself follows the normal distribution, then the function must be linear or almost linear. The reason is that the distribution of a function of normal random variables will be normal, only if the function is linear, as discussed in Section 3.2.5.

Therefore, the relationship of the PT-CT gap with the basic variables (creep factor and end slope) must be examined as a part of the model validation exercise. If the PT-CT gap turns out to be a highly non-linear function of X_1 and/or X_2, then the basic tenet of CE-PM would become questionable.

The gap prediction model under the assumptions of CE-PM can be represented as

$$Y_i(a_{ik}, t) = b_0(a_{ik}, t) + b_1(a_{ik}, t)X_1 + b_2(a_{ik}, t)\, X_2 \tag{4.4}$$

Since the basic variables, X_1 and X_2 are time and location invariant, the model coefficients, $b_j(a_{ik}, t)$, j = 0 – 2, must include the dependence on time and location. The reason is that there is no other place in the model but these coefficients to incorporate this dependence.

The model coefficients are like influence functions, which can be determined by running the C-DEPTH code for different values of time and location.

The assumptions 3 and 4 (on page 27), further simplify the PT-CT gap function into a univariate function. In a fuel channel with 4 fixed spacers (e.g. Bruce B, unit 8), there are two end spans (1 and 5), and three internal spans (2, 3 and 4).

For the internal spans according to CE-PM report, the effect of end slopes (X_2) can be ignored, such that the PT-CT gap can be modelled as a function of the creep factor only:

$$Y_{INTi}(a_{ik}, t) \approx b_{01}(a_{ik}, t) + b_{11}(a_{ik}, t)X_1 \tag{4.5}$$

Using Eq.(4.3), the measured gap can be written as

$$Y_{INTi}^{M}(a_{ik}, t) \approx b_{01}(a_{ik}, t) + b_{11}(a_{ik}, t)X_1 + W_M \tag{4.6}$$

In the end spans according to the CE-PM report, the effect of creep factor (X_1) can be ignored, such that the measured gap can be modelled as

$$Y_{ENDi}^{M}(a_{ik}, t) \approx b_{02}(a_{ik}, t) + b_{12}(a_{ik}, t)X_2 + W_M \tag{4.7}$$

Since the PT-CT gap model is reduced to a linear model as a function of a single normally distributed random variable, the statistical estimation can be done in a fairly simple and direct manner. In this case, standard statistical methods like the maximum likelihood method can be used for the estimation of unknown distribution parameters.

Note that the mean and variance of the measurement error (W_M) are estimated through a data normalization procedure described in the CE-PM report. Although the normalization procedure may be a standard practice in the industry, the data and estimation procedure should conform to the requirements of the statistical method. For example, the estimation of mean and variance of a random variable is typically based on a sample of independent values. The data used to estimate measurement error must meet this requirement.

4.5 Remarks: Fuel Channels with Loose Spacers

The overall deformation of a fuel channel and hence the PT-CT gap profile is strongly dependent on the location of spacers. In case of channels with loose spacers, the dependence of gap on the creep factor and end slopes will be complicated by locations of loose spacers, which tend to move over the service life of the channel. For example, after a SLAR maintenance campaign if spacers were placed into a new configuration, the future gap profile in the subsequent period could change significantly. Even after SLAR, there remains a possibility of further spacer movement. In summary, the gap profile would become strongly dependent on the SLAR history of a channel. The CE-PM report does not provide a clear explanation about the calibration of the creep factor and end slopes in fuel channels with loose spacers.

4.6 Closing Remarks

It is rather strange that the CE-PM report does not discuss at all the dependence of the PT-CT gap on the basic variables in a deterministic sense. For example, in the absence of any randomness, how does the gap vary with end slopes? How does the gap vary with the creep factor? If the dependence is indeed linear or approximately linear, as implied by the modeling assumptions of the CE-PM report, then the statistical estimation would become quite a manageable task.

Thus, the relationship between the PT-CT gap and basic variables (creep factor, end slope, and spacer locations) must be examined at the outset of any probabilistic assessment model. If the PT-CT gap turns out to be a highly non-linear function of X_1 and/or X_2, then the basic tenet of CE-PM would become questionable.

If the assumptions of the CE-PM report are indeed valid, then this Section of the review shows that the gap model can be calibrated in a direct manner using a standard estimation method. Apparently, there is no need for an iterative Monte Carlo method for model calibration as proposed by the CE-PM report. If, however, the main assumption of the CE-PM report, namely, PT-CT gap follows a normal distribution, is not valid, the model calibration problem must be revisited.

4.7 Points of Clarification

In a probabilistic assessment, a suitable performance criterion can be represented as a function of random variables. The nature (linear versus nonlinear) of the performance function should be carefully examined, since this will govern the derivation of the distribution and its parameters. The essence of this review is that a probabilistic model adopted for the assessment must be consistent with basic principles of the analysis of a function of random variables. Based on the discussion presented in this Chapter, CE analysts are requested to clarify the following points.

4.1 It is clear from the previous discussion that the most significant assumption or assertion of the CE-PM report is that the PT-CT gap follows the normal distribution. The only evidence provided in support of this assumption is a histogram presented in Figure 7 of [1]. Can you elaborate more about the source of the data plotted in this Figure? For instance, which channels and spans the data come from, and what is the EFPH that the data correspond to? Since the minimum gap varies with time, data for different spans and EFPH values should not be pooled without the consideration of potential heterogeneity in the data. For example, the minimum gap data from different locations and time points cannot be pooled, since they do not come from the same population.

4.2 Since the Candu industry has gathered a substantial amount of gap measurement data from past fuel channel inspections, the assumption of the normality of the gap distribution should be easy to verify. The CE analyst is recommended to compile actual gap measurement data from different fuel channels and evaluate the probability distribution of gap data by a standard statistical method, e.g., K-S test, Chi-Square test and the probability paper plot. The normality of gap distribution should be evaluated at different axial locations and time points at which the data were collected.

4.3 The cumulative distribution function in Figure 8 of [1] is plotted on a linear scale, which makes it difficult to evaluate the goodness of fit in the tail regions of the distribution. The data must be plotted on the normal probability paper to verify the assumption of normality.

4.4 The CE-PM report proposes that PT-CT gap in fuel channels is primarily influenced by two normal random variables, the creep factor and the end slope. It is further assumed that the PT-CT gap also follows the normal distribution. This assumption has a profound implication in the statistical estimation method developed in the CE-PM report. An immediate implication of these assumptions is that the PT-CT gap must be a linear (or almost linear) function of the creep factor and end slope. The reason is that the distribution of a function of normal random variables will be normal, only if the function is linear, as discussed in Section 3.2.5. The PT-CT gap being a linear function of creep factor and end slope must be verified using the computer model C-DEPTH to provide a mechanistic basis to the modeling assumptions.

4.4.1 CE-PM assumes that PT-CT gap in internal spans is a univariate function of the creep factor, since the effect of end slope is negligible. It should be easy to verify this proposition. For a typical fuel channel, for a given time (e.g., 150,000 EFPH) and location (centre of an internal span), the PT-CT gap should be plotted as a function of the creep factor. This plot should be almost linear, in order to be consistent with the assumptions of CE-PM report. This relationship should be explored for other EFPH values, just to make sure that the linear relationship is valid for a wide ranging time and locations. The CEI analyst is requested to provide such an example for some typical fuel channel.

4.4.2 The analysis of the previous step (4.4) should be repeated to explore the relationship between PT-CT gap in end spans and the end slope. To make it clear, for a given time (e.g., 150,000 EFPH) and location (centre of an end span), the PT-CT gap should be plotted as a function of the end slope. This plot should be almost linear, in order to be consistent with the assumptions of CE-PM report.

4.5 If the PT-CT gap indeed turns out to be linear with respect to the creep factor and end slopes, then the statistical estimation can be based on much stronger footing by using well established statistical methods like the maximum likelihood method. The iterative Monte Carlo method uses long heuristic arguments to justify a fitting criterion, which seems to be unnecessary, and perhaps an unreliable criterion. The CE analyst is requested to compare the results of statistical estimation obtained from standards methods (MoM or MLM) versus the proposed iterative Monte Carlo method.

4.6 If the PT-CT gap does NOT turn out to be a linear function of the creep factor and end slope, then the basic tenet of the CE-PM report cannot be considered valid in a fundamental sense. The CE-PM assumptions of normality of gap data can then be considered purely incidental, and the results of the analysis will become questionable. After the PT-CT contact, the spread of the contact is possibly a non-linear function. Is this aspect included in the analysis?

4.7 The limitation of the Simplex method (see Section 4.3) suggests that a single creep factor cannot match the observed gap profile with a high degree of accuracy. Does it not mean that the use of a constant creep factor for all 5 spans of the pressure tube is an inappropriate choice?

4.8 Equation (5.1) of the CE-PM report, $(\sigma_{bench})^2 = \left(\sigma_{creep}\right)^2 + (\sigma_{mes})^2$, does not seem to be correct. This equation is valid only if the PT-CT gap is a simple sum of the creep factor and the measurement error of the form: $gap = creep\ factor + mes\ error$. This relation is not true, as it ignores various multiplying factors as shown by Eq.(4.6). The CE analyst must clarify this issue.

4.9 There are several minor issues related to intriguing and unsubstantiated statements scattered throughout the CE-PM report. Some examples are given below:

4.9.1 The Central Limit Theorem is cited as a justification for Equation 5.1 of the CE-PM report. Why and how?

4.9.2 A statement above Figure 7: "the fact that a population is normally distributed is a characteristic of the population….". In what sense normality becomes a characteristic?

4.10 There is a subtle point about the independence of the gap data that is seemingly ignored throughout the CE-PM report. The reason to raise this point is that when the gap data are pooled in the distribution fitting and parameter estimation, all data points must be independent. Let's clarify this further. In a given fuel channel, can the minimum gap values in the internal spans be considered independent of each other? The gap inspection of a single fuel channel will provide three minimum gap values from the three internal spans. It is not clear as to how these values are treated in statistical estimation. These three values are not truly independent values, since they are related to a single creep factor. This issue of independence requires a thorough discussion in the CE-PM report. The reason is that the use of correlated values in statistical estimation will produce biased estimates.

4.11 The standard deviation of measurement error is purported to be 0.25 mm for fixed spacers. The CE analysts must provide a better documentation about the estimation of this important quantity. The reason is also related to the previous point (4.10) of independence. If the measurement error analysis is based on pooling of correlated data, then the resulting estimate will be biased. The role of measurement error is quite important in a subtle sense. The variability of gap data is decomposed between the model variability and measurement error. By attributing more variability to measurement error, the model variability can be effectively reduced. A reduction in the model variability will also reduce variability in the prediction stage, i.e., for predicting the probability of contact.

4.12 The gap prediction model for fuel channels with loose spacers is not clearly described. As discussed in Section 4.5, the PT-CT gap profile would be strongly dependent on the history of spacer movements. Therefore, the following questions need to be addressed:

4.12.1 The gap profile would become a function of the history of spacer movement. In this case, will the gap profile follow a normal distribution? Even if the gap follows a particular distribution, isn't it true that the distribution parameters will become a function of spacer movement history?

4.12.2 Since the gap profile would change after movements of spacers, how will the distribution of the creep factor change with spacer movements? Similarly, the effect of end slopes will also depend on spacer movements. The CE-PM report must provide a clear explanation about the model calibration and reliability prediction for fuel channels with loose spacers.

4.13 The blister cracking model presented in Section 6.2 is a logistic regression equation as a function of the blister depth and the stress. Although the estimation of regression model is a straightforward exercise, its application to the calculation of probability of cracking described in Section 6.3 is not clear. The CE analyst should elaborate on the rationale of the convolution method of calculation. The steps of the method should be described in a technical manner along with appropriate equations. Since the CNSC does not accept the blister cracking as an acceptable state of performance, this topic may as well be removed from the assessment methodology.

4.14 A basic assumption of the CE-PM report is that the creep ratio has negligible influence on the gap in the end spans, and hence the end span gap data only need to be used in the calibration of the end slope distribution. Figure 27 of the CE-PM report is useful to evaluate this assumption. Figure 27 shows the sensitivity of the creep ratio to the PT-CT gap in the channel P08. As seen by the width of the bounds, the gap in the end spans seems to be fairly sensitive to the creep ratio. In fact, the order of this sensitivity is similar to that of the sensitivity of end slopes as shown by Figure 28.

Results presented in Figures 27 and 28 appear to contradict the basic premise of statistical analysis given in the CE-PM report. Please clarify this issue by providing additional examples.

4.15 Similarly, the CE-PM report assumes that the effect of end slopes on the gap in internal spans can be ignored. Comparing the results given in Figures 27 and 28 for the two internal spans (excluding the central span), the following observation can be made. The sensitivity of the creep ratio to gap is of lower order than that of the end slope. Since the end slope does seem to have a measurable influence on the gap in these spans, why should this be ignored from the analysis? Please clarify this issue.

5. Prediction of the PT Failure Frequency

5.1 Background

The overall goal of CE-PM is to estimate the PT failure frequency in the reactor core and compare it with the acceptance limit specified in the CSA Standard N285.8 [2]. For example, an acceptable PT failure frequency, as per CSA Standard, for the Bruce-8 reactor core is 4.13×10^{-3} failures per year.

In the CE-PM report, the pressure tube limit state is defined as a combination of two events, namely, PT-CT contact and blister formation followed by blister cracking. A blister is expected to form at that contact location where the equivalent hydrogen concentration (H_{eq}) exceeds a threshold, known as the blister formation threshold (BFT). Thus, blister formation can be considered as the first limit state, which is followed by the second limit state of blister cracking. Here, the term "failure" signifies that a pressure tube has reached one of these two limit states.

<u>It must be stressed that the CNSC staff considers the "contact leading to blister formation (i.e., Heq > BFT)" as the ultimate failure state.</u>

Section 7 of the CE-PM report describes a hybrid method for calculating the PT failure frequency in the reactor core. The probability of PT-CT contact is estimated by the Monte Carlo simulation method, whereas H_{eq} is predicted by an empirical regression model. The probability of blister cracking is calculated by a numerical integration method. The hybrid approach leads to the cumulative distribution of time to failure of each pressure tube, which is then converted into the failure frequency. The failure frequency is defined as the average or expected number of PT failures per year (≈ 8000 EFPH). Failure frequencies of all PTs in the reactor core are added to calculate an overall failure frequency for the entire reactor core.

Presumably the purpose of a probabilistic assessment is to demonstrate that risk associated with the operation of a component or system is less than a minimum acceptable value as specified by a code or standard. A first fundamental question then arises as to what is the relevant reliability metric that should be adopted in the PT failure problem. Is the failure frequency as defined and calculated in CE-PM a relevant measure of reliability? The answer to this question depends on the type of reliability analysis to which the PT failure problem belongs.

In the reliability literature, the reliability analysis problem is classified into two broad categories: non-repairable and repairable equipment (or system or component) reliability problems [4].

In the non-repairable or "first-failure" problem, the mission reliability and probability of failure are relevant metrics of reliability, whereas the failure rate (or failure frequency) is a relevant risk metric in the repairable system problem [4]. Every reliability analysis must begin with identifying the type of problem, repairable or non-repairable, that it intends to address in the probabilistic assessment.

In the CE-PM report, failure frequency is adopted as a reliability measure without any justification and discussion about the type of the PT failure problem it analyzes. This point must be addressed, at least at a conceptual level, to make the technical basis document a sound piece of work.

In essence, the reliability measure reported in a probabilistic assessment must be commensurate with the type of limit state (serviceability or ultimate), nature of failure mode (self-announced or latent), rate of progression of limit states and the maintainability of the system. The failure rate and failure probability serve as a meaningful reliability measure depending on the definition of the reliability analysis. This issue is explored in depth in Section 5.4.

The conceptual distinction between the repairable and non-repairable reliability problems cannot be described in a few paragraphs. Therefore, this Section presents key concepts of the time dependent reliability theory with simple illustrative examples. Based on this discussion, the importance of using a correct measure can be better understood.

It is recognized that in this Chapter a considerable emphasis is placed on using a correct reliability metric (i.e., failure rate versus failure probability) in a probabilistic assessment. The reason for this in-depth discussion is to pave the way for correct applications of the reliability theory to the risk-informed decision making in the nuclear industry.

5.2 Reliability of a non-repairable component

5.2.1 Basic Concepts

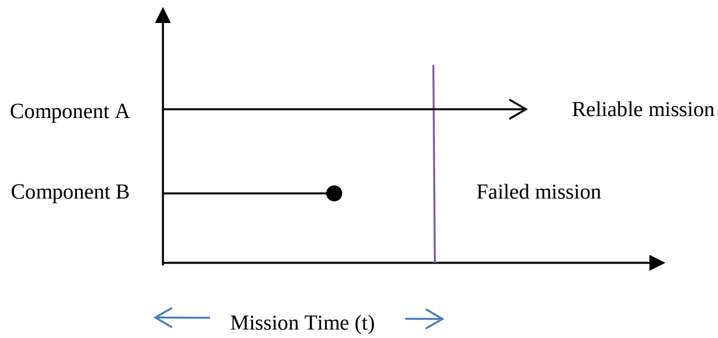


Figure 5.1: Concept of reliability of non-repairable components

The non-repairable problem means that if the component fails during the operational period (or mission), there is no opportunity to repair and restore the component. In other words, the first occurrence of component failure amounts to the failure of the mission. An example of this type of problem is a failure of an aircraft engine during a flight resulting in ending engine's life and perhaps ending the flight itself.

Figure 5.1 illustrates the concept of reliability of non-repairable components. Since the component A survives the mission duration (*t*), it is called a reliable mission. In contrast, component B fails during the mission causing a mission failure.

The mission means that operating interval in which a reliable operation of component (or system or equipment) is sought. In case of a nuclear reactor, this "mission" can be a time interval between two maintenance outages, or a future inspection interval.

It is in the context of non-repairable problems that the classical definition of reliability was put forward as *the probability of a system functioning within specified limits for a specified time under postulated conditions*. The risk is defined as the expected consequence of a failure, which is a product of the probability of failure with its consequence.

The mission reliability and mission probability of failure are quantified using the lifetime distribution of the component in relation to the mission duration.

5.2.2 Lifetime Distribution

The lifetime distribution describes in probability the ability of a component to survive to a certain age. The lifetime distribution is a regular probability distribution, like gamma, Weibull or exponential, which provides the best-fit to available lifetime (or time to failure) data. The lifetime is denoted as X, which has the probability density, $f_X(x)$, and distribution, $F_X(x)$. The component reliability at any age x is given by the reliability function, $R_X(x) = 1 - F_X(x)$. The likelihood of failure of a component at any age $X = x_1$ is reflected by the hazard rate, $h(x_1)$, which is given as:

$$h_X(x_1) = \frac{f_X(x_1)}{1 - F_X(x_1)} = \frac{f_X(x_1)}{R_X(x_1)} \qquad (5.1)$$

The denominator in the above equation comes from the condition that the component of age x_1 has survived the time interval, $(0, x_1)$.

The probability that the component lifetime (X) will be in a range: $x_1 \leq X < x_2$ is given as

$$p_{12} = F_X(x_2) - F_X(x_1) \qquad (5.2)$$

This probability also implies that in a group of n components, $p_{12}$is a fraction of components which will fail on average during the time interval $(x_2 - x_1)$.

5.2.3 Failure Frequency

In a group of n components, the number of failures in a time interval, (x_1, x_2), $x_1 < x_2$, is denoted as $N(x_1, x_2)$. It is a random variable with an expected (or average) value of

$$E[N(x_1, x_2)] = n\big(F_X(x_2) - F_X(x_1)\big) = n\, p_{12} \qquad (5.3)$$

In the CE-PM report, failure frequency is calculated for a one year interval at age t, which is essentially the following quantity:

$$E[N(t, t+1)] = n(F_X(t+1) - F_X(t)) \qquad (5.4)$$

A more precise definition of the expected number of failures per unit time at an age t can be given as

$$\frac{d}{dt} E[N(t)] = \theta(t) = n \lim_{dt \to 0} \frac{\big(F_X(t + dt) - F_X(t)\big)}{dt} = n\, f_X(t) \qquad (5.5)$$

Thus, the failure frequency in a non-repairable problem is equivalent to the probability density scaled by the original number of components in the reactor core. In the long term, this failure frequency will approach to zero since $f_X(x) \to 0$ as time $x \to \infty$. In other words, over the long term operation of the reactor, the failure frequency is expected to diminish, which must not be confused with an increase in reliability.

It is important to note that neither the expected number of failures nor the failure frequency is a measure of reliability of a component in the time interval, (x_1, x_2). This issue is explained in the next Section.

5.2.4 Mission Reliability

The reliability of a component in the time interval (x_1, x_2) essentially means the fraction of components of age x_1 that would survive an operational interval $(x_2 - x_1)$. The mission reliability is technically defined as the following conditional probability:

$$R_X(x_1, x_2) = P[X > x_2 | X > x_1] = \frac{1 - F_X(x_2)}{1 - F_X(x_1)} = \frac{R_X(x_2)}{R_X(x_1)} \quad (5.6)$$

Thus, the mission reliability is a ratio of the reliability at the end of the mission date to the reliability at the starting date of the mission.

A complement of the mission reliability is the mission probability of failure given as

$$P_f(x_1, x_2) = 1 - R_X(x_1, x_2) = \frac{F_X(x_2) - F_X(x_1)}{1 - F_X(x_1)} = \frac{p_{12}}{R_X(x_1)} \quad (5.7)$$

It is very important to recognize that the mission probability of failure ($P_f(x_1, x_2)$) of a component in the operational interval of $(x_2 - x_1)$ in general is NOT equal to the fraction (p_{12}) of components that will have life time in the rage: $x_1 \leq X < x_2$. The difference, as shown by Equations (5.2) and (5.7), lies in the conditioning argument used in the calculation of the mission probability of failure.

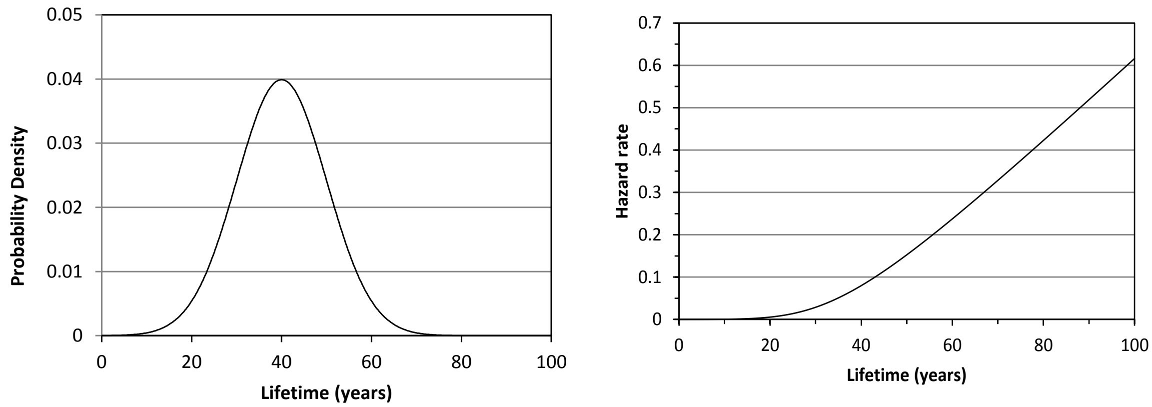
The CE-PM report uses p_{12} as a reliability measure, whereas $P_f(x_1, x_2)$ is a measure of reliability as per the mathematical theory of reliability.

Only in a special case of $F_X(x_1) \approx 0$, the mission probability of failure and probability of lifetime being in that interval are approximately equal in a numerical sense, i.e., $P_f(x_1, x_2) \approx$

p_{12}. Although this can happen in early life of a highly reliable component, this approximation can be quite erroneous in the mid to late life of the component.

5.2.5 Illustration

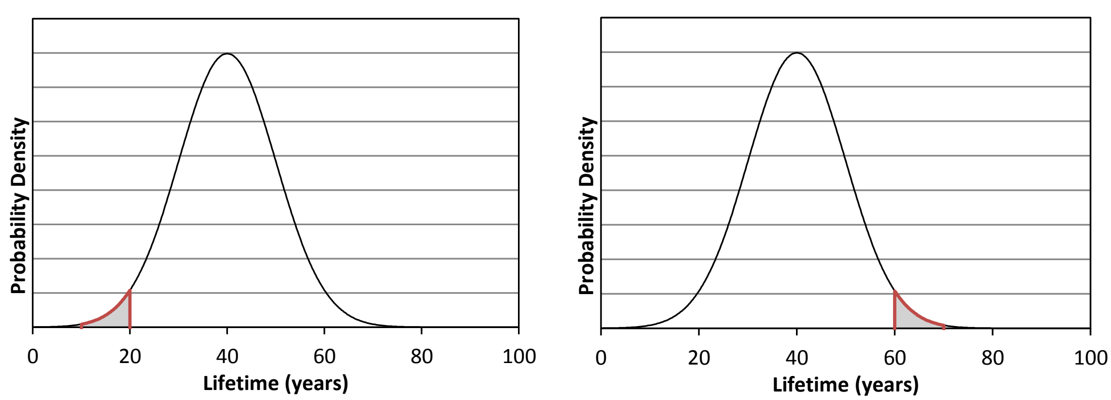
The distinction between the concepts related to failure frequency and failure probability can be further clarified by a simple example.



(a) (b)

Figure 5.2(a): Lifetime distribution, and (b) the hazard function: Non-repairable component

Consider that the lifetime distribution of a component has a mean life of 40 years and a COV of 0.25, and it follows some distribution, $f_X(x)$, as shown in Figure 5.2(a). The hazard rate of this distribution is also plotted in Figure 5.2(b).



(a) (b)

Figure 5.3: The fraction of components having lifetime (a) between 10 and 20 years, and (b) between 60 and 70 years.

The fraction of components with lifetime between 10 and 20 years is calculated as 0.021 (or 2.1%), which is the shaded area in Figure 5.3 (a). The fraction of components with lifetime between 60 and 70 years is also the same as 0.021, as shown in Figure 5.3 (b). Thus, the expected number of failures observed in the age interval 10 – 20 and 60 – 70 is the same, 2.1%.

Does it mean that the reliability of a 10 year and 60 year old component is also the same? The answer is clearly NO.

For this very reason, the expected number of failure and the failure frequency are not a correct risk metric for non-repairable equipment. The correct metrics is the probability of failure over the evaluation period, which in this case is 10 years. The probability of failure of a 10-year mission starting at age 10 is 0.02, which is calculated using Eq. (5.7). The failure probability increases to 0.94, if the mission were to start at age 60. The hazard rate at age 10 is 0.0004, and at age 60 it is 0.23 (see Figure 5.2 b), which also illustrates an increasing likelihood of failure with aging. This example shows that the failure probability correctly reflects an adverse effect of aging on reliability, whereas the failure frequency in a non-repairable problem is not indicative of reliability in an absolute sense.

So what is the use of the information about the failure frequency or the expected number of failures in a given time interval? This quantity is indicative of the average number of future replacements in a given time interval, which is useful in planning and budgeting of a maintenance or replacement project. For example, if we start with a fleet of 1000 components with lifetime distribution shown in Figure 5.2, on average 21 components will need to be replaced in the interval 10 – 20 and 60 – 70 year.

5.3 Reliability of a repairable component

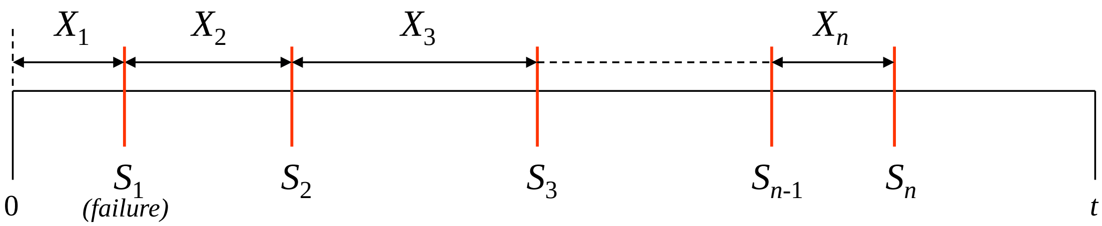


Figure 5.4: The service life of a repairable component with recurring failures

The concept of the failure rate or frequency, i.e., the expected number of failures per unit time, is meaningful as a reliability measure for repairable components only.

A portion of service life of a repairable component is shown in Figure 5.4. Here, S_1 is the occurrence of first failure, after which the failed component is replaced (almost immediately) by an identical new component. This model assumes that the component failure is immediately detected, such as a fuse blowing up in a household. If the repair time followed by a failure is quite long (and a random variable), then the resulting problem must be modelled as an "alternating renewal process". This model, due to added complexity, is beyond the scope of the current discussion.

In an operational time interval, $(0, t)$, failures occur randomly at times $S_2, S_3, \cdots$and the time between failures, $X_1, X_2, \cdots$, are component lifetimes, which are independent and identically distributed random variables with distribution $F_X(x)$.

In a technical sense, the theory of stochastic renewal process is needed to model the reliability of a repairable system [7].

In a repairable system, the attention is focussed on the number of failures, $N(t)$, which is also a random variable. Since the distribution of $N(t)$ is not easy to obtain, except in a few special cases, the reliability analysis focuses on the expected number of failures, $E[N(t)]$.

5.3.1 Failure Rate

The failure rate is defined as the expected number of failures per unit time in a repairable system. The failure rate, $\eta(t)$, at time *t* can be calculated from the following recursive, integral equation written in terms of the distribution of the time between failures (X) as

$$\eta(t) = f_X(t) + \int_0^t \eta(t-x) f_X(x) dx \tag{5.8}$$

Here $f_X(t)$ denotes the probability density of the lifetime (X) at time *t*.

If a failure is detected only by inspection and testing, like in stand-by safety systems, the renewal equation will have to be modified to account for the inspection interval and the time of detection.

The failure rate has a steady state value, which is asymptotically equal to the reciprocal of the mean time between failures (μ_X) [4]:

$$\eta(t) \to \frac{1}{\mu_X} \qquad \text{as } t \to \infty \tag{5.9}$$

It is interesting to note that in case of non-repairable components the failure rate approaches zero as the time approaches to infinity, as discussed below Eq. (5.5), whereas the same quantity approaches to a finite limit in case of a repairable system as shown by Eq. (5.9).

As seen from equations (5.5) and (5.8), the failure rate ($\eta(t)$) in a repairable system and that in a non-repairable system are different quantities, i.e., $\eta(t) \neq \theta(t)$.

5.3.2 Homogeneous Poisson Process (HPP) Model

The HPP model is the most widely used model of recurring events in the nuclear industry as well as other areas of engineering and science. In the HPP model, the time between failure (X) is an exponential random variable and the distribution of the number of failures, $N(t)$, follows the Poisson distribution given as

$$P[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{5.10}$$

The expected number of failures is given as $E[N(t)] = \lambda t$, and the failure rate is a constant, i.e.,

$$\eta(t) = \lambda \tag{5.11}$$

Since the exponential distribution cannot model the aging effect, the HPP model also lacks this ability. Note that the aging effect means the hazard rate is increasing with the age of the component. The HPP model basically describes random occurrences of failures without any time trend.

The reliability in a time interval $(0, t)$, i.e., probability of no failure, can be calculated from Eq. (5.10) with $k = 0$ as $e^{-\lambda t}$, and probability of failure as, $P_f(t) = 1 - e^{-\lambda t}$. In case of highly reliable equipment, λ tends to be so small that the probability of failure can be approximated as $P_f(t) \approx \lambda t$. For this reason, the failure rate can be used as a reliability measure of repairable equipment.

In PSA, the homogeneous Poisson process is used to model initiating events that can lead to core damage. Given the occurrence of an initiating event, the probability of core damage, p_d, is calculated from a suitable fault tree analysis. From the decomposition property of HPP, the core damage also becomes an HPP with rate $\lambda\, p_d$. Since core damage is a singular event which will mark the end of life of a reactor, there is no point in talking about recurring core damages. Nevertheless, the Poisson process model allows us to calculate the probability of (first) core damage as $e^{-\lambda t p_d}$, which is approximately, $P_{fd}(t) \approx \lambda p_d t$. Thus, the annual core damage frequency, λp_d with t=1, becomes synonymous with the annual core damage probability.

In practice such subtle distinctions between repairable and non-repairable problems are regrettably glossed over resulting in frequent uses of incorrect terminology. In case of the HPP model with $\lambda \ll 1$, this abuse of terminology can be forgiven. However, the misuse/abuse of terminology has now spread over to cases where the HPP model is inapplicable.

It appears that the use of failure frequency as a reliability metric in risk assessments of nuclear systems has crept over from the PSA literature where terms like core damage frequency and frequency of initiating events are routinely used.

5.3.3 An Example

In this Section, the key differences between repairable and non-repairable equipment reliability problems are further illustrated by a simple example.

Consider that the lifetime (X) of a component is modelled as a gamma distributed random variable with the shape and scale parameters of α = 2 and λ = 0.04, respectively. This particular distribution is chosen, since all relevant quantities can be derived in analytical forms. The mean lifetime is 50 years and the coefficient of variation (COV) of the lifetime is 0.70.

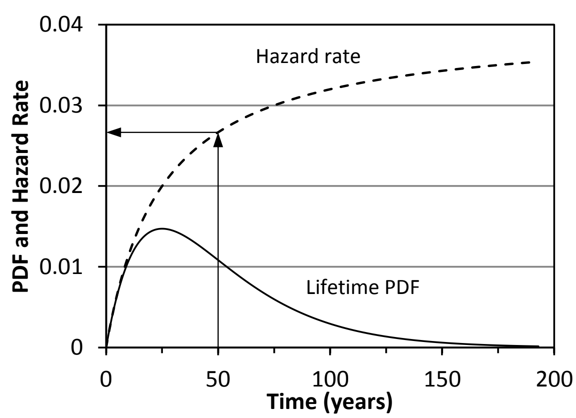
Other properties of this gamma lifetime distributions are given below:

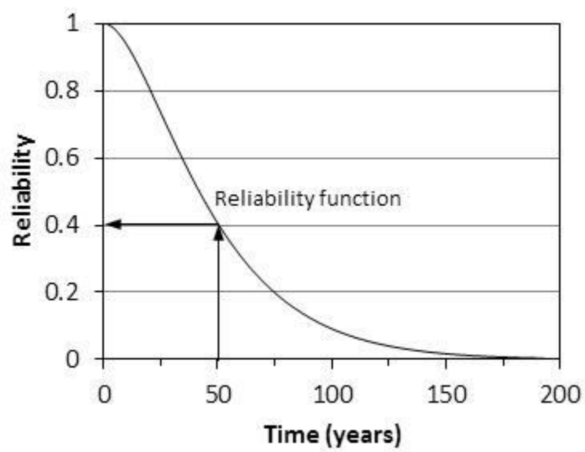
$$\text{Probability density function: } f_X(t) = \lambda^2 t\, e^{-\lambda t}$$

$$\text{Cumulative distribution function:} F_X(t) = 1 - (1 + \lambda t)e^{-\lambda t}$$

$$\text{Hazard rate: } h_X(t) = \frac{\lambda^2 t}{(1+\lambda t)} \tag{5.12}$$

$$\text{Reliability function: } R_X(t) = 1 - F_X(t) = (1 + \lambda t)e^{-\lambda t}$$





(a) (b)

Figure 5.5: (a) Lifetime distribution and the hazard rate, and (b) reliability function of the gamma distribution

First consider a non-repairable equipment reliability problem. The gamma lifetime distribution and hazard rate are plotted in Figure 5.5 (a). The hazard rate at age 50 is 0.026, and reliability at age 50 is 0.4 (Figure 5.5 (b)). It means that 60% of a fleet of such equipment is not expected to survive up to age of 50 years.

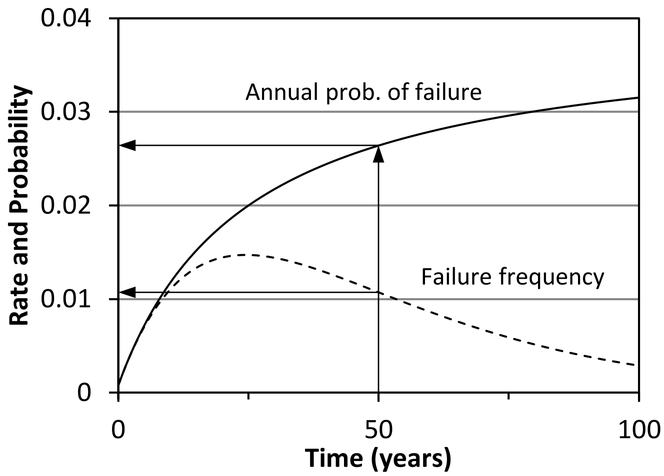


Figure 5.6: Difference between annual probability of failure and failure frequency: Non-repairable problem

In case of a single non-repairable component, the failure frequency is calculated using Eq. (5.4) with $n = 1$, whereas the annual probability of failure is calculated using Eq. (5.7) with $x_2 = x_1 + 1$. As shown by Figure 5.6, the annual probability of failure in an age interval 50 – 51 year is 0.026, which is more than twice the failure frequency (0.011) in this interval.

Although Figure 5.6 shows that both measures are approximately equal up to an age of 10 years, this numerical proximity cannot serve as a basis to justify the use of incorrect terminology and formulas in any reliability assessment.

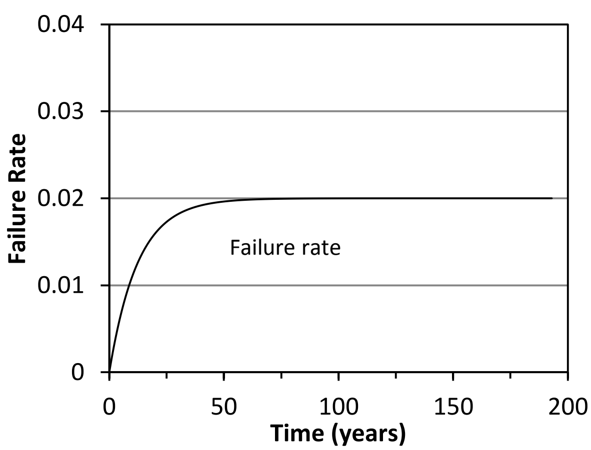


Figure 5.7: Failure rate with gamma distributed time between failures: Repairable problem

The failure rate for the gamma distributed time between failures can be analytically derived as [7]

$$\eta(t) = \frac{\lambda}{2}(1 - e^{-2\lambda t}) \tag{5.13}$$

The failure rate versus time relation plotted in Figure 5.7 is rather interesting. The failure rate first increases and reaches to a constant steady state value equal to $(\lambda/2\ )$ as a result of beneficial effect of replacement.

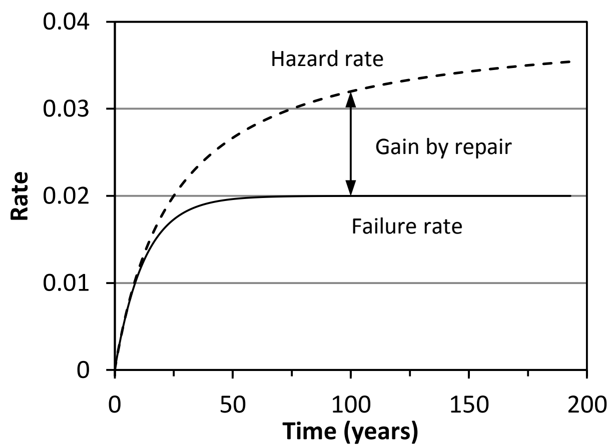


Figure 5.8: Comparison of failure rate of repairable equipment with the hazard rate of a non-repairable equipment

The comparison of failure rate and hazard rate shown in Figure 5.8 is really interesting, as it shows an important qualitative difference between these two quantities. The hazard rate increases over time indicating that the non-repairable equipment becomes increasingly vulnerable to failure. In contrast, the failure rate reaches a constant steady state value implying that in the long term the process of failures and renewals reaches to a stationary equilibrium.

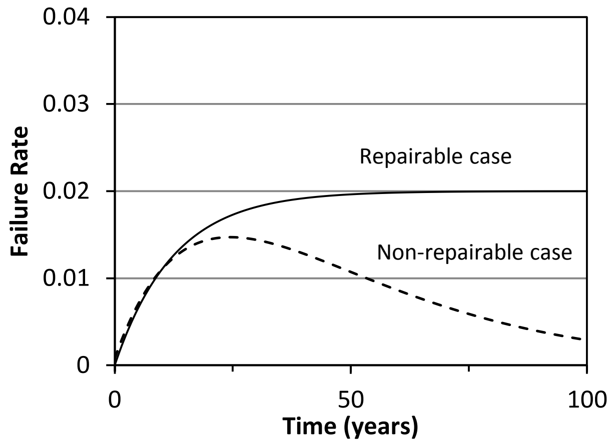


Figure 5.9: Comparison of the failure rate of repairable and non-repairable equipment: Example of the gamma distribution

Lastly, Figure 5.9 compares the failure calculated in repairable and non-repairable problems assuming the same gamma lifetime distribution. It is remarkable to see a completely different behaviour of the failure rate in these two cases. If the failure rate of a repairable system is calculated using the formula for non-repairable problem, and vice versa, it would lead to grossly incorrect results.

5.4 Reliability Analysis and Limit States of the Pressure Tube: A discussion

5.4.1 Background

As discussed previously, the failure frequency is an appropriate measure of reliability of repairable systems only. The mission (interval) probability of failure is a valid measure of reliability of non-repairable systems.

The CSA Standard [2] describes several limits states of the performance of the fuel channel, such as critical flaw sizes, minimum fracture toughness and PT-CT contact. However, the Standard does not provide any guidance about the classification of such limit states into repairable or non-repairable failures. The Standard also omits the relation and relevance of the chosen reliability measure, i.e., failure frequency, to the reliability with respect to postulated failure modes. Therefore, it is important to understand the nature of limit states of pressure tube performance, classify them correctly into repairable and/or non-repairable reliability problems and then quantify the applicable reliability measure.

It is also important to discuss the relation among types of limit states, types of failure modes, progression of limit states and maintainability in defining a reliability problem.

A *limit state* means that state (or condition) of the system which divides the system performance into acceptable and non-acceptable domains. For example, the blister formation threshold is a limit state of deuterium ingress at a point of PT-CT contact. Here, H_{eq} < BFT means acceptable performance, and H_{eq} > BFT means an unacceptable condition.

In structural reliability theory, limit states are broadly classified into two groups depending on the severity of the consequence of a failure. They are, namely, serviceability limit states and ultimate limit states. A serviceability limit state indicates a significant departure in the structure's condition from the original design state. Nevertheless, this departure from the design state has benign consequences, i.e., it does not compromise the system safety and functionality. The serviceability limit state is like an alarm which prompts the engineer to initiate mitigating actions. In contrast, the ultimate limit state means a failure that would severely impair the system safety and functionality with potentially severe consequences. For example, excessive deflection of a floor slab is classified as a serviceability limit state, whereas the formation of a plastic mechanism leading to collapse is an ultimate limit state of the floor slab. In summary, the classification of a failure mode into serviceability or ultimate limit state largely depends on the severity of the consequences of the system failure.

In some reactor components, serviceability and ultimate limit states might be a set of closely connected events, such as the *progression* of a serviceability limit state over time into an ultimate limit state. This situation must be carefully modeled in the reliability analysis. For example, in case of flow accelerated corrosion of a feeder bend, wall thickness loss up to a certain limit can be considered as a serviceability limit state. At the same time, the wall thinning can ultimately cause a feeder failure. Thus, the rate of progression from a serviceability to an ultimate limit state is a critical issue in the reliability analysis.

In the context of reliability analysis, *failure modes* are of two types: (1) self-announced failure, and (2) latent failure. The self-announced failure means that the occurrence of a failure is (almost) immediately detected by the operator/user of the equipment. For example, the loss of electric supply to a household is noticed by the occupants. A plane falling from the sky is another extreme example of this type of failure mode.

The latent failure, as the name implies, means the occurrence of a system failure is not detectable until a detailed inspection is carried out. In case of a safety system, the occurrence of a failure can be latent due to the standby nature of the operation. Some degradation failures may belong to this category, since entering into a degradation state does not interrupt the function of a component to a degree that it becomes immediately noticeable. For example, the exact time at which the engine oil in a car reaches to a degraded state is not known. Therefore, the driver is advised to inspect the condition of oil periodically.

The maintainability refers to the degree to which a system is amenable to repair (or replace) after a failure, such that its operation can be restored to a safe and functional state. The maintainability is typically achieved by design, i.e., by building those features in the system layout that provide access to the system and allow modular repair work in a safe and efficient manner. It should be stressed that maintainability by itself does not determine the type of system reliability analysis to be repairable or non-repairable. In the repairable system reliability analysis, it is assumed that an opportunity to repair exists right after a failure. For example, an aircraft engine is designed to have high maintainability. A failure to start the engine on the ground is repairable, whereas a run-time engine failure in the air is non-repairable.

5.4.2 Repairable versus Non-Repairable Reliability Problems

How do we know that a limit state is repairable or non-repairable? In practice, this classification is not a cut -and –dry situation, rather it depends on several factors discussed in the previous Section. Based on this discussion, some general guiding principles for the classification can be constructed as follows.

A prime consideration in the reliability analysis is the consequence of failure, i.e., risk, which is implicitly embodied in classification of limit states. The type of limit state is therefore a dominating consideration.

A self-announced, serviceability limit state can be modelled as the repairable problem provided that the system is maintainable.

It is most appropriate to model an ultimate limit state as a non-repairable reliability problem.

A latent failure mode is typically in the realm of a non-repairable reliability problem. However, this problem can be converted into the repairable type by introducing inspections. An inspection program can be designed to detect and remove latent failures. In this case however, the system availability and failure frequency become a function of the inspection interval. Analysis of these parameters requires more complex stochastic process models. Since the equipment condition is not known without an inspection, the probability of failure over the inspection interval is a meaningful reliability measure, at least for safety critical systems. A sufficiently low probability of failure would ensure the operator and the regulator that an unreasonable number of failures will not be discovered at the time of next inspection.

A latent, serviceability limit state of static nature can be modelled as a repairable problem. However, if this limit state can rapidly evolve into an ultimate limit state (self-announced or latent), then it is prudent to model this as a non-repairable problem.

In conclusion, the reader must be reminded that the classification of a reliability problem is very much dependent on the operational considerations. A failure to start an aircraft engine on the ground is repairable, whereas a run-time engine failure in the air is a non-repairable problem. Therefore, a probabilistic assessment must identify the class of reliability problems it is intended to address.

5.4.3 Limit States of Pressure Tubes

The condition of a fuel channel is not directly visible to the operator. Only the inspection and gauging tools provide data about the state of deformation of the fuel channel. For example, pressure tube elongation, diametral creep and a contact with the calandria tube are detected by inspection tooling. There are different types of degradation related limit states of PT failure, as given in the CSA Standard [2].

In the context of the CE-PM report, the limit state of PT-CT contact leading to blister formation must be clearly identified. Is it a serviceability limit state or an ultimate limit state? The classification would depend upon the consequence of failure. Blister formation by itself may be a benign event. However, if the blister growth rapidly leads into the pressure tube cracking and rupture, then it should be considered as an ultimate limit state.

The limit state of blister formation is certainly a latent failure mode, since its occurrence can only be detected by an inspection. Is it a repairable or non-repairable failure mode? In principle, a contact can be mitigated by the SLAR tooling after the detection of a contact. In this sense, the PT-CT contact is a repairable failure. However, since the failure mode is latent and only detectable during an inspection, the PT—CT contact cannot entirely be treated as a repairable problem. Another issue is that the likelihood of spacer movements in post-SLAR life cannot be completely eliminated. Thus, if this mode were to be treated as a repairable problem, the calculation of the failure rate must account for the frequency and scope of the fuel channel inspection program. Otherwise, it is prudent to evaluate the probability of PT-CT contact and blister formation over the inspection interval of the pressure tube.

The frequency and scope of fuel channel inspections play an important role in classifying the reliability problem in the following manner. A relation between the inspection interval and the rate at which a blister can grow to the size that leads into cracking and rupture (or leakage) is a deciding factor.

Suppose ALL fuel channels in a reactor are inspected at a 2 year interval. A knowledge base exists to allow us to say with high confidence that a blister is unlikely to cause a PT rupture within two years of its formation. Then, this mode can be treated as a repairable problem. The failure frequency calculated considering 2 year inspection interval would be a suitable measure of reliability.

If the current knowledge base suggests that the blister formation is likely to cause a PT rupture (or loss of integrity) in less than 2 years, then this problem should certainly be treated as a non-repairable reliability problem. In this case, the probability of failure in a 2 year interval is a suitable measure of reliability.

Another consideration is the sample size for pressure tube inspections. If the sample size is much smaller than the population size, then some tubes may not be inspected for a long time. In such uninspected tubes, a PT—CT contact and blisters can remain undetected for a long time. For such channels, an annualized measure of reliability (i.e., failure frequency) is rather meaningless.

In summary, the reliability measure calculated in a probabilistic assessment must be commensurate with the type of limit state (serviceability or ultimate), nature of failure mode (self-announced or latent), rate of progression of limit states and the maintainability of the system. The failure rate and failure probability serve as a meaningful reliability measure depending on the context of the reliability analysis.

5.5 Closing Remarks

The CE-PM report describes a simulation based method to predict the expected number of PT failures in one reactor year intervals (8000 EFPH assumed as one reactor year) over the operating life of the reactor. The predicted failure frequency is then compared with the acceptance limit specified in the Table C.1 of the CSA Standard N285.8 [2]. Since this analysis lies in the field of time dependent reliability analysis, the analysis method and the reliability measure chosen in CE-PM must be consistent with basic principles of the reliability analysis. This aspect is thoroughly examined in this Chapter of the review, and reliability concepts related to non-repairable and repairable equipment are explained in detail.

The failure frequency defined as the expected number of failures per unit time is not a universal measure of reliability. It is a meaningful measure in the reliability analysis of repairable systems. Using the failure frequency as a reliability measure of a non-repairable system is conceptually inaccurate.

The CE-PM report estimates the PT lifetime distribution, i.e., the distribution of the time of blister formation (and blister cracking) for each pressure tube in the reactor core. Since CE-PM stops after the "first" failure of a PT, and it does not consider any repair or replacement followed by a failure in the simulation, it is concluded that PT failure has NOT been analyzed as a repairable reliability problem. Since CE-PM has implicitly analyzed a non-repairable reliability problem, the probability of failure over a suitable evaluation period is a meaningful reliability measure. This measure must be correctly calculated and reported in the probabilistic assessment.

For the sake of conceptual and procedural accuracy of the probabilistic assessment, it is recommended that the CE-PM report include a section that describes the nature of PT limit states evaluated in the assessment and accordingly justify the chosen reliability measure.

5.6 Points of Clarifications

5.1 The CE-PM report should replace all colloquial terms by correct technical terms. For example, integrated probability should be replaced by the cumulative probability. The aggregate frequency is in essence the system failure frequency.

5.2 Based on the discussion presented in this Chapter (Sections 5.2 and 5.3), it is recommended that the CE-PM authors clearly state the type of limit state, the type of failure mode (latent or self-announced), rate of progression of the failure mode, and maintainability of the pressure tube with respect to the failure modes considered in the probabilistic assessment. Then, the CE-PM report must state whether they consider the problem as a repairable or non-repairable reliability analysis with due explanation.

5.3 The PT failure problem (i.e., blister formation/cracking) in the CE-PM report is implicitly analyzed as a non-repairable equipment reliability problem. This conclusion is based on the following facts: (1) the failure frequency calculated in the CE-PM report uses the formula that is applicable to a non-repairable problem, as shown in Section 5.2. (2) The CE-PM simulation method also stops after the first failure of a PT, and it does not consider any repair or replacement followed by a failure in the simulation, as required in a repairable problem (see Section 5.3). (3) The failure frequency for a repairable system is calculated using an integral equation or using analytical formulae in some special cases (e.g., the HPP model or gamma process). This approach is not used in the CE-PM report. Thus, it is concluded that the CE-PM report does not analyze the blister formation as a repairable reliability analysis problem. Therefore, it is recommended that CE-PM calculate and report the probability of PT failure over the inspection interval.

5.4 In Section 7, paragraph 5, CE-PM provides an incorrect response to a concern raised about the calculated failure frequency. If the PT failure problem belongs to a non-repairable equipment category, the expected number of failures in the reactor would diminish over time, since the PT population would shrink due to failures accumulating over time. It is not a question of "engineering approximation", rather it is a reality of the problem. For this very reason, failure frequency adopted by CE-PM is not indicative of reliability, as discussed throughout this Section. The meaning of the statement that "*such a case is of no engineering interest since it is a long way beyond the acceptance criterion"* is debatable. The reason is that the failure frequency and failure probability begin to differ way before the end of life, as shown by examples presented in this Chapter.

6. Other Issues and Points of Clarification

6.1 Introduction

This chapter summarizes the comments and points of clarification related to remaining topics, such as the simulation sample size, convergence, sensitivity analysis and spacer movement models. Since these points are fairly concise and specific to the CE-PM report, they are discussed in a point-wise form in the following section.

6.2 Points of Clarification

6.1 The CE analyst should provide a more quantitative and statistical rationale for justifying the number of simulations used in the analysis, which is typically 1000 simulations.

The number of simulations required in the Monte Carlo method depends on the desired relative accuracy of the estimate as well as the order of the probability that is being estimated.

Suppose in the estimation of a probability, p, a relative error of v_x is considered acceptable, then the number of simulations required, N_{sim}, can be estimated as

$$N_{sim} \approx \frac{1}{v_X^2 p} \qquad (6.1)$$

If a 10% relative error is tolerable, i.e., $v_X = 0.1$, then the required number of simulations would be roughly $100/p$. The CE analyst should state the tolerable statistical error (v_x) in the analysis and accordingly show that the chosen number of simulations are adequate in the estimation.

6.2 In order to estimate the probability of an order of 10^{-3} with 10% relative accuracy, approximately 10^5 simulations are required. Since an interest in the probabilistic assessment is in finding the time at which the CSA acceptance criterion is violated, which is in the order of 4.13×10^{-3}, the simulation sample size must be adequate to estimate this probability.

It is therefore puzzling to read the following statement on page 61: “1000 simulations are justified on the basis that at risk channels have frequency of the order of 10^{-1}. If a channel already has failure frequency of the order of 10^{-1}, it means the reactor is way out in the unacceptable domain, a situation that is not of interest in the analysis. What is of interest is the time at which the reactor violates an acceptance criterion, so that inspection and maintenance can be planned. The CE analyst is requested to address this issue.

6.3 On page 62 in Appendix B, a statement like “1000 simulations are more conservative is the error bounds are taken into account……”, and “….since the lower bound is used to calculate the aggregate frequency, the assessment is more conservative", are difficult to understand. Such statements do not have a quantitative basis, and terms like the lower bound is not defined explicitly.

The CE analysts is requested to define the meaning of the lower bound and provide mathematical formulae used in the calculation.

6.4 A statement before Table 7 (page 2) that "Using 1000 simulations leads to the error bounds … 2000 EFPH wider (2σ) ..." is not clear. Let's examine this in a bit more detail. The mean and standard deviation of the time to contact (T) for P08 is estimated as μ_T = 338,170 EFPH and σ_T = 41,229 EFPH. The standard deviation of the estimated mean is given as $\sigma_T/\sqrt{N_{sim}}$. For $N_{sim} = 1000$, this turns out to be 1304 hours. In essence, the CE-PM report is providing error bounds on the average time to contact (μ_T), which has a little value in the reliability analysis. The prediction limits on the time to contact (T) are meaningful in the reliability analysis. Therefore, the CE-PM analyst is requested to plot the mean and prediction limits of the time to PT-CT contact.

6.5 According to Table 7 of the CE-PM, the lognormal distribution was fitted to the time to contact data. It also makes a reference to "censored" data. Which statistical estimation method was used to estimate the parameters of the lognormal distribution?

6.6 The lognormal distribution is not considered ideal for modeling the lifetime distribution. The reason is that its hazard rate first increases and then decreases with time. Since the probability of PT-CT contact should increase with time in a monotonic manner, the lognormal distribution may not be an ideal choice. Did the CE analyst try to fit any other alternate distributions, such as the Weibull distribution?

6.7 In Figure 18, what is the definition of 5% CL and how is it calculated? Since the probability of contact changes with time, the 5% CL should also reflect this change. Isn't it? Please provide mathematical formulae used to calculate this quantity.

6.8 Appendix E presents the estimation of the spacer movement model based on Gentilly-2 data. Figure 21 shows the Weibull cumulative distribution of the absolute ratio. This should be plotted on the Weibull probability paper.

6.9 Appendix F describes the Bruce spacer movement model. On page 76, it is stated that the Weibull distribution is suitable to model asymmetric data. This is not fully accurate, since the lognormal distribution is the most appropriate to model long tail data.

6.10 With reference to Figure 23, the data and the Weibull distribution must be plotted on the Weibull probability paper in order to show the goodness of fit. Plotting data and cumulative distribution on an arithmetic scale is rather meaningless.

6.11 Figure 23 (page 77) shows a discontinuity in the cumulative distribution of the displacement ratio. Does it not mean that the positive and negative spacer movement belong to separate statistical population? Isn't it more appropriate to fit separate distributions to the positive and negative movements?

6.12 In Figure 24, a comparison of cumulative distribution of simulated and observed movements is somewhat meaningless, as it does not show the goodness of fit in the tail regions. It is suggested to make this plot on a log scale. This comment also applies to Figure 25.

6.13 Appendix G illustrates the importance of different input parameters on PT-CT contact assessment. With reference to Figure 27, please provide definitions of the lower and upper bounds. Figure 27 shows the effect of variation of the creep ratio on the PT-CT gap. What is the effect of creep ratio on the time of PT-CT contact? This comment also applies to Figure 28.

6.14 The sensitivity analysis presented in Appendix G is merely a qualitative analysis of an informal nature. In a formal sense, the purpose of the sensitivity analysis is to show the effect of an input parameter, a random variable (e.g., creep ratio), on the variability of the output.

A formal method used in the probabilistic analysis is called the global sensitivity analysis [6], which quantifies the percentage contribution of a random variable to the variance of the output. For example, it is more meaningful to know the contribution of the creep ratio of the variance of PT-CT gap or the time of PT-CT contact. This analysis would provide a rationale for collecting additional inspection data for accurate modeling of variables with high sensitivity to the result.

The CE analysis is suggested to consider such a formal sensitivity analysis.

6.15 An important aspect of the risk assessment is also to understand the extent of variability in the system under investigation. In this particular case, the PT-CT contact is the main initiator of an undesirable event, namely, blister formation. So what is the range of mean time to contact across the fuel channel population of a reactor core? Similarly, what is the range of standard deviation of the time contact in the reactor core? This information may be useful to show how wide spread is the effect of the degradation mechanism in the reactor core. Since the simulation results are already available for all the fuel channels, this information can be easily extracted and displayed.

7. References

[1] Candu Energy, "Technical Basis for Probabilistic Assessment of Pressure Tube and Calandria Tube Contact," 2014.

[2] CSA N285.8-10, "Technical requirements for in-service evaluation of zirconium alloy pressure tubes in CANDU reactors," Canadian Standards Association, Toronto, 2010.

[3] E. Nadeau, "Derivation of input distributions for probabilistic evaluation using in-reactor measurements.," in *Proc. ASME 2012 PVP conference*, Toronto, ON, July 15-19, 2012.

[4] S. Rigdon and A. Basu, Statistical Methods for the Reliability of Repairable Systems, NY: John Wiley & Sons Inc., 2000.

[5] S. Conti, J. Gosling, J. Oakley and A. O'Hagan, "Gaussian process emulation of dynamic computer codes.," *Biometrika*, vol. 96, pp. 663-676, 2009.

[6] X. Zhang and M. Pandey, "An effective approximation for variance-based global sensitivity analysis," *Reliability Engineering and System Safety*, vol. 121, no. 1, pp. 164-174, 2014.

[7] E. Parzzen, Stochastic Processes, San Francisco: Holden-Day, 1962.