

Electronic Thesis and Dissertation Repository

8-27-2014 12:00 AM

Probabilistic Modeling and Bayesian Inference of Metal-Loss Corrosion with Application in Reliability Analysis for Energy Pipelines

Hao Qin
The University of Western Ontario

Supervisor
Wenxing Zhou
The University of Western Ontario

Graduate Program in Civil and Environmental Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science
© Hao Qin 2014

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Civil Engineering Commons](#), and the [Structural Engineering Commons](#)

Recommended Citation

Qin, Hao, "Probabilistic Modeling and Bayesian Inference of Metal-Loss Corrosion with Application in Reliability Analysis for Energy Pipelines" (2014). *Electronic Thesis and Dissertation Repository*. 2246. <https://ir.lib.uwo.ca/etd/2246>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

**PROBABILISTIC MODELING AND BAYESIAN INFERENCE OF METAL-
LOSS CORROSION WITH APPLICATION IN RELIABILITY ANALYSIS FOR
ENERGY PIPELINES**

(Thesis format: Integrated Article)

by

Hao Qin

Graduate Program in Engineering Science
Department of Civil and Environmental Engineering

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Engineering Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Hao Qin 2014

ABSTRACT

The stochastic process-based models are developed to characterize the generation and growth of metal-loss corrosion defects on oil and gas steel pipelines. The generation of corrosion defects over time is characterized by the non-homogenous Poisson process, and the growth of depths of individual defects is modeled by the non-homogenous gamma process (NHGP). The defect generation and growth models are formulated in a hierarchical Bayesian framework, whereby the parameters of the models are evaluated from the in-line inspection (ILI) data through the Bayesian updating by accounting for the probability of detection (POD) and measurement errors associated with the ILI data. The Markov chain Monte Carlo (MCMC) simulation in conjunction with the data augmentation (DA) technique is employed to carry out the Bayesian updating. Numerical examples that involve both the simulated and actual ILI data are used to validate the proposed Bayesian formulation and illustrate the application of the methodology.

A simple Monte Carlo simulation-based methodology is further developed to evaluate the time-dependent system reliability of corroding pipelines in terms of three distinctive failure modes, namely small leak, large leak and rupture, by incorporating the corrosion models evaluated from the Bayesian updating methodology. An example that involves three sets of ILI data for a pipe joint in a natural gas pipeline located in Alberta is used to illustrate the proposed methodology. The results of the reliability analysis indicate that ignoring generation of new defects in the reliability analysis leads to underestimations of the probabilities of small leak, large leak and rupture. The generation of new defects has the largest impact on the probability of small leak.

Keywords

Pipeline; Metal-loss corrosion; Stochastic process; Hierarchical Bayesian; Measurement error; Probability of detection; Missing data; Markov chain Monte Carlo; Data augmentation; System reliability

DEDICATION

To my parents

ACKNOWLEDGMENTS

First of all, I would express all my gratitude and appreciation to my supervisor Dr. Wenxing Zhou, for his continuous support and patient guidance throughout this study. His expertise and academic attitude always help and encourage me to overcome the difficulties faced in the fulfillment of this thesis. It has been my honour to pursue a Master's degree under his supervision.

I would also like to thank Dr. Hanping Hong for his kind help and constructive advice throughout my study at Western University. I am grateful to my committee members - Dr. Michael Bartlett, Dr. Hanping Hong and Dr. Takashi Kuboki for being my examiners and giving sound advices to this thesis. The research reported in this thesis was jointly supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and TransCanada Pipelines Limited through the Collaborative Research and Development Program. These financial supports are very much appreciated. The financial support received from Western University and Dr. Wenxing Zhou is greatly appreciated.

I would like to thank my colleagues of our research group for their friendship and assistance, especially for Dr. Shenwei Zhang's effective help in this study. Lastly, I am truly grateful to my family for their endless love and support.

TABLE OF CONTENTS

ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS.....	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS AND SYMBOLS	xi
Chapter 1 Introduction	1
1.1 Background.....	1
1.2 Bayesian Methodology	3
1.2.1 Overview.....	3
1.2.2 Bayesian inference.....	3
1.2.3 Markov chain Monte Carlo simulation	4
1.3 Objective and Research Significance.....	5
1.4 Scope of the Study	5
1.5 Thesis Format.....	6
References.....	6
Chapter 2 Probabilistic Modeling and Bayesian Inference of Generation and Growth of Corrosion Defects on Pipelines Based on Imperfect Inspection Data	8
2.1 Introduction.....	8
2.2 Uncertainties in the ILI Tool.....	10
2.2.1 Measurement error	10
2.2.2 Probability of detection and probability of false call.....	11
2.3 Probabilistic Models for Defect Generation and Growth	12

2.3.1 Defect generation	12
2.3.2 Defect growth.....	14
2.4 Bayesian Updating of Defect Generation and Growth Models	16
2.4.1 Overview	16
2.4.2 Likelihood functions	16
2.4.2.1 Likelihood function for ILI-reported depths.....	16
2.4.2.2 Likelihood function for the number of detected defects.....	17
2.4.3 Prior distributions.....	17
2.4.4 Posterior distributions, MCMC simulation and missing data	18
2.5 Illustrative Examples	20
2.5.1 Example 1	20
2.5.2 Example 2	29
2.6 Summary and Conclusions	39
References.....	40
 Chapter 3 Reliability Analysis of Corroding Pipelines Considering the Generation and Growth of Corrosion Defects.....	 44
3.1 Introduction.....	44
3.2 Corrosion Generation and Growth Models.....	45
3.2.1 Defect generation and growth modeling.....	45
3.2.2 Bayesian updating of the defect generation and growth models	47
3.3 Time-dependent System Reliability Analysis.....	48
3.3.1 Limit state functions	48
3.3.2 Burst and rupture pressure capacity models	49
3.3.3 Basic assumptions and analysis procedures.....	50
3.4 Example	53
3.4.1 General information.....	53
3.4.2 Results.....	54

3.4.2 Sensitivity analysis.....	55
3.5 Summary and Conclusions	60
References.....	61
Chapter 4 Conclusions and Recommendations for Future Study	64
4.1 Probabilistic Modeling and Bayesian Inference of Metal-loss Corrosion	64
4.2 Time-dependent System Reliability Analysis of Corroding Pipelines	65
4.3 Recommendations for Future Study	66
References.....	66
Appendix 2A Derivations of Full Conditional Posterior Distributions of Model Parameters.....	68
Appendix 2B Procedures for Generating Defect Initiation Times from a Non- homogenous Poisson Process	71
References.....	71
Appendix 2C Procedures for Simulating Corrosion Data with the Simplifying Assumption for Defect Generation	72
Appendix 2D Procedures for Simulating Corrosion Data without the Simplifying Assumption for Defect Generation	74
CURRICULUM VITAE.....	76

LIST OF TABLES

Table 2.1 Summary of the simulated inspection data	21
Table 2.2 Posterior statistics of model parameters for Example 1	22
Table 2.3 Summary of the simulated inspection data corresponding three different POD curves	26
Table 2.4 Summary of the ILI-reported defect information	30
Table 2.5 Posterior statistics of model parameters for Example 2	31
Table 3.1 Probabilistic characteristics of the basic random variables	54

LIST OF FIGURES

Figure 1.1 Typical external metal-loss corrosions on buried steel pipelines	2
Figure 1.2 An ILI tool being retrieved from the pipeline after the inspection.....	2
Figure 2.1 POD curves with different sets of values of q and x_{th}	12
Figure 2.2 Mean, 2.5-percentile, 97.5-percentile and a given realization of NHPP.....	13
Figure 2.3 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II	24
Figure 2.4 Comparison of the predicted and actual depths at year 20	25
Figure 2.5 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II based on more realistic corrosion inspection data	28
Figure 2.6 Comparison of the two assumed POD curves	30
Figure 2.7 Predicted number of defects as a function of time corresponding to the high and relatively low detectability assumptions	32
Figure 2.8 Predicted growth paths of selected defects corresponding to the high and relatively low detectability assumptions.....	35
Figure 2.9 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II for the relatively low detectability assumption.....	36
Figure 2.10 Comparison of the growth paths corresponding to the base case and Scenario I and II for the relatively low detectability assumption	39
Figure 3.1 Cumulative probabilities of small leak, large leak and rupture.....	55
Figure 3.2 Impact of generation of new defects on the system reliability corresponding to the base case, Scenario I and II.....	58

Figure 3.3 Comparison of probabilities of small leak and burst corresponding to the base case, Scenario I and II..... 60

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations

COV	coefficient of variation
DA	data augmentation
HPP	homogenous Poisson process
iid	independent and identically distributed
ILI	in-line inspection
MCMC	Markov chain Monte Carlo
MFL	Magnetic Flux Leakage
M-H	Metropolis-Hastings
NHGP	non-homogenous gamma process
NHPP	non-homogenous Poisson process
PDF	probability density function
PMF	probability mass function
POD	probability of detection
POFC	probability of false call
SMYS	specified minimum yield strength

Symbols

\mathbf{a}	vector of the constant biases associated with ILI tools
a_i	constant bias associated with the ILI tool in the i^{th} inspection
\mathbf{b}	vector of the non-constant biases associated with ILI tools
b_j	non-constant bias associated with the ILI tool in the i^{th} inspection
ε_{ij}	random scattering error associated with the ILI-reported depth of the j^{th} defect at the i^{th} inspection
ρ_{il}	correlation coefficient of the random scattering errors associated with the i^{th} and l^{th} inspection

Σ_{E_j}	covariance matrix of the random scattering errors associated with the j^{th} defect
y_{ij}	the measured depth of the j^{th} defect at the i^{th} inspection
q	constant that characterizes the inherent detecting capability of the ILI tool
x_{th}	the detection threshold
w_t	wall thickness
t	time elapsed since the installation of pipeline
t_i	elapsed time from the installation data up to the i^{th} inspection
$f_P(\cdot)$	probability density function of the Poisson distribution
$f_G(\cdot)$	probability density function of the gamma distribution
$f_N(\cdot)$	probability density function of the normal distribution
$f_{IG}(\cdot)$	probability density function of the inverse gamma distribution
$m(t)$	expected number of defects generated over the time interval $(0, t]$
m_i	expected number of defects generated over the time interval $(t_{i-1}, t_i]$
$N(t)$	total number of defects generated within a time interval $(0, t]$
N_i	total number of corrosion defects in the i^{th} inspection
N_i^o	number of defects that have initiated prior to the $(i-1)^{\text{th}}$ inspection
N_i^s	number of defects that initiated between the $(i-1)^{\text{th}}$ and i^{th} inspections
N_i^{sd}	number of defects that initiated between the $(i-1)^{\text{th}}$ and i^{th} inspections and have been detected by inspection tools
N_i^{su}	number of defects that initiated between the $(i-1)^{\text{th}}$ and i^{th} inspections and have not been detected by inspection tools
$\overline{\text{POD}}_i$	average POD with respect to the N_i^s defects
$X(t)$	depth of a given defect at time t
$I_{(0,\infty)}(\cdot)$	indicator function
$\alpha(t)$	time-dependent shape parameter of the gamma process
β	time-independent scale parameter of the gamma process
$\Gamma(\cdot)$	gamma function
t_{sr}	initiation time of defect r
ξ_r	random effect of defect r

x_{ir}	actual depth of the r^{th} defect at the time of the i^{th} inspection
Δx_{ir}	the growth of actual depth of the r^{th} defect between the $(i-1)^{\text{th}}$ and i^{th} inspection
Δa_{ir}	shape parameter of the gamma distribution of Δx_{ir}
$L(\cdot)$	likelihood function
\mathbf{x}_j	vector of actual depth of defect j associated with all inspections
\mathbf{y}_j	vector of measured depth of defect j associated with all inspections
\mathbf{N}^{gd}	vector of the number of newly detected defects
$\Delta \mathbf{x}_j$	vector of depth increment of defect j associated with all inspections
$\Delta \mathbf{x}$	vector of depth increment associated with all defects and inspections
t_n	the most recent inspection of the pipeline
T	total forecasting period
g_1	limit state function for the corrosion defect penetrating the pipe wall
g_2	limit state function for plastic collapse under internal pressure
g_3	limit state function for the unstable axial extension of the through-wall
p	internal pressure of the pipeline
r_b	burst pressure capacity of the pipe at a given defect
r_{rp}	rupture pressure capacity of the pipe at a given defect
D	outside diameter of the pipeline
L_j	length of corrosion defect j
M	Folias factor or bulging factor
$P_{sl}(\tau)$	probability of small leak
$P_{ll}(\tau)$	probability of large leak
$P_{rp}(\tau)$	probability of rupture
σ_y	yield strength of the pipe material
σ_f	flow strength of the pipe material
ζ_b	model error associated with the burst capacity model

Chapter 1 Introduction

1.1 Background

Pipelines are widely recognized as the safest and most effective means to transport large quantities of hydrocarbons over long distances. According to the Canadian Energy Pipeline Association, there are approximately 115,000 km of natural gas and liquids transmission pipelines in Canada; Canada exported approximately \$83.5 billion worth of crude oil and natural gas in 2012, most of which was transported by pipelines, and the Canadian pipeline operators spent about \$1.1 billion in 2012 to monitor and maintain the vast pipeline network across Canada.

Metal-loss corrosion is a common threat to the structural integrity of steel pipelines. Figure 1.1 shows typical external metal-loss corruptions on buried steel pipelines. The periodical inspection of pipelines using high-resolution in-line inspection (ILI) tools is widely employed in the pipeline industry and a key component of the pipeline corrosion management practice. Figure 1.2 shows an ILI tool that has just completed the inspection of a pipeline and is being retrieved at the receiving end. The data obtained from an ILI on a given pipeline include the locations and sizes of corrosion features (i.e. defects) on the pipeline, which provide a snapshot of the condition of corrosion, whereas the data obtained from multiple ILI carried out at different times on the same pipeline allow one to infer the progress of the corrosion condition over time. The main focus of the study reported in this thesis is to develop methodologies to make inference of the state and progress of corrosion on a given pipeline based on multiple sets of ILI data.

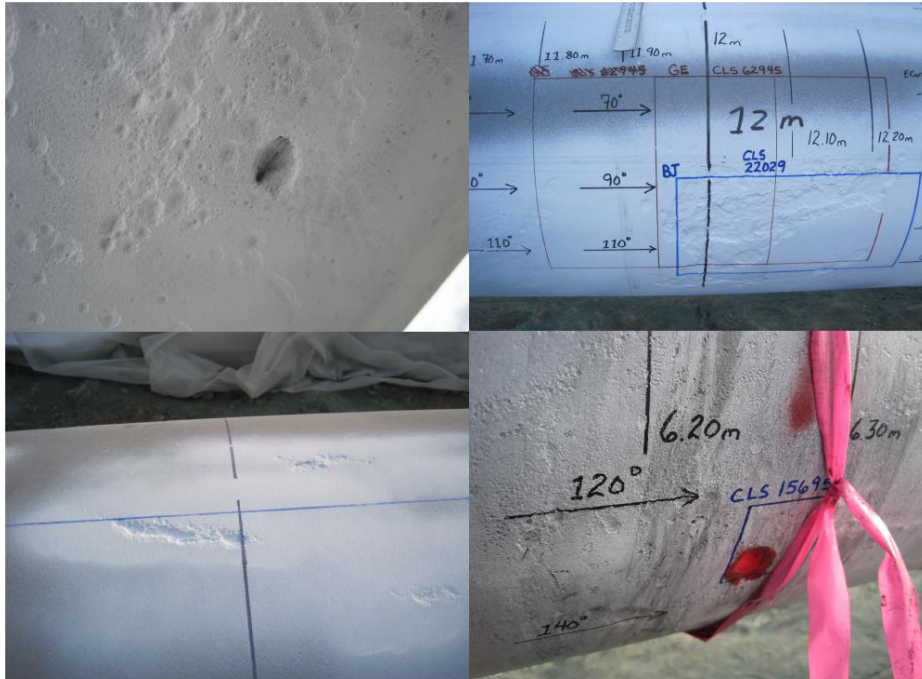


Figure 1.1 Typical external metal-loss corruptions on buried steel pipelines



Figure 1.2 An ILI tool being retrieved from the pipeline after the inspection

The corrosion process, which involves the generation of new defects and growth of existing defects over time, is by nature highly uncertain, and the ILI data are imperfect due to the limited detectability and measurement errors associated with the ILI tools. In light of this, the Bayesian methodology was selected as the main vehicle to achieve the objective of the study. The Bayesian methodology has been widely used to carry out the condition assessment of aging structures and infrastructures (e.g. Zheng and Ellingwood 1998; Enright and Frangopol 1999; Kuniewski et al. 2009; Zhang and Zhou 2013). It provides an ideal framework to combine existing knowledge and/or experience about the condition of a structure with the new information contained in the inspection data to develop the updated knowledge of the condition of the structure. The methodology can also deal with uncertainties from different sources in a straightforward manner. A brief description of the Bayesian methodology as well as the computational techniques involved in the methodology is presented in the next section.

1.2 Bayesian Methodology

1.2.1 Overview

The essential characteristic of the Bayesian methodology is its explicit use of probability for quantifying uncertainties in inferences based on statistical data analysis (Gelman et al. 2004). The application of the Bayesian methodology can be divided into the following three steps:

1. Set a full probability model, namely a joint probability distribution for all observable and unobservable quantities in a scientific problem.
2. Calculate and interpret the joint posterior distribution based on the observed data.
3. Evaluate the fit of the model and implications of the resulting posterior distribution.

1.2.2 Bayesian inference

Based on Bayes' rule, the joint posterior distribution of a vector of parameters, θ , in a scientific model conditional on the observed data, \mathbf{y} , can be written as follows:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\boldsymbol{\theta},\mathbf{y})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} \quad (1.1)$$

where $p(\boldsymbol{\theta}|\mathbf{y})$ is the joint posterior distribution of $\boldsymbol{\theta}$, $p(\mathbf{y}|\boldsymbol{\theta})$ is the likelihood function of the observed data \mathbf{y} and $p(\boldsymbol{\theta})$ is the prior distribution of the model parameters $\boldsymbol{\theta}$. $p(\mathbf{y})$ is an integral of the product $p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})$ over all values of $\boldsymbol{\theta}$ and can be regarded as a normalizing constant to ensure that $p(\boldsymbol{\theta}|\mathbf{y})$ is a proper density. This means that Eq. (1.1) can be further expressed as

$$p(\boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (1.2)$$

where “ \propto ” represents proportionality.

The usual Bayesian structure given in Eq. (1.2) can be extended to a hierarchical model if multiple parameters from a hierarchy of multiple levels are involved. For example, a two-level hierarchical Bayesian model is given by

$$p(\boldsymbol{\theta}|\mathbf{y},\boldsymbol{\Omega}) \propto p(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\Omega})\pi(\boldsymbol{\Omega}) \quad (1.3)$$

where $\boldsymbol{\Omega}$ are the prior parameters of $\boldsymbol{\theta}$; $\pi(\boldsymbol{\theta}|\boldsymbol{\Omega})$ is the likelihood function of the first-level model parameters $\boldsymbol{\theta}$ conditional on the second-level model parameters $\boldsymbol{\Omega}$, and $\pi(\boldsymbol{\Omega})$ is the prior distribution of the model parameters $\boldsymbol{\Omega}$.

1.2.3 Markov chain Monte Carlo simulation

Because of the computational difficulties involved in the evaluation of the joint posterior distribution in the Bayesian updating, the Markov chain Monte Carlo (MCMC) simulation techniques are commonly used to numerically evaluate the joint posterior distribution. In the MCMC simulation, random samples of the parameters $\boldsymbol{\theta}$ are drawn sequentially, with the probability distribution of the current sampled draws depending on the values of the samples drawn in the previous step. This forms a Markov chain. After an initial sequence of iterations (i.e. the so-called burn-in period (Gelman et al. 2004)), the random samples drawn from the subsequent iterations converge to the target distribution, which is the joint posterior distribution. If the number of sequences is large enough, the samples drawn after the burn-in period can then be used to evaluate the

probabilistic characteristics (e.g. mean and standard deviation) of the posterior distribution. Many MCMC sampling algorithms have been reported in the literature, e.g. the celebrated Metropolis-Hasting (M-H) algorithm (Gelman et al. 2004), Gibbs sampler (Gelman et al. 2004) and slice sampling approach (Neal 2003). A comprehensive review of the MCMC algorithms can be found in Liang et al. (2010).

1.3 Objective and Research Significance

The study reported in this thesis is part of a Collaborative Research and Development (CRD) program funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada and TransCanada Pipelines Limited. The objectives of this study were to 1) develop a Bayesian framework to make statistical inferences of the metal-loss corrosion process, which includes the growth of existing defects and generation of new defects over time, based on imperfect data collected from multiple ILIs, and 2) develop methodologies to evaluate the time-dependent system reliability of corroding pipelines by incorporating the corrosion models obtained through the Bayesian updating methodology.

The proposed Bayesian framework provides a rational and consistent approach to make quantitative inferences of the corrosion process on pipelines while taking into consideration the inherent uncertainties associated with the corrosion process and uncertainties associated with the inspection data. The research outcome will assist pipeline integrity engineers in developing defensible maintenance strategies for corroding pipelines subjected to the safety and resource constraints. The probabilistic corrosion models and Bayesian framework developed in this study can also be extended to other aging structures and infrastructures subjected to localized deterioration.

1.4 Scope of the Study

This study consists of two main topics that are presented in Chapters 2 and 3, respectively. Chapter 2 presents the stochastic process-based models to characterize the generation and growth of individual corrosion defects on steel pipelines. The generation and growth models are formulated and statistically inferred from the inspection data in a hierarchical Bayesian framework by taking into account the uncertainties in the

inspection data. The MCMC simulation techniques in conjunction with the data augmentation (DA) are employed to evaluate the model parameters. The developed models and the proposed Bayesian methodology are illustrated and validated by both the simulated and real inspection data. Chapter 3 presents a simulation-based methodology to evaluate the time-dependent system reliability of corroding pipelines by simultaneously considering the generation and growth of corrosion defects. This methodology provides a tool to incorporate the defect generation and growth models obtained from the Bayesian updating to evaluate the system reliability of corroding pipelines.

1.5 Thesis Format

This thesis is prepared in an Integrated-Article Format as specified by the School of Graduate and Postdoctoral Studies at Western University, London, Ontario, Canada. A total of four chapters are included in the thesis. Chapter 1 presents a brief introduction of the background, objective and scope of this study. Chapters 2 and 3 form the main body of the thesis, each of which is presented in an integrated-article format without an abstract, but with its own references. The summary, conclusions and recommendations for future research are given in Chapter 4.

Several simulation algorithms and Bayesian formulations developed and derived in this study are given in appendices, which follow the last chapter. Each appendix is given an identification that consists of a number and a letter. The number indicates the chapter that the appendix is associated with, and the letter indicates the sequence of the appendix appearing in that chapter. For example, Appendix 2A is the first appendix associated with Chapter 2.

References

Enright, M. P., & Frangopol, D. M. (1999). Condition prediction of deteriorating concrete bridges using Bayesian updating. *Journal of Structural Engineering*, 125(10), 1118-1125.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004). *Bayesian Data Analysis*, (2nd edition). Chapman & Hall/CRC.

Kuniewski, S. P., van der Weide, J. A., & van Noortwijk, J. M. (2009). Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection. *Reliability Engineering & System Safety*, 94(9), 1480-1490.

Liang, F., Liu, C. and Chuanhai, J. (2010). *Advanced Markov Chain Monte Carlo Methods*. Wiley Online Library.

Neal, R. M. (2003). Slice sampling. *The Annals of Statistics*, 31(3): 705-767.

Zhang, S., & Zhou, W. (2013). System reliability of corroding pipelines considering stochastic process-based models for defect growth and internal pressure. *International Journal of Pressure Vessels and Piping*, 111, 120-130.

Zheng, R., & Ellingwood, B. R. (1998). Role of non-destructive evaluation in time-dependent reliability analysis. *Structural Safety*, 20(4), 325-339.

Chapter 2 Probabilistic Modeling and Bayesian Inference of Generation and Growth of Corrosion Defects on Pipelines Based on Imperfect Inspection Data

2.1 Introduction

Metal-loss corrosion involves two processes, namely the growth of existing defects and the generation of new defects. Both processes involve significant inherent uncertainties. A rational probabilistic approach to characterize these two processes can facilitate various tasks (e.g. reliability evaluation and determination of optimal maintenance strategies) involved in the corrosion management of oil and gas pipelines.

The stochastic processes, e.g. the gamma process (e.g. Maes et al. 2009a; Maes et al. 2009b) and Markov chain (e.g. Timashev et al. 2008; Caleyó et al. 2009), have been employed in the context of modeling the growth of corrosion defects. Recently, the gamma process-based corrosion growth models in conjunction with the hierarchical Bayesian methodology have been developed based on the inspection data obtained from multiple in-line inspections (ILIs) (Zhang and Zhou 2013; Zhang et al. 2014). The gamma process has non-negative and independent gamma-distributed increments over disjoint (non-overlapping) time increments, and is suitable to characterize the monotonic corrosion growth process and account for the temporal variability of the corrosion growth. However, the above-mentioned studies only considered the growth of existing defects but ignored the generation of new defects. Such a simplification may adversely impact the accuracy of the integrity assessment and maintenance decision-making of corroding pipelines.

A corrosion defect can initiate randomly in space and time. The Poisson processes, including the homogeneous and non-homogenous Poisson process (HPP and NHPP), have been widely used to model the defect generation (e.g. Hong 1999; Valor et al. 2007). Hong (1999) employed HPP to characterize the generation of new defects and considered the impact of newly generated defects on the evaluation of the failure probability of corroding pipelines. Valor et al. (2007) employed NHPP to model the generation of corrosion pits, and the Markov chain to model the pit growth. Note that the above studies

did not address the evaluation of parameters of the HPP and NHPP models based on the corrosion inspection data, which involve uncertainties as a result of the imperfect detectability of the inspection tool.

The periodic inspections of pipelines provide valuable information pertaining to the condition of the corrosion on pipelines. The ILI data include the sizes of individual defects measured by the ILI tool as well as the number of defects detected by the ILI tool at the time of inspection. The former is subjected to the sizing uncertainty (i.e. the measurement errors) (Kariyawasam and Peterson 2008), whereas the latter is subjected to the detecting uncertainty as reflected in the probability of detection and probability of false call. It is of high practical value to make statistical inferences of the generation and growth of corrosion defects simultaneously based on the inspection data, while taking into account both the sizing and detecting uncertainties. Such studies are however scarce in the literature. Kuniewski et al. (2009) developed a sampling-inspection strategy for the reliability evaluation of corroding structures and proposed a Bayesian methodology to update the NHPP-based defect generation model based on the sampling inspection data. The probability of detection was considered in the updating, but the measurement errors were ignored. Although the gamma process-based growth of corrosion defects was considered in the reliability analysis, the parameters of the growth model were assumed to be known; therefore, the updating of the growth model based on the inspection data was not addressed in their study.

The objective of the work reported in this chapter is to develop a probabilistic model to characterize the growth of existing defects and generation of new defects based on the imperfect inspection data. The growth modeling was focused on the defect depth (i.e. in the through-pipe wall thickness direction), as this is the most critical defect dimension. The model was formulated in a Bayesian framework, which accounts for the inherent variability involved in the corrosion process as well as the sizing and detecting uncertainties associated with the ILI tool. To this end, the non-homogeneous gamma process was used to model the growth of defect depths, and the non-homogeneous Poisson process was employed to model the generation of new defects. The Markov chain Monte Carlo (MCMC) simulation techniques in conjunction with the data augmentation

algorithm for dealing with the missing data were used to carry out the Bayesian updating to evaluate the probabilistic characteristics of the model parameters. Numerical examples involving both hypothetical and real inspection data were used to illustrate the proposed model.

The rest of this chapter is organized as follows. Section 2.2 describes the uncertainties involved in the ILI data; Section 2.3 presents the probabilistic models for the defect generation and growth adopted in this study; the Bayesian methodology for evaluating the defect generation and growth models based on the inspection data is described in Section 2.4, and illustrated using numerical examples in Section 2.5, and conclusions are presented in Section 2.6.

2.2 Uncertainties in the ILI Tool

Two categories of uncertainties associated with the ILI tool were considered in this study, namely the measurement error and imperfect detectability. The former includes the biases and random scattering error, whereas the latter is characterized by the probability of detection (POD) and probability of false call (POFC).

2.2.1 Measurement error

The measured depth of the j^{th} defect at the i^{th} inspection, y_{ij} , ($i = 1, 2, \dots, j = 1, 2, \dots$) can be related to the corresponding actual depth, x_{ij} , through the following equation (Fuller 1987; Jaech 1985):

$$y_{ij} = a_i + b_i x_{ij} + \varepsilon_{ij} \quad (2.1)$$

where a_i and b_i denote the constant and non-constant biases, respectively, associated with the ILI tool used in the i^{th} inspection, and ε_{ij} denotes the random scattering error associated with the ILI-reported depth of the j^{th} defect at the i^{th} inspection, and is assumed to be normally distributed with a zero mean and standard deviation σ_i . It is further assumed that for a given inspection i , ε_{ij} and ε_{ik} ($j \neq k$) (i.e. the random scattering errors associated with the ILI-reported depths of the j^{th} and k^{th} defects) are independent, whereas for a given defect j , ε_{ij} and ε_{lj} ($i \neq l$) may be correlated with a correlation coefficient of ρ_{il}

(Al-Amin et al. 2012). Let $\mathbf{E}_j = (E_{1j}, E_{2j} \dots E_{nj})'$ denote the vector of random scattering errors associated with n inspections for defect j , with “'” representing transposition. It follows from the above assumption that \mathbf{E}_j is multivariate normal-distributed and has a probability density function (PDF) given by

$$f_{\mathbf{E}_j}(\boldsymbol{\varepsilon}_j) = \frac{1}{(\sqrt{2\pi})^{n/2} |\boldsymbol{\Sigma}_{\mathbf{E}_j}|^{1/2}} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_j' \boldsymbol{\Sigma}_{\mathbf{E}_j}^{-1} \boldsymbol{\varepsilon}_j\right) \quad (2.2)$$

where $\boldsymbol{\Sigma}_{\mathbf{E}_j}$ denotes the $n \times n$ variance-covariance matrix of \mathbf{E}_j with the element at the i^{th} row and l^{th} column equal to $\rho_{il}\sigma_i\sigma_l$. In this study, a_i , b_i and $\boldsymbol{\Sigma}_{\mathbf{E}_j}$ were assumed to be known quantities whose values can be evaluated by comparing the ILI-reported and corresponding field-measured depths for a set of benchmark defects (Al-Amin et al. 2012) or inferred from the vendor-supplied specifications for the accuracy of the ILI tools.

2.2.2 Probability of detection and probability of false call

POD represents the ability of an ILI tool to detect a true corrosion defect. It is typically a function of the size of the defect and a set of parameters indicating the inherent detecting capability of the ILI tool. The following exponential POD function (Zheng and Ellingwood 1998) was adopted in this study:

$$\text{POD}(x) = \begin{cases} 1 - e^{-q(x-x_{th})} & x \geq x_{th} \\ 0 & x < x_{th} \end{cases} \quad (2.3)$$

where x denotes the actual depth of a given defect; x_{th} denotes the detection threshold, i.e. the smallest defect size that can be detected, and q is a constant that characterizes the inherent detecting capability of the ILI tool. Figure 2.1 shows the POD curves given by Eq. (2.3) with three sets of values of q ($1/\%w_t$) and x_{th} ($\%w_t$), where $\%w_t$ represents the percentage of the pipe wall thickness (w_t). This figure indicates that the detectability of the tool increases as the value of q increases and/or the detection threshold x_{th} decreases.

The probability of false call (POFC) is the probability of an ILI tool obtaining an indication of a defect that does not exist in reality. For the high-resolution ILI tools commonly used to inspect oil and gas pipelines, POFC is typically negligibly small.

Therefore, POFC was ignored in the present study; in other words, all the ILI-reported corrosion defects were assumed to be true corrosion defects.

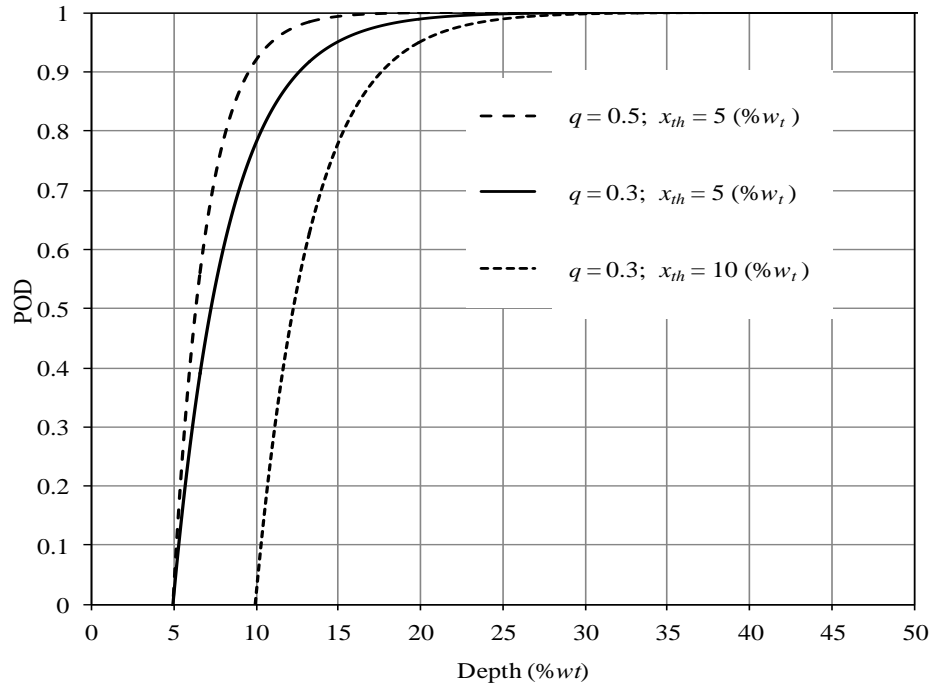


Figure 2.1 POD curves with different sets of values of q and x_{th}

2.3 Probabilistic Models for Defect Generation and Growth

2.3.1 Defect generation

The non-homogeneous Poisson process (NHPP) was employed to characterize the generation of new defects, as the model has been widely used in the literature (e.g. Valor et al. 2007; Kuniewski et al. 2009). According to this model, the total number of defects, $N(t)$, generated within a time interval $[0, t]$ (e.g. $t = 0$ denotes the time of installation of the pipeline) over a given segment of the pipeline follows a Poisson distribution with a probability mass function (PMF), $f_P(N(t)|m(t))$, defined as (Beichelt and Fatti 2002)

$$f_P(N(t)|m(t)) = \frac{(m(t))^{N(t)} e^{-m(t)}}{N(t)!} \quad (t \geq 0) \quad (2.4)$$

where $m(t)$ denotes the expected number of defects generated over the time interval $[0, t]$, and is assumed in this study to follow a power-law function of time (Kuniewski et al. 2009): $m(t) = \int_0^t \lambda \delta s^{\delta-1} ds = \lambda t^\delta$ with the parameters λ and δ ($\lambda, \delta > 0$) to be quantified based on the inspection data. Figure 2.2 depicts the means, 2.5- and 97.5-percentile values, and realizations of $N(t)$ over 20 years based on Eq. (2.4) corresponding to two sets of assumed values of λ and δ . Note that $N(t)$ degenerates to a homogeneous Poisson process (HPP) if δ equals unity.

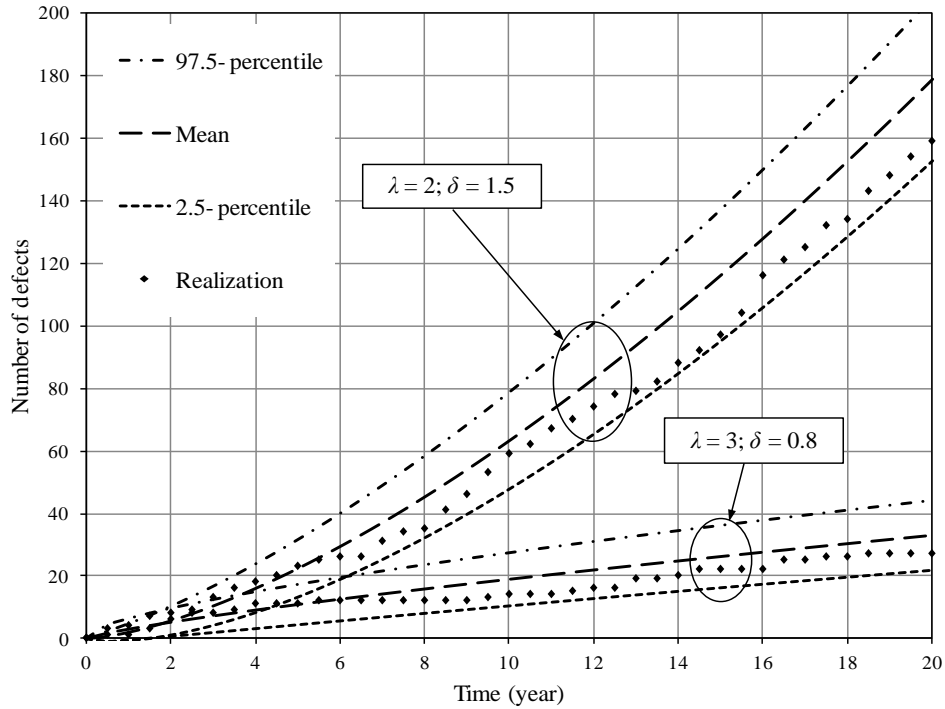


Figure 2.2 Mean, 2.5-percentile, 97.5-percentile and a given realization of NHPP

Suppose n inspections have been carried out for a given pipeline segment over a certain period of time. It is assumed that each inspection is able to identify new and existing corrosion defects by tracking their spatial positions. This assumption is consistent with the corrosion inspection practice for oil and gas pipelines (Al-Amin et al. 2012). At the time of the i^{th} inspection ($i = 1, 2, \dots, n$), t_i , the total number of corrosion defects on the pipeline segment, N_i , can be divided into those defects that have initiated prior to the $(i-1)^{\text{th}}$ inspection, N_i^o , and those defects that initiated between the $(i-1)^{\text{th}}$ and

i^{th} inspections, N_i^g . The quantity N_i^g then follows a Poisson distribution with a PMF given by

$$f_P(N_i^g | \lambda, \delta) = \frac{(m_i)^{N_i^g}}{N_i^g!} \exp(-m_i) \quad (2.5)$$

where $m_i = \int_{t_{i-1}}^{t_i} \lambda \delta s^{\delta-1} ds = \lambda(t_i^\delta - t_{i-1}^\delta)$, and $t_0 \equiv 0$.

Because of the imperfect detectability of the ILI tool, the detected number of defects is in general less than the actual number of defects. Let N_i^{gd} and N_i^{gu} denote the detected and undetected portions of N_i^g , respectively, i.e. $N_i^g = N_i^{gd} + N_i^{gu}$. Based on the Poisson splitting property (Kulkarni 1995), N_i^{gd} and N_i^{gu} follow Poisson distributions with the corresponding PMFs as follows:

$$f_P(N_i^{gd} | \lambda, \delta) = \frac{(\overline{\text{POD}}_i m_i)^{N_i^{gd}}}{N_i^{gd}!} \exp(-\overline{\text{POD}}_i m_i) \quad (2.6)$$

$$f_P(N_i^{gu} | \lambda, \delta) = \frac{[(1 - \overline{\text{POD}}_i) m_i]^{N_i^{gu}}}{N_i^{gu}!} \exp[-(1 - \overline{\text{POD}}_i) m_i] \quad (2.7)$$

where $\overline{\text{POD}}_i$ is the average POD with respect to the N_i^g defects. $\overline{\text{POD}}_i$ can be calculated as $\overline{\text{POD}}_i = \int \text{POD}(x) f_{X_i^g}(x) dx$, where $f_{X_i^g}(x)$ is the PDF of the depths of N_i^g defects at time t_i .

2.3.2 Defect growth

The non-homogeneous gamma process (NHGP) (Zhang et al. 2014) was employed to characterize the growth of depths of corrosion defects. It follows that the depth of a given defect at time t , $X(t)$, is gamma distributed with the PDF given by:

$$f_G(x(t) | \alpha(t), \beta) = \frac{\beta^{\alpha(t)} x(t)^{\alpha(t)-1} \exp[-x(t)\beta]}{\Gamma(\alpha(t))} I_{(0, \infty)}(x(t)) \quad (2.8)$$

where $\alpha(t)$ is the time-dependent shape parameter and assumed to be a power-law function of time, i.e. $\alpha(t) = \varphi_1 (t - t_s)^{\varphi_2}$ ($t > t_s$), with t_s denoting the defect initiation time

(the time at which a defect initiates and starts growing); β ($\beta > 0$) is the time-independent rate parameter (or inverse of the scale parameter) (Ang and Tang 2007); $\Gamma(\cdot)$ denotes the gamma function, and $I_{(0,\infty)}(x(t))$ is the indication function, which equals unity if $x(t) > 0$ and zero otherwise. The mean, variance and coefficient of variation (COV) of $X(t)$ equal $\alpha(t)/\beta$, $\alpha(t)/\beta^2$ and $1/(\alpha(t))^{0.5}$, respectively. The quantity φ_1/β represents the mean of the depth at the first unit increment of time since t_s ; φ_2 reflects the slope of the mean growth path of the defect with $\varphi_2 > 1$, $\varphi_2 < 1$ and $\varphi_2 = 1$ representing an accelerating, a decelerating and a linear mean growth path, respectively. Furthermore, $\varphi_2 = 1$ corresponds to a homogeneous gamma process.

In this study, the parameters φ_1 and φ_2 were assumed to be common for all the defects, whereas t_s and β were assumed to be defect-specific to account for the spatial variability of the defect growth. Let t_{sr} and β_r denote the time of initiation and rate parameter for the r^{th} defect ($r = 1, 2, \dots$), respectively. It should be emphasized here that the index r is used to enumerate all defects, including detected and undetected defects, to distinguish from the index j that is used to enumerate detected defects. The parameter β_r was further assumed to be an exponential function of the random effect parameter ζ_r , i.e. $\beta_r = e^{\zeta_r}$. The advantage of expressing β_r as an exponential function is that it ensures β_r to be positive and can easily incorporate local covariates, if any, to more accurately characterize β_r .

It follows from the above-described assumptions that the depth increment of defect r over the time interval between the $(i-1)^{\text{th}}$ and i^{th} inspections, denoted by Δx_{ir} , is gamma-distributed with a shape parameter $\Delta\alpha_{ir}$ and a rate parameter β_r , where $\Delta\alpha_{ir}$ is given by

$$\Delta\alpha_{ir} = \begin{cases} \varphi_1(t_i - t_{sr})^{\varphi_2}, & i = 1 \\ \varphi_1(t_i - t_{sr})^{\varphi_2} - \varphi_1(t_{i-1} - t_{sr})^{\varphi_2}, & i = 2, 3, \dots, n \end{cases} \quad (2.9)$$

The depth of defect r at the i^{th} inspection, x_{ir} , is then the summation of the depth at the $(i-1)^{\text{th}}$ inspection and Δx_{ir} ; that is, $x_{ir} = x_{i-1,r} + \Delta x_{ir}$. Note that x_{ir} at $t = t_{sr}$ is assumed to equal zero.

2.4 Bayesian Updating of Defect Generation and Growth Models

2.4.1 Overview

The Bayesian updating was employed to make statistical inferences of the parameters of the defect generation and growth models described in Section 2.3 based on the ILI data. Through Bayes' theorem, the Bayesian updating combines the previous knowledge about uncertain model parameters with the new information contained in the observed data to lead to updated knowledge about these parameters. The previous knowledge is reflected in the prior distributions; the new information in the observed data is incorporated in the likelihood functions, and the updated knowledge of the parameters is reflected in the posterior distributions. The formulations of the prior and posterior distributions as well as the likelihood functions for the defect generation and growth models are described in the following sections.

2.4.2 Likelihood functions

2.4.2.1 Likelihood function for ILI-reported depths

Consider that a set of defects have been detected in a total of n inspections. Suppose that defect j is first detected in the l^{th} ($l = 1, 2, \dots, \text{or}, n$) inspection. Let $\mathbf{y}_j = (y_{lj}, y_{l+1,j}, \dots, y_{l+k,j}, \dots, y_{nj})'$ denote the vector of the ILI-reported depths for defect j . Further let $\mathbf{x}_j = (x_{lj}, x_{l+1,j}, \dots, x_{l+k,j}, \dots, x_{nj})'$ denote the vector of the actual depths of defect j corresponding to the ILI-reported depths. Given the measurement error model described in Section 2.2.1, the likelihood of \mathbf{y}_j conditional on \mathbf{x}_j can be expressed as

$$L(\mathbf{y}_j|\mathbf{x}_j) = (2\pi)^{-\frac{n-l+1}{2}} |\boldsymbol{\Sigma}_{\mathbf{E}_j}|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{y}_j - \mathbf{a} - \mathbf{b}\mathbf{x}_j)' \boldsymbol{\Sigma}_{\mathbf{E}_j}^{-1} (\mathbf{y}_j - \mathbf{a} - \mathbf{b}\mathbf{x}_j) \right] \quad (2.10)$$

where $\mathbf{a} = (a_l, a_{l+1}, \dots, a_n)'$, and \mathbf{b} is an $n-l+1 \times n-l+1$ diagonal matrix with the k^{th} element equals to b_{l+k} . The above formulation assumes that once a defect is detected for the first time, it will be detected in all subsequent inspections. It should be noted that this assumption can be relaxed in the analysis; that is, the defect is not necessarily detected in all subsequent inspections. In this case, the ILI-reported depths corresponding to the inspections that do not detect the defect can be considered as the missing data and

handled using the multiple imputation technique (Rubin 2009), which has been implemented in widely used Bayesian updating software such as OpenBUGS (Lunn et al. 2009).

2.4.2.2 Likelihood function for the number of detected defects

To simplify the likelihood functions for the number of defects, it is assumed that the defects detected for the first time (referred to as the newly detected defects) in the i^{th} inspection are generated between the $(i-1)^{\text{th}}$ and i^{th} inspections. This assumption ignores the possibility that some of the newly detected defects in the i^{th} inspection may in fact initiate prior to the $(i-1)^{\text{th}}$ inspection but remain undetected until the i^{th} inspection. The assumption results in overestimation of the intensity of the defect generation. It follows from the assumption that the newly detected defects in the i^{th} inspection can be denoted by N_i^{gd} as defined in Section 2.3.1. Based on this assumption and Eq. (2.6), the likelihood function for the set of newly detected defects in n inspections, i.e. N_i^{gd} ($i = 1, 2, \dots, n$), is given by

$$L(N_1^{gd}, N_2^{gd}, \dots, N_n^{gd} | \lambda, \delta) = \prod_{i=1}^n \frac{[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]^{N_i^{gd}}}{N_i^{gd}!} \exp\{-[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]\} \quad (2.11)$$

Note that the evaluation of $\overline{\text{POD}}_i$ involves the depths of both detected and undetected defects in the i^{th} inspection. The depths of undetected defects were treated as the missing data and handled using the data augmentation (DA) technique (Tanner and Wong 1987), which is described in Section 2.4.4. It follows that $\overline{\text{POD}}_i$ serves as a link between the defect generation and growth models in the Bayesian updating.

2.4.3 Prior distributions

The gamma distribution was selected as the prior distribution for parameters λ and δ of the NHPP-based defect generation model, and for parameters φ_1 and φ_2 of the NHGP-based defect growth model, based on the consideration that the gamma distribution

ensures these parameters to be positive and can be conveniently constructed to be non-informative. Consistent with the assumption stated in Section 2.4.2.2, the prior distribution for the initiation time of a given detected defect j , t_{sj} , was selected to be uniformly distributed with the corresponding upper bound (ub_j) equal to the time of the inspection that detects the defect for the first time and the lower bound (lb_j) equal to the time of the immediate previous inspection. The prior distributions for t_{sj} for different defects were further assumed to be mutually independent. The random effect parameter ζ_r corresponding to different defects were assumed to follow independent identical (iid) normal prior distributions with a mean of zero and a common uncertain variance σ_ξ^2 . The hierarchical structure for ζ_r facilitates the generation of the depths of undetected defects as required by the data augmentation analysis. Finally, the prior distribution for $1/\sigma_\xi^2$ was assigned a gamma distribution (i.e. σ_ξ^2 follows an inverse-gamma distribution) as commonly suggested in the literature (e.g. Ntzoufras 2011), which leads to the conjugate posterior distribution for $1/\sigma_\xi^2$ and can improve the computational efficiency. The shape (rate) parameters of the gamma prior distributions for φ_1 , φ_2 , λ , δ and $1/\sigma_\xi^2$ are denoted by c (d), e (f), g (h), κ (η) and γ (ω), respectively.

2.4.4 Posterior distributions, MCMC simulation and missing data

Because it is not possible to analytically derive the complex joint posterior distribution of the parameters of the defect generation and growth models, the Markov chain Monte Carlo (MCMC) simulation techniques (Gilks 2005) were employed to numerically evaluate the joint posterior distribution of the model parameters. A hybrid algorithm combining the Metropolis-Hastings (M-H) algorithm and Gibbs sampling (Tierney 1994) was implemented in MatlabTM to carry out the MCMC simulation. The derivations of full conditional posterior distributions of the model parameters as required by the hybrid algorithm are included in Appendix 2A. It is emphasized that both the detected and undetected defects were incorporated in the Bayesian updating. The actual depths of the detected defects are related to the ILI-reported depths through the likelihood function given by Eq. (2.10), whereas the actual depths of the undetected defects were treated as the missing data and imputed using the DA technique (Tanner and Wong 1987). Therefore, the joint posterior distribution of the model parameters was evaluated from the

depths of the overall defect population as opposed to the depths of the detected defect population only.

DA is an iterative process and can be straightforwardly incorporated in the MCMC simulation. A given DA iteration includes two steps, namely the imputation step and the posterior step (Little and Rubin 2002). The former is used to generate the samples of the missing data from its corresponding probabilistic distribution conditional on the current state of model parameters, and the latter is used to generate a new set of samples of model parameters from their corresponding posterior distributions conditional on both the observed and missing data. Details of the DA technique can be found in Little and Rubin (2002).

A step-by-step procedure for generating samples of the model parameters as well as samples of the missing data (i.e. depths of undetected defects) in the k^{th} ($k = 1, 2, \dots$) MCMC simulation sequence is described as follows, where the notation $\bullet_{(k)}$ is used to denote the value of \bullet obtained in the k^{th} simulation sequence.

1. Impute depths of undetected defects (i.e. missing data).

1.1) Generate the number of undetected defects initiated between the $(i-1)^{\text{th}}$ and i^{th} ($i = 1, 2, \dots, n$) inspections, i.e. $N_i^{gu}(k)$, from the Poisson PMF given by Eq. (2.7) with λ , δ and $\overline{\text{POD}}_i$ replaced by $\lambda_{(k-1)}$, $\delta_{(k-1)}$ and $\overline{\text{POD}}_{i_{(k-1)}}$, respectively.

1.2) Generate depths of $N_i^{gu}(k)$ undetected defects, $x_{v(k)}^{iu}$, ($v = 1, 2, \dots, N_i^{gu}(k)$) as follows.

1.2.1) Set $v = 0$;

1.2.2) generate a random effect parameter $\xi_{(k)}$ from the normal distribution with a zero mean and a variance of $\sigma_{\xi}^2_{(k-1)}$, and then calculate $\beta_{(k)} = e^{\xi_{(k)}}$;

1.2.3) generate a defect initiation time $t_{s(k)}$ based on the procedure described in Appendix 2B.

1.2.4) generate a defect depth x from the gamma PDF given by Eq. (2.8) with φ_1 ,

φ_2 , t_s and β equal to $\varphi_{1(k-1)}$, $\varphi_{2(k-1)}$, $t_{s(k)}$ and $\beta_{(k)}$, respectively;

1.2.5) if x is less than the detection threshold of the ILI tool, i.e. x_{th} , accept x as the depth of an undetected defect; otherwise, accept x with a probability of $1 - \text{POD}(x)$;

1.2.6) set $v = v + 1$ if x is accepted, and

1.2.7) repeat Steps 1.2.2) through 1.2.6) until $v = N_i^{gu(k)}$.

2. For the set of N_i^{gd} defects ($i = 1, 2, \dots, n$), i.e. the newly detected defects in the i^{th} inspection, generate the corresponding depths at the i^{th} and all subsequent inspections, $x_{lj(k)}^{id}$ ($l = i, i + 1, \dots, n; j = 1, 2, \dots, N_i^{gd}$) as follows:

2.1) generate the increments of the depth between consecutive inspections for each of the N_i^{gd} defects, $\Delta x_{lj(k)}^{id}$, from the full conditional posterior distribution listed in Appendix 2A using the M-H algorithm, and

2.2) calculate $x_{lj(k)}^{id} = \sum_{s=i}^l \Delta x_{sj(k)}^{id}$ for $j = 1, 2, \dots, N_i^{gd}$.

3 Calculate $\overline{\text{POD}}_{i(k)}$ for $i = 1, 2, \dots, n$ as follows:

$$\overline{\text{POD}}_{i(k)} = \frac{\sum_{j=1}^{N_i^{gd}} \text{POD}(x_{ij(k)}^{gd}) + \sum_{v=1}^{N_i^{gu(k)}} \text{POD}(x_{v(k)}^{gu})}{N_i^{gd} + N_i^{gu(k)}} \quad (2.12)$$

4 Sample $\lambda_{(k)}$, $\delta_{(k)}$, $\varphi_{1(k)}$, $\varphi_{2(k)}$, $t_{sj(k)}$, $\xi_{j(k)}$ ($j = 1, 2, \dots, \sum_{i=1}^n N_i^{gd}$) and $\sigma_{\xi}^2_{(k)}$ from the corresponding full conditional posterior distributions listed in Appendix 2A using either the M-H algorithm or Gibbs sampling.

2.5 Illustrative Examples

2.5.1 Example 1

In the first example, we used hypothetical (i.e. simulated) inspection data to illustrate and validate the proposed Bayesian methodology. The parameters λ , δ , φ_1 and φ_2 of the

defect generation and growth models were set to be deterministic quantities with $\lambda = 2$, $\delta = 1.2$, $\varphi_1 = 3$ and $\varphi_2 = 0.9$, and the random effect parameter ξ for different defects was assumed to follow iid normal distributions with a zero mean and $\sigma_\xi^2 = 0.36$. It is assumed that three inspections were carried out after the installation of the pipeline ($t = 0$) with $t_i = i \times 5$ years ($i = 1, 2$ and 3). For simplicity, the constant and non-constant biases included in the measurement error model given by Eq. (2.1) were set to equal zero and unity, respectively, for all the inspections, and the random scattering errors associated with different inspections were assumed to be mutually independent with the same standard deviation of unity. Finally, the POD functions associated with all the inspections were assumed to be identical, with the parameters q and x_{th} set to be $q = 0.30$ ($1/\%w_t$) and $x_{th} = 1$ ($\%w_t$). Table 2.1 summarizes the simulated inspection data, whereby the simulation procedure is described in Appendix 2C. Note that the simulation is based on the assumption stated in Section 2.4.2.2, i.e. the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections.

Table 2.1 Summary of the simulated inspection data

Time of Inspection		Year 5	Year 10	Year 15
Number of Detected Defects		9	19 (10)	31 (12)
Measured Depth ($\%w_t$)	Mean	8.0	14.3 (6.5)	19.5 (8.8)
	Standard deviation	3.9	9.7 (3.5)	12.8 (3.8)

Note: The information for newly detected defects in years 10 and 15 year is in brackets.

The Bayesian updating was carried out to evaluate the parameters of the defect generation and growth models based on the simulated inspection data. The shape and scale parameters of the gamma prior distributions for λ , δ , φ_1 and φ_2 were set to be unity, and the shape and scale parameters of the inverse-gamma prior distribution for σ_ξ^2 were set to be 10. A total of 100,000 MCMC simulation sequences were generated following the procedure described in Section 2.4.4, with the first 10,000 sequences considered as the burn-in period (Gelman et al. 2004) and discarded. The samples in the rest of the

sequences were used to evaluate the probabilistic characteristics of the model parameters. The means, medians and standard deviations of the posterior marginal distributions of the model parameters that are common to all the defects are summarized in Table 2.2, where $\overline{\text{POD}}_1$, $\overline{\text{POD}}_2$ and $\overline{\text{POD}}_3$ denote the average POD for the defects generated prior to year 5, between years 5 and 10, and between years 10 and 15, respectively. The results in Table 2.2 suggest that the posterior mean and median values of λ , δ , φ_1 and φ_2 are in good agreement with the corresponding actual values.

Table 2.2 Posterior statistics of model parameters for Example 1

Parameter	Generation (NHPP)		Growth (NHGP)		$\overline{\text{POD}}_1$	$\overline{\text{POD}}_2$	$\overline{\text{POD}}_3$
	λ	δ	φ_1	φ_2			
Mean	1.79	1.26	2.99	0.86	0.66	0.57	0.71
Median	1.79	1.25	2.91	0.91	0.66	0.57	0.71
Standard Deviation	0.31	0.08	0.35	0.09	0.08	0.06	0.06

To investigate the impact of undetected defects on the outcome of the Bayesian updating, two additional scenarios were considered. Scenario I assumes perfect detectability associated with all three inspections (i.e. no undetected defects), whereas Scenario II considers POD but includes only the detected defects (i.e. ignoring the missing data) in calculating $\overline{\text{POD}}_i$ ($i = 1, 2$ and 3) and updating the growth model. In contrast, the results summarized in Table 2.2 are referred to as the base case. The MCMC simulation was carried out for Scenarios I and II. For the base case and Scenarios I and II, the mean values of the number of generated defects, i.e. $m(t) = \lambda t^\delta$, were then calculated for $0 \leq t \leq 20$ years, where the values of λ and δ were set to the corresponding posterior medians. The results are shown in Fig. 2.3. For comparison, $m(t)$ evaluated from the actual values of λ and δ , the simulated total numbers of defects (including the detected and undetected defects) at the times of the three inspections, as well as the simulated numbers of defects detected by the three inspections are also shown in Fig. 2.3. As

indicated in the figure, $m(t)$ corresponding to the base case is practically identical to the actual mean and agree well with the total number of defects, whereas both Scenarios I and II lead to underestimated $m(t)$ values with the degree of underestimation increasing with time. The values of $m(t)$ corresponding to Scenario I at $t = 5, 10$ and 15 years agree well with the inspection data. This is expected because perfect detectability is assumed for Scenario I. The $m(t)$ curve corresponding to Scenario II lies in between those corresponding to Scenario I and the base case. This is because although POD is accounted for in Scenario II, $\overline{\text{POD}}_i$ is overestimated as a result of ignoring the missing data in the calculation.

For the base case and Scenarios I and II, the depths of the detected defects at year 20 were predicted and compared with the corresponding actual defect depths. The predicted depth for a given defect was selected as the mean depth predicted from the NHGP-based growth model, with values of the model parameters (i.e. φ_1 , φ_2 , t_{sj} and ξ_j) set to the corresponding posterior medians. The results are shown in Fig. 2.4. Figure 2.4 suggests that all three cases predict the growth of corrosion defects reasonably well: the predicted depths for 90% of the 31 detected defects fall between the two bounding lines of actual depth $\pm 10\%w_t$. The differences between the predictions corresponding to the three cases are marginal: the predictions for relatively shallow defects (say, depth $\leq 30\%w_t$) corresponding to Scenarios I and II tend to be slightly higher than those corresponding to the base case.

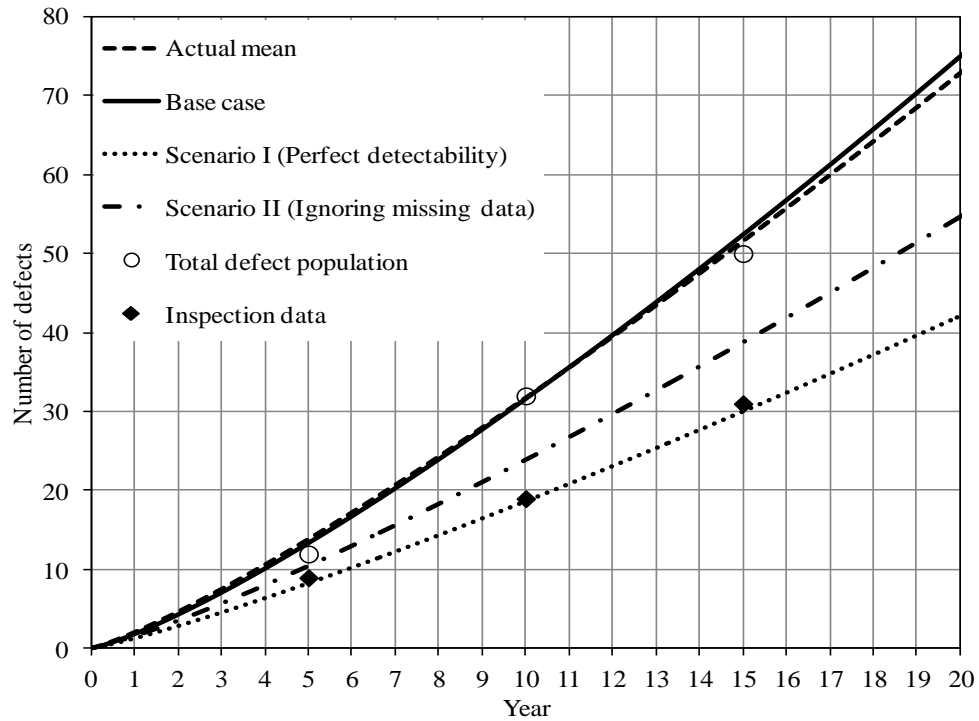


Figure 2.3 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II

As described in Section 2.4.2.2, the Bayesian methodology developed in this study involves the simplifying assumption that the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections. To investigate the impact of this assumption on the predictive capability of the proposed methodology, we further simulated more realistic corrosion data considering the possibility that some of the newly detected defects in the i^{th} inspection may in fact initiate prior to the $(i-1)^{\text{th}}$ inspection but remain undetected until the i^{th} inspection. These data were then used to update the corrosion generation and growth models and make predictions.

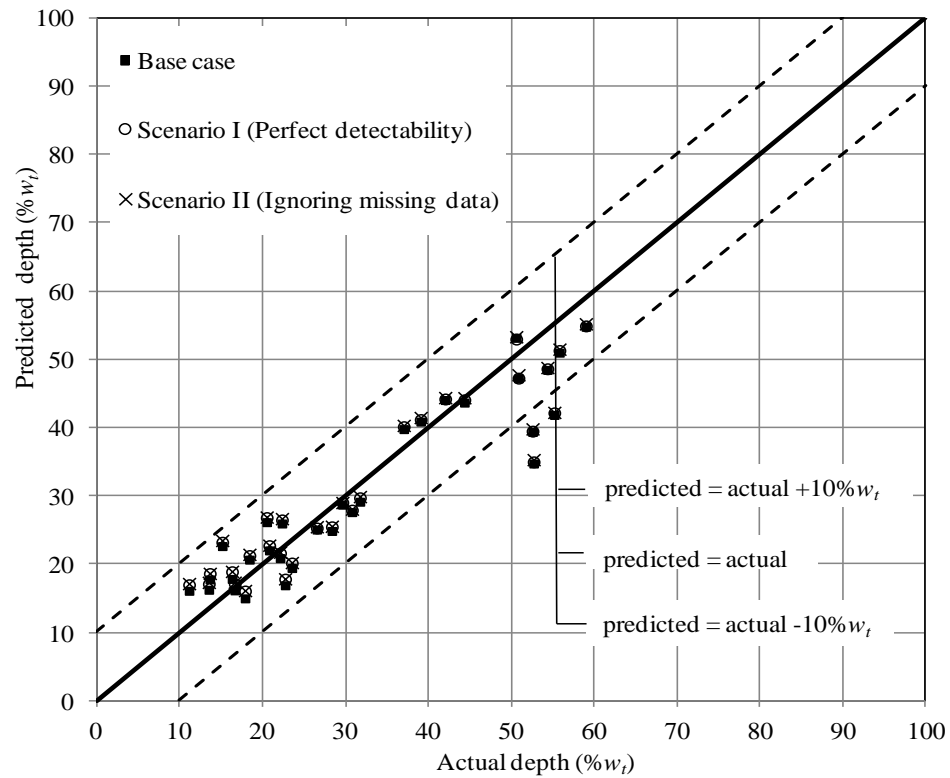


Figure 2.4 Comparison of the predicted and actual depths at year 20

The corrosion inspection data at years 5, 10 and 15 were simulated based on the same set of parameters as those used to generate the data summarized in Table 2.1 except for the POD curve. Three different POD curves were considered in this case, corresponding to POD of 90%, 70% and 50%, respectively, for a defect depth of $5\% w_t$ with a detection threshold of $1\% w_t$. The simulated inspection data corresponding to the three different POD curves are summarized in Table 2.3. The simulation procedure is described in Appendix 2D.

The Bayesian updating was then carried out corresponding to the base case, Scenarios I (i.e. perfect detectability) and II (i.e. ignoring the missing data in evaluating the average POD). The mean values of the number of generated defects, i.e. $m(t) = \lambda t^\delta$, were calculated for $0 \leq t \leq 20$ years, where the values of λ and δ were set to the corresponding posterior medians. The results are shown in Figs. 2.5 in a similar fashion as those shown in Fig. 2.3.

Table 2.3 Summary of the simulated inspection data corresponding three different
POD curves

(a) POD = 90% for a defect depth of $5\%w_t$

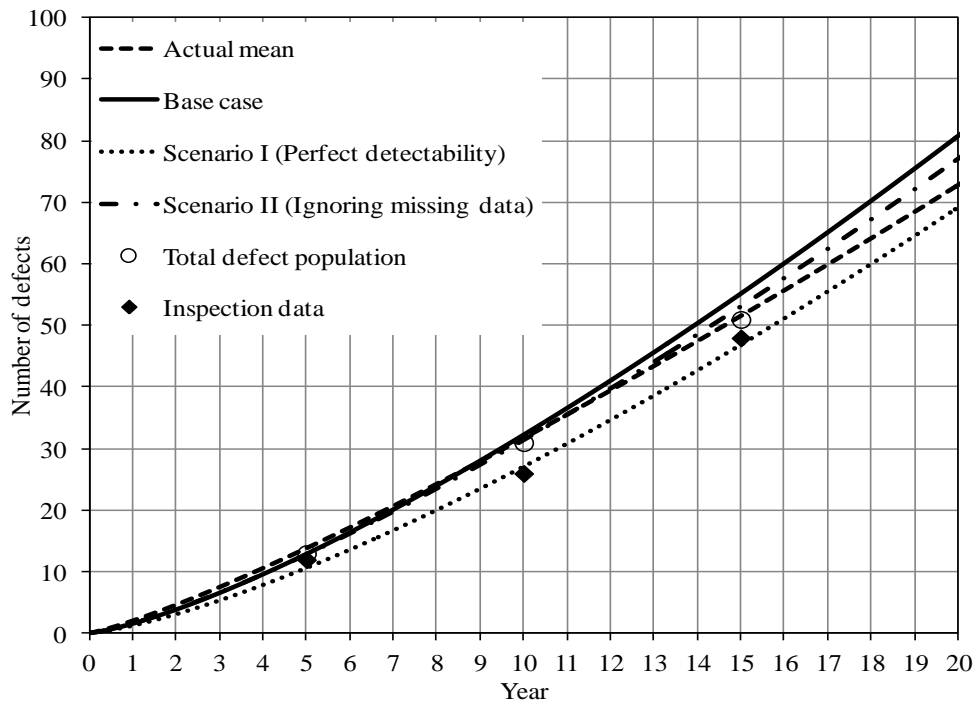
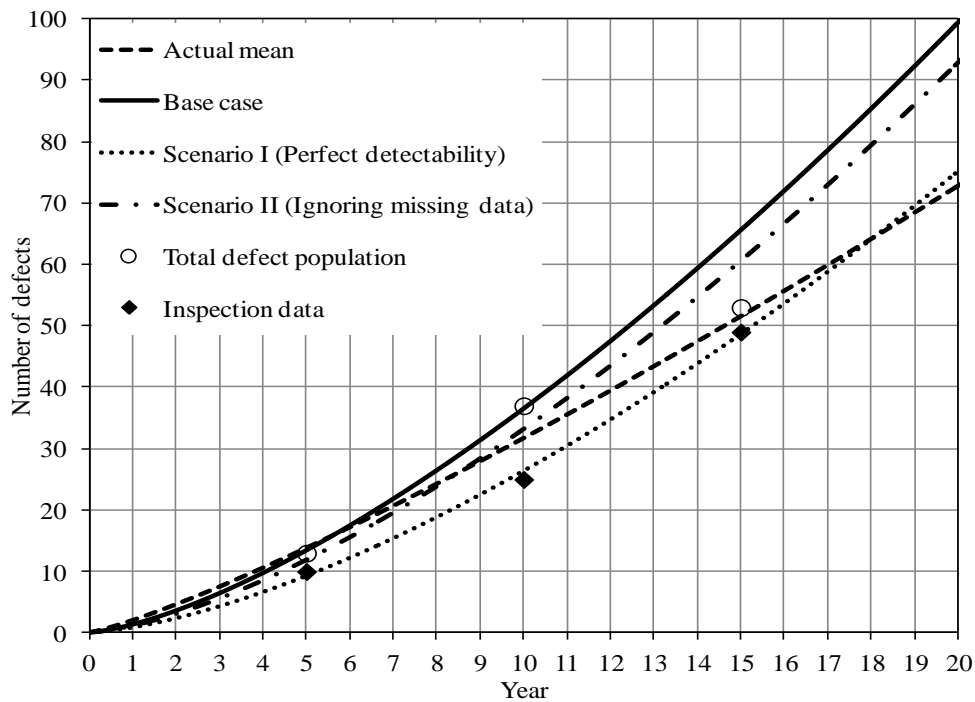
Time of Inspection		Year 5	Year 10	Year 15
Number of Detected Defects		12	26 (14)	48 (22)
Measured Depth ($\%w_t$)	Mean	5.9	12.4(9.2)	17.8 (10.0)
	Standard deviation	4.2	8.2 (6.4)	11.7 (7.2)

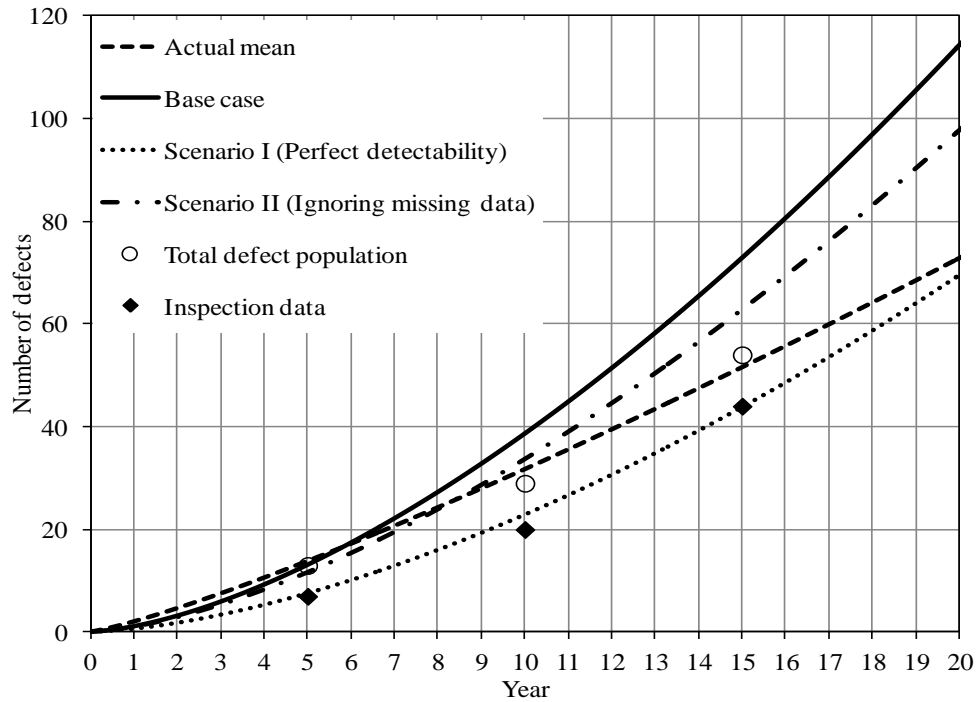
(b) POD = 70% for a defect depth of $5\%w_t$

Time of Inspection		Year 5	Year 10	Year 15
Number of Detected Defects		10	25 (15)	49 (24)
Measured Depth ($\%w_t$)	Mean	8.0	14.1 (10.4)	18.8 (12.5)
	Standard deviation	4.9	9.0 (7.8)	12.0 (7.6)

(c) POD = 50% for a defect depth of $5\%w_t$

Time of Inspection		Year 5	Year 10	Year 15
Number of Detected Defects		7	20 (13)	44 (24)
Measured Depth ($\%w_t$)	Mean	7.0	14.1(11.7)	18.3 (12.1)
	Standard deviation	5.3	8.3 (6.4)	11.4 (6.8)

(a) POD of 90% for a defect depth of $5\% w_t$ (b) POD of 70% for a defect depth of $5\% w_t$



(c) POD of 50% for a defect depth of $5\% w_t$

Figure 2.5 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II based on more realistic corrosion inspection data

As indicated in Fig. 2.5, the $m(t)$ curves corresponding to the base case overestimate the total number of defects. The degree of overestimation decreases as the detectability of the inspection tool increases. This is because the higher is the detectability of the inspection tool, the smaller portion of the newly detected defects are previously undetected defects and the smaller impact does the simplifying assumption have on the prediction. It is interesting to note that the $m(t)$ curves corresponding to Scenario II agree with the total defect population better than those corresponding to the base case and Scenario I. This is because ignoring the missing data in calculating the average POD in Scenario II leads to overestimation of the average POD and underestimation of the total number of defects, which somewhat offsets the overestimation of the total number of defects due to the simplifying assumption that the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections. It should be pointed out that the POD assumptions corresponding to Figs. 2.5(a) and 2.5(b) are more

representative of the inspection tools commonly used in the pipeline industry than that corresponding to Fig. 2.5(c). For the former two assumptions, the conservatism in the predictions associated with the base case is relatively small.

2.5.2 Example 2

In the second example, real ILI data collected from a pipe joint (approximately 13.6m long) in a natural gas pipeline located in Alberta were used to illustrate the proposed methodologies. The pipeline was constructed in 1972 and inspected by high-resolution magnetic flux leakage (MFL) tools in 2004, 2007 and 2009. Note that the pipeline had also been inspected prior to 2004; however, the corresponding inspection data are not available to the present study. The numbers of defects on the pipe joint considered and the statistics of the corresponding defect depths reported by the ILIs in 2004, 2007 and 2009 are summarized in Table 2.6.

The measurement errors associated with the three ILI tools as well as the correlation between the random scattering errors associated with different ILI tools were quantified using the Bayesian methodology in a previous study (Al-Amin et al. 2012). The calibrated biases, the random scattering errors as well as the correlations between the random scattering errors are as follows: $a_1 = 2.04$ ($\%w_t$), $a_2 = -15.28$ ($\%w_t$), $a_3 = -10.38$ ($\%w_t$), $b_1 = 0.97$, $b_2 = 1.40$, $b_3 = 1.13$; $\sigma_1 = 5.97$ ($\%w_t$), $\sigma_2 = 9.05$ ($\%w_t$) and $\sigma_3 = 7.62$ ($\%w_t$); $\rho_{12} = 0.70$, $\rho_{13} = 0.72$ and $\rho_{23} = 0.78$ (Al-Amin et al. 2012), where the subscripts '1', '2' and '3' denote the parameters associated with the ILI tools used in 2004, 2007 and 2009, respectively. The above-mentioned measurement errors were quantified based on 128 defects that were located on several pipe joints in the same pipeline considered in this example, but were mitigated and ceased growing prior to 2000.

The actual POD functions associated with the ILI tools are unavailable. We therefore assumed the three ILI tools to have the identical exponential POF function given by Eq. (2.3). The detection threshold x_{th} in Eq. (2.3) was assumed to be 1 ($\%w_t$), whereas the parameter q was characterized for both the high and relatively low detectability assumptions. The former corresponds to a POD of 90% for a defect depth of 5% w_t , resulting in $q = 0.58$ ($1/\%w_t$), and the latter corresponds to a POD of 70% for the depth of

$5\%w_i$, resulting in $q = 0.30$ ($1/\%w_i$). The two assumed POD curves are compared in Fig. 2.6.

Table 2.4 Summary of the ILI-reported defect information

Time of Inspection		2004	2007	2009
Number of detected defects		67	78 (11)	99 (21)
Measured	Mean	17	11 (7)	11 (8)
Depth ($\%w_i$)	Standard deviation	7.6	10.0 (1.2)	7.8 (1.2)

Note: The information for newly detected defects in 2007 and 2009 is in brackets.

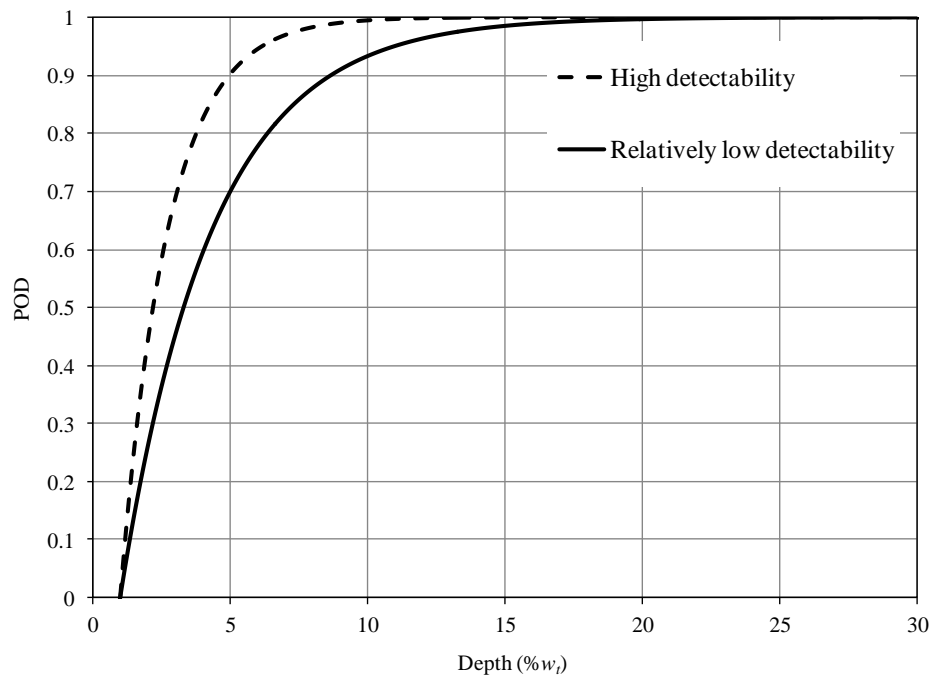


Figure 2.6 Comparison of the two assumed POD curves

The shape and scale parameters of the gamma prior distributions for λ , δ , φ_1 , φ_2 and $1/\sigma_\xi^2$ were set to be 10. For a given POD curve, a total of 100,000 MCMC simulation sequences were generated with the first 10,000 sequences discarded. The samples in the rest of the sequences were used to evaluate the posterior characteristics of the model parameters. The posterior means, medians and standard deviations of the model

parameters corresponding to the high and low detectability assumptions are summarized in Table 2.5, where $\overline{\text{POD}}_1$, $\overline{\text{POD}}_2$ and $\overline{\text{POD}}_3$ denote the calculated average PODs for the defects initiated within the periods of 1972-2004, 2004-2007 and 2007-2009, respectively.

The mean, 2.5- and 97.5-percentile values of the number of defects generated as a function of time according to NHPP are shown in Fig. 2.7. The mean value was obtained as $m(t) = \lambda t^\delta$, with λ and δ equal to their corresponding posterior median values. The 2.5 and 97.5- percentile values of the number of defects were obtained by assuming that the total number of defects at a given time approximately follows a normal distribution with a mean of λt^δ and a standard deviation of $\sqrt{\lambda t^\delta}$ according to the central limit theorem. For comparison, the numbers of defects reported by ILI tools in 2004, 2007 and 2009, respectively, are also plotted in Fig. 2.7. Figure 2.7 indicates that the predicted mean numbers of defects in 2004, 2007 and 2009 are greater than those reported by ILI tools. This is expected because the number of defects predicted by the NHPP model includes both the detected and undetected defects, whereas the ILI-reported defects are detected defects only. The mean numbers of defects corresponding to the relatively low detectability assumption are higher than those corresponding to the high detectability assumption. This also makes sense because the total number of defects increases as the detectability of the tool decreases, if the number of detected defects remains the same.

Table 2.5 Posterior statistics of model parameters for Example 2

Parameter		Generation (NHPP)		Growth (NHGP)			$\overline{\text{POD}}_1$	$\overline{\text{POD}}_2$	$\overline{\text{POD}}_3$
		λ	δ	φ_1	φ_2	σ_ξ^2			
Mean	H ¹	0.70	1.41	3.62	0.58	0.21	0.98	0.76	0.76
	L ²	0.69	1.44	3.65	0.57	0.21	0.93	0.60	0.58
Median	H	0.67	1.41	3.60	0.58	0.20	0.98	0.77	0.76
	L	0.66	1.44	3.62	0.56	0.21	0.93	0.60	0.58
Standard Deviation	H	0.24	0.10	0.53	0.05	0.03	0.01	0.10	0.08
	L	0.24	0.10	0.60	0.05	0.03	0.02	0.09	0.08

1. High detectability. 2. Relatively low detectability

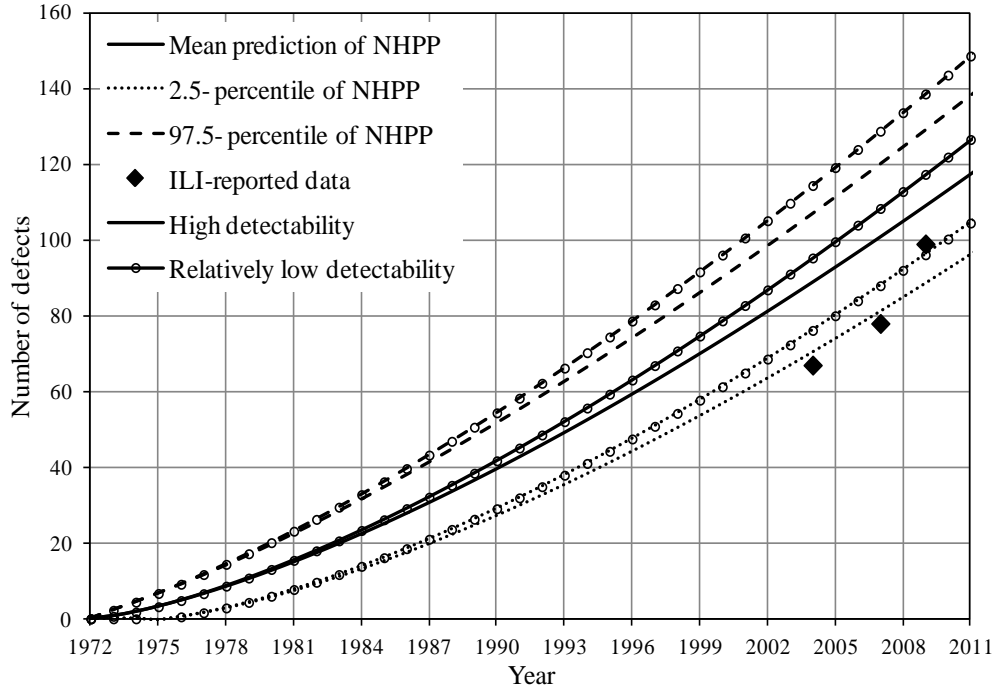
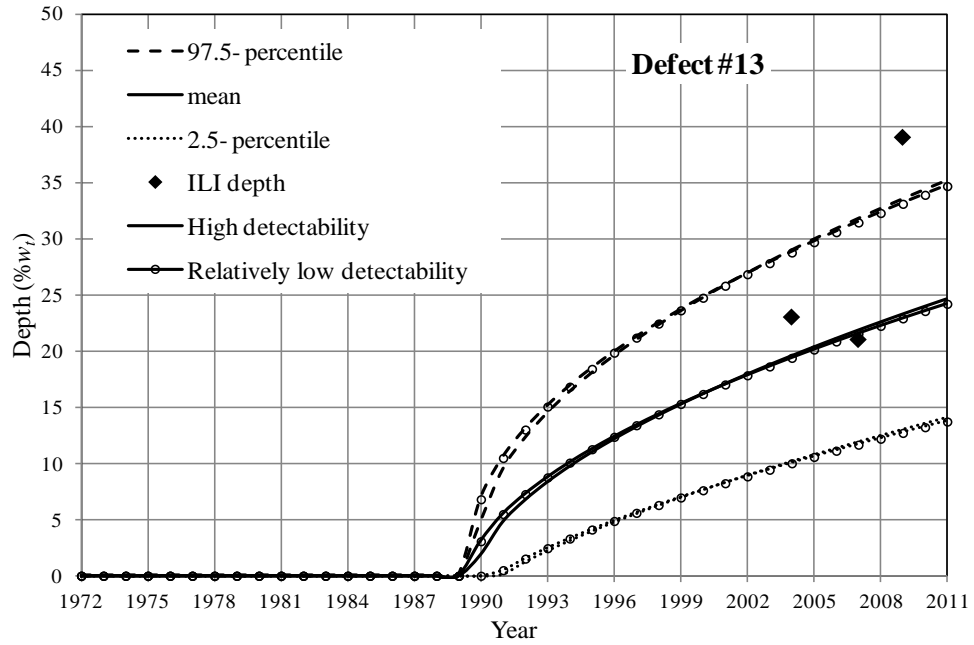
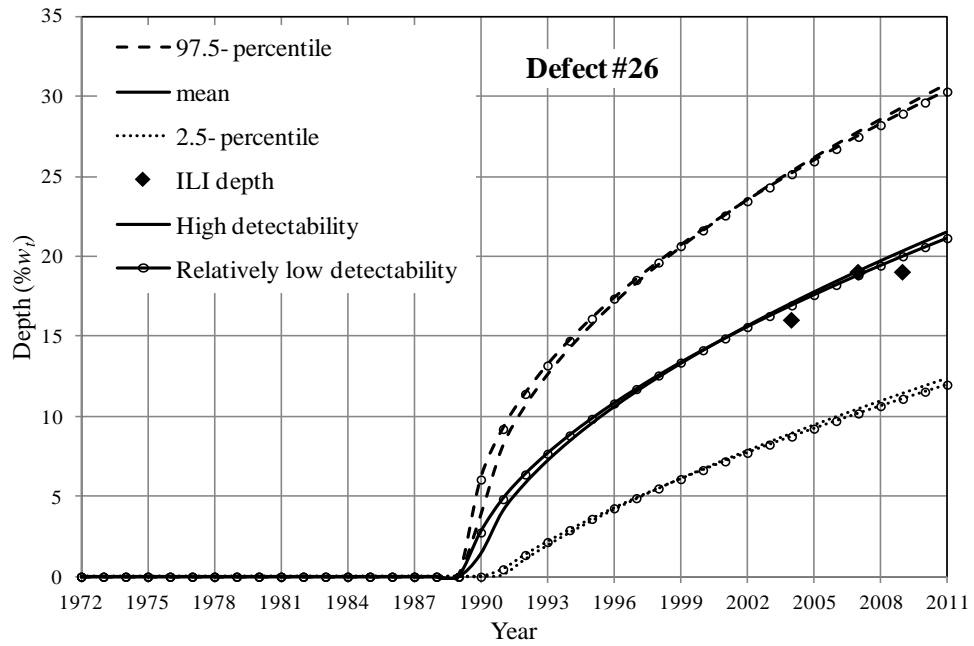


Figure 2.7 Predicted number of defects as a function of time corresponding to the high and relatively low detectability assumptions

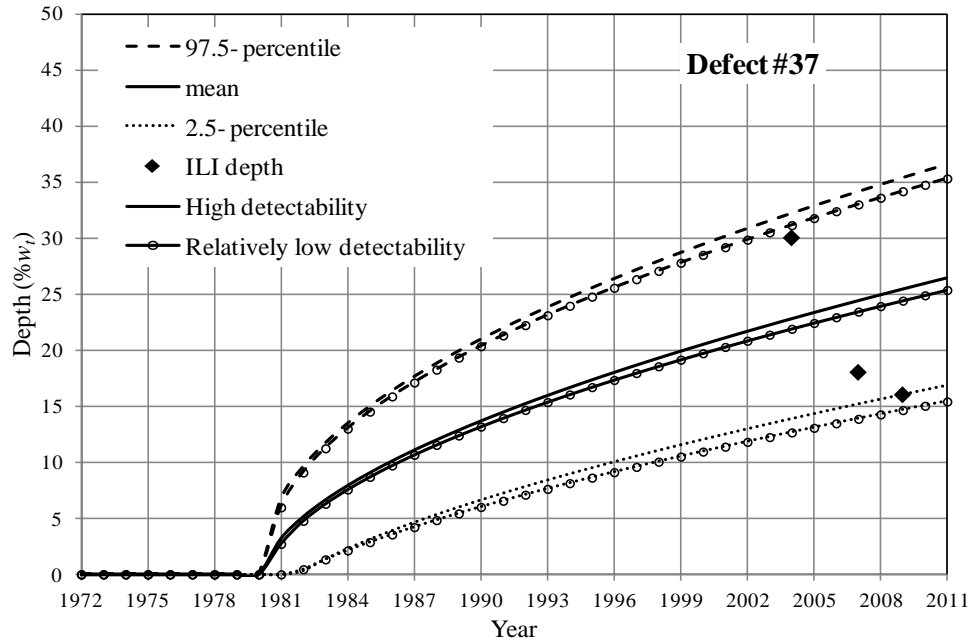
The mean, 2.5- and 97.5- percentile values of the predicted growth paths for five selected detected defects, Defects #13, #26, #37, #61 and #64, are plotted in Figs. 2.8(a) through 2.8(e), respectively. For a given defect, the mean, 2.5- and 97.5- percentile values were evaluated based on the fact that the defect depth at a given time t ($t \geq t_{sj}$) follows a gamma distribution with the shape parameter equal to $\varphi_1(t - t_{sj})^{\varphi_2}$ and the rate parameter equal to e^{ζ_j} . The values of φ_1 , φ_2 , t_{sj} and ζ_j were set to their corresponding posterior medians. For comparison, the corresponding ILI-reported depths in 2004, 2007 and 2009 are also plotted in the same figures. The results indicate that the predicted growth path differs from defect to defect, which is expected because of the defect-specific initiation time t_{sj} and random effect parameter ζ_j . The predicted depths corresponding to the high detectability assumption are almost the same as those corresponding to the relatively low detectability assumption, suggesting that the detectability of the ILI tool has a negligible impact on the predicted growth paths of the detected defects.



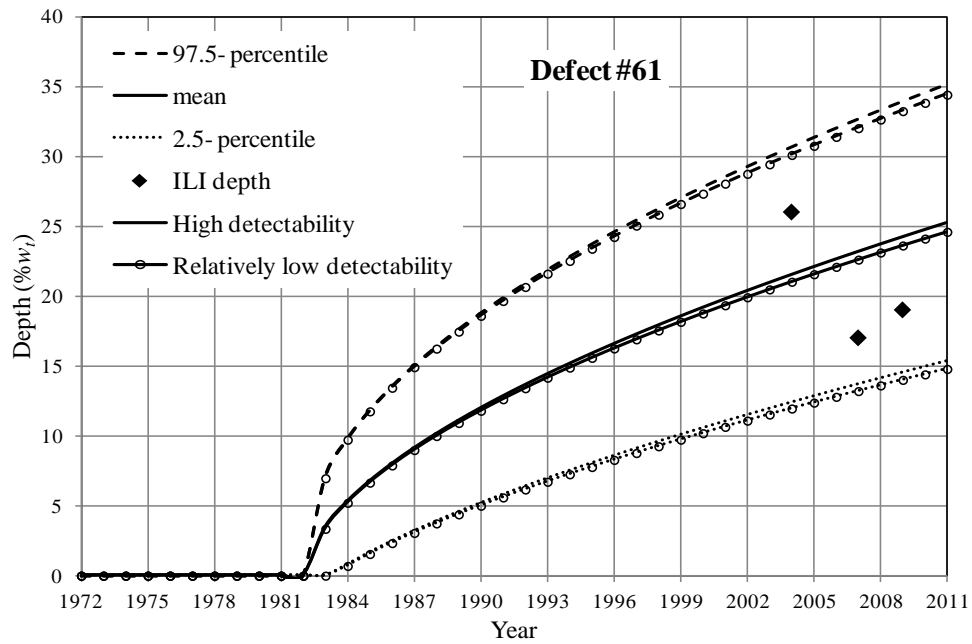
(a) Defect #13



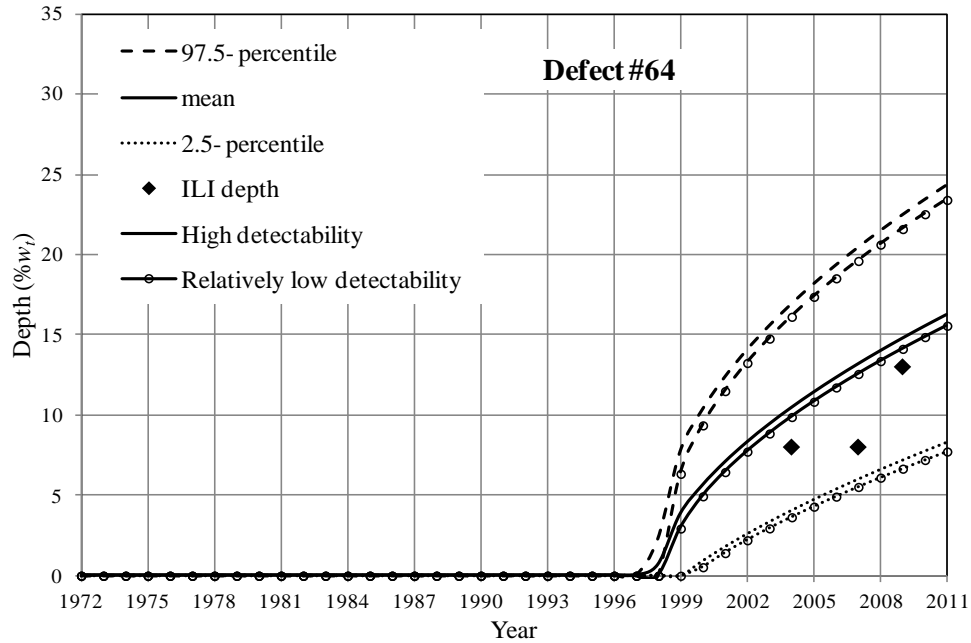
(b) Defect #26



(c) Defect #37



(d) Defect #61



(e) Defect #64

Figure 2.8 Predicted growth paths of selected defects corresponding to the high and relatively low detectability assumptions

The impact of POD and the undetected defects (i.e. missing data) on the outcome of the Bayesian updating was further investigated. In addition to the base case presented in Table 2.7, two additional scenarios, similar to the two scenarios considered for Example 1, were considered: Scenario I assumes perfect detectability for all three inspections, whereas Scenario II accounts for POD but ignores the missing data in calculating $\overline{\text{POD}}_i$ ($i = 1, 2$ and 3) and updating of the growth model. The mean values of the predicted number of defects, $m(t)$, corresponding to the base case and Scenarios I and II are shown in Fig. 2.9. Note that all the predictions shown in Fig. 2.9 are based on the POD curve corresponding to the relatively low detectability assumption. For comparison, the numbers of defects reported by the three inspections are also shown in the same figure. As indicated in Fig. 2.9, $m(t)$ corresponding to the base case are greater than the corresponding ILI-reported numbers of defects as explained in the previous section. On the other hand, $m(t)$ corresponding to Scenario I agree with the ILI-reported numbers of defects better than the other two cases, and $m(t)$ values corresponding to Scenario II are

greater than the ILI data and those corresponding to Scenario I but less than those corresponding to the base case.. The explanation for these observations is the same as those for Fig. 2.3.

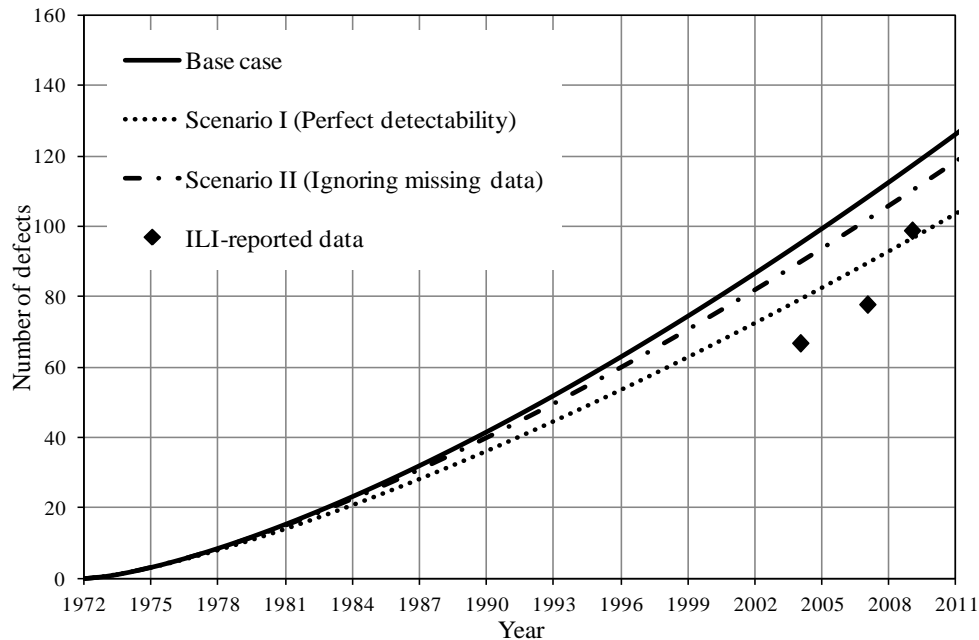
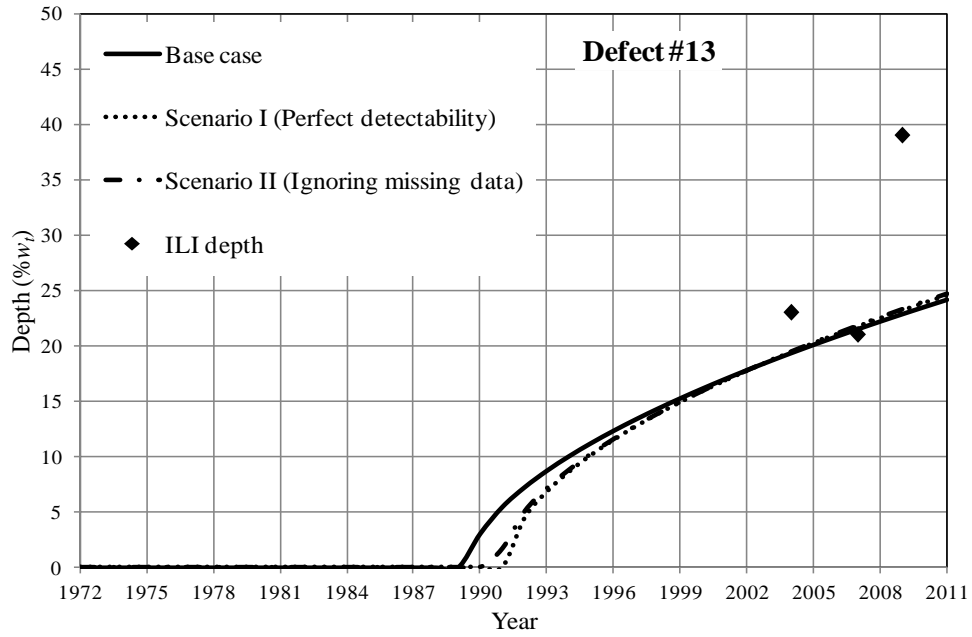
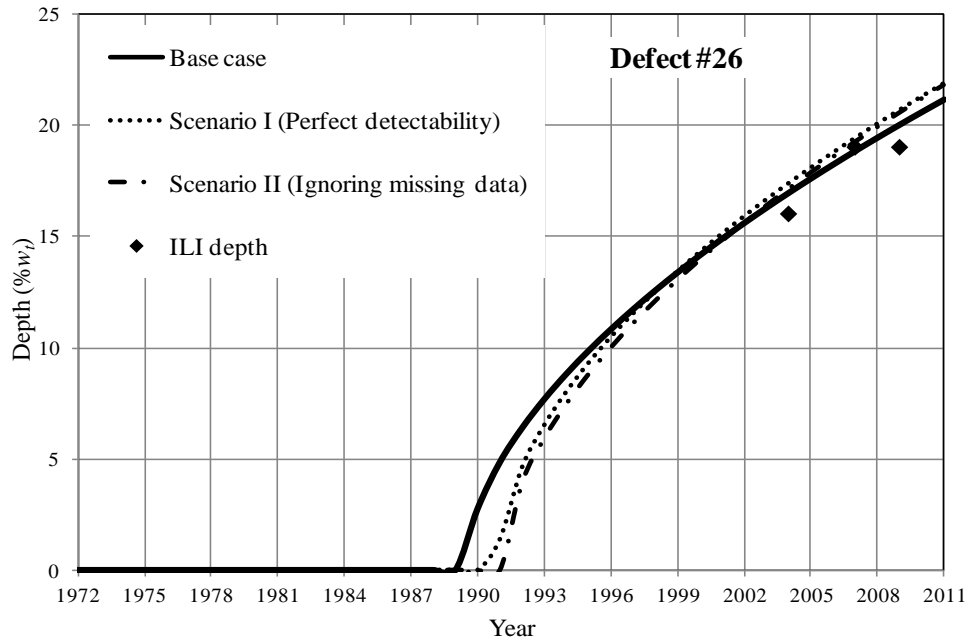


Figure 2.9 Comparison of predicted numbers of defects corresponding to the base case and Scenarios I and II for the relatively low detectability assumption

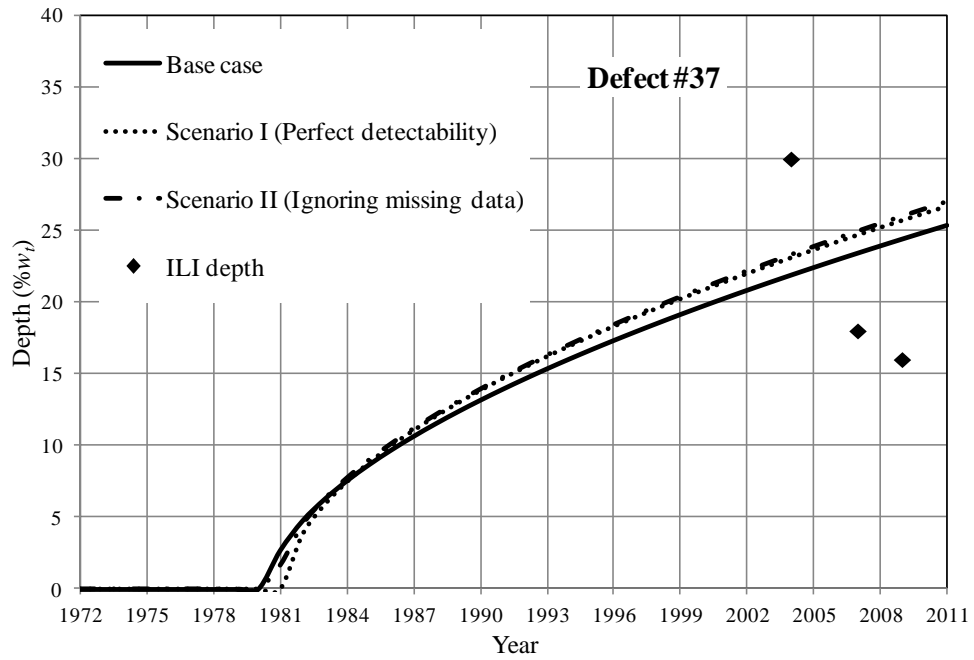
The mean growth paths for Defects #13, #26, #37, #61 and #64 predicted from the base case and Scenarios I and II corresponding to the relatively low detectability assumption are plotted in Figs. 2.10(a) through 2.10(e), respectively. The results indicate that the predicted growth paths corresponding to Scenario I and II are in general slightly higher than those corresponding to the base case. This is because the undetected defects (i.e. missing data), which are typically shallow defects, were included in the Bayesian updating of the parameters of the growth model in the base case. Ignoring the undetected defects in the Bayesian updating leads to that the values of the defect-common parameters (i.e. φ_1 and φ_2) tend to be higher than those of the same parameters by considering the missing data; therefore, the predicted defect growth paths corresponding to Scenario I and II tend to be somewhat higher than those corresponding to the base case for the same defect.



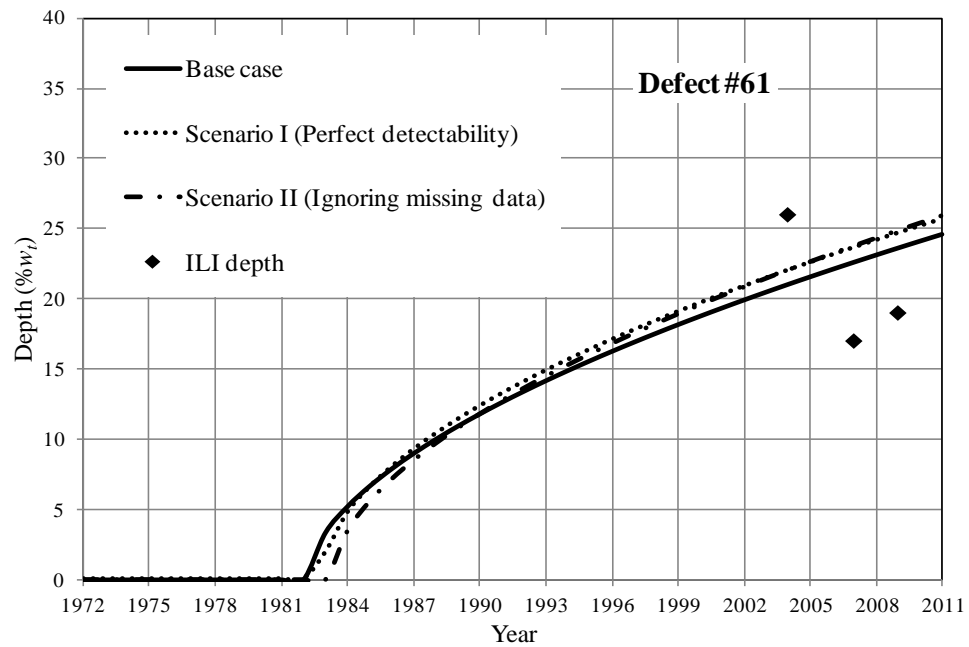
(a) Defect #13



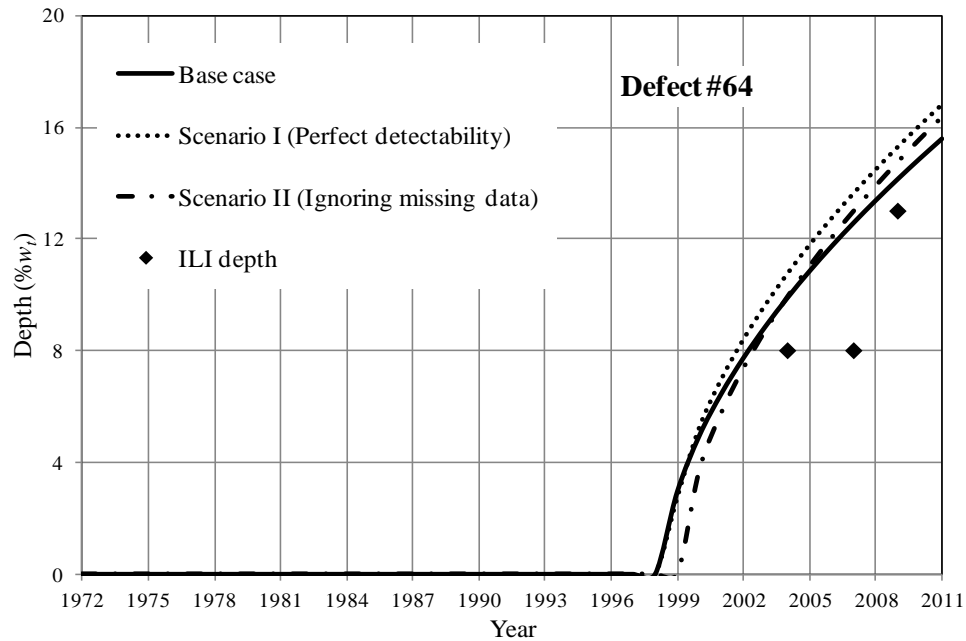
(b) Defect #26



(c) Defect #37



(d) Defect #61



(e) Defect #64

Figure 2.10 Comparison of the growth paths corresponding to the base case and Scenario I and II for the relatively low detectability assumption

2.6 Summary and Conclusions

A stochastic process-based hierarchical Bayesian methodology was proposed to characterize the generation and growth of metal-loss corrosion defects on oil and gas pipelines. The defect generation was characterized by the non-homogenous Poisson process, and the growth of the defect depth was modeled by the non-homogenous gamma process with a time-dependent shape parameter and a time-independent scale parameter. All the model parameters were treated as uncertain variables and evaluated from the Bayesian updating based on the imperfect inspection data obtained from multiple ILI runs. The imperfect detectability of the ILI tool as characterized by POD as well as the measurement errors associated with the ILI data were accounted for in the Bayesian updating. The Markov chain Monte Carlo (MCMC) simulation together with the data augmentation (DA) algorithm to deal with the undetected defects (i.e. missing data) was

employed to evaluate the posterior distributions of the parameters involved in the defect generation and growth models.

The application of the proposed model was illustrated using the simulated inspection data in Example 1 and real inspection data in Example 2. For both examples, three different scenarios were considered, namely including the missing data in calculating the average POD for each inspection, ignoring the missing data in calculating the average POD and assuming perfect detectability for all the inspections (i.e. ignoring POD). Based on the inspection data simulated in accordance with the simplifying assumption involved in the proposed model, namely the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections, the results of the Bayesian updating indicate that the last two scenarios lead to underestimated overall defect population, which is consistent with the model formulations and validates the proposed model. Based on the more realistic simulated inspection data, the results of the Bayesian updating indicate that the first scenario leads to overestimation of the overall defect population because of the simplifying assumption involved in the proposed model. However, the degree of overestimation is marginal as long as the detectability of the inspection tool is high or relatively high, which is typically the case for high-resolution inspection tools employed on oil and gas pipelines nowadays. The overestimation of the defect population due to the simplifying assumption is somewhat compensated by the overestimation of the average POD in the second scenario; as a result, the predicted overall defect population agrees well with the actual population. The last scenario leads to underestimated overall defect population. On the other hand, the growth paths of the detected defects predicted by the three scenarios differ only slightly, with the predictions corresponding to the latter two scenarios being slightly higher than those corresponding to the first scenario for shallow defects.

References

Al-Amin, M., Zhou, W., Zhang, S., Kariyawasam, S., & Wang, H. (2012). Bayesian Model for Calibration of ILI Tools. In: Proceedings of the 9th International Pipeline Conference (pp. 201-208), ASME, New York.

Ang, A. H. S., & Tang, W. H. (2007). *Probability Concepts in Engineering Planning and Design: Emphasis on Applications in Civil & Environmental Engineering*. John Wiley & sons.

Beichelt, F., & Fatti, L. P. (2002). *Stochastic Processes and Their Applications*. Taylor & Francis.

Caleyo, F., Velázquez, J. C., Valor, A., & Hallen, J. M. (2009). Markov chain modelling of pitting corrosion in underground pipelines. *Corrosion Science*, 51(9), 2197-2207.

Fuller W.A. (1987). *Measurement Error Models*. John Wiley & sons.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (2004) *Bayesian Data Analysis*, (2nd edition). Chapman & Hall/CRC.

Gilks, W. R. (2005). *Markov Chain Monte Carlo*. John Wiley & sons.

Hong, H. P. (1999). Inspection and maintenance planning of pipeline under external corrosion considering generation of new defects. *Structural Safety*, 21(3), 203-222.

Jaech J.L. (1985). *Statistical Analysis of Measurement Errors*. John Wiley & sons.

Kariyawasam, S., & Peterson, W. (2008). Revised corrosion management with reliability based excavation criteria. In: *Proceedings of the 7th International Pipeline Conference* (pp. 489-500), ASME, New York.

Kulkarni, V. G. (1995). *Modeling and Analysis of Stochastic Systems*. Chapman & Hall/CRC.

Kuniewski, S. P., van der Weide, J. A., & van Noortwijk, J. M. (2009). Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection. *Reliability Engineering & System Safety*, 94(9), 1480-1490.

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd edition). Chapman & Hall/CRC.

Lunn, D., Spiegelhalter, D., Thomas, A. and Best, N. (2009). The BUGS project: evolution, critique and future directions (with discussion). *Statistics in Medicine*, 28: 3049-82.

Maes, M.A., Faber, M.H and Dann M.R. (2009a). Hierarchical Modeling of Pipeline defect growth subject to ILI uncertainty, In: *Proceedings of the 27th Offshore Ocean and Arctic Engineering Conference*, Honolulu, US, June, OOAEE2009-79470, 12pp.

Maes, M.A., Dann, M.R. and Breitung, K.W. (2009b). Reliability-based integrity assessment of structural systems subject to heterogeneous deterioration and large-scale inspection uncertainty, In: *International Conference on Structural Safety and Reliability (ICOSSAR)*, Osaka, Japan, September 2009.

Ntzoufras, I. (2011). *Bayesian modeling using WinBUGS*. John Wiley & sons.

Rubin, D. B. (2009). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & sons.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398), 528-540.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22, 1701-1728.

Timashev, S. A., Malyukova, M. G., Poluian, L. V., & Bushinskaya, A. V. (2008). Markov description of corrosion defects growth and its application to reliability based inspection and maintenance of pipelines. In: *Proceedings of the 7th International Pipeline Conference* (pp. 525-533), ASME, New York.

Valor, A., Caleyó, F., Alfonso, L., Rivas, D., & Hallen, J. M. (2007). Stochastic modeling of pitting corrosion: a new model for initiation and growth of multiple corrosion pits. *Corrosion science*, 49(2), 559-579.

Zhang, S., Zhou, W., Al-Amin, M., Kariyawasam, S., & Wang, H. (2014). Time-Dependent Corrosion Growth Modeling Using Multiple ILI Data. *Journal of Pressure Vessel Technology*, 136, 041202, 1-7.

Zhang, S., & Zhou, W. (2013). System reliability of corroding pipelines considering stochastic process-based models for defect growth and internal pressure. *International Journal of Pressure Vessels and Piping*, 111, 120-130.

Zheng, R., & Ellingwood, B. R. (1998). Role of non-destructive evaluation in time-dependent reliability analysis. *Structural Safety*, 20(4), 325-339.

Chapter 3 Reliability Analysis of Corroding Pipelines Considering the Generation and Growth of Corrosion Defects

3.1 Introduction

Metal-loss corrosion poses a significant threat to the structural integrity of oil and gas pipelines. The corrosion process involves both the generation of new defects and growth of existing defects over time. Both aspects should be taken into consideration in the evaluation of the time-dependent reliability of corroding pipelines.

The majority of reliability evaluations of corroding pipelines reported in the literature employed the random variable-based growth models for corrosion defects (e.g. Ahammed 1998; Caleyó et al. 2002; Amirat et al. 2006; Zhou 2010, 2011). Note that the random variable-based growth model cannot capture the temporal variability involved in the corrosion growth process (Pandey et al. 2009). This drawback of the random variable-based corrosion growth model can be overcome by the stochastic process-based growth model (van Noordwijk and Frangopol 2004; Frangopol et al. 2004). Hong (1999) evaluated the system reliability of corroding pipelines in the context of developing the optimal inspection and maintenance schedule subjected to the reliability constraint. The Markov process was employed to characterize the growth of corrosion defects, and the homogeneous Poisson process was used to model the generation of new defects. Parametric analyses were carried out to illustrate the impact of the probability transition matrix of the Markov process on the system reliability and optimal inspection schedule, whereas the updating of the defect growth and generation models based on the inspection data was not discussed in Hong's study. Recently, a homogeneous gamma process-based growth model was employed in the system reliability evaluation of corroding pipelines (Zhang and Zhou 2013). A Bayesian approach was applied to evaluate the parameters involved in the gamma process-based growth model using the inspection data obtained from multiple in-line inspections (ILIs). The updated parameters were used to predict the growth of depths (i.e. in the through-pipe wall thickness direction) of individual corrosion defects. The predicted depths of the corrosion defects were then employed in the limit state functions to evaluate the time-dependent system reliability of corroding pipelines.

However, the generation of new defects was ignored in Zhang and Zhou's study; that is, the number of corrosion defects in the pipeline was assumed to be fixed over time. It follows that this assumption leads to overestimation of the system reliability of the corroding pipeline.

This chapter presents a methodology to evaluate the time-dependent system reliability of corroding pipelines by incorporating the Bayesian updating of the defect generation and growth models based on the imperfect inspection data. The gamma process and non-homogeneous Poisson process as presented in Chapter 2 were employed to characterize the growth of existing defects and generation of new defects on the pipeline. The parameters of the growth and generation models were evaluated through the Bayesian updating based on the imperfect in-line inspection (ILI) data. The Markov chain Monte Carlo (MCMC) simulation techniques were employed to carry out the updating. The updated defect growth and generation models were then incorporated in the simple Monte Carlo simulation to carry out the system reliability analysis of corroding pipelines in terms of three distinctive failure modes, namely small leak, large leak and rupture. A numerical example is used to illustrate the proposed methodology.

The rest of this chapter is organized as follows. Section 3.2 describes the defect generation and growth models; Section 3.3 presents the Monte Carlo simulation-based methodology to evaluate the time-dependent system reliability of corroding pipelines by considering the generation and growth of corrosion defects; the illustrative numerical example is described in Section 3.4, and the conclusions are presented in Section 3.5.

3.2 Corrosion Generation and Growth Models

3.2.1 Defect generation and growth modeling

The non-homogeneous Poisson process (NHPP) and non-homogenous gamma process (NHGP) as described in Chapter 2 were employed to model the generation and growth of corrosion defects, respectively. Let t (years) ($t = 1, 2, \dots$) denote the time elapsed since the installation of a given pipeline ($t = 0$ indicating the time of installation) and t_{sr} denote the initiation time of defect r ($r = 1, 2, \dots$). Further let t_n denote the time of the most

recent inspection of the pipeline. Pipeline integrity engineers are typically interested in evaluating the system reliability of the pipeline for a given forecasting period T (years) (e.g. $T = 5$ or 10 years) starting from t_n .

It follows from Eq. (2.5) in Chapter 2 that the increase in the number of defects from years $t - 1$ to t , denoted by ΔN_t , follows a Poisson distribution with a probability density function (PDF) given by

$$f_P(\Delta N_t | \lambda, \delta) = \frac{[m(t) - m(t-1)]^{\Delta N_t}}{\Delta N_t!} \exp [m(t) - m(t - 1)] \quad (3.1)$$

where $m(t)$ denotes the expected number of defects generated over the time interval $[0, t]$ and is commonly expressed as $m(t) = \int_0^t \lambda v(s) ds$. The expression $\lambda v(t)$ is the so-called intensity function, with λ and $v(t)$ being the proportionality constant and shape function, respectively. As described in Section 2.3.1 of Chapter 2, we assumed a power-law intensity function: $\lambda v(t) = \lambda \delta t^{\delta-1}$. It follows that $m(t) = \lambda t^\delta$. Both λ and δ were treated as uncertain parameters and updated through the Bayesian inference based on the ILI data. Substituting $m(t) = \lambda t^\delta$ into Eq. (3.1) results in

$$f_P(\Delta N_t | \lambda, \delta) = \frac{[\lambda t^\delta - \lambda (t-1)^\delta]^{\Delta N_t}}{\Delta N_t!} \exp \{[\lambda t^\delta - \lambda (t - 1)^\delta]^{\Delta N_t}\} \quad (3.2)$$

The total number of defects generated from the time of the most recent inspection (t_n) to year $t_n + \tau$ ($\tau = 1, 2, \dots, T$), i.e. the τ^{th} year within the forecasting period, is given by

$$N(t_n + \tau) = N(t_n + \tau - 1) + \Delta N_\tau = \sum_{t=t_n+1}^{t_n+\tau} \Delta N_t \quad (3.3)$$

It follows from Eq. (2.8) in Chapter 2 that the growth of the depth of defect r ($r = 1, 2, \dots$) from years $t - 1$ to t , Δx_{tr} , follows a gamma distribution with PDF given by

$$f_G(\Delta x_{tr} | \Delta \alpha_{tr}, \beta_r) = \frac{\beta_r^{\Delta \alpha_{tr}} \Delta x_{tr}^{\Delta \alpha_{tr}-1} \exp[-\Delta x_{tr} \beta_r]}{\Gamma(\Delta \alpha_{tr})} I_{(0,\infty)}(\Delta x_{tr}) \quad (3.4)$$

where $\Delta \alpha_{tr}$ and β_r are the shape and rate parameters of the corresponding gamma distribution, respectively. As described in Section 2.3.2 of Chapter 2, $\Delta \alpha_{tr}$ was assumed

to equal $\varphi_1(t - t_{sr})^{\varphi_2} - \varphi_1(t - t_{sr} - 1)^{\varphi_2}$ ($t \geq t_{sr} + 1$) and β_r equals to e^{ξ_r} , where ξ_r is the random effect parameter.

The depth of defect r at year $t_n + \tau$ ($\tau = 1, 2, \dots, T$), $x_r(t_n + \tau)$, i.e. the depth at the τ^{th} year within the forecasting period, is given by

$$x_r(t_n + \tau) = x_r(t_n + \tau - 1) + \Delta x_{tr} = \sum_{t=1}^{t_n + \tau} \Delta x_{tr} \quad (3.5)$$

Note that $x_r(t) \equiv 0$ and $\Delta x_{tr} \equiv 0$ for $t \leq t_{sr}$.

3.2.2 Bayesian updating of the defect generation and growth models

The Bayesian updating was employed to make statistical inferences of the parameters of the defect generation and growth models based on the imperfect ILI data, as described in Section 3.2.1. Through the Bayesian updating, the previous knowledge about the corrosion models and the new information contained in the ILI data can be combined to update the knowledge about the defect generation and growth models. The previous knowledge is reflected in the prior distributions of the model parameters; the new information in the ILI data is incorporated in the likelihood functions, and the updated knowledge of the model parameters is reflected in the posterior distributions.

The inspection data obtained from each ILI run consists of both the number of detected defects and measured depths of detected defects. The probability of detection (POD) and measurement error associated with the ILI tool were incorporated in the Bayesian updating. More specifically, the likelihood function for the number of detected defects incorporates the average POD of each ILI run, whereas the likelihood function for the ILI-reported depths incorporates the measurement error.

Because of the computational difficulties involved in the evaluation of the joint posterior distribution of the model parameters in the Bayesian updating, the Markov chain Monte Carlo (MCMC) simulation techniques were used to numerically evaluate the joint posterior distribution. In this study, a hybrid algorithm combining the Metropolis–Hastings (M-H) algorithm and Gibbs sampling was implemented in MatlabTM to carry out the MCMC simulation. The data augmentation (DA) technique (Tanner and Wong 1987)

was also incorporated in the MCMC simulation to deal with the undetected defects (i.e. missing data). Detailed descriptions of the Bayesian updating including the selection of the prior distribution, derivation of the full conditional posterior distributions required by the MCMC simulation algorithm and a step-by-step procedure to combine the MCMC simulation with DA are given in Chapter 2.

3.3 Time-dependent System Reliability Analysis

3.3.1 Limit state functions

A pressurized pipeline at a corrosion defect may fail by two distinctive failure mechanisms, namely the small leak and burst (Zhou 2010). For natural gas pipelines (as opposed to liquid pipelines), a burst can be further classified as a large leak or rupture (Zhou 2010) depending on if the through-wall defect resulting from the burst extends unstably along the longitudinal (axial) direction of the pipeline.

The limit state function for a given corrosion defect penetrating the pipe wall at time t , $g_1(t)$, is

$$g_1(t) = 0.8w_t - x(t) \quad (3.6)$$

where w_t denotes the wall thickness of the pipeline, and $x(t)$ denotes the depth of the corrosion defect at time t . The use of $0.8w_t$ as opposed to w_t in the above equation is consistent with typical industry practice, as a remaining ligament thinner than $0.2w_t$ is considered prone to developing cracks that could lead to leaks (Al-Amin and Zhou 2014).

The limit state function for plastic collapse under the internal pressure at the defect at time t , $g_2(t)$, is given by

$$g_2(t) = r_b(t) - p \quad (3.7)$$

where $r_b(t)$ denotes the burst pressure capacity of the pipe at the defect at time t , and p is the internal pressure of the pipeline and assumed to be time-independent in this study.

The limit state function for the unstable axial extension of the through-wall defect that results from the burst, $g_3(t)$, is given by

$$g_3(t) = r_{rp}(t) - p \quad (3.8)$$

where $r_{rp}(t)$ is the pressure capacity of the pipeline at the location of the through-wall defect resulting from the burst at time t . A burst is classified as a rupture if $g_3(t) \leq 0$; otherwise, it is defined as a large leak.

Given Eqs. (3.6), (3.7) and (3.8), the cumulative probabilities of small leak, large leak and rupture within a time interval $(0, t]$, denoted by $P_{sl}(t)$, $P_{ll}(t)$ and $P_{rp}(t)$, respectively, are defined as follows (Al-Amin and Zhou 2014)

$$P_{sl}(t) = \text{Prob}[g_1(t) \leq 0 \cap g_2(t) > 0] \quad (3.9a)$$

$$P_{ll}(t) = \text{Prob}[g_1(t) > 0 \cap g_2(t) \leq 0 \cap g_3(t) > 0] \quad (3.9b)$$

$$P_{rp}(t) = \text{Prob}[g_1(t) > 0 \cap g_2(t) \leq 0 \cap g_3(t) \leq 0] \quad (3.9c)$$

where $\text{Prob}[\bullet]$ denotes the probability of an event, and “ \cap ” represents the intersection (i.e. joint event). Note that the probability of a small leak occurring first then followed by a burst was considered extremely small and therefore ignored in the analysis.

3.3.2 Burst and rupture pressure capacity models

In this study, we adopted the B31G modified criterion (Kiefner and Vieth 1989) to evaluate the burst pressure of a pipeline containing a single corrosion defect. The burst pressure capacity r_b is given by

$$r_b = \zeta_b \frac{2w_t(\sigma_y + 68.95)}{D} \left(\frac{1 - 0.85 \frac{x}{w_t}}{1 - 0.85 \frac{x}{Mw_t}} \right) \quad \frac{x}{w_t} \leq 0.8 \quad (3.10a)$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \frac{L^2}{Dw_t} - 0.003375 \frac{L^4}{(Dw_t)^2}} & \frac{L^2}{Dw_t} \leq 50 \\ 3.293 + 0.032 \frac{L^2}{Dw_t} & \frac{L^2}{Dw_t} > 50 \end{cases} \quad (3.10b)$$

where σ_y is the yield strength of the pipe material; $\sigma_y + 68.95$ (MPa) (i.e. $\sigma_y + 10$ ksi) is the flow stress; ζ_b is the model error; D is the outside diameter of the pipeline; L is the defect length, i.e. in the longitudinal direction of the pipeline, and M is the Folias factor or bulging factor.

The flow stress-dependent failure criterion for a through-wall flaw developed by (Kiefner et al. 1973) was employed in this study to calculate r_{rp} as follows:

$$r_{rp} = \frac{2w_t\sigma_f}{MD} \quad (3.11)$$

where σ_f is the flow stress and equals $\sigma_y + 68.95$ (MPa), and M is the Folias factor given by Eq. (3.10b). Due to a lack of information in the literature, the model error associated with Eq. (3.11) was ignored in the analysis.

3.3.3 Basic assumptions and analysis procedures

A pipeline containing multiple corrosion defects is a series system because failure at any defect causes the failure of the pipeline. For simplicity, only the defect depth and the number of defects were treated as time-dependent in this study, whereas all the other uncertain parameters (e.g. the defect length and internal pressure) involved in the limit state functions were assumed to be time-independent. Furthermore, the spatial variability of the pipe geometry (i.e. diameter and wall thickness) and material property (i.e. yield strength), internal pressure and model error associated with burst capacity model was ignored; that is, these parameters at different defects were assumed to be fully correlated. Finally, the defect lengths associated with different defects were assumed to be independent. It then follows that the probability of small leak is governed by the maximum value of the depths of all the defects, and the probability of burst is governed by the minimum value of the burst pressure capacities associated with all the defects.

The growth of the depth of individual defect was characterized by the NHGP and the generation of new defects was modeled by the NHPP described in Section 3.2.1. The posterior median values of the model parameters involved in the generation and growth models were evaluated from the ILI data through the Bayesian updating described in

Section 3.2.2, and further employed in a simple Monte Carlo simulation to evaluate the time-dependent system reliability of a corroding pipeline.

Consider that a total of N_e defects have been detected and sized by the inspection tools on a given pipeline. A simple Monte Carlo simulation-based approach was developed to evaluate the reliability of the pipeline after the most recent inspection considering the growth of existing defects and generation of new defects. The analysis procedure involved in a given simulation trial is described as follows. Note that the depths of the undetected defects generated prior to the most recent inspection are generally small compared to the detected defects. Therefore, these undetected defects were ignored in the reliability analysis because the system reliability is governed by the critical defects with relatively large depths and/or high growth rates.

1. Generate samples for σ_y , w_t and D , lengths of the j^{th} existing defects L_j ($j = 1, 2, \dots, N_e$), the internal pressure p and the model error for the burst capacity model ζ_b from the corresponding probabilistic distributions.

- 2 Start from the forecasting time $\tau = 1$ (year)

- 2.1) Obtain the depth of the j^{th} existing defect at year $t_n + \tau$, $x_j(t_n + \tau)$, based on Eq. (3.5). In Eq. (3.5), Δx_{t_j} is generated from the gamma distribution given by Eq. (3.4) with the values of φ_1 , φ_2 , t_{sj} and ζ_j equal to the corresponding median values of their posterior distributions updated from the ILI data through the Bayesian methodology.

- 2.2) Obtain the number of new defects generated from year t_n to year $t_n + \tau$, $N(t_n + \tau)$, based on Eq. (3.3). In Eq. (3.3), ΔN_τ is generated from the Poisson distribution given by Eq. (3.2) with the values of λ and δ equal to their corresponding posterior median values resulting from the Bayesian updating.

- 2.3) For simplicity and to be slightly conservative, set the initiation time t_{sk} of the k^{th} ($k = 1, 2, \dots, \Delta N_\tau$) new defect generated within year $t_n + \tau$ ($\tau = 1, 2, \dots, T$) equal to $t_n + \tau - 1$.

2.4) Generate the random effect parameter, ξ_k , of the k^{th} ($k = 1, 2, \dots, \Delta N_t$) new defect generated within the year $t_n + \tau$ ($\tau = 1, 2, \dots, T$) from the normal distribution $f_N(0, \sigma_\xi^2)$ with σ_ξ^2 equal to the corresponding posterior median value obtained from the Bayesian updating.

2.5) Generate the depth of the k^{th} new defect ($k = 1, 2, \dots, N(t_n + \tau)$) at year $t_n + \tau$, $x_k(t_n + \tau)$, based on Eq. (3.5). In Eq. (3.5), $\Delta x_{\tau k}$ is generated from the gamma distribution given by Eq. (3.4) with the values of φ_1 and φ_2 equal to the corresponding posterior median values, and the values of t_{sk} and ξ_k were obtained from steps 2.3) and 2.4), respectively.

2.6) Generate the lengths for all the newly generated defects from the corresponding probabilistic distribution.

2.7) Calculate $g_1 = 0.8w_t - \max\{\mathbf{x}(t_n + \tau)\}$, where $\mathbf{x}(t_n + \tau)$ is the vector of depths of all the existing and newly generated defects. Set the depth $x_i(t_n + \tau)$ ($i = 1, 2, \dots, N(t_n + \tau) + N_e$) to be $80\%w_t$ if $x_i(t_n + \tau) > 80\%w_t$.

2.8) Substitute the values of w_t , D , σ_y , \mathbf{L} , ζ_b and $\mathbf{x}(t_n + \tau)$ into the B31G modified model given by Eq. 3.10(a) to calculate the burst pressure capacities at different defects, where \mathbf{L} is the vector of lengths of all the existing and newly generated defects. Calculate $g_2 = \min\{\mathbf{r}_b\} - p$, where \mathbf{r}_b denotes the vector of burst pressure capacities at the considered defects.

2.9) If $g_1 > 0$ and $g_2 > 0$, set $\tau = \tau + 1$. Go to step 2.1) if $\tau \leq T$, and start a new simulation trial otherwise.

2.10) If $g_1 \leq 0$ and $g_2 > 0$, set $sl(\tau) = sl(\tau) + 1$, where $sl(\tau)$ is the counter for small leak. Set $\tau = \tau + 1$. Go to step 2.1) if $\tau \leq T$, and start a new simulation trial otherwise.

2.11) If $g_2 \leq 0$, calculate $g_3 = r_{rpm} - p$, where r_{rpm} is the rupture pressure at the defect that has the lowest burst pressure at year $t_n + \tau$. If $g_3 > 0$, set $ll(\tau) = ll(\tau) + 1$, where $ll(\tau)$ is the counter for large leak; if $g_3 \leq 0$, set $rp(\tau) = rp(\tau) + 1$, where $rp(\tau)$ is the counter for rupture. Set $\tau = \tau + 1$. Go to step 2.1) if $\tau \leq T$, and start a new simulation trial otherwise.

By repeating steps 1 to 2 for the desired number of simulation trials, N_{sim} , the probabilities of small leak, large leak and rupture up to year τ ($\tau = 1, 2, \dots, T$), can be evaluated as follows:

$$P_{sl}(\tau) \approx \frac{\sum_{t=1}^{\tau} sl(t)}{N_{sim}} \quad (3.12a)$$

$$P_{ll}(\tau) \approx \frac{\sum_{t=1}^{\tau} ll(t)}{N_{sim}} \quad (3.12b)$$

$$P_{rp}(\tau) \approx \frac{\sum_{t=1}^{\tau} rp(t)}{N_{sim}} \quad (3.12c)$$

3.4 Example

3.4.1 General information

In this section, the same natural gas pipe joint described in Section 2.5.2 of Chapter 2 is used to illustrate the proposed system reliability evaluation methodology for corroding pipelines. The pipe joint was made from API 5L X52 steel with a specified minimum yield strength (SMYS) of 359 MPa and a specified minimum tensile strength (SMTS) of 455 MPa, and has an outside diameter of 508 mm, an operating pressure of 5.66 MPa and a nominal wall thickness of 5.56 mm. The joint is 13.6 m long and has been inspected by high-resolution ILI tools in 2004, 2007 and 2009. After the ILI in 2009, a total of 99 defects have been detected and sized by the inspection tools. A summary of the depths of the detected defects is shown in Table 2.6 of Chapter 2. The measurement errors associated with the three ILIs employed in 2004, 2007 and 2009 are presented in Section 2.5.2. The probabilistic characteristics of the random variables involved in the reliability analysis are listed in Table 3.1. Note that the lengths (L) of individual defects were assumed to be independent and identically distributed with a mean of 50 mm and a coefficient of variation (COV) of 15.6% (Zhou 2010).

Table 3.1 Probabilistic characteristics of the basic random variables

Parameter	Unit	Distribution type	Nominal value	Mean	COV (%)	Source
D	mm	Normal	508	508	0.06	CSA (2007)
w_t	mm	Normal	5.56	5.56	4.5	Jiao et al. (1995)
L	mm	Normal	50	50	15.6	Zhou (2010)
σ_y	MPa	Normal	359	398	3.4	Al-Amin and Zhou (2014)
p	MPa	Gumbel	5.66	5.93	2.0	CSA (2007)
ζ_b	--	Gumbel	1.00	1.297	25.8	Zhou and Huang (2012)

3.4.2 Results

The posterior median values of the parameters of the defect generation and growth models have been evaluated from the Bayesian updating based on the three sets of ILI data as described in Chapter 2. These values are summarized in Table 2.5 of Chapter 2 corresponding to the assumptions of the high and relatively low detectability of the ILI tools. They were then employed in the simple Monte Carlo simulation procedure described in Section 3.3.3 to evaluate the time-dependent system reliability of the pipe joint. For the purpose of illustration, we only employed the posterior median values of the model parameters corresponding to the high detectability assumption (i.e. a POD of 90% for a defect depth of $5\%w_t$ with a detection threshold of $1\%w_t$) to evaluate the system reliability.

A total of 1,000,000 simulation trials were carried out to evaluate the probabilities of small leak, large leak and rupture over an 11-year forecasting period since the most recent inspection, i.e. from years 2010 to 2020. The probabilities of small leak, large leak and rupture over the forecasting period considering the growth of existing defects as well as the generation and growth of new defects were calculated, and the results are shown in Fig. 3.1. Figure 3.1 indicates that the probability of small leak increases the fastest with time compared with the probabilities of large leak and rupture. The probability of large leak is the highest of those of the three failure modes.

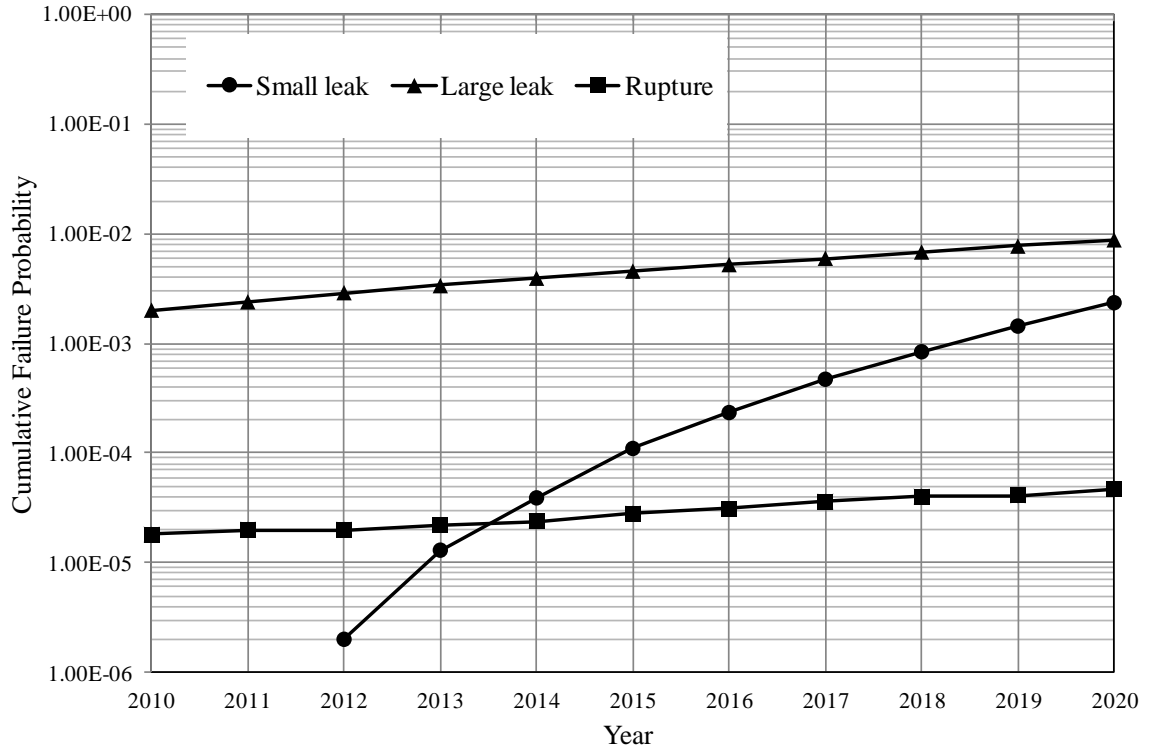
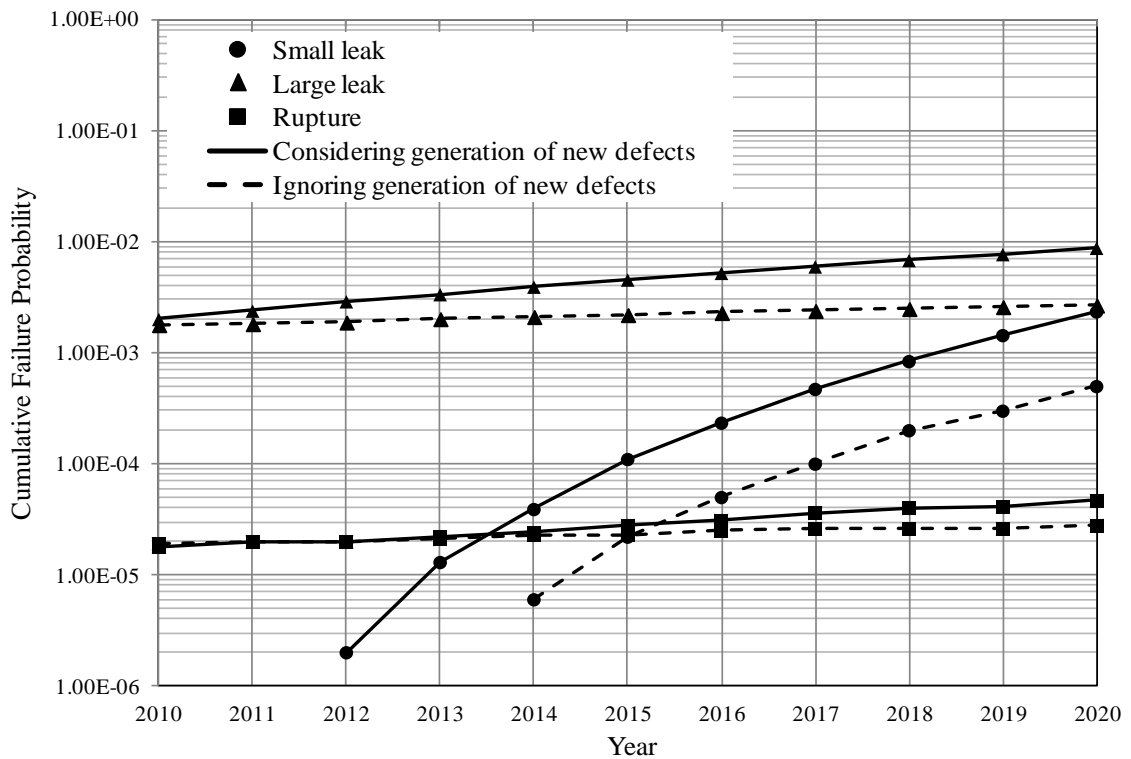


Figure 3.1 Cumulative probabilities of small leak, large leak and rupture

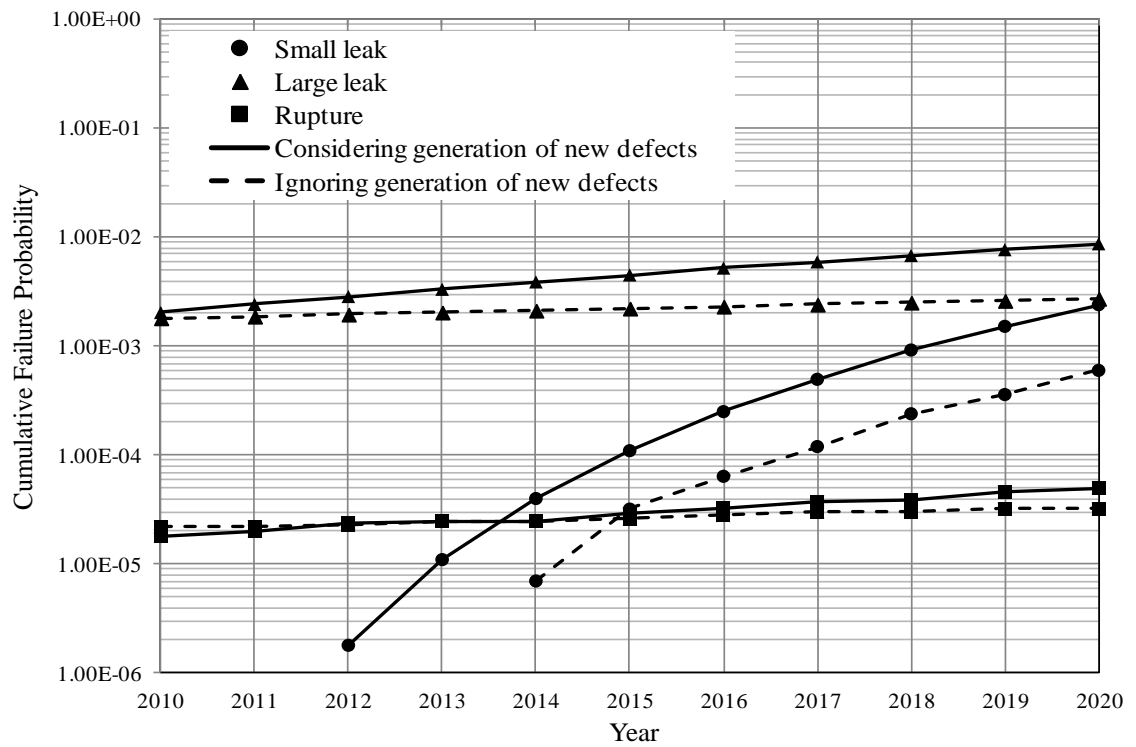
3.4.2 Sensitivity analysis

During the forecasting period, the depths of the existing defects increase with time. At the same time, new defects initiate and start to grow once initiated. Both the existing and new defects therefore impact the system reliability evaluated. In addition to the base case scenario where both POD and the undetected defects (i.e. missing data) were considered in the Bayesian updating of the model parameters, two additional scenarios were considered to investigate the impact of POD and the undetected defects on the system reliability. The two scenarios are the same as those considered in Section 2.5 of Chapter 2; that is, Scenario I assumes perfect detectability for all three inspections, whereas Scenario II accounts for POD but ignores the missing data in calculating the average POD in each inspection. The posterior median values of the model parameters corresponding to the base case, Scenario I and II were then used to carry out the reliability analysis.

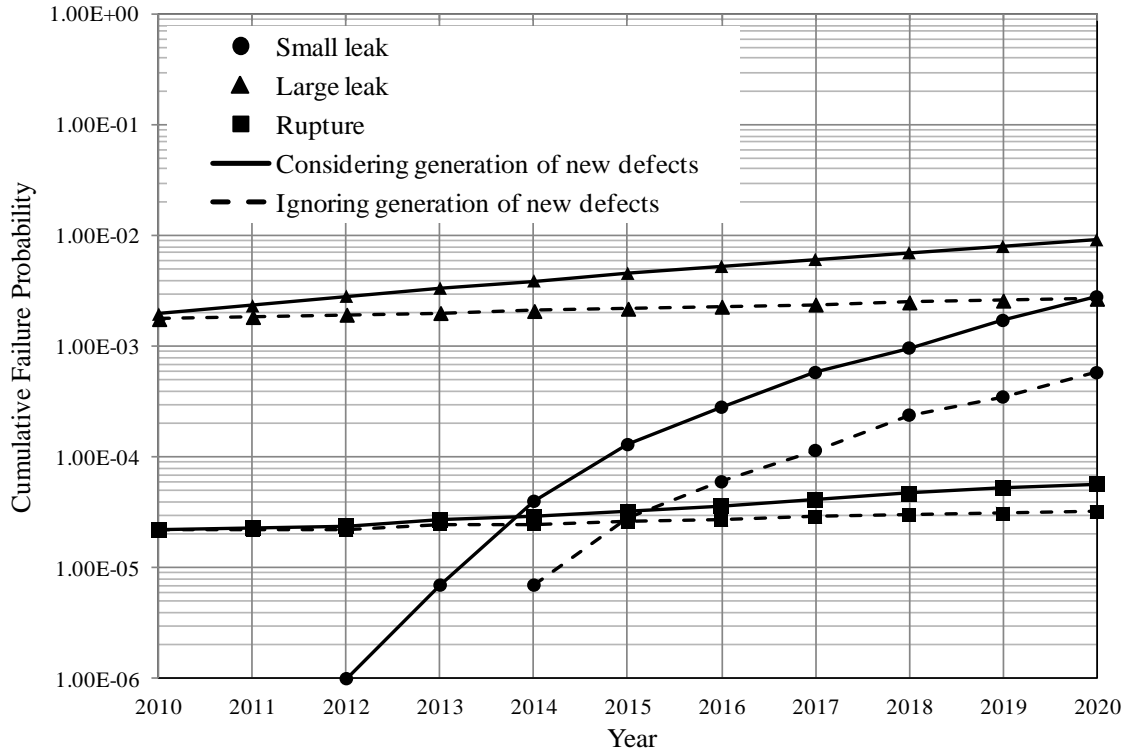
The probabilities of small leak, large leak and rupture corresponding to the base case, Scenario I and II were evaluated by considering and ignoring the generation of new defects, and are compared in Fig. 3.2. Figure 3.2 indicates, as expected, that considering the generation of new defects results in higher probabilities of small leak, large leak and rupture than ignoring the generation of new defects. The new defects have a greater impact on the probability of small leak than on the probabilities of large leak and rupture. This is because the probability of small leak is governed by only two variables (i.e. the pipe wall thickness and defect depth) as shown in the limit state function g_1 , and is therefore more sensitive to the uncertainty in the depths of newly generated defects.



(a) Base case



(b) Scenario I (perfect detectability)



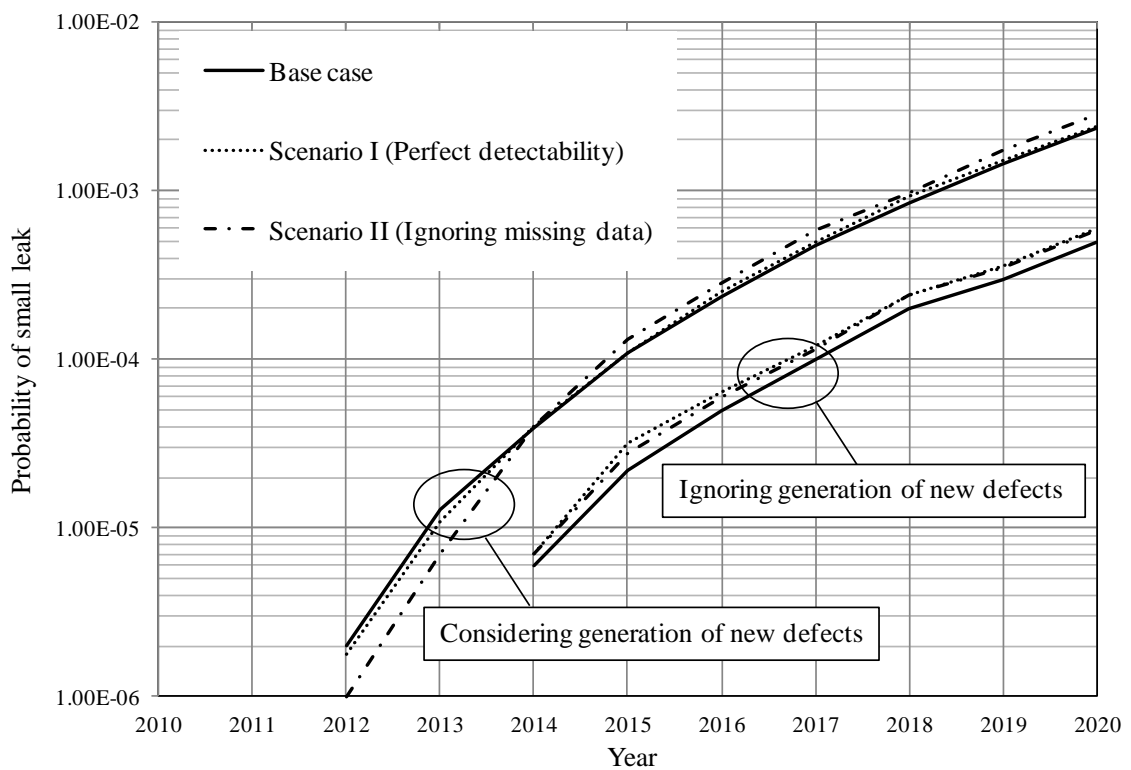
(c) Scenario II (ignoring missing data)

Figure 3.2 Impact of generation of new defects on the system reliability corresponding to the base case, Scenario I and II

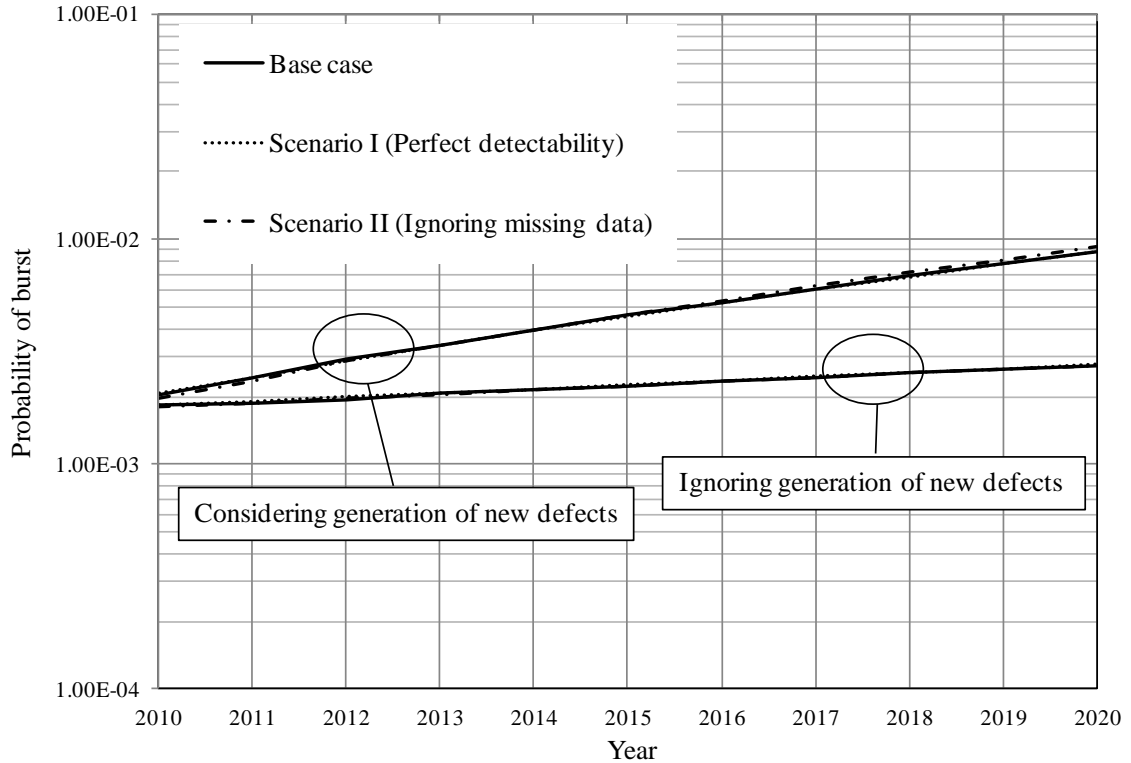
To better compare the failure probabilities associated with the base case and Scenarios I and II, the probabilities of small leak and burst (i.e. the summation of probabilities of large leak and rupture) corresponding to the three cases are depicted in Figs. 3.3(a) and 3.3(b), respectively. In each of the two figures, the probabilities of small leak or burst were evaluated by ignoring and considering the generation of new defects.

Figure 3.3 indicates that the probabilities of burst corresponding to the three scenarios are practically identical, whereas the probabilities of small leak corresponding to the three scenarios are marginally different. The fact that different analysis scenarios have a larger impact on the probabilities of small leak than the probabilities of burst is mainly attributed to that the probability of small leak is governed by the uncertainties in only two variables (i.e. wall thickness and defect depth) and therefore more sensitive to the differences in the defect generation and growth models corresponding to the three

scenarios. If the generation of new defects is considered, Scenario II leads to the highest probabilities of small leak for the majority of the forecasting period. This is because that Scenario II results in on one hand a higher defect generation intensity than Scenario I and on the other hand slightly higher predicted depths of both existing and new defects than the base case. If the generation of new defects is ignored, the base case leads to the lowest probabilities of small leak for the forecasting period. This is because the predicted depths of existing defects corresponding to the base case are slightly lower than those predicted by Scenario I and II as observed in Chapter 2.



(a) Probabilities of small leak



(b) Probabilities of burst

Figure 3.3 Comparison of probabilities of small leak and burst corresponding to the base case, Scenario I and II

3.5 Summary and Conclusions

This chapter presents a methodology to evaluate the time-dependent system reliability of corroding pipelines over a given forecasting period starting from the most recent inspection by simultaneously considering the generation and growth of corrosion defects. The generation of corrosion defects was modeled by the non-homogenous Poisson process, whereas the growth of defects was characterized by the non-homogenous gamma process. The parameters involved in the generation and growth models were evaluated from the imperfect ILI data through the Bayesian updating, and the corresponding posterior median values of these model parameters were then used in a simple Monte Carlo simulation to evaluate the system reliability of corroding pipelines considering three different failure modes, i.e. small leak, large leak and rupture.

An example that involves three sets of ILI data for a pipe joint within a natural gas pipeline located in Alberta was used to illustrate the proposed methodology. The results of the reliability analysis indicate that ignoring the generation of new defects leads to underestimations of the probabilities of small leak, large leak and rupture. The generation of new defects has the highest impact on the probability of small leak. The impact on the evaluated system reliability due to three different approaches to deal with POD associated with the inspection tool in the Bayesian updating of the defect generation and growth models was also examined. The analysis results suggest that these three approaches have a marginal impact on the probabilities of small leak, but practically no impact on the probabilities of burst. The methodology developed in this study could be implemented in a decision support tool to facilitate the pipeline corrosion management.

References

Ahamed, M. (1998). Probabilistic estimation of remaining life of a pipeline in the presence of active corrosion defects. *International Journal of Pressure Vessels and Piping*, 75(4), 321-329.

Al-Amin, M., & Zhou, W. (2014). Evaluating the system reliability of corroding pipelines based on inspection data. *Structure and Infrastructure Engineering*, 10(9), 1161-1175.

Amirat, A., Mohamed-Chateaneuf, A., & Chaoui, K. (2006). Reliability assessment of underground pipelines under the combined effect of active corrosion and residual stress. *International Journal of Pressure Vessels and Piping*, 83(2), 107-117.

Caleyo, F., Gonzalez, J. L., & Hallen, J. M. (2002). A study on the reliability assessment methodology for pipelines with active corrosion defects. *International Journal of Pressure Vessels and Piping*, 79(1), 77-86.

CSA. CSA-Z662 Oil and Gas Pipeline Systems. Canadian Standards Association, Rexdale, Ontario 2007.

Frangopol, D. M., Kallen, M. J., & van Noortwijk, J. M. (2004). Probabilistic models for life-cycle performance of deteriorating structures: review and future directions. *Progress in Structural Engineering and Materials*, 6(4), 197-212.

Hong H. P. (1999). Inspection and maintenance planning under external corrosion considering generation of new defects. *Structural Safety*, 21, 203-222.

Jiao, G., Sotberg, T., & Igland, R. (1995). SUPERB 2M—statistical data: basic uncertainty measures for reliability analysis of offshore pipelines. Superb Joint Industry Project Restricted Report, Report No. STF70 F, 95212.

Kiefner, J. F., Maxey, W. A., Eiber, R. J., & Duffy, A. R. (1973). Failure stress levels of flaws in pressurized cylinders. *ASTM Special Technical Publication*, (536), 461-481.

Kiefner, J. F., & Vieth, P. H. (1989). A modified criterion for evaluating the remaining strength of corroded pipe (No. PR-3-805). Battelle Columbus Div., OH (USA).

Pandey, M. D., Yuan, X. X., & van Noortwijk, J. M. (2009). The influence of temporal uncertainty of deterioration on life-cycle management of structures. *Structure and Infrastructure Engineering*, 5(2), 145-156.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398), 528-540.

van Noortwijk, J. M., & Frangopol, D. M. (2004). Two probabilistic life-cycle maintenance models for deteriorating civil infrastructures. *Probabilistic Engineering Mechanics*, 19(4), 345-359.

Zhang, S., & Zhou, W. (2013). System reliability of corroding pipelines considering stochastic process-based models for defect growth and internal pressure. *International Journal of Pressure Vessels and Piping*, 111, 120-130.

Zhou, W. (2010). System reliability of corroding pipelines. *International Journal of Pressure Vessels and Piping*, 87(10), 587-595.

Zhou, W. (2011). Reliability evaluation of corroding pipelines considering multiple failure modes and time-dependent internal pressure. *Journal of Infrastructure Systems*, 17(4), 216-224.

Zhou, W., & Huang, G. X. (2012). Model error assessments of burst capacity models for corroded pipelines. *International Journal of Pressure Vessels and Piping*, 99, 1-8.

Chapter 4 Conclusions and Recommendations for Future Study

4.1 Probabilistic Modeling and Bayesian Inference of Metal-loss Corrosion

To account for the temporally and spatially variable nature of the metal-loss corrosion on steel oil and gas pipelines, stochastic process-based models are presented in Chapter 2 to characterize the generation and growth of corrosion defects. A Bayesian framework was established to make inference of the corrosion models based on imperfect in-line inspection (ILI) data.

The defect generation was characterized by the non-homogenous Poisson process (NHPP), and the growth of the defect depth was modeled by the non-homogenous gamma process (NHGP) with a time-dependent shape parameter and a time-independent scale parameter. The defect generation and growth models were formulated in a hierarchical Bayesian framework, which allows the consideration of the probability of detection (POD) and measurement errors associated with the ILI data in the statistical inference of the corrosion models. A simplifying assumption was employed in the Bayesian formulation; that is, the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections. The Bayesian updating of these model parameters were evaluated using the Markov chain Monte Carlo (MCMC) simulation based on the imperfect ILI data. The data augmentation (DA) technique was employed in conjunction with the MCMC simulation to deal with the undetected defects (i.e. missing data) in the Bayesian updating resulting from the imperfect detectability of the ILI tool.

The application of the proposed model was validated using the simulated inspection data in Example 1 of Chapter 2 and further illustrated using real ILI data in Example 2 of Chapter 2. For both examples, three different scenarios were considered, namely including the missing data in calculating the average POD for each inspection, ignoring the missing data in calculating the average POD and assuming perfect detectability for all the inspections (i.e. ignoring POD). Based on the simulated inspection data with the simplifying assumption involved in the proposed model, the results indicate that the last

two scenarios lead to underestimated overall defect population, which validates the proposed Bayesian framework. Based on the more realistic simulated inspection data (i.e. part of the newly detected defects may be previously undetected defects), the results indicate that the first scenario leads to overestimation of the overall defect population because of the simplifying assumption involved in the proposed model. However, the degree of overestimation is marginal as long as the detectability of the ILI tool is high or relatively high, which is typically the case for high-resolution ILI tools employed on oil and gas pipelines nowadays. The overestimation of the defect population due to the simplifying assumption is somewhat compensated by the overestimation of the average POD in the second scenario; as a result, the predicted overall defect population agrees the best with the actual defect population among the three scenarios. In Example 2, it is observed that the growth paths of the detected defects predicted by the three scenarios differ only slightly, with the predictions corresponding to the latter two scenarios being slightly higher than those corresponding to the first scenario.

4.2 Time-dependent System Reliability Analysis of Corroding Pipelines

A simple Monte Carlo simulation-based methodology is described in Chapter 3 to evaluate the time-dependent system reliability of corroding pipelines containing multiple active corrosion defects for a given forecasting period starting from the most recent inspection. This methodology incorporates the corrosion generation and growth models evaluated from the imperfect ILI data through the Bayesian updating as described in Chapter 2. The time-dependent system reliability of corroding pipelines is evaluated by simultaneously considering the generation and growth of corrosion defects in terms of three distinctive failure modes, namely small leak, large leak and rupture.

The proposed methodology was illustrated using the same pipe joint as described in Example 2 of Chapter 2. The results of the reliability analysis indicate that ignoring the generation of new defects in the reliability analysis leads to underestimation of the system reliability and has the largest impact on the probability of small leak. Furthermore, the three different approaches to consider POD in the Bayesian updating, i.e. including the missing data in calculating the average POD, ignoring the missing data

in calculating the average POD and ignoring POD, somewhat impact the probability of small leak. Ignoring the missing data in calculating the average POD in the Bayesian updating of the corrosion growth and generation models generally leads to slightly higher probabilities of small leak than the other two approaches.

4.3 Recommendations for Future Study

The recommended future studies are described as follows.

First, local covariates such as the type and condition of the coating on the pipe, the effectiveness of the cathodic protection, and moisture level and corrosivity of the surrounding soils, if available, can be explicitly taken into account in the corrosion generation and growth models to improve the predictive capabilities of these models.

Second, the potential spatial correlations between different corrosion defects are worth investigating. To this end, the application of the random field theory (e.g. Khoshnevisan 2002; Bensi 2010; Bensi et al. 2011) could be explored.

Third, a reliability/risk-based decision support tool can be developed to incorporate the Bayesian models and reliability analysis methodologies reported in this study to assist pipeline integrity engineers in developing defensible inspection and maintenance strategies for corroding pipelines.

Finally, the stochastic process-based Bayesian degradation models can be applied to other structures and infrastructure systems subjected to localized deteriorations.

References

Bensi, M. T. (2010). A Bayesian Network Methodology for Infrastructure Seismic Risk Assessment and Decision Support. PhD Thesis, UC Berkeley: Civil and Environmental Engineering.

Bensi, M., Der Kiureghian, A., & Straub, D. (2011). Bayesian network modeling of correlated random variables drawn from a Gaussian random field. *Structural Safety*, 33(6), 317-332.

Khoshnevisan (2002). Multiparameter Processes - An Introduction to Random Fields.
Springer.

Appendix 2A Derivations of Full Conditional Posterior Distributions of Model Parameters

Parameter	Conditional posterior distribution	
$\Delta \mathbf{x}_j$	$p(\Delta \mathbf{x}_j \mathbf{y}_j, \boldsymbol{\Sigma}_{E_j}, \mathbf{a}, \mathbf{b}, t_i, \Delta \boldsymbol{\alpha}, \xi_j, t_{s_j}) \propto f_G(\Delta x_{ij}) * L(\mathbf{y}_j \mathbf{x}_j) = \frac{\beta_j^{\Delta \alpha_{ij}} \Delta x_{ij}^{\Delta \alpha_{ij} - 1} \exp(-\Delta x_{ij} \beta_j)}{\Gamma(\Delta \alpha_{ij})} *$ $(2\pi)^{-\frac{n-l+1}{2}} \boldsymbol{\Sigma}_{E_j} ^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{y}_j - \mathbf{a} - \mathbf{b} \sum_{i=l}^{n-l+1} \Delta x_{ij})^T \boldsymbol{\Sigma}_E^{-1} (\mathbf{y}_j - \mathbf{a} - \mathbf{b} \sum_{i=l}^{n-l+1} \Delta x_{ij})\right] \propto$ $\exp\left[-\frac{1}{2}\left(\mathbf{y}_j - \mathbf{a} - \mathbf{b} \sum_{i=l}^{n-l+1} \Delta x_{ij}\right)^T \boldsymbol{\Sigma}_E^{-1} \left(\mathbf{y}_j - \mathbf{a} - \mathbf{b} \sum_{i=l}^{n-l+1} \Delta x_{ij}\right)\right]$	(A.1)
λ	$p(\lambda \mathbf{N}^{gd}, \delta) \propto f_G(\lambda) * L(\mathbf{N}^{gd} \lambda, \delta) =$ $\frac{h^g \lambda^{g-1} \exp(-\lambda h)}{\Gamma(g)} * \prod_{i=1}^n \frac{[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]^{N_i^{gd}}}{N_i^{gd}!} \exp\{-[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]\} \propto$ $\lambda^{g + \sum_{i=1}^n N_i^{gd} - 1} * \exp\left\{-\lambda \left[\sum_{i=1}^n \overline{\text{POD}}_i (t_i^\delta - t_{i-1}^\delta) + h\right]\right\} \propto$ $f_G(g + \sum_{i=1}^n N_i^{gd}, h + \sum_{i=1}^n \overline{\text{POD}}_i (t_i^\delta - t_{i-1}^\delta))$	(A.2)
δ	$p(\delta \mathbf{N}^{gd}, \lambda) \propto f_G(\delta) * L(\mathbf{N}^{gd} \lambda, \delta) =$ $\frac{\kappa^\eta \delta^{\eta-1} \exp(-\delta \eta)}{\Gamma(\kappa)} * \prod_{i=1}^n \frac{[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]^{N_i^{gd}}}{N_i^{gd}!} \exp\{-[\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]\} \propto$ $\delta^{\kappa-1} \exp(-\delta \eta) * \exp\left[-\lambda \sum_{i=1}^n \overline{\text{POD}}_i (t_i^\delta - t_{i-1}^\delta)\right] * \prod_{i=1}^n [\overline{\text{POD}}_i \lambda (t_i^\delta - t_{i-1}^\delta)]^{N_i^{gd}}$	(A.3)

φ_1	$p(\varphi_1 \Delta\mathbf{x}, \varphi_2, \xi, \mathbf{t}_s) \propto f_G(\varphi_1) * L(\Delta\mathbf{x} \mathbf{t}_s, \varphi_1, \varphi_2, \xi) = \frac{c^d \varphi_1^{c-1} \exp(-\varphi_1 d)}{\Gamma(c)} *$ $\prod_{i=1}^n \prod_{r=1}^{\sum_{l=1}^i (N_l^{gd} + N_l^{gu})} \frac{[\exp(\xi_r)]^{\Delta\alpha_{ir}} (\Delta x_{ir})^{\Delta\alpha_{ir}-1} \exp[-\Delta x_{ir} \exp(\xi_r)]}{\Gamma[\Delta\alpha_{ir}]} \propto$ $\varphi_1^{c-1} \exp(-\varphi_1 d) * \prod_{i=1}^n \prod_{r=1}^{\sum_{l=1}^i (N_l^{gd} + N_l^{gu})} \frac{[\exp(\xi_r)]^{\Delta\alpha_{ir}} \Delta x_{ir}^{\Delta\alpha_{ir}-1}}{\tau[\Delta\alpha_{ir}]}$	(A.4)
φ_2	$p(\varphi_2 \Delta\mathbf{x}, \varphi_1, \xi, \mathbf{t}_s) \propto f_G(\varphi_2) * L(\Delta\mathbf{x} \mathbf{t}_s, \varphi_1, \varphi_2, \xi) = \frac{e^f \varphi_2^{e-1} \exp(-\varphi_2 f)}{\Gamma(e)} *$ $\prod_{i=1}^n \prod_{r=1}^{\sum_{l=1}^i (N_l^{gd} + N_l^{gu})} \frac{[\exp(\xi_r)]^{\Delta\alpha_{ir}} (\Delta x_{ir})^{\Delta\alpha_{ir}-1} \exp[-\Delta x_{ir} \exp(\xi_r)]}{\Gamma[\Delta\alpha_{ir}]} \propto$ $\varphi_2^{e-1} \exp(-\varphi_2 f) * \prod_{i=1}^n \prod_{r=1}^{\sum_{l=1}^i (N_l^{gd} + N_l^{gu})} \frac{[\exp(\xi_r)]^{\Delta\alpha_{ir}} \Delta x_{ir}^{\Delta\alpha_{ir}-1}}{\tau[\Delta\alpha_{ir}]}$	(A.5)
t_{sj}	$p(t_{sj} \Delta\mathbf{x}_j, \varphi_1, \varphi_2, \xi_j) \propto f_U(lb_j, ub_j) * L(\Delta\mathbf{x}_j \mathbf{t}_s, \varphi_1, \varphi_2, \xi_j) =$ $\frac{1}{ub_j - lb_j} * \prod_{i=l}^n \frac{[\exp(\xi_j)]^{\Delta\alpha_{ij}} \Delta x_{ij}^{\Delta\alpha_{ij}-1} \exp(-\Delta x_{ij} \exp(\xi_j))}{\tau[\Delta\alpha_{ij}]} \propto \prod_{i=l}^n \frac{[\exp(\xi_j)]^{\Delta\alpha_{ij}} \Delta x_{ij}^{\Delta\alpha_{ij}-1}}{\tau[\Delta\alpha_{ij}]}$	(A.6)
ξ_j	$p(\xi_j \Delta\mathbf{x}_j, \varphi_1, \varphi_2, t_{sj}, \sigma_\xi^2) \propto f_N(0, \sigma_\xi^2) * L(\Delta\mathbf{x}_j \mathbf{t}_s, \varphi_1, \varphi_2, \xi_j) =$ $\frac{1}{\sigma_\xi \sqrt{2\pi}} \exp\left(-\frac{\xi_j^2}{2\sigma_\xi^2}\right) * \prod_{i=l}^n \frac{[\exp(\xi_j)]^{\Delta\alpha_{ij}} \Delta x_{ij}^{\Delta\alpha_{ij}-1} \exp(-\Delta x_{ij} \exp(\xi_j))}{\tau[\Delta\alpha_{ij}]} \propto$ $\exp\left(-\frac{\xi_j^2}{2\sigma_\xi^2}\right) * \prod_{i=l}^n [\exp(\xi_j)]^{\Delta\alpha_{ij}} * \exp(-\Delta x_{ij} \exp(\xi_j))$	(A.7)

σ_ξ^2	$p(\sigma_\xi^2 \xi, \gamma, \omega) \propto f_{IG}(\gamma, \omega) * L(\xi \sigma_\xi^2) =$ $\frac{\omega^\gamma}{\tau[\gamma]} (\sigma_\xi^2)^{-\gamma-1} \exp\left(-\frac{\omega}{\sigma_\xi^2}\right) * \prod_{r=1}^n (N_i^{gd} + N_i^{gu}) \frac{1}{\sqrt{2\pi\sigma_\xi^2}} \exp\left(-\frac{\xi_r^2}{2\sigma_\xi^2}\right) \propto$ $(\sigma_\xi^2)^{-\gamma-1} \exp\left(-\frac{\omega}{\sigma_\xi^2}\right) * (\sigma_\xi^2)^{-\frac{\sum_{r=1}^n (N_i^{gd} + N_i^{gu}) \xi_r^2}{2}} \exp\left(-\frac{\sum_{r=1}^n (N_i^{gd} + N_i^{gu}) \xi_r^2}{2\sigma_\xi^2}\right) =$ $(\sigma_\xi^2)^{-\gamma - \frac{\sum_{r=1}^n (N_i^{gd} + N_i^{gu}) \xi_r^2}{2} - 1} * \exp\left(-\frac{\omega + \frac{\sum_{r=1}^n (N_i^{gd} + N_i^{gu}) \xi_r^2}{2}}{\sigma_\xi^2}\right) \propto$ $f_{IG}\left(\gamma + \frac{\sum_{i=1}^n (N_i^{gd} + N_i^{gu})}{2}, \omega + \frac{\sum_{r=1}^n (N_i^{gd} + N_i^{gu}) \xi_r^2}{2}\right)$	(A.8)
----------------	--	-------

Appendix 2B Procedures for Generating Defect Initiation Times from a Non-homogenous Poisson Process

It is noted that an NHPP, $N(t)$, with the mean value function $m(t) = \lambda t^\delta$ ($t > 0$) between the time interval $(t_{i-1}, t_i]$ can be transformed to a corresponding HPP, $N(z)$, with the mean value function $m(z) = \lambda z$ between the time interval $(0, z_i]$ by letting $z = t^\delta - t_{i-1}^\delta$, where $z_i = t_i^\delta - t_{i-1}^\delta$. Given that r defects have initiated within the time interval $(t_{i-1}, t_i]$ or equivalently the transformed time interval $(0, z_i]$, it follows from the property of the HPP (Parzen 1999) that the unordered transformed initiation times, z_1, z_2, \dots, z_r , are independent and uniformly distributed within the interval of 0 to z_i . This allows the initiation times, t_k ($k = 1, 2, \dots, r$), to be generated from the following two steps:

1) sample z_k from a uniform distribution between 0 and z_i , and

2) calculate $t_k = \sqrt[\delta]{z_k + t_{i-1}^\delta}$.

References

Parzen, E. (1999). Stochastic processes. SIAM, PA.

Appendix 2C Procedures for Simulating Corrosion Data with the Simplifying Assumption for Defect Generation

The following is the procedure for simulating the corrosion data and corresponding inspection results assuming that the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections.

1) Sample the number of newly generated defects N_i^g in the i^{th} ($i = 1, 2, \dots, n$) inspection from the Poisson PMF given by Eq. (2.5).

2) Sample the initial times t_{sr} ($r = 1, 2, \dots, N_i^g$) for the N_i^g newly generated defects based on the procedure described in Appendix 2B.

3) Sample the random effect parameters ξ_r ($r = 1, 2, \dots, N_i^g$) for the N_i^g defects from the normal distribution with a mean of zero and a variance of σ_ξ^2 .

4) Generate the inspected depth of detected defects

4.1) For the i^{th} inspection, set $r = 1$.

4.2) Generate the depth increment, Δx_{ir}^g , from the gamma distribution given by Eq. (2.8) with the shape and scale parameters equal to $\varphi_1(t_i - t_{sr})^{\varphi_2}$ and e^{ξ_r} , respectively. Note that $x_{ir}^g = \Delta x_{ir}^g$ for newly generated defects in the i^{th} inspection interval.

4.3) If $x_{ir}^g > x_{th}$, accept the defect as a detected defect with a probability of $\text{POD}(x_{ir}^g)$. If the defect is accepted as a detected defect, go to Step 4.4); otherwise, set $r = r+1$ and go to Step 4.2).

4.4) Generate the depth increment associated with the k^{th} inspection ($k = i+1, i+2, \dots, n$) for the accepted defect, Δx_{kr} , from the gamma distribution given by Eq. (2.8) with the shape and scale parameters equal to $\varphi_1(t_k - t_{sr})^{\varphi_2} - \varphi_1(t_{k-1} - t_{sr})^{\varphi_2}$ and e^{ξ_r} , respectively. Calculate $x_{kr} = \sum_{l=i}^k \Delta x_{lr}$ as the actual depth of the detected defect r at the time of the k^{th} inspection.

4.5) Generate the ILI-reported depth for the detected defect from the i^{th} inspection to the n^{th} inspection from a multivariate normal distribution $f_E(\mathbf{a} + \mathbf{x}_r \cdot \mathbf{b}, \Sigma_E)$.

4.6) Set $r = r + 1$ and repeat steps 4.2) through 4.5) until $r = N_i^g + 1$.

4.7) Repeat Steps 4.1) through 4.6) for all the inspections, $i = 1, 2, \dots, n$.

Appendix 2D Procedures for Simulating Corrosion Data without the Simplifying Assumption for Defect Generation

The following is the procedure for simulating the corrosion data and corresponding inspection results without the assumption that the newly detected defects in the i^{th} inspection are all generated between the $(i-1)^{\text{th}}$ and i^{th} inspections, i.e. some of the newly detected defects in the i^{th} inspection may in fact initiate prior to the $(i-1)^{\text{th}}$ inspection but remain undetected until the i^{th} inspection.

1) Sample the number of newly generated defects N_i^g in the i^{th} ($i = 1, 2, \dots, n$) inspection from the Poisson PMF given by Eq. (2.5).

2) Sample the initial times t_{sr} ($r = 1, 2, \dots, N_i^g$) for the N_i^g newly generated defects based on the procedure described in Appendix 2B.

3) Sample the random effect parameters ξ_r ($r = 1, 2, \dots, N_i^g$) for the N_i^g defects from the normal distribution with a mean of zero and a variance of σ_ξ^2 .

4) Generate the inspected depth of detected defects initiated between the $(i-1)^{\text{th}}$ and i^{th} inspection.

4.1) For the i^{th} inspection, set $r = 1$.

4.2) Generate the depth increment, Δx_{ir}^g , from the gamma distribution given by Eq. (2.8) with the shape and scale parameters equal to $\varphi_1(t_i - t_{sr})^{\varphi_2}$ and e^{ξ_r} , respectively. Note that $x_{ir}^g = \Delta x_{ir}^g$ for newly generated defects in the i^{th} inspection interval.

4.3) Generate the depth increment associated with the k^{th} inspection ($k = i+1, i+2, \dots, n$) for defect r , Δx_{kr} , from the gamma distribution given by Eq. (2.8) with the shape and scale parameters equal to $\varphi_1(t_k - t_{sr})^{\varphi_2} - \varphi_1(t_{k-1} - t_{sr})^{\varphi_2}$ and e^{ξ_r} , respectively. Calculate $x_{kr} = \sum_{l=i}^k \Delta x_{lr}$ as the actual depth of defect r at the time of the k^{th} inspection.

4.4) If $x_{ir}^g > x_{th}$, accept the defect as a detected defect with a probability of $\text{POD}(x_{ir}^g)$.

- If the defect is accepted as a detected defect, go to Step 4.5); otherwise, accept it as an undetected defect and record its depth x_{kr} ($k = i+1, i+2, \dots n$). Then set $r = r+1$ and go to Step 4.2).
- 4.5) Generate the ILI-reported depth for the detected defect from the i^{th} inspection to the n^{th} inspection from a multivariate normal distribution $f_{MVN}(\mathbf{a} + \mathbf{x}_r \cdot \mathbf{b}, \Sigma_E)$.
- 4.6) Set $r = r+1$ and repeat steps 4.2) through 4.5) until $r = N_i^g + 1$.
- 5) Generate the inspected depth of detected defects initiated prior to the $(i-1)^{\text{th}}$ inspection.
- 5.1) If $i = 1$, go to step 4); otherwise, set $r = 1$.
- 5.2) Check the depth, x_{ir} , for the previously undetected defects. If $x_{ir} > x_{th}$, accept the defect as a detected defect with a probability of $\text{POD}(x_{ir})$. If the defect is accepted as a detected defect, go to Step 5.3); otherwise, accept it as an undetected defect and record its depth x_{kr} ($k = i + 1, i + 2, \dots n$). Then set $r = r+1$ and repeat Step 5.2).
- 5.3) Generate the ILI-reported depth for the detected defect from the i^{th} inspection to the n^{th} inspection from a multivariate normal distribution $f_{MVN}(\mathbf{a} + \mathbf{x}_r \cdot \mathbf{b}, \Sigma_E)$.
- 5.4) Set $r = r+1$ and repeat steps 5.2) through 5.3) until $r = N_{(i-1)}^u + 1$, where $N_{(i-1)}^u$ is the number of defect initiate prior to the $(i-1)^{\text{th}}$ inspection but remain undetected until the i^{th} inspection.
- 6) Repeat Steps 4) and 5) for all the inspections, $i = 1, 2, \dots, n$.

CURRICULUM VITAE

Name: Hao Qin

Post-secondary Education and Degrees: Harbin Institute of Technology
Harbin, China
2008-2012, Bachelor of Engineering (B. Eng)
(Civil Engineering)

The University of Western Ontario
London, Ontario, Canada
2012-2014, Master of Engineering Science (M.E.Sc)
(Structural Engineering)

Honours and Awards: Western Graduate Research Scholarship (WGRS)
2012- 2014

Graduate Student Teaching Assistantship
2012- 2014

People's Scholarship
2008-2011

Related Work Experience: Research and Teaching Assistant
The University of Western Ontario
2012-2014

Publications:

Qin, H., Zhang, S. and Zhou, W. (2013) "Inverse Gaussian process-based corrosion growth modeling and its application in the reliability analysis for energy pipelines", *Frontiers of Structural and Civil Engineering*, 7(3): 276-287.

Zhang, S., Zhou, W. and **Qin, H.** (2013) "Inverse Gaussian process-based corrosion growth model for energy pipelines considering sizing error in inspection data", *Corrosion Science*, 73: 309-320.

Qin, H. and Zhou, W. (2014) "Reliability analysis of corroding pipelines considering the growth and generation of corrosion defects", *The 10th International Pipeline Conference*, Calgary, Canada.

Qin, H. and Lu, D. (2014) "Bayesian stochastic modeling of pitting corrosion and inference using MCMC simulation", *The 4th International Symposium on Life-Cycle Civil Engineering*, Tokyo, Japan.