PROBABILISTIC MODELLING OF SPATIO-TEMPORAL UNCERTAINTIES IN DEGRADATION

by

Adetola Rotimi Adegbola

Bachelor of Science, University of Ibadan, 2007

Master of Science, Newcastle University, 2010

A dissertation

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the program of

Civil Engineering

Toronto, Ontario, Canada, 2019

Author's Declaration

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

Probabilistic Modelling of Spatio-temporal Uncertainties in Degradation

Doctor of Philosophy, 2019

Adetola Rotimi Adegbola

Civil Engineering

Ryerson University

Abstract

Deterioration is a major problem facing engineering structures, systems and components (SSCs). To maintain the structural integrity and safe operation of such SSCs all through their service life, it is important to understand how degradation phenomena progress over time and space. Hence degradation modelling has been increasingly used to model existing deterioration, predict future deterioration as well as provide input for infrastructure management in terms of inspection and maintenance decision making. As deterioration is known to be random, modelling of spatial and temporal uncertainty remains a crucial challenge for infrastructure asset professionals.

The main objective of the thesis is to develop sophisticated models for characterizing spatial and temporal uncertainties in deterioration modelling with a view to enhancing decision making under uncertainty. The thesis proposes a two-dimensional copula-based gamma distributed random field for the spatial uncertainties, and a copula-based multivariate gamma process model to characterize stochastic dependence of multiple degradation phenomena. Techniques for estimating the model parameters and

simulating the field or process, prediction of the remaining lifetime distribution as well as condition-based maintenance optimization are also presented.

To study the extreme value distribution of the random field, the thesis also presents a numerical method based on the Karhunen-Loève expansion for evaluating extrema of both one- and two-dimensional homogeneous random fields. The simulation results are benchmarked against existing analytical models for special cases. In addition, the study also investigates the effect of parameter (epistemic) uncertainty on the extreme value distribution of the field. Finally, the thesis presents a practical application of the proposed copula-based gamma field by treating the wall profile of a feeder pipe as one- and two-dimensional gamma fields. The thesis demonstrates a practical application of the multivariate gamma process model to rutting, cracking, and surface roughness of highway pavements.

In summary, the proposed models have advanced the knowledge and techniques of stochastic deterioration modelling in the engineering field.

Acknowledgements

First, I would like to thank my supervisor, Dr. Xianxun (Arnold) Yuan, who introduced me to the research area and gave me academic, financial and moral support. In the last few years, his guidance and knowledge of the research area have made me to complete this thesis.

I would also like to express my sincere appreciation to the members of my dissertation examination committee: Dr. Bhagwant Persaud, Dr. Min Wang, Dr. Sharareh Taghipour and the external examiner, Dr. Wenxing Zhou of Western University, for taking time out to serve on the committee. My thanks also go to Dr. Min Wang of CANDU Energy Inc., Mississauga for his technical input and for providing some of the data used in this research.

Furthermore, I thank all professors, from whom I took their courses, in the Departments of Civil Engineering, Industrial Engineering and Statistical Sciences. My special thanks also go to fellow graduate students and post-doctoral fellows who I met during my studies at Ryerson University, especially those in the Ryerson Institute for Infrastructure Innovation (RIII). It was a privilege for me to study in the department.

Last but not the least, I would like to thank my family, most importantly my father, siblings and fiancée for their love and support throughout the duration of my doctoral program.

Table of Contents

List of Tables

List of Figures

List of Symbols, Notations and Abbreviations

α	Shape parameter
β	Scale parameter of a gamma distribution
$\gamma(\cdot)$	Semivariogram function/model
Γ	Gamma function
Δ	Difference or change in a certain quantity
ζ	Maintenance threshold
θ	Correlation length
$\boldsymbol{\theta}$	Vector of parameters
λ	Eigenvalue; kriging weight
μ	Mean
ξ	Standard normal variates
$\prod$	Product
ρ	Correlation coefficient
σ	Standard deviation
$\sum$	Summation
ϕ	Standard normal probability density function
Φ	Standard normal cumulative distribution function
C	Copula cumulative distribution function; autocovariance function
CDF	Cumulative distribution function
COV	Coefficient of variation
EV	Extreme value
$f(\cdot)$	Probability density function; eigenfunction
$F(\cdot)$	Cumulative distribution function
FAC	Flow-accelerated corrosion
FFS	Fitness for service
$G(\cdot)$	Gamma cumulative distribution function

Ga	Gamma distribution
H	Joint distribution function
KL	Karhunen-Loeve
L	Length of random field
MCMC	Markov Chain Monte Carlo Simulation
MLE	Maximum likelihood estimation
N_{sim}	Number of simulations
P	Probability
PDF	Probability density function
$\boldsymbol{R}$	Correlation matrix
$X(s)$	Random field
$X(t), X_t$	Stochastic process
$Z(\cdot)$	Gaussian field

1. Introduction

1.1 Background

Adequate infrastructure in form of roads, bridges, airports, power plants, waterways, ports and dams to name a few, no doubt, remains the backbone of economic development of every country. However, with continued rise in population coupled with migration of people, infrastructure is constantly under increasing pressure and deterioration becomes inevitable over time. According to the Canadian Infrastructure Report Card (2016), one-third of municipal infrastructure is in Fair to Very Poor condition due to deterioration and untimely maintenance. One approach to addressing the problem of deteriorating infrastructure is to continually seek better ways to model and manage deterioration of the infrastructure asset. This is a challenge for infrastructure asset practitioners, managers and researchers because deterioration is known to be a random phenomenon and no one can predict with certainty how it is going to progress over time (temporally) and space (spatially).

The Wiley dictionary (1997), compiled by L.F. Webster, defines deterioration as "physical manifestation of failure" of a structure, material or component. Such components in civil engineering include concrete columns, steel girders, carbon steel pipes, and asphalt road pavements. The failure manifestation may be in the form of fracture, fatigue cracking, pitting corrosion, flaking, rutting, etc., which ultimately lowers the quality or value of the material or structure. In infrastructure asset management, deterioration or degradation modelling is a key component that needs to be studied and understood. Despite many models that have been proposed in the literature for deterioration of physical infrastructure, "more sophisticated stochastic" models are required to deal with uncertainties (Le Son et. al. 2013). Most especially, characterization of spatial dependency and dependency among different degradation phenomena need to be addressed.

The probabilistic degradation models in the literature can be broadly classified into two main groups in terms of the type of uncertainty they address. Stochastic process based models focus on capturing temporal uncertainty in deterioration. However, such models

fail to address the issue of spatial uncertainty of dependence in degradation. Several stochastic processes such as gamma process, inverse Gaussian process, geometric Brownian motion and Markov chains have been proposed to model degradation. For example, discrete-time Markov chains were used to model deterioration processes (Sharabah et al. 2006, Guida and Pulchini 2011). Homogeneous Markov chain was employed to model temporal degradation of reinforced concrete (Possan and Andradec 2014) and a nonhomogeneous Markov chain model was proposed for degradation (Jin and Mukherjee 2014). Regarding the inverse Gaussian process for degradation modelling, Zhang (2014) proposed a model for corrosion defect growth while Ye et al. (2014) incorporated random effects. In the literature, Park and Padgett (2005) and Elsayed and Liao (2004) have proposed degradation models based on the geometric Brownian motion. Also, Zhang (2014) extended Elsayed and Liao (2004)'s model by formulating a defect-specific hierarchical Bayesian corrosion growth model. Gamma processes have been widely used to model temporal variability of deterioration (Pandey et al. 2009; Zhang 2014; Shemehsavar 2014). Lawless and Crowder (2004) and Tsai et al. (2012) incorporated random effects into their gamma process models.

The other group of degradation models, which are random field-based, are able to capture stochastic dependence in the spatial domain but not much in the temporal uncertainty. For example, Na et. al. (2012) and Sudret (2008), proposed lognormal random fields to model spatial variabilities in material properties, geometries or degradation parameters of reinforced concrete structures. Peng and Stewart (2013) proposed truncated normal and lognormal random fields while Shafei and Alipour (2015) proposed multidimensional non-Gaussian random fields. Aryai and Mahmoodian (2017) and Zhao and Yun (2018) applied the Karhunen-Loeve expansion to represent random fields for degradation modelling.

Narrowing down to nuclear infrastructure, the primary heat transport system (PHTS) of a typical CANDU reactor has components such as a reactor core, turbine, generators, hundreds of feeder pipes, etc. These feeders, made of carbon steel, carry coolants needed to cool the reactor core and thereby preventing overheating of the core. It has been observed that as they age the feeders experience several degradation phenomena,

among which is wall thinning due to flow-accelerated corrosion (FAC). Also, the extent of wall thinning varies from location to location along the feeder pipes. Therefore, we have both spatial and temporal uncertainties. If the extent of local wall loss exceeds the limits specified for the design of the piping component, the feeder has to be replaced to ensure continuing operational safety of the power plant.

FAC is a type of steel corrosion exacerbated by coolant flow in feeder pipes at high velocities and temperatures. Factors known to affect the rate of FAC include mass transfer, coolant flow velocity, pipe material composition in terms of chromium content, stress levels such as temperature and pressure, bend angle, coolant pH and oxidizing species (Duan et al. 2009; Chung 2010; Lee et. al. 2016; Slade and Gendron 2005).

A feeder consists of a number of bends, each bend being further divided into four zones: extrados, intrados, left cheek, and right cheek (*Figure 1.1*). Ultrasonic technology has been widely used to measure wall thickness for nuclear pipes. The ultrasonic scanning systems used in CANDU industry typically consists of 6 or 14 probes, arranged within an array or bracelet. To cover the full circumference, a feeder pipe is typically scanned several runs. It was generally believed that thinning due to FAC was most severe on the extrados of a bend. This can be attributed to the fact that during bending of straight new feeder pipes with uniform wall thickness, the extrados of a bend gets thinner while the intrados becomes thicker. Another reason may be due to turbulence on the extrados of a bend as flow of heavy water hits the feeder pipes. However, this belief was later challenged by some field observations. For example, highly-localized thinning has been observed close to a grayloc weld, well away from bends of a feeder removed from the Pickering station (Jin and Awad 2011) and along a straight pipe located in a downstream of a check valve in a Korean nuclear power plant (Lee et. al. 2016). This further shows the need to develop a more sophisticated model that fully considers the overall profiles of feeder pipes.
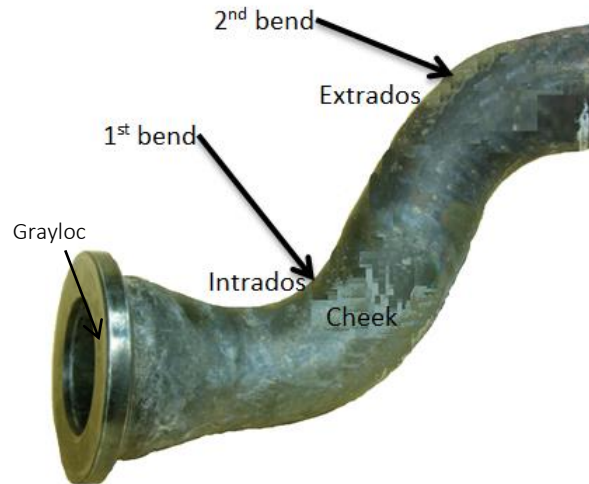
Figure 1.1: A typical feeder segment taken from a nuclear plant.

Still on the deterioration of feeders, the current state of practice in the nuclear industry, as found in the feeder fitness for service guidelines (FFSG), recommends two types of assessments: Condition Monitoring Assessment (CMA) and Operational Assessment (OA). CMA assesses the past performance of feeder pipes by comparing the previous OA with the actual feeder performance. In contrast, OA assesses the future state over the next operating cycle. With respect to wall thinning, FFS assessment predicts future minimum thicknesses (especially at the next inspection) as well as remaining service life. The idea is to ensure that the minimum wall thickness at the next inspection period is greater than the minimum thickness allowed in the guidelines for CANDU reactor feeders. According to CSA N285.4, when the inspection result does not satisfy the acceptance criteria, it must be demonstrated to the regulatory body i.e. Canadian Nuclear Safety Commission (CNSC) that the predicted component integrity will be sufficient at the end of the next inspection window. Furthermore, an inspection window of 3 years is recommended by the CSA. In case of disagreement between FFSG and N285.4, the latter supersedes. A drawback of this method is that the assessment considers only temporal uncertainty with regards to FAC without accounting for spatial uncertainty. Quantifying spatial uncertainty, which boils down to dependence modelling, could address this issue.

Besides asset deterioration in the nuclear industry, physical assets in other infrastructure sectors such as pavements, bridges, buildings, water mains, sewer pipes, and so on, experience deterioration over time as they age. According to a report by Gill and Lawson (2013), annualized cost of operating and maintaining roads in Ontario alone was $2.7 billion with a substantial part of the amount invested in rebuilding and restoring highways across the province. As stated in the Ministry of Transportation Ontario (MTO) (2013)'s manual of pavement design and rehabilitation, MTO carries out inspections of pavement sections and provides reports on pavement deterioration rates, distress types and their causes. This information is used to select appropriate maintenance and rehabilitation strategy for the concerned sections. Using flexible pavements as a case study, deterioration in flexible pavements include different forms of cracking, surface deformation such as rutting (Figure *1.2*) and disintegration in the form of potholes to name a few. These multiple degradation phenomena are often dependent on one another because they have common underlying causes such as traffic load and environmental conditions. Another probable reason for their dependence is that the phenomena may interact with one another because of their proximity. Therefore, there is a need to capture the dependence and model the degradation modes together.

Figure 1.2: Parallel ruts on a highway (Source: https://www.cbc.ca/news/canada/newfoundland-labrador/worsening-ruts-government-experimenting-with-polymer-pavement-1.3196025).

Case studies considered in this thesis are wall thinning in nuclear feeder pipes caused by flow accelerated corrosion (FAC) and multiple degradation mechanisms in flexible pavements. It is expected that the proposed methodologies will be applicable to degradation mechanisms in other civil engineering systems and components.

1.2   Objectives and Scopes

The overall goal of the dissertation is to develop advanced dependence models for the modelling of spatial and temporal uncertainties in deterioration and examine their effects on decision making. The research is divided into a number of parts with a view to answering the following research questions:

- What is the extreme value distribution of random fields? We answer this question by simulating Gaussian, gamma and lognormal random fields using the K-L expansion and then compare the results with semi-analytical Ditlevsen function and Poisson approximation.

- How do aleatory (model) and epistemic (parameter) uncertainties affect the extreme value distribution of one- and two-dimensional random fields? These questions are addressed by Monte Carlo simulation-based methods.

- How is spatial uncertainty in degradation modelled? What is the benefit of a spatial model? In many previous studies, emphasis has been on temporal modelling of degradation as a stochastic process. Markov chains, gamma, Wiener, Poisson and Inverse Gaussian processes have been developed and used to temporally model degradation. A major drawback of this approach is that spatial uncertainty is not taken into consideration. To address this problem, we propose a method based on the random field theory. Specifically, a gamma random field is proposed and illustrated with a case study on wall thinning in nuclear feeders due to flow-accelerated corrosion.

- How are multiple degradation phenomena in components and systems modelled and how does it affect lifetime distribution and maintenance decisions? These questions are answered by adapting the gamma random field model to a

6

multivariate gamma process model. The model is illustrated with an example of multiple degradation in flexible pavements.

1.3   Research Methodology and Significance

All structures experience degradation over time and space in form of corrosion, cracking, creep, fatigue, to name but a few. Advanced degradation modelling is therefore required to model uncertainties, not only in the temporal domain, but also in the spatial domain. This dissertation proposes a data-driven methodology for deterioration modelling. Furthermore, when deterioration phenomena are modelled with random fields, extreme values (EV) of such random fields become important quantities in risk and reliability analyses. As a benchmark, the empirical extreme value distribution in one dimension will be compared with existing semi-analytical and analytical extreme value distributions. We will show in later chapters that the gamma random field presented therein is able to capture the spatial uncertainty; and the proposed multivariate gamma process, which is a variation of the gamma random field, captures the temporal uncertainty and dependence among degradation modes. Parameter estimation will be done by the maximum likelihood method because the gamma field model and its variation are mathematically tractable.

On the issue of EV of random fields, Ditlevsen has proposed a distribution for EV of one-dimensional standard Gaussian fields within a finite interval with an unknown parameter G. However, the exact EV distribution of non-Gaussian fields within a finite region remains unknown. The thesis proposes a methodology based on Monte Carlo simulation to calibrate parameter G and then use double transformations of the Gaussian field to any translation fields of interest to study EV of non-Gaussian fields within a finite region. It will be demonstrated later that EV distributions of Gaussian and non-Gaussian random fields can be evaluated using an empirical G and parameters of the fields, without simulation.

On the issue of wall thinning due to flow-accelerated corrosion (FAC) in nuclear feeders, large amounts of wall thickness data are collected during pipe inspections. However, only the minimum wall thickness corresponding to maximum deterioration is used for temporal

modelling of FAC and making decisions, while all other measurements are discarded (Yuan et al. 2008). The model does not take into consideration spatial dependence and possible interaction of the corrosion defects. Furthermore, Jyrkama and Pandey (2012) noted that it is not certain the minimum thickness occurs at the same location every time measurements are taken. For instance, the minimum thickness normally occurs on the extrados of a new feeder pipe. As the feeder pipe experiences FAC, the exact position of the minimum thickness changes with time. Last, as there are an infinite number of locations in the feeder pipe but only a finite number of points are scanned. Therefore, there is a possibility that the minimum measured value is not the true minimum. The methodology employed in this thesis uses the entire wall profile for modelling with a view to capturing the spatial uncertainty and finding a distribution for the maximum deterioration.

Regarding multiple degradation phenomena in flexible pavements, the major issues are capturing temporal uncertainty and dependence among competing degradation phenomena. It is suspected that modelling each phenomenon separately may lead to overestimation or underestimation of the remaining lifetime and possible suboptimal inspection and maintenance strategy. Given the current states of each degradation mode and the estimated parameters of the model, future degradation states may be predicted. Furthermore, the remaining lifetime distribution may be estimated if the failure thresholds of each degradation mode is known. Finally, information obtained from the multivariate deterioration modelling may be employed to make informed decisions regarding optimization of inspection and maintenance strategy. These benefits will be demonstrated with a case study.

1.4   Organization

The dissertation is organized as follows. Chapter 2 is the literature review which discusses concepts such as dependence modelling, random fields and their extreme values, stochastic process-based modelling of degradation, inspection and maintenance optimization and methods of parameter estimation. Chapter 3 presents gamma random

field model for degradation. Also presented in Chapter 3 are theoretical and numerical extreme value distributions of one- and two-dimensional standard normal, gamma and lognormal random fields on closed intervals. The simulation procedure employs Karhunen-Loeve expansion for Gaussian fields with quadratic exponential, exponential and triangular autocorrelation functions. Chapter 4 deals with the case study of flow accelerated corrosion in nuclear feeders. The feeder wall profile is treated as a one- and two-dimensional gamma fields and their corresponding extreme value distributions are evaluated. Also presented in Chapter 4 are a kriging-based method of estimating missing data and effect of missing data on extreme value distribution. Chapter 5 focusses on a multivariate gamma process model for multiple degradation modelling. The results of the model are used to predict the remaining lifetime distribution and optimize condition-based maintenance. Chapter 6 reports the conclusions of the dissertation and future research directions.

2.   Literature Review

2.1   Dependence Modelling

In statistics and probability, dependence refers to a relationship between random variables. The most common method of measuring dependence is the Pearson's product moment correlation which measures linear correlation between random variables. It, however, gives no information about nonlinearity between the variables. Other measures such as Spearman's rank correlation and Kendall's correlation can capture nonlinear dependence between random variables. In this thesis, we chose copula to model dependence because of its ease of construction and ability to capture dependence structure between variables without the marginal distribution effect.

2.1.1 Copula

There are several ways of constructing multivariate distributions. Among them is the copula method. A copula is a function linking marginal distributions into a multivariate distribution. Given marginal distributions of random variables, copulas can be used to link the marginal distribution together to form the joint distribution. Nelson (2006) defines a copula of dimension n (or an n-copula) as a function C with the following properties:

i. C is grounded and n-increasing,

ii.   The domain of C is I^n where I^n is an n-dimensional unit cube.

iii.   For every $\boldsymbol{u} = u_1, \dots, u_n$ in I^n, $C(\boldsymbol{u}) = 0$ if at least one u_i is zero and $C(u) = u_j$ if all $u_i's$ are 1 except u_j,

iv.   For every $\boldsymbol{u}$ and $\boldsymbol{v}$ in I^n and given that $\boldsymbol{u} \leq \boldsymbol{v}$, the volume $V_C(\boldsymbol{u}, \boldsymbol{v}) \geq 0$.

A special case of copula is the product copula $\prod$ that characterizes independent copulas i.e.

$$\prod = C(u_1, \dots, u_n) = u_1 \dots u_n \qquad\qquad 2.1$$

According to Sklar's theorem (Sklar, 1973), for an n-dimensional multivariate distribution H with marginal distributions $F_1, F_2, \ldots F_n$, there exists a copula C such that

$$H(x_1, \ldots x_n) = C(F_1(x_1), \ldots, F_n(x_n)) \qquad 2.2$$

Conversely, if C is an n-dimensional copula and $F_1, F_2, \ldots, F_n$ are distribution functions, then H is a joint distribution function. Another important concept in copula is the Frechét-Hoeffding Bounds. Basically, the Frechét-Hoeffding bounds describe the lower and upper bounds of a copula. Any multivariate copula C satisfies the inequality below:

$$max(u_1 + \ldots + u_n - n + 1, 0) \leq C(u_1, \ldots, u_n) \leq min(u_1, \ldots, u_n) \qquad 2.3$$

Examples of copulas are Gaussian, student t and Archimedean families of copula. Figure 2.1 shows the bivariate version of some of these copulas. Archimedean copulas are one- or two-parameter families of copulas. They are very useful due to their relative ease of construction and properties such as symmetry and associativity. Families of Archimedean copulas in the literature include Clayton, Ali-Mikhail-Haq, Gumbel and Frank copulas.

Table 2.1 summarises the forms and generators of some families of one-parameter multivariate Archimedean copulas as well as valid ranges of the parameter (Nelsen 2006).
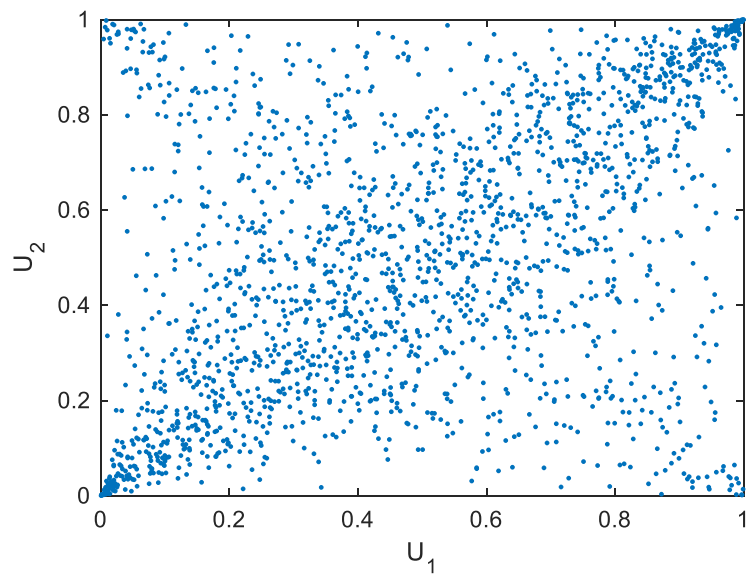
(a)
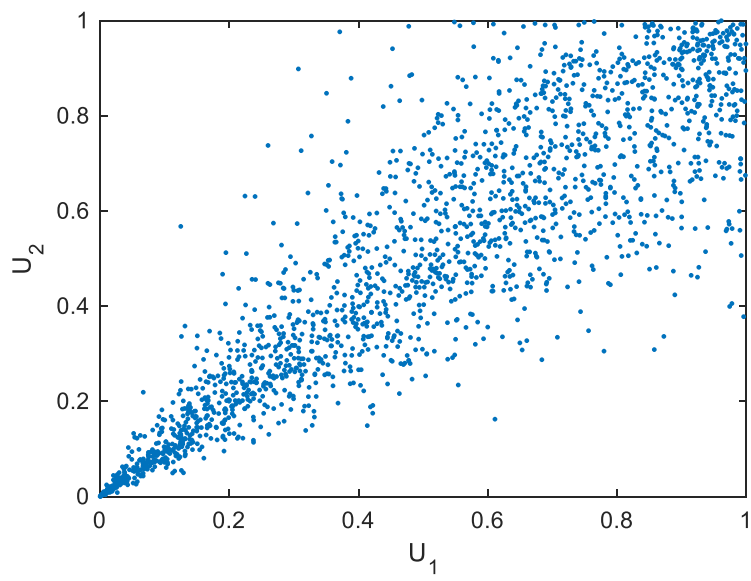
(b)
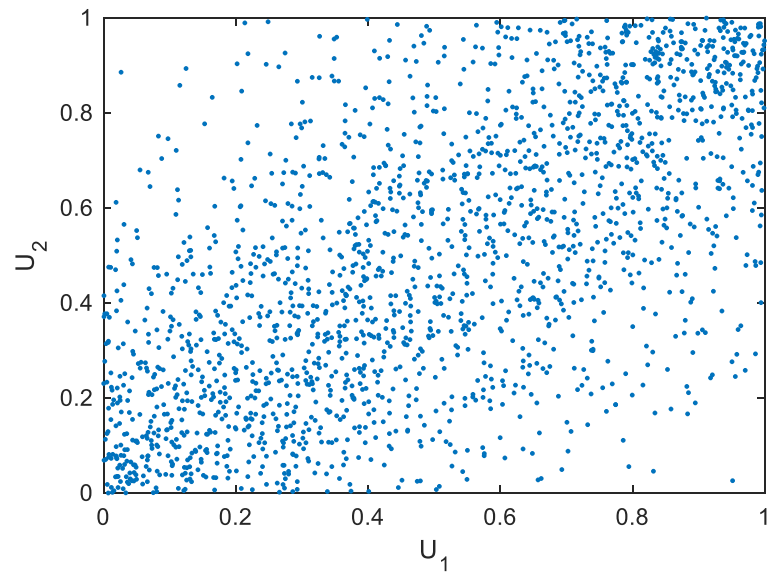
(c)
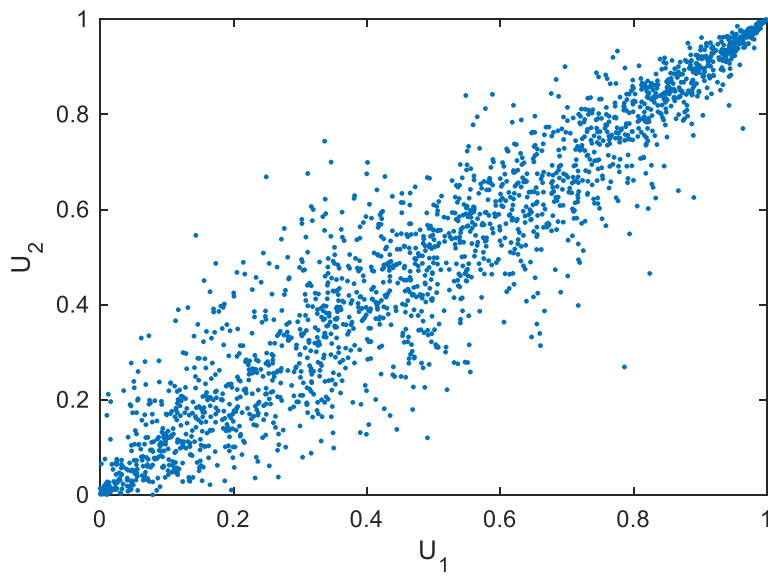
(d)

Figure 2.1 Simulated bivariate copulas (a) t-copula ρ=0.5, (b) Clayton θ=5, (c) Frank θ=5, (d) Gumbel θ=5).

Table 2.1 Archimedean copulas – generators and parameter ranges. Adapted from Nelsen (2006)

Copula	$C_\theta(u_1, u_2, \ldots, u_n)$	$\psi_\theta(t)$	$\theta \in$
Clayton	$\left[max\left(u_1^{-\theta} + \cdots + u_n^{-\theta} - n + 1, 0\right)\right]^{-\frac{1}{\theta}}$	$\dfrac{1}{\theta}(t^{-\theta} - 1)$	$[-1, \infty)$
Ali-Mikhail-Haq	$\dfrac{u_1 u_2 \ldots u_n}{1 - \theta(1 - u_1)(1 - u_2)\ldots(1 - u_n)}$	$ln\,\dfrac{1 - \theta(1 - t)}{t}$	$[-1, 1)$
Gumbel	$exp(-[(-ln\,u_1)^\theta + (-ln\,u_2)^\theta + \ldots + (-ln\,u_n)^\theta]^{\frac{1}{\theta}})$	$(-ln\,t)^\theta$	$[1, \infty)$
Frank	$-\dfrac{1}{\theta} ln\left\{1 + \dfrac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)\ldots\left(e^{-\theta u_2} - 1\right)}{(e^{-\theta} - 1)^{n-1}}\right\}$	$-ln\,\dfrac{e^{-\theta t} - 1}{e^{-\theta} - 1}$	$(-\infty, \infty)$ $\backslash\{0\}$

Another type of copula commonly used in the literature is the Gaussian copula. Gaussian copulas are distributions over $[0\ 1]^n$ where n is the dimension of the copula. For a multivariate Gaussian copula,

$$C(u_1, u_2, \ldots, u_n) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\ldots, \Phi^{-1}(u_n)) \qquad 2.4$$

$$C(u_1, u_2, \ldots, u_n)$$
$$= \int_{-\infty}^{\Phi^{-1}(u_1)} \ldots \int_{-\infty}^{\Phi^{-1}(u_n)} \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{0.5}} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{R}^{-1}\mathbf{x}\right) dx_1 \ldots dx_n \qquad 2.5$$

where Φ^{-1} and Φ are the inverse cdf and joint cdf respectively and $\mathbf{R}$ is the matrix of correlation among individual univariate distributions. More information about the Gaussian copula will be discussed in the next section.

Bivariate Gaussian copulas, as a special case of multivariate Gaussian copulas, can be simulated by generating random variates from multivariate normal distribution and then obtaining the standard normal CDF for each variate. Examples of bivariate copulas generated by this method are shown in Figure *2.2*. Furthermore, a trivariate copula obtained through the same method is shown in Figure *2.3*.
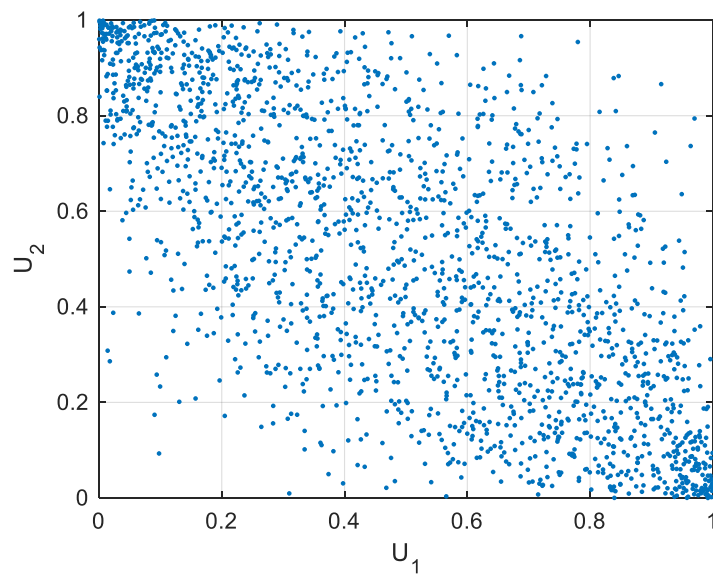
(a)
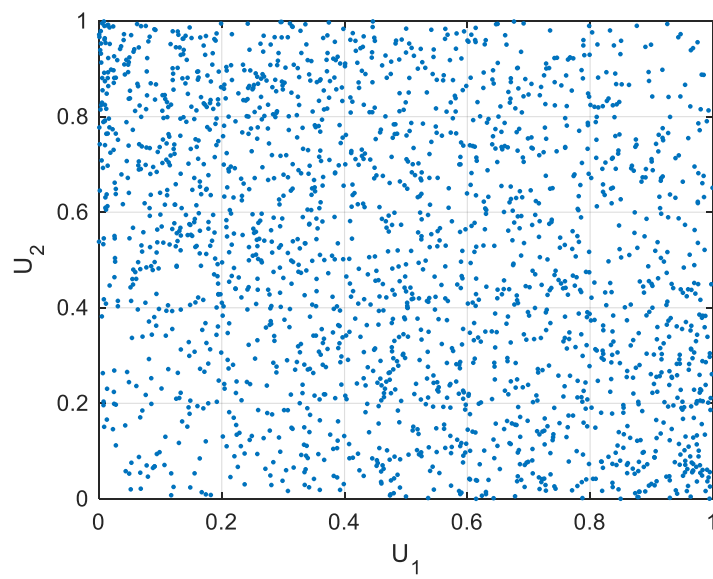
(b)
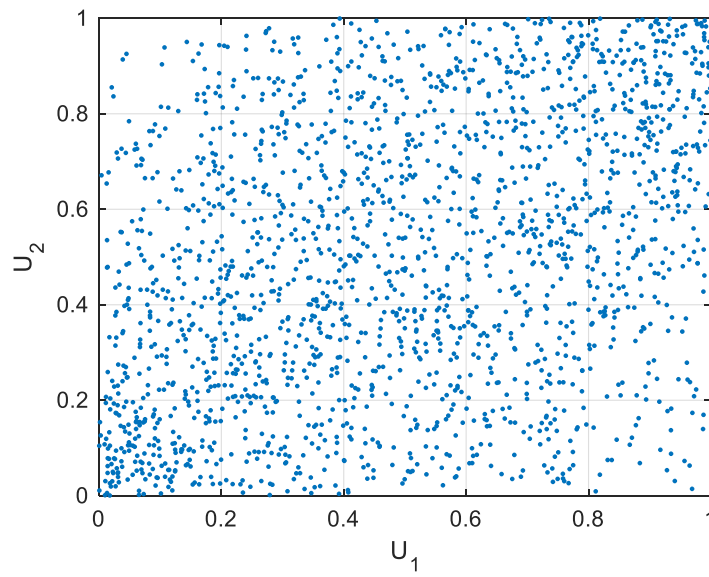
(c)
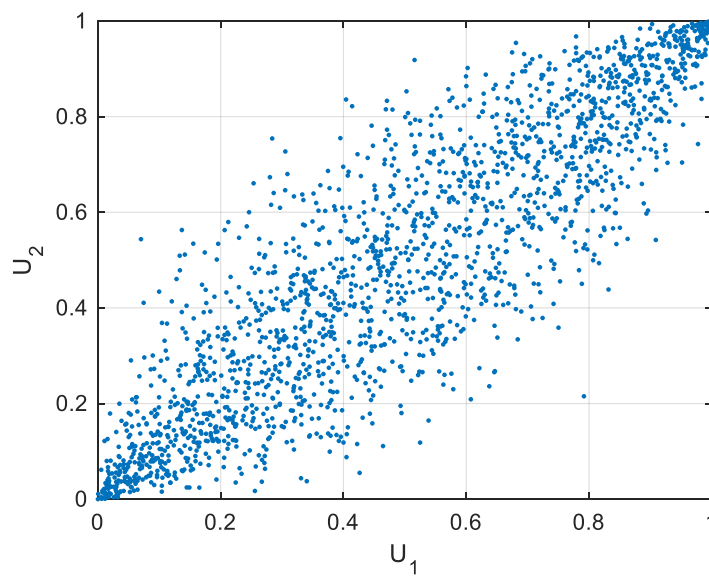
(d)

Figure 2.2 Simulated bivariate Gaussian copulas (a) ρ=-0.7, (b) ρ=-0.3, (c) ρ=0.4, (d) ρ=0.9.
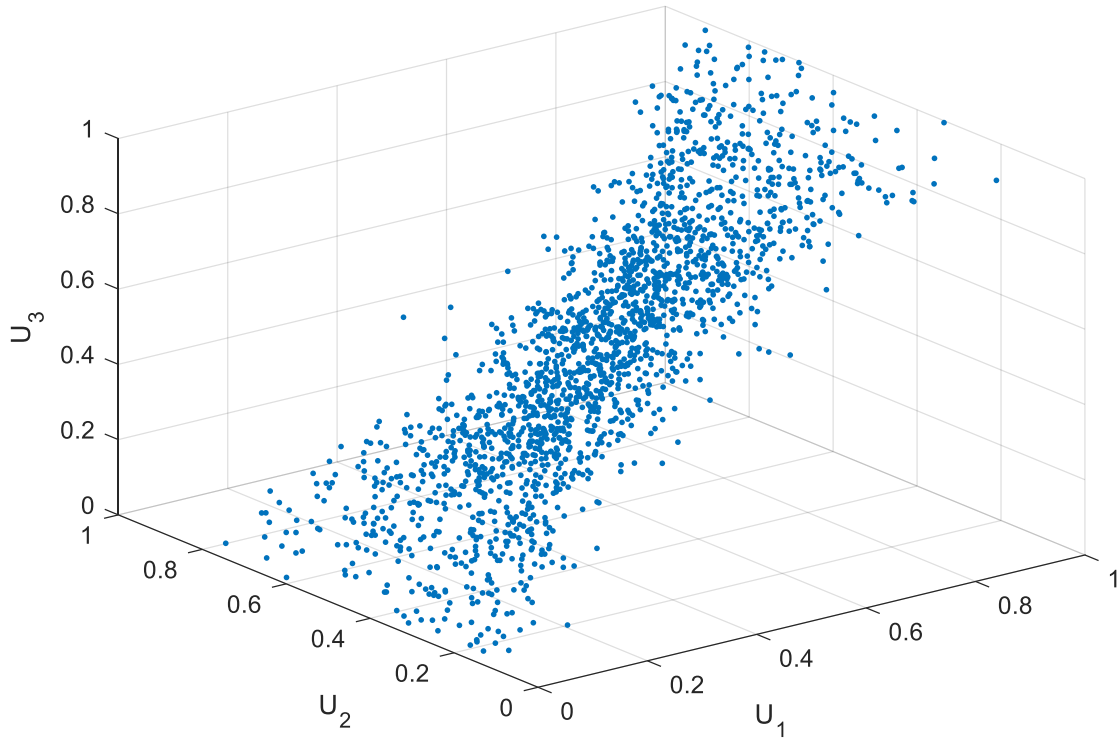
Figure 2.3 Simulated trivariate Gaussian copulas ($\rho_{12} = 0.7, \rho_{13} = 0.3, \rho_{23} = -0.4$).

There is a sizeable amount of literature on the use of copulas to study the dependence structure with respect to degradation modelling and reliability analysis. For example, Zhou et al. (2012) used a copula function to study effect of dependent defect growth on reliability of corroding pipelines. Similarly, Hao and Su (2014) employed copula to describe the dependence of two performance characteristics governing the degradation of a product. Li and Xue (2014) employed a multivariate copula function to describe dependence among multiple performance characteristics of a dormant system. Literature discussing construction of multivariate copulas can be found in Joe (2001) and Nelsen (2006). Still, constructing multivariate copulas of higher dimensions remains a challenging mathematical task. A simple way of constructing multivariate copulas is the inversion method, which works as follows. Given the univariate marginal distributions $F_1, F_2, \ldots, F_n$ with a joint distribution H, the copula C can be constructed by

$$C(u_1, u_2, \dots, u_n) = H(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n),) \qquad \text{2.6}$$

where $F_i^{-1}(u_i)$ is the inverse of the marginal distribution.

One of the major questions that arise when modelling with copulas is the choice of copula. This is akin to fitting data to different random variables to see which probability distribution fits best in some way. This question has been addressed by Genest and Rivest (1993). Maximum likelihood estimation and the Akaike information criterion (AIC) has been shown to be an effective method in this regard (Zhang and Singh 2006).

2.1.2 Multivariate Gamma Distribution

As the name suggests, a multivariate gamma distribution is a multivariate distribution with gamma distributions as its marginals. The multivariate gamma distribution involves a double transformation of the standard normal distribution to gamma distribution using the multivariate normal density. Hence, it is regarded as Gaussian copula-based. The gamma distribution mentioned here is a two-parameter continuous probability distribution whose probability density function (PDF) is expressed as

$$g(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)} \qquad \text{2.7}$$

for $x \geq 0$, where $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively. The cumulative distribution function (CDF) is expressed as

$$G(x; \alpha, \beta) = \frac{\Gamma(\alpha, x/\beta)}{\Gamma(\alpha)} \qquad \text{2.8}$$

where $\Gamma(p, q) = \int_0^q x^{p-1} e^{-x} du$ is called the lower incomplete gamma function, and $\Gamma(p) = \Gamma(p, \infty)$ the complete gamma function, which is used in Eq. (2.7) as well. Because the distribution functions involve the gamma function, the distribution is called a gamma distribution. A gamma distributed random variable is written in shorthand as $X \sim Ga(\alpha, \beta)$.

Several important properties of the gamma distribution are worth noting. First, the mean and variance of the gamma random variable are $\alpha\beta$ and $\alpha\beta^2$, respectively. The coefficient of variation is $1/\sqrt{\alpha}$, independent of the scale parameter β. The moment generating function is $(1 - \beta x)^{-\alpha}$. Second, as the name suggests, the shape parameter α dictates the shape of the PDF curve. In particular, for $0 < \alpha \leq 1$, the PDF decays immediately after at $x = 0$, whereas for $\alpha > 1$, the PDF increases from 0 at $x = 0$ to a peak, or the mode, which is equal to $\beta(\alpha - 1)$, and then decays off afterwards. When $\alpha = 1$, the gamma distribution reduces to an exponential distribution. Moreover, the excess kurtosis (i.e., the conventional kurtosis minus 3) of a gamma distribution is $6/\alpha$, again independent of β. Since the kurtosis decreases as α increases, this means that small-α gamma distributions exhibit longer and fatter upper tail than big-α distributions. Meanwhile, it also shows that a gamma distribution, regardless of α and β, has a fatter tail than a Gaussian distribution because the excess kurtosis is a positive number. Thus, the gamma distribution is leptokurtic. This observation will be applied to interpretations on some of the results obtained in the study later.
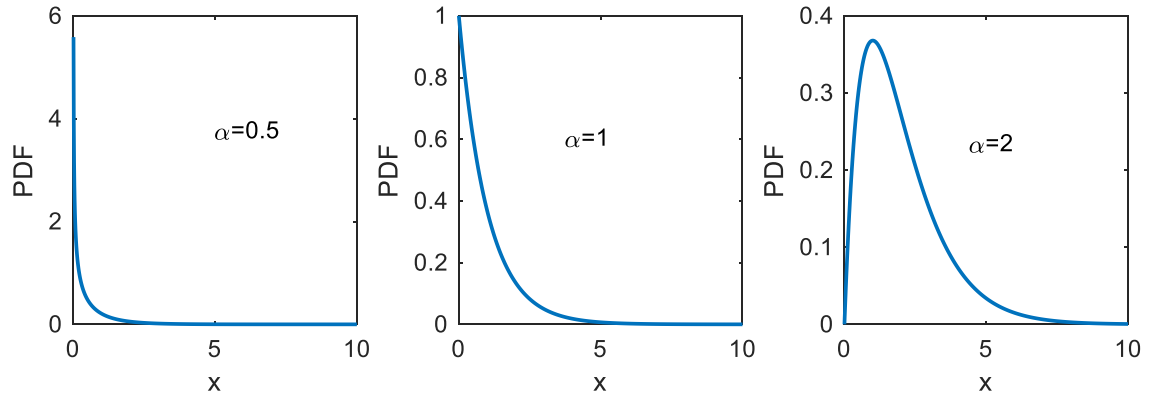
Figure 2.4 Comparison of the PDFs of gamma distributions with different shape parameters (β=1)

Of considerable practical importance are the scaling and additive properties of gamma distribution. Consider a gamma distribution $X \sim Ga(\alpha, \beta)$ multiplied by a scalar constant k, the resulting distribution is a gamma distribution with a factored scale parameter i.e. $kX \sim Ga\left(\alpha, \frac{\beta}{k}\right)$. In a similar vein, consider a number of gamma distributions $X_i \sim Ga(\alpha_i, \beta)$,

where $i = 1, \ldots, n$, then the sum of all the distributions is a gamma distribution. Mathematically,

$$\sum_{i=1}^{n} X_i \sim Ga\left(\sum_{i=1}^{n} \alpha_i, \beta\right) \qquad 2.9$$

Unlike the univariate gamma distribution, the multivariate gamma (MG) distributions have several non-equivalent definitions. The reason for this is that many approaches can be used to construct a dependent, multivariate probability distribution from a given marginal distribution. Early constructions were mainly derived from a natural extension of the univariate probability density function, characteristic function or moment generating function. For example, Kibble (1941) and Moran (1969) developed a bivariate gamma distribution by generalizing the univariate probability density function using Laguerre polynomials, whereas Krishnamoorthy and Parthasarathy's multivariate gamma distribution was a direct extension of the moment generating function from its univariate form $m(t) = (1 - \beta t)^{-\alpha}$ to the matrix form $m(t_1, \ldots, t_n) = [\det(\mathbf{I} - \mathbf{R}t)]^{-\alpha}$, where $\mathbf{I}$, $\mathbf{R}$, $\mathbf{t}$ represent the identity matrix, a positively definite matrix, and a diagonal matrix with diagonal elements $t_1, \ldots, t_n$, respectively. Later, other approaches such as mixing, compounding, and conditioning were used, often inspired from practical modelling needs (for a systematic treatment of constructing multivariate distributions, refer to Joe (2001) and Kotz et al. (2000)). As a result, more than twenty versions of bivariate gamma distribution and eight versions of multivariate gamma distribution were reviewed in Balakrishnan and Lai (2009) and Kotz et al. (2000). A common problem of majority of the multivariate gamma distributions reviewed there is that they allow for only positive correlation. To loosen this constraint, a copula-based approach has been widely applied recently (Rodríguez-Picón et al. 2017, Wang et al. 2015). This study adopts the copula approach because of its ease in concept and computation as well as its flexibility in dependence modelling. Specifically, the Gaussian copula is used.

Therefore, the Gaussian copula-based multivariate gamma distribution is defined as follows. Suppose that $X_1, \ldots, X_n$ each is a gamma random variable, i.e., $X_i \sim Ga(\alpha_i, \beta_i), i = 1, \ldots, n$. The joint cumulative distribution function of $X_1, \ldots, X_n$ is defined as

$$G_n(\boldsymbol{x}) = \Phi_n(z_1, \ldots, z_n; \mathbf{R}) \hspace{3cm} \text{2.10}$$

where $\Phi_n(\boldsymbol{z}; \mathbf{R})$ denotes the joint cumulative distribution function of a standard multivariate normal distribution with a correlation matrix $\mathbf{R}$; $z_i = \Phi^{-1}(u_i)$; $u_i = G(x_i; \alpha_i, \beta_i)$, and $\Phi^{-1}(u)$ denotes the inverse of the standard normal CDF at probability u. It is readily shown that the joint probability density function of the multivariate gamma distribution is expressed as

$$g_n(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{z}^T\mathbf{R}^{-1}\boldsymbol{z}\right) \prod_{i=1}^{n} \frac{g(x_i; \alpha_i, \beta_i)}{\phi(z_i)} \hspace{2cm} \text{2.11}$$

Note that the correlation matrix $\mathbf{R} = \{r_{ij}\}$ in Eqs. (2.10) $-$ (2.11) is not defined for the original gamma variables, but for the transformed standard normal variables. In other words, the correlation coefficient r_{ij} should not be understood as the coefficient between X_i and X_j (or x-correlation coefficient hereafter). Rather, it is between Z_i and Z_j (or z-correlation coefficient), where $Z_i = \Phi^{-1}[G(X_i; \alpha_i, \beta_i)]$. Denote by ρ_{ij} the x-correlation coefficient of X_i and X_j. The exact relationship between ρ_{ij} and r_{ij} is difficult to assess analytically. For illustration purpose, simulation results of the $\rho - r$ relationship for a bivariate gamma distributions with different marginal distribution $Ga(\alpha, 1)$ are shown in Figure 2.5 The figure shows a nonlinear relationship between ρ and r, confirming the nonlinear dependence defined by the copula function. The sensitivity analyses show that the relationship is a function of the shape parameters of the bivariate gamma distributions and independent of the scale parameters. As the shape parameters of the distributions become smaller, the graph diverges more at the negative end of the correlation range. When α increases, however, the relationship becomes almost linear. For $\alpha = 100$, the curve is almost a 45-degree straight line. This is reasonable because in this case the univariate gamma distribution is close to a normal distribution, and the bivariate gamma distribution very close to a bivariate normal distribution.
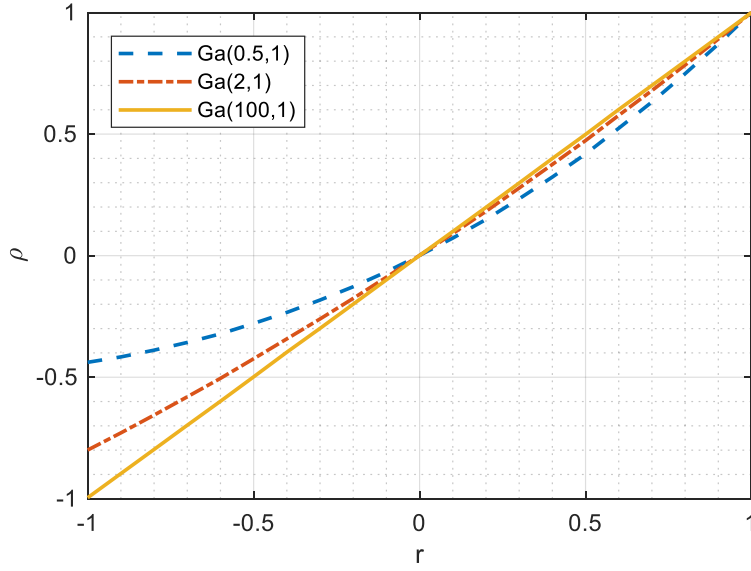
Figure 2.5 Relationship between x- and z-correlation coefficients for different marginal distributions Ga(α,1).

Tail dependence is another notion of stochastic dependence (Joe 2001). It concerns particularly the dependence at the lower or upper tail of a multivariate distribution. In this regard, the Gaussian copula has zero tail dependence. This means that for any two globally, Gaussian copula-dependent random variables X_i and X_j, their lower or upper quantiles, regardless of the value of ρ_{ij}, are asymptotically independent. Here, we take the Gaussian copula as an assumption of the gamma field model and leave other copula options for future study.

2.2   Random Fields and Extreme Value Issue

This section reviews random fields and extreme values in relation to deterioration modelling. As we will see later that a random field is a function defined in space, its extreme values are of central importance in the context of reliability and safety engineering. Although, an extreme value may refer to a maximum or minimum value of a function, the emphasis here is on maximum values. First, we start with homogenous and nonhomogeneous random fields. Then, a literature review of Gaussian and non-Gaussian random field modelling for degradation is presented. Finally, extreme values of random fields are reviewed.

2.2.1 Homogeneous Random Fields

A random or stochastic field $X(s)$, is a function whose values are random variables in space (Vanmarcke, 2010). Random fields may exist in one-, two- or three-dimensional space, i.e., s may be a scalar or a vector of location. In addition, a random field may include time $X(s, t)$. This is known as a space-time process or a spatiotemporal model. There are several key properties of a random field such as homogeneity, isotropy, ergodicity and separability. These concepts will be discussed in the following paragraphs.

Although stationarity and homogeneity are often used interchangeably to characterize some sort of stability of fluctuation, we use stationarity to describe time-related stability, and homogeneity for spatially related stability. Mathematically speaking, stationarity and homogeneity are dealt with in the same way. In other words, homogeneity describes the state of random fields with respect to space. Therefore, a random field $X(s)$ is said to be homogeneous if the joint probability density functions depend on the relative locations of points. Furthermore, a homogeneous random field may be strongly or weakly homogeneous. A random field is strongly homogeneous if its probability distribution is homogeneous. On the other hand, in a weakly homogeneous field, only the first few moments are constant (Fenton and Griffiths, 2008). A good example of a weakly homogenous field is a second-order homogeneous field whose covariance function depends on the spatial lag between locations and independent of the absolute locations.

A homogeneous random field $X(s)$ may be characterized by its constant mean, variance and correlation length. The mean is defined as the expectation of the random field while variance is defined as:

$$\sigma^2(s) = \mathrm{E}[X^2(s)] - \{\mathrm{E}\{X(s)\}\}^2 \qquad\qquad 2.12$$

The coefficient of correlation function, which is a dimensionless quantity, is obtained by dividing the covariance of random variables at two points s_i and s_j separated by h_{ij}, where h_{ij} is a certain distance measured between the two points s_i and s_j, by the product

of the standard deviations of the variables at both locations. Mathematically, the correlation between two points in a homogeneous random field $X(s)$, is

$$\rho(h_{ij}) = \frac{\mathrm{Cov}[X(s_i), X(s_j)]}{\sigma(s_i)\sigma(s_j)} \qquad 2.13$$

where Cov is the covariance and $\sigma(.)$ is the standard deviation. Correlation length θ, defined as a distance between two random field locations over which correlation tends to zero, is an important parameter of homogeneous random fields.

Talking about isotropy, it may be regarded as a subset of homogeneity. In an isotropic random field, the covariance or correlation between any points depends only on the distance between the points. In other words, the joint probability density functions and covariance functions remain the same when an isotropic random field is rotated. Mathematically, the covariance function $C(s_i, s_j) = C(h_{ij})$, where h_{ij} is the Euclidean distance between s_i and s_j. This means all isotropic random fields are homogenous random fields because isotropic fields are a class of homogenous fields. However, the reverse is not the case. On the other hand, anisotropy is the opposite of isotropy. An anisotropic random field is a field in which the correlation lengths vary with directions, for example, there is $\theta = [\theta_x, \theta_y]$ in a two-dimensional field.

Of considerable importance is ergodicity of random fields. Roughly speaking, a random field is said to be ergodic if a realization of the field provides all information, including statistical properties, about its joint probability distributions (Vanmarcke, 2010). For an ergodic random field, its average over the space is the same as its average over the probability space.

In a multidimensional random field, the correlation function may be separable or inseparable. A correlation function in an n-dimensional random field is fully separable if the correlation function is the product of all individual one-dimensional correlation functions i.e. $\rho(h_{ij}) = \rho(h_1) \dots \rho(h_n)$. This property of random field has several practical implications especially in spatiotemporal modelling where the correlation function for time is separated from the one for space $\rho(s, t) = \rho(s)\rho(t)$. This is an example of a partially

separable correlation function in which the correlation function is the product of correlation functions of lower dimensions. The separability property of correlation functions will be applied to two-dimensional random fields in Chapter 4.

2.2.2 Nonhomogeneous Random Fields

The opposite of a homogeneous field is a nonhomogeneous field. For a nonhomogeneous field, the statistics of the random field change with location, e.g. the autocovariance function depends on the locations. For a nonhomogeneous random field, Vanmarcke (2010) noted that the random field can be transformed to an "approximately" weakly homogeneous random field with a constant mean and variance of zero and one respectively by normalizing it. This transformation eliminates nonhomogeneity in the first-order statistics but nonhomogeneity still exists in higher order moments.

While a lot of statistical and analytical tools have been developed to characterize homogeneous random fields and stationary processes, there is yet no mathematical structure developed that fits all classes of nonhomogeneous random fields. In view of this, some researchers have proposed ways of dealing with nonhomogeneous random fields. For example, Denis and Cremoux (2002) proposed a method to divide a nonhomogeneous random field into smaller homogeneous fields. In considering spatial variation of soil properties for reliability analysis, Wu et al. (2012) constructed a nonhomogeneous lognormal random field by taking the product of the homogeneous random field and a depth-dependent function. Azuri et al. (2013) accounted for spatial nonhomogeneity in the rainfall intensity by using a trend and a residual. In addition, the use of moving window kriging-based random field models for nonhomogeneous fields have been demonstrated (Harris et al., 2010; Haas, 1990).

2.2.3 Random Field Modelling for Degradation

Most problems encountered in engineering such as deterioration (Adegbola et al. 2015; Stewart et. al. 2007; Yuan & Pandey 2009), wave heights (Mori & Yasuda 2002) and soil

properties (Na et al. 2009) involves modelling of spatial uncertainties, which often manifest a Gaussian or non-Gaussian probabilistic behavior. As such, stochastic fields have been found suitable for modelling spatial uncertainty and variability in deterioration.

Few studies have modelled deterioration with Gaussian and non-Gaussian random fields. For example, Shafei and Alipour (2015) applied multidimensional Gaussian and non-Gaussian stochastic fields to study spatial and temporal variability of deterioration parameters. To model variability in material properties and geometry, Peng and Stewart (2013) used truncated normal and lognormal random fields and Na et. al. (2012) used two-dimensional lognormal fields obtained by translation of the Gaussian fields. Further, Sudret (2008) characterized the extent of deterioration in reinforced concrete beams by describing parameters of the degradation model as homogeneous lognormal random fields.

The Karhunen-Loeve (KL) expansion is one of the most commonly used representation of stochastic fields. Recently, Zhao and Yun (2018) and Aryai and Mahmoodian (2017) used the KL expansion to model spatial random fields with regards to deterioration in a reinforced-concrete bridge slab and cast-iron pipes, respectively. Meanwhile, Na et. al. (2012) simulated non-Gaussian fields by translating Gaussian fields generated with a power spectral density function.

Another well-established method in spatio-temporal modelling is the use of separable fields. Oumouni et. al. (2019) proposed a separable random field model for degradation modelling as a product of a gamma process and a spatial random field.

2.2.4 Extreme Values of Random Fields

A big challenge of the spatial uncertainty modelling is the quantification of the extreme value (EV). This quantity is of fundamental importance in system reliability and safety analysis. For example, to assess overall risk during structural integrity assessment, the second order statistics of the stochastic field (mean, standard deviation and correlation length) are insufficient, and it is often desirable to obtain the full probability distribution of

the extreme value. Extensive literature review indicates that there is no analytical solution for the extreme value distribution, except some asymptotic solutions from the classical extreme value theory and the outcrossing theory.

This section reviews a few key results gleaned from literature that are relevant to the solution of EV of the random field in a finite region. We start with the classical EV theory, which suggests that both Gaussian and gamma distributions fall into the Type I or Gumbel domain of attraction. Asymptotic results for stationary processes or fields focus on normal cases. Finally, the Ditlevsen results for the EV distribution in a finite region are reviewed.

2.2.4.1.   Extreme Values of Random Variables

The classical EV theory concerns the asymptotic distribution of the extreme value (either the maximum or minimum) of a sequence of independent and identically distributed (i.i.d.) random variables $\{X_1, \dots, X_n\}$ as n goes to infinity. Let $M_n = \max\{X_1, \dots, X_n\}$. Then $\Pr\{M_n \leq x\} = [F(x)]^n$, where $F(x)$ is the CDF of X_i. For any given x, clearly the probability approaches to zero. However, the classical EV theory aims to find a proper set of normalization parameters a_n and b_n such that $\Pr\{a_n(M_n - b_n) \leq x\}$ converges to a distribution function $H(x)$ for all x. Indeed, it has been found that depending upon the tail behavior of $F(x)$, the normalized M_n converges to one of the three types of probability distributions, namely: Gumbel, Frechet and Weibull distributions, or Type I to III EV distributions, respectively. Details of the EV theory can be found in, e.g., Leadbetter et.al (1983).

As two special cases, both Gaussian and gamma sequences converge asymptotically to the Gumbel (Type I) distribution. For the standard normal sequence, the normalization factors are expressed as $a_n = \sqrt{2 \log n}$, and $b_n = a_n - (\log\log n + \log 4\pi)/2a_n$. For the gamma sequence $\{X_i \sim Ga(\alpha, \beta)\}$, $a_n = 1/\beta$ and $b_n = \beta(\log n + (\alpha - 1)\log\log n - \log\Gamma(\alpha))$. Detailed derivation of these results can be found in Embrechts et al. (1997).

In practice, one may be interested in using the asymptotic EV distribution as an approximation to the exact solution, particularly when the latter is unknown, as we shall

see soon in the random field cases. As a benchmark, Figure *2.6* and Figure *2.7* compare the asymptotic and exact distributions for the standard normal distribution and a gamma distribution Ga(2, 0.5), respectively. In both figures, the CDFs are compared in the upper panels, in which the blue solid lines represent the exact solution $F^n(x)$ whereas the red broken lines represent the asymptotic solution $H(a_n(x - b_n))$, where $H(x) = \exp(-\exp(-x))$. For the normal case, the asymptotic distribution almost always underestimates the probability. For the gamma case, however, the asymptotic distribution always overestimates the probability. For both cases, it is shown that the rate of convergence along n is very slow as the error plots in the right panels show.
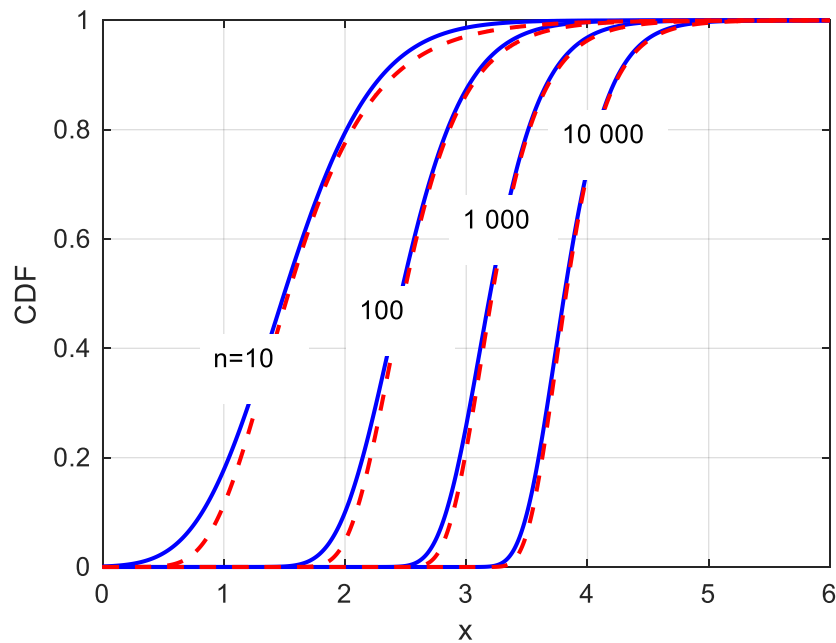
(a)
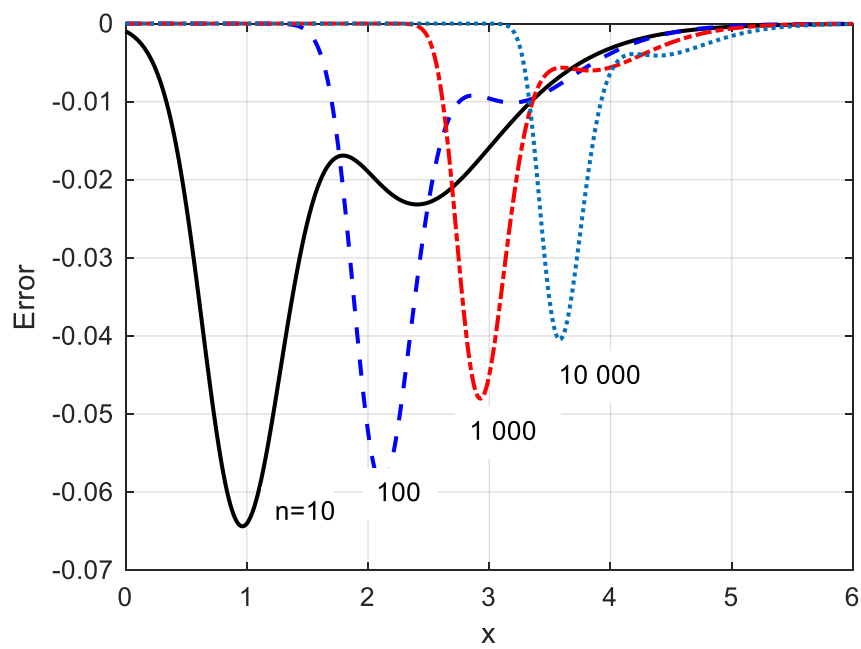
(b)

Figure 2.6 Comparison of the exact and asymptotic distributions for standard normal distribution

(a)

(b)
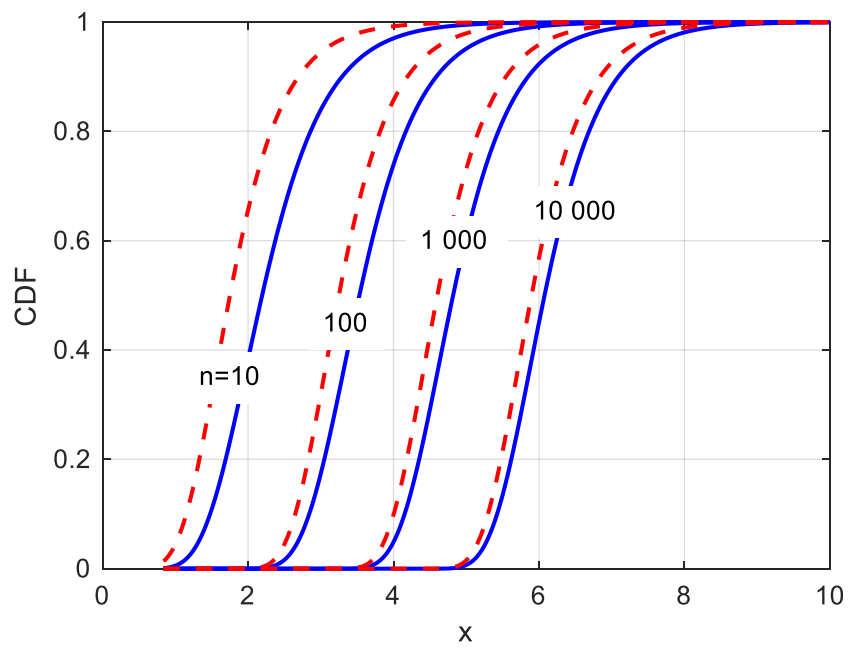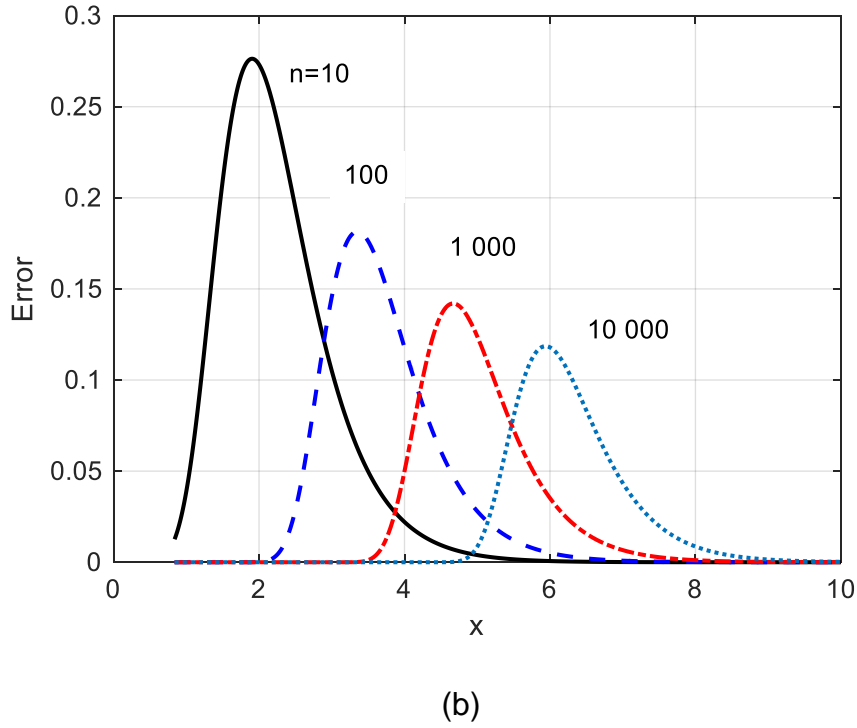
Figure 2.7 Comparison of the exact and asymptotic distributions for gamma distribution Ga(2,0.5)

2.2.4.2.   Extreme Values of Stationary Stochastic Processes

Over decades, researchers have tried to extend the beautiful results of the classical EV theory to general situations such as dependent random sequences, stationary processes, non-stationary processes (Leadbetter, et al. 1983), and multi-dimensional random fields (Adler & Taylor 2007). Unlike the i.i.d. sequence, there is no exact solution available for the EV of a stationary field within a finite interval, even for a normal field. However, several interesting approximate results are collected here.   First, under a certain regularity condition, the asymptotic extreme value distribution of a stationary Gaussian process still follows a Type I distribution. For a mean-square differentiable, stationary Gaussian field $Z(x)$ with zero mean and unit variance, its maximum value over the region $[0, L]$, $M_L = \max\{Z(x), 0 \leq x \leq L\}$, converges to Type I distribution $H(x) = \exp(-\exp(-a_T(x - b_T)))$ as $L \to \infty$, in which $a_L = \sqrt{2 \log L}$ and $b_L = a_L + \log\left(\sqrt{-r''(0)}/2\pi\right)/a_L$, where $r''(0)$ is the second derivative of the correlation function $r(h)$ evaluated at $h = 0$ (Leadbetter et al.

30

1983, p.171). Clearly, this solution applies only if $r''(0) < 0$, which is exactly the condition for mean-square differentiability. The quadratic exponential correlation model mentioned in Section 3.1 satisfies the condition. For nondifferentiable Gaussian fields (such as those with exponential and triangular correlation models), the asymptotic Gumbel distribution still holds true with the same a_L and yet with a different b_L which is a function of the decaying order of $r(h)$. In fact, for both the exponential and triangular correlation models, $b_L = a_L + (\log \log L + \log(C^2/\pi))/2a_L$, in which $C = 3/\theta$ for the exponential model and $1/\theta$ for the triangular model (Leadbetter et al. 1983, p.217).
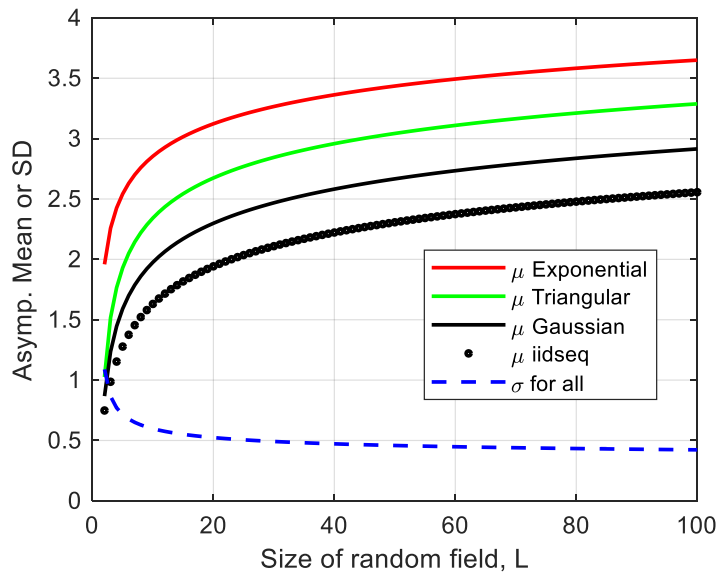
Figure 2.8: Asymptotic statistics of the EV of standard normal field under different correlation models ($\theta = 1$ for all cases).

Figure 2.8 shows the means and standard deviations of the asymptotic Gumbel distribution for the abovementioned four correlation models. For a comparison purpose, the EV statistics of the i.i.d. normal sequence are also shown. The standard deviation decreases with the size of random field whereas the mean EVs are monotonically increasing. Among the four correlation models, the mean EV of the exponential model is the highest and that of the spherical model lowest. The mean EV of the iid random sequence is even lower than that of the spherical model. This is expected because all

four models have zero or close to zero correlation when $h = 1$. Random variables of the random field within $h = 1$, although correlated, still contribute to push the EV up higher.

Unlike the i.i.d. sequence, there is no exact solution available for the EV of a stationary field within a finite interval, even for a normal field. The work by Ditlevsen (1966) seems to be the only attempt in this issue.

Another approximation is through the notion of mean upcrossing rate using Poisson assumption (Leadbetter et al. 1983). Suppose that one can find the mean upcrossing rate of a homogeneous random field over a given crossing level x, denoted by $v(x)$. Then the EV distribution can be evaluated as

$$F_M(x; L) = \Pr(M_L \leq x) = \Pr(N_x(L) = 0) \approx \exp\big(-Lv(x)\big)$$

2.14

where $N_x(L)$ denotes the number of upcrossing over the threshold level x within the interval $[0, L]$. For a mean-square differentiable standard normal process, the mean rate is

$$v(x) = \frac{1}{2\pi} \sqrt{-\rho''(0)} \exp\left(-\frac{x^2}{2}\right)$$

2.15

where $\rho''(0)$ is the second derivative of the correlation function evaluated at $h = 0$. Grigoriu (1984) further provided the solution for non-Gaussian translation process as

$$v(x) = \frac{1}{2\pi} \sqrt{-\rho''(0)} \exp\left(-\frac{u^2(x)}{2}\right)$$

2.16

where $u(x) = \Phi^{-1}[G(x)]$ for the gamma field.

2.2.4.3.   Ditlevsen Distribution

Ditlevsen (1966) might be the very first researcher who attempted to derive an exact distribution for the EV of a finite random field. Starting from a bivariate connection function $C(s_1, s_2)$, Ditlevsen was able to find the exact EV distribution of a standard normal field of a quadratic exponential correlation model as defined in Section 3.1 within the interval of $[0, L]$ as

$$F_M(x; L) = [\Phi(x)]^{1+G} \exp\left(-\sqrt{3/\pi}\frac{L}{\theta} \cdot \frac{\phi(x)}{\Phi(x)}\right)$$ 2.17

where $\phi(\cdot)$ is the standard normal PDF. The only unknown parameter is the G term, which is defined as $G = \int_0^L \partial C(s_1, 0)/\partial s_2 \, ds_1$, where $C(s_1, s_2) = F_M(x; s_1 + s_2)/$ $[F_M(x; s_1)F_M(x; s_2)]$. The exact value of G is unknown. But clearly, G is a function of L. In theory, for a given x and L, the G value could be solved numerically at least. However, the computation is complicated. Ditlevsen (2004) later claimed that the $G(L)$ term is 'often of minor importance,' and for this reason, he suggested to drop the G term in Eq. (2.17) to make an approximate estimation of the EV distribution. We call this approximate Ditlevsen model. Note that for the exponential and triangular correlation models, the random field is not mean-square differentiable and thus the Ditlevsen solution does not exist.

In order to solve for G numerically, the PDF of the Ditleven's distribution is needed in the maximum likelihood function. Given the Ditlevsen's distribution in Eq. (2.17), the derivation of the PDF, which is the first derivative of the CDF with respect to x, is presented below.

By definition, $f_M(x) = \frac{d}{dx}F_M(x; L)$. Let $\gamma = \frac{1}{\theta}\sqrt{\frac{3}{\pi}}$. Applying the product rule,

$f_M(x) = \exp\left[-\frac{\phi(x)\gamma L}{\Phi(x)}\right]\frac{d}{dx}\Phi(x)^{1+G} + \Phi(x)^{1+G}\frac{d}{dx}\{\exp\left[-\frac{\phi(x)\gamma L}{\Phi(x)}\right]\}$. The final expression for the PDF is found to be

$$f_M(x) = \exp\left[-\frac{\phi(x)\gamma L}{\Phi(x)}\right]\Phi(x)^G\left\{[1+G]\phi(x) + \gamma L\left[\frac{\phi^2(x)}{\Phi(x)} - \phi'(x)\right]\right\}$$ 2.18

Note that $\phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ is the standard normal probability density function. So, its derivative, $\phi'(x) = \frac{-x}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) = -x\phi(x)$.
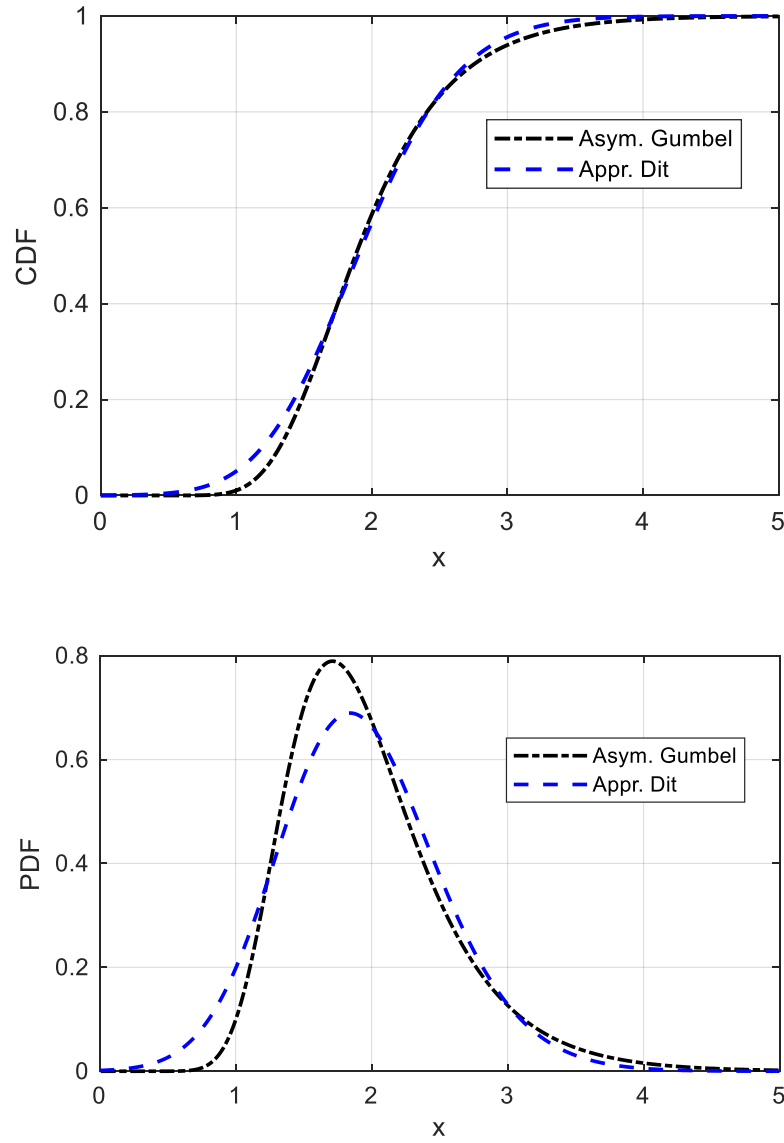
Figure 2.9: Comparison of the asymptotic Gumbel distribution with the approximate Ditlevsen distribution for stationary standard normal field with Gaussian correlation model $(\theta = 1, \text{L} = 10)$.

It is important to recognize that the Ditlevsen results (both exact and approximate solutions) can be easily translated to the gamma field. All one needs to do is to replace x with $\Phi^{-1}\big(G(x)\big)$ in the right hand side of Eqs. (2.17).

2.3   Stochastic Process Modelling of Degradation

Roughly speaking, stochastic processes may be viewed as a special class of random fields. In this case, the field is one-dimensional and time-indexed. Due to the random nature of degradation in civil engineering structures, several stochastic models have been developed and used to study degradation. In the literature, Markov chains, Gamma, Wiener, Poisson and inverse Gaussian processes are some of the common stochastic models used for modelling degradation over time. Variations and extensions of these processes have also been applied to degradation. These involve including random effects to account for heterogeneity in degradation and incorporating explanatory variables such as stress levels. A few of the stochastic processes that are widely used in degradation modelling are discussed in the following subsections.

2.3.1   Gamma Process

A gamma process can be stationary or nonstationary. First, we discuss the stationary gamma process in terms of its definition, properties and use in the literature. This is then followed by a discussion on nonstationary gamma process.

A stationary gamma process $\{X(t), t \geq 0\}$ is a continuous-time stochastic process satisfying the following:

- $X(0) = 0$ with probability one;

- $X(t)$ has independent increments i.e. for $0 \leq t_1 < t_2 \leq t_3 < t_4$, $X(t_2) - X(t_1)$ and $X(t_4) - X(t_3)$ are independent;

- The increment $X(t_2) - X(t_1)$ follows a gamma distribution $Ga(\alpha(t_2 - t_1), \beta)$ for any $0 \leq t_1 < t_2$.

The probability density function of the independent increments mentioned in the definition of the stationary gamma process is given as

$$g(\Delta x(t); \alpha\Delta t, \beta) = \frac{\Delta x(t)^{\alpha\Delta t-1} e^{-\frac{\Delta x(t)}{\beta}}}{\beta^{\alpha\Delta t}\Gamma(\alpha\Delta t)} \qquad 2.19$$

for $x(t) \geq 0$, where $\alpha\Delta t > 0$ is the time-dependent shape parameter while $\beta > 0$ is the scale parameter. A stationary gamma process is characterized by its mean $\alpha\beta t$ and variance $\alpha\beta^2 t$. This shows that the mean and variance of a stationary gamma process increase linearly with time. The ratio of the standard deviation of the gamma process to its mean is known as the coefficient of variation COV.

$$COV = 1/\sqrt{\alpha t} \qquad 2.20$$

This ratio is independent of the scale parameter. In deterioration modelling particularly, gamma process has been widely applied to model the temporal uncertainty (e.g., Yuan & Pandey 2009; Shemehsavar 2014; Zhang & Zhou 2014). A comprehensive survey of the application of Gamma process in maintenance engineering can be found in Van Noortwijk (2007).

The gamma process is a pure-jump process suitable for modelling monotonic degradation because of its mathematical tractability. In the literature, gamma processes have been employed to model temporal variability of deterioration (Pandey et al. 2009; Zhang 2014). An extension to gamma process models involves incorporating fixed and random effects. Random effects represent the effects due to unobserved heterogeneity. On the other hand, fixed effects refer to the influence of covariates such as stress levels, bend angle, etc. Lawless and Crowder (2004) were the first to propose a gamma process model incorporating covariates and random effects to account for "unexplained differences" in deterioration rates of different units. Tsai et al. (2012) also incorporated random effects into the rate parameter of a gamma process model with a view to determining optimal design for degradation tests. A score test for the presence of random effects has also been discussed (Lawless and Crowder 2004; Yuan 2007).

A nonstationary gamma process $\{X(t), t \geq 0\}$ is a continuous-time stochastic process satisfying the following:

- $X(0) = 0$ with probability one;

- $X(t)$ has independent increments i.e. for $0 \leq t_1 < t_2 \leq t_3 < t_4$, $X(t_2) - X(t_1)$ and $X(t_4) - X(t_3)$ are independent;

- The increment $\Delta X(t) = X(t_2) - X(t_1)$ follows a gamma distribution $Ga(\Delta\alpha(t), \beta)$ for any $0 \leq t_1 < t_2$.

For a nonstationary gamma process, the shape parameter is not linear with time. Suppose the shape function of the degradation process follows a power law

$$\alpha(t) = at^c \qquad\qquad 2.21$$

where a and c are nonnegative constants, the parameters of the process (i.e. a, c and β) can be obtained by the maximum likelihood method. When the parameter $c > 1$, the degradation process is accelerating and when $c < 1$, the process is slowing down. If $c = 1$, the process reduces to stationary gamma process. Results of stationary gamma processes in terms of expectation, variance and coefficient of variation are applicable to nonstationary gamma processes.

2.3.2 Inverse Gaussian Process

Another important stochastic process commonly used to model degradation is the inverse Gaussian process. In this section, we present its definition, properties and use in the literature for degradation modelling. An inverse Gaussian process $\{X(t), t \geq 0\}$ is a stochastic process with monotonic degradation path satisfying

- $X(0) = 0$ almost sure;

- $X(t)$ has independent increments i.e. for $0 \leq t_1 < t_2 \leq t_3 < t_4$, $X(t_2) - X(t_1)$ and $X(t_4) - X(t_3)$ are independent;

- The increment $X(t_2) - X(t_1)$ follows an inverse Gaussian distribution for any $0 \leq t_1 < t_2$.

The inverse Gaussian distribution is a two-parameter distribution characterized by its mean μ ($\mu > 0$) and shape parameter α ($\alpha > 0$). The probability density function (PDF) is given as

$$f(x; \mu, \alpha) = \left(\frac{\alpha}{2\pi x^3}\right)^{0.5} \exp\left\{-\frac{\alpha(x-\mu)^2}{2\mu^2 x}\right\}$$ 2.22

for $x > 0$ and 0 for $x \leq 0$. The cumulative distribution function (CDF) is

$$F(x; \mu, \alpha) = \Phi\left[\left(\frac{\alpha}{x}\right)^{0.5}\left(\frac{x}{\mu} - 1\right)\right] + \exp\left(\frac{2\alpha}{\mu}\right)\Phi\left[-\left(\frac{\alpha}{x}\right)^{0.5}\left(\frac{x}{\mu} + 1\right)\right]$$ 2.23

where $\Phi(.)$ is the standard normal CDF.

Some important properties of the inverse Gaussian distribution are worth pointing out. First, the mean and variance of the distribution are μ and μ^3/α, respectively. The coefficient of variation is $\sqrt{\mu/\alpha}$. Second, the excess kurtosis is $15\mu/\alpha$. As the excess kurtosis is a positive number, the inverse Gaussian distribution has a fatter tail than any univariate Gaussian distribution. Third, the first passage time for a fixed positive level of a Brownian motion with a drift follows an inverse Gaussian distribution. Fourth, the inverse Gaussian distribution has scaling and additive properties similar to those of the gamma distribution. For instance, if an inverse Gaussian distribution X with parameters μ and λ is multiplied by a scalar k, the resulting distribution is an inverse Gaussian distribution with parameters $k\mu$ and $k\lambda$.

The probability density functions for the inverse Gaussian distribution with specified parameters are shown in Figure *2.10*.
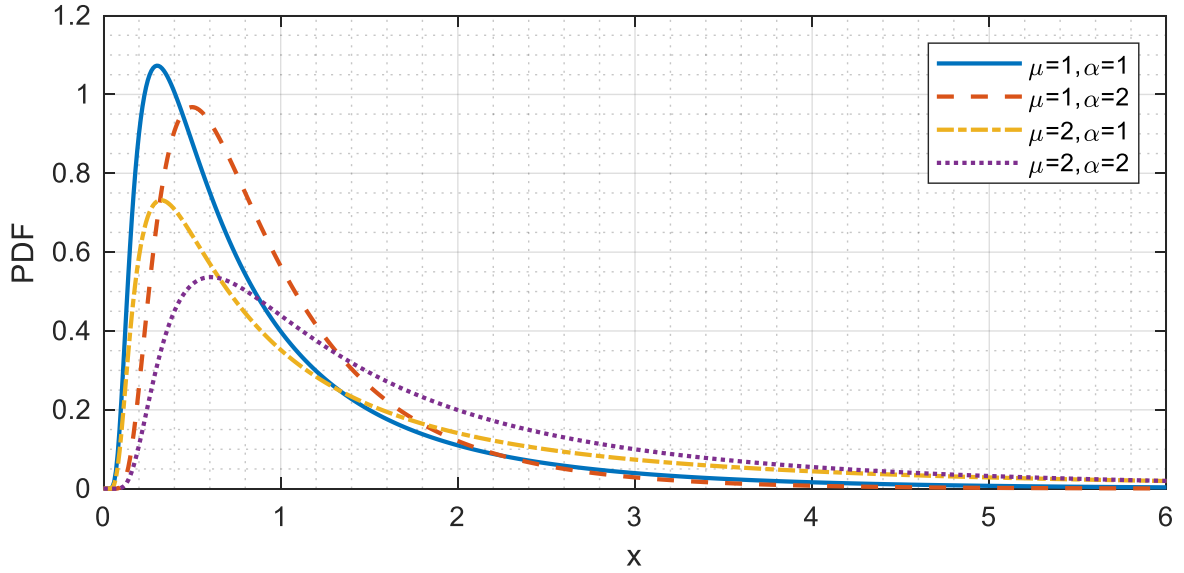
Figure 2.10 Comparison of the PDFs of inverse Gaussian distributions with different parameters

With respect to degradation modelling, Zhang (2014) proposed an inverse Gaussian process to model corrosion defect growth in buried energy pipelines using imperfect inspection data. Few researchers have studied incorporating random effects and covariates into inverse Gaussian process models with illustrations to demonstrate its applicability to degradation modelling. Case studies have been performed to illustrate application of inverse Gaussian process with random effects to plan accelerated degradation test experiments (Ye et al., 2014) and fit laser data (Peng, 2015; Wang and Xu, 2010; Ye and Chen, 2014). Furthermore, Peng et al. (2014) demonstrated the applicability of the Bayesian method for degradation modelling with inverse Gaussian process.

2.3.3 Geometric Brownian Motion

Before discussing geometric Brownian motion, we would like to discuss the Brownian motion. This is because the Brownian motion is incorporated into most versions of the geometric Brownian motions proposed in the literature for deterioration modelling. The

standard Brownian motion or Wiener process $\{B_t, t \geq 0\}$ is a continuous time stochastic process with the following properties

- $B(0) = 0$ almost sure;

- B_t has independent increments i.e. for $0 \leq t_1 < t_2 \leq t_3 < t_4, B(t_2) - B(t_1)$ and $B(t_4) - B(t_3)$ are independent;

- $B(t)$ follows a Gaussian distribution with mean 0 and variance t for all $t \geq 0$.
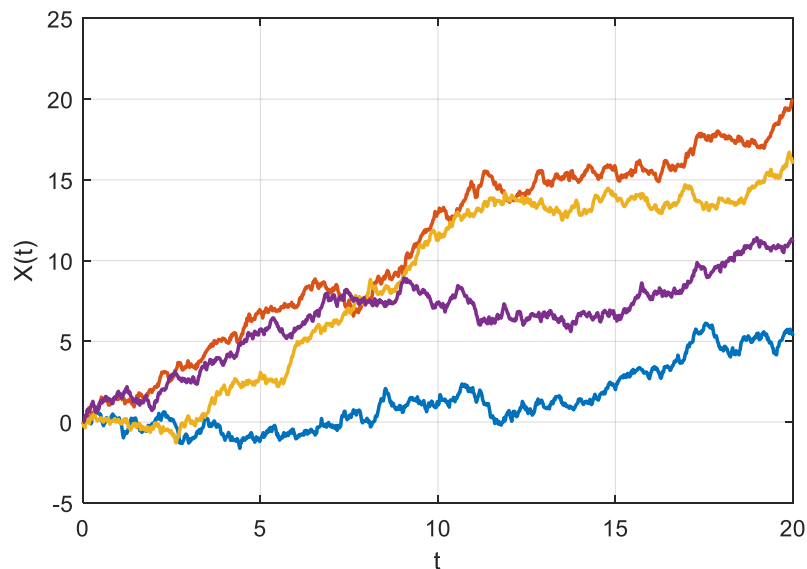
Figure 2.11 Four realizations of a 10000-step Brownian motion X(t) = 0.75t + B(t)

As observed in Figure 2.11, a major drawback of the Brownian motion for degradation modelling is that it does not ensure monotonic (non-decreasing) degradation growths. As a result, the use of geometric Brownian motion processes to model degradation in engineering has been reported in the literature.

A geometric Brownian motion $\{X(t), t > 0\}$ is a continuous-time stochastic process whose logarithm of the random part follows a Brownian motion with drift. Basically, the geometric Brownian motion was proposed as a variation of the Brownian motion to ensure positive degradation paths. In a paper by Park and Padgett (2005), the authors proposed an

accelerated lifetime test model based on geometric Brownian motion. In their model, $X(t)$ is a geometric Brownian motion process defined as

$$\log X(t) - \log X_o = B(t) \tag{2.24}$$

where $X(t)$ and X_o are the instantaneous degradation process and initial degradation respectively. This model, which may otherwise be rewritten as $X(t) = X_o e^{B(t)}$, does not ensure non-decreasing degradation growths. Elsayed and Liao (2004) proposed a geometric Brownian motion model for instantaneous degradation rate to ensure an increasing monotonic degradation pattern as

$$X(t) = X_o e^{bt} e^{\sigma B(t)} \tag{2.25}$$

where b is the drift parameter and σ is the diffusion parameter. Zhang (2014) extended their model by formulating a defect-specific hierarchical Bayesian corrosion growth model that includes inspection data measurement error in the model.

2.3.4 Markov Chains

Degradation may be assumed to be a Markov chain. For a Markov chain, the future state depends on the present (or the most recent state) and independent of the past. This class of Markov chains is sometimes described as memoryless. Markov chains can be discrete or continuous time. A discrete-time Markov process $X(t)$ is a Markov chain if

$$P\{X(t+1)|X(1), \dots, X(t)\} = P\{X(t+1)|X(t)\} \tag{2.26}$$

It is characterized by the transition probability defined by $p_t(i,j) = P\{X(t) = j | X(t-1) = i\}$ where i and j are transition states. When this probability does not depend on time, the Markov chain is said to be homogeneous, otherwise nonhomogeneous. States in Markov chains can be either discrete or continuous. Figure *2.12* is an illustration of discrete state Markov chain. Apart from the so call "memoryless" Markov chains, there is a class of Markov chains in which the future state of the process

depends on some of the past states. A Markov chain of order m is a stochastic process satisfying

$$P(X_{n+1}|X_n, X_{n-1}, \dots, X_1) = P(X_{n+1}|X_n, X_{n-1}, \dots, X_{n-m}) \qquad 2.27$$

That is, the future state of the process depends on the previous m states.
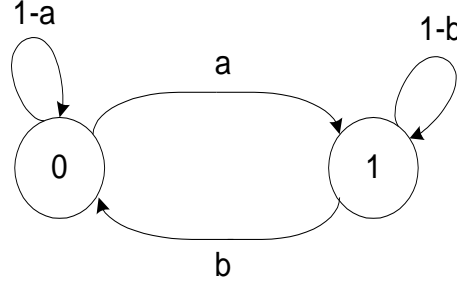
Figure 2.12 A 2-state Markov chain with transition matrix $P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$ where $0 \leq a, b \leq 1$.

For a continuous time Markov chain, $X_t, t > 0$,

$$P\left(X_{t+s}|X_s, X_{s_n}, \dots, X_{s_o}\right) = P(X_{t+s}|X_s) \qquad 2.28$$

for any $0 \leq s_o < s_1 \dots < s_n < s$. This definition still means that the future is predicted based on the given present state and independent of all the past states.

In the literature, many researchers have proposed Markov chain models for deterioration modelling. For instance, case studies have been performed to model deterioration of buildings based on discrete-time Markov chain (Sharabah et al. 2006) and degradation of reinforced concrete over time using homogeneous Markov chain (Possan and Andradec, 2014). Jin and Mukherjee (2014) presented a nonhomogeneous Markov chain model for degradation using a MC sampling approach. In the paper, the authors proposed a method of estimating empirical transition probabilities. While many Markov chain degradation models proposed in the literature are based on the independent increment assumption, Guida and Pulchini (2011), on the other hand, proposed a discrete-time

Markov model for degradation process with dependent increments. Similarly, Giorgio et al. (2011) proposed a Markov chain degradation model with transition probabilities between states dependent on the current age and state of the system.

2.3.5 Multivariate Degradation Modelling

Many previous studies focused on modelling degradation phenomena as independent stochastic processes. However, many structures or components experience multiple degradation phenomena which are dependent on one another. This may be due to the degradation phenomena having common underlying causes or the phenomena interacting because of proximity to one another. Assuming independence, in this case, may underestimate or overestimate lifetime prediction of such structures or components.

There have been previous studies on stochastic modelling of dependent degradation. As bivariate degradation modelling is the basis and a special form of multivariate degradation modelling, the former is reviewed first. In the early days of bivariate degradation modelling, Whitmore et al. (1998) proposed a two-dimensional Wiener process to model degradation. Their model comprises two processes - the component, which is directly observable, is the marker while the other component, which is unobservable, determines the failure time. Both components are correlated and have a bivariate Gaussian distribution. In recent years, more researchers have proposed bivariate degradation models based on inverse Gaussian processes (Peng et. al. 2016; Duan et. al 2018) with dependence between the performance characteristics described by copula functions. Pan et. al. (2013) and Xu et. al. (2018) proposed bivariate degradation models based on Wiener processes. In both models, copula functions were used to capture the dependency among the performance characteristics and parameter estimation done using Bayesian Markov Chain Monte Carlo (MCMC). Hao and Su (2014) also proposed a degradation model for two performance characteristics using the Frank copula function and MCMC for parameter estimation because the model is "complicated and analytically intractable". Several variants of bivariate degradation models have been proposed such as models in which performance characteristics have different marginal distribution

functions (Sari et. al. 2009) or different stochastic processes govern the behaviour of each performance characteristic (Rodriguez-Picon et. al., 2017).

Fewer publications have dealt with the issue of multivariate degradation modelling. Liu et al. (2014) proposed a model for multiple degradation processes with marginal inverse Gaussian processes. Li and Xue (2014) proposed a multivariate degradation model based on multivariate Wiener process. Xu et. al. (2016) and Xu et. al. (2018) proposed multivariate degradation models with vine copulas and drift Brownian motions to describe relationship between more than two dependent performance degradation processes. Peng et. al. (2016) proposed a Bayesian multivariate degradation model with a multivariate copula function and inverse Gaussian marginals.

Over the last decade, deterioration modelling with gamma processes has been extended to bivariate and multivariate gamma processes. Wang et al. (2015) proposed a bivariate nonstationary gamma degradation process. Their model assumed that a product state could be described by two dependent performance characteristics whose degradation mechanisms both follow nonstationary gamma processes. Also, a copula function was used to characterize the dependence structure. An earlier paper by Pan and Balakrishnan (2011) proposed a bivariate stationary gamma degradation model for reliability analysis of products with two dependent performance characteristics. More recently, Shemehsavar (2014) proposed a monotonically increasing bivariate gamma model with latent component and marker. Similar to Whitmore et. al. (1998), the latent process cannot be observed and determines the failure time while the second (i.e. the marker) can be observed. Both processes follow Kibble's bivariate gamma distribution with the same positive shape parameter and a scale parameter of 1. Variants of bivariate gamma process models have been proposed by Caballé et al. (2015) and Castro et al. (2015) by modelling multiple degradation growths and sudden shocks in a system using gamma processes with initiation times following a nonhomogeneous Poisson process. Both competing degradation growths and sudden shocks were treated as dependent, but the degradation processes were assumed to be independent of one another. Pan et. al. (2016) extended the work of Pan and Balakrishnan (2011) by proposing a multivariate

gamma process model, which uses a multivariate Birnbaum-Saunders distribution for lifetime distribution.

There have been fewer studies on multivariate degradation modelling for two reasons. First, modelling dependence among multiple processes requires constructing suitable multivariate distributions to describe dependence among the processes. This challenge is usually tackled by employing multivariate copulas (Li and Xue 2014; Liu et. al. 2014; Xu et. al. 2018). Second, parameter inference from multivariate process models are usually complicated or computationally difficult to deal with. An approach to address this issue involves using a 2-stage MLE-based method for parameter estimation (Liu et. al. 2014), Bayesian Markov Chain Monte Carlo simulation (Pan et. al. 2016; Peng et. al. 2016) or deriving parameters from transformation and modified likelihood functions (Li and Xue 2014). To address both challenges, we propose a copula-based model, with a well-behaved likelihood function, for dependence modelling.

2.4   Inspection and Maintenance Optimization

As structural components experience deterioration over time, inspections and maintenance need to be carried out from time to time to safely keep them in service. This section presents a review a widely used maintenance strategy in the literature known as condition-based maintenance.

2.4.1 Condition-Based Maintenance

In recent times, maintenance strategy is increasingly shifting towards condition-based maintenance (CBM) methods. Under the CBM strategy, the decision to repair or replace a component/system or to do nothing is based on the outcome of the inspection of the component or system. There has been a lot of research done to answer questions as to when and how often inspection and maintenance should be carried out and what should be the basis of these decisions.

In the literature, inspection and maintenance decisions are based on criteria such as expected cost minimization, availability or reliability maximization or a combination. In this

section, we review few of the models that have been proposed in the last two decades. Guillaumot et al. (2003) proposed a cost minimization model based on Markovian decision process that incorporates measurement errors during inspections. Onoufriou and Frangopol (2002) presented applications of different reliability-based inspection and maintenance optimization techniques for offshore structures and bridges. These techniques were based on minimizing life-cycle costs subject to a target reliability being met. Similarly, Rangel-Ramírez and Sørensen (2012) proposed a risk-based inspection planning optimization to minimize the service life costs subject to a threshold reliability level. Other models include Rackwitz and Joanni (2009) who proposed a cost-benefit maintenance optimization model and Pandey et al. (2009) who proposed a cost rate minimization model with inspection interval and preventive maintenance threshold as decision variables. With respect to multi-objective optimization, an optimization methodology involving cost minimization and safety maximization can be found in Podofillini et al. (2006). For more comprehensive reviews of the condition-based maintenance strategy, interested readers may refer to Ahmad and Kamaruddin (2012) and Alaswad and Xiang (2017).

2.4.2 Lifetime Prediction

A major benefit derivable from degradation modelling of a structure is being able to estimate reliability and predict the lifetime of the structure. Generally, a structure, component or system is considered to have failed when the cumulative degradation in it reaches a predetermined failure threshold ζ. This means that failure does not have to be catastrophic. The failure is characterized by a lifetime distribution which basically is a probability density function defined over a range of time. Its cumulative density function (CDF) $F(t)$ gives the probability that the component will fail by a given time t. The CDF is defined as $F(t) = P(T \leq t) = P(X(t) \geq \zeta)$. For more information on lifetime distribution, see Van Noortwijk (2007) and Yu et al. (2008).

Many models based on several stochastic process have been employed to estimate remaining useful life of components and systems. On one hand, Wiener, inverse Gaussian and gamma process-based models are used to model continuous time states.

A major advantage of using the Wiener process for deterioration modelling is that the PDF of the first passage time of the process has an analytical function which is the inverse Gaussian distribution. Tseng et.al. (2003) proposed a method for assessing the lifetime distribution of passed units of LED lamps based on the Wiener process. More recently, based on the Wiener process and inverse Gaussian distribution, Hu et. al. (2018) proposed a method to predict the real-time remaining useful life of wind turbine bearings. Wang & Xu (2010) obtained the failure time distribution using an inverse Gaussian process model. For discrete time degradation models, on the other hand, Markov chain-based models are used. These include models based on hidden Markov model (Zhou et. al. 2010) and hidden semi-Markov Model (Yu, 2009), to name a few.

Gamma process has been used to model degradation, predict reliability, and compute lifetime and remaining lifetime distribution of components (Yuan, 2007). Still on the component level, Wei and Xu (2014) presented a method to estimate remaining useful life of batteries using a gamma process. In their paper, Monte Carlo simulation was used to obtain lifetime distribution while failure threshold was a combination of both a constant and a Gaussian random variable. In a similar vein, Nystad et al. (2012) proposed a nonstationary gamma process to model a degradation phenomenon with gamma-distributed failure threshold and estimate the remaining useful life. On the system level, Khorasgani et. al. (2016) developed a stochastic simulation and inverse-FORM (first order reliability method) approaches for predicting remaining useful life of subsystems and systems with several components experiencing deterioration.

2.5   Statistical Estimation of Parameters

Parameter estimation involves using sample data to estimate parameters of a model. As it is practically impossible to obtain an entire population, these estimated parameters give information about the population. In this subsection, we will discuss two methods of estimating parameters of models namely the maximum likelihood method and the Markov chain Monte Carlo simulation.

2.5.1 Maximum Likelihood Method

This thesis adopts the maximum likelihood (ML) method for parameter estimation. The idea behind the ML estimation is to find the value of a parameter or parameters that maximize(s) the likelihood of obtaining a set of observations. With respect to the multivariate gamma distribution presented in Section 2.1.2, the joint distribution function of the function is given in Eq. 2.11 of Section 2.1.2. The log-likelihood function for the joint distribution function is

$$l(\alpha, \beta, \theta) = \log g_n(\boldsymbol{x_i}, \alpha, \beta, \theta) \qquad\qquad 2.29$$

The constant terms can be safely dropped from the log-likelihood function and the maximum likelihood estimates obtained numerically. This is equivalent to solving the set of equations

$$\frac{\partial l}{\partial \alpha_i}, \frac{\partial l}{\partial \beta_i}, \frac{\partial l}{\partial \theta} = 0; i = 1, \dots, n \qquad\qquad 2.30$$

Moving on to a brief discussion of some properties of the estimator. Under regularity conditions, a maximum likelihood estimator possesses a number of important properties namely

- Consistency
- Asymptotic unbiasedness
- Asymptotic normality
- Efficiency
- Invariance

The maximum likelihood estimator is consistent if the estimator converges in probability to its true value as the sample size goes to infinity. The estimator is asymptotically unbiased and normally distributed i.e. $(\widehat{\Theta} - \theta) \sim N(0, \boldsymbol{I}^{-1})$, where $\boldsymbol{I}^{-1}$ is the Fisher information matrix. It is also known that the maximum likelihood estimator is asymptotically efficient and achieves the Cramer-Rao lower bound when the sample size approaches infinity. This bound refers to the sample's Fisher information matrix inverse

evaluated at the true value of the estimator. Finally, the estimator is invariant. In other words, $f(\widehat{\Theta})$ is an estimator of $f(\theta)$ if $\widehat{\Theta}$ is an estimator of θ.

Due to their independent increment property, gamma, inverse Gaussian and Brownian motion processes are said to be Markovian processes. The independent increment property of these stochastic processes makes parameter estimation by maximum likelihood possible and prediction of future degradation uncomplicated. For a gamma process model with covariates and random effects, Lawless and Crowder (2004) discusses parameter estimation using the MLE method. Guida and Pulcini (2011) and Giorgio et al. (2011) both employed the maximum likelihood method to estimate parameters of their age- and state-dependent Markov chain models. The former model involves random effects. In a similar vein, a few researchers have used the MLE method to estimate parameters of their geometric Brownian motion-based degradation models, among whom are Elsayed and Liao (2004), Ye et al. (2014) and Ye and Chen (2014). Wang and Xu (2010) proposed the Expectation-Maximization algorithm to find the MLE iteratively when direct estimation of the parameters using the MLE method is computationally difficult.

So far, the discussion has been on using the MLE method to estimate parameters of univariate stochastic processes. Maximum likelihood has been used for statistical inference with respect to bivariate and multivariate random processes. For example, Liu et al. (2014) used a two-stage method to estimate parameters of their multiple degradation processes. First, the authors estimated the marginal distribution parameters and then the copula parameters from the likelihood function using the results from the first stage. Statistical inference using the MLE method will be discussed in detail in Chapter 5.

2.5.2 Markov Chain Monte Carlo Simulation

Another versatile method used for parameter estimation is Markov Chain Monte Carlo (MCMC) simulation, which is based on Bayesian statistics. MCMC simulation is used to sample from a multivariate continuous distribution by constructing a Markov chain that

ultimately converges to a desired distribution. In using MCMC, some initial iterations of the Markov chain are discarded. This is known as the "burn-in" period. After the burn-in period, the more the number of iterations used in the MCMC simulation, the higher the chance of obtaining the desired distribution. MCMC samples are correlated because a typical iteration depends on the previous iteration. In the ordinary Monte Carlo simulation, however, the samples are independent and identically distributed.

A number of algorithms have been used in MCMC simulation. These include the Metropolis-Hastings (M-H) algorithm and the Gibbs sampler among others. MCMC simulation was invented by Metropolis et al. (1953) to investigate state equations of interacting particles in Physics. The simulation algorithm was later improved upon by Hastings (1970). Basically, the M-H algorithm involves initializing parameters, usually denoted as $\boldsymbol{\theta}$. Then, the algorithm uses what is known as the proposal distribution $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$ to generate random seeds for the next iteration based on the current values of the parameters. The proposed move resulting from the generated seeds is accepted if it satisfies $u < \min\left(1, \frac{p(\boldsymbol{\theta}^*)q(\boldsymbol{\theta}_i|\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}_i)q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_i)}\right)$, where u and $p(\boldsymbol{\theta})$ are a uniformly distributed random variate and a distribution proportional to the target posterior distribution. Otherwise, the Markov chain remains in the same location. For a symmetric proposal distribution where $q(\boldsymbol{\theta}_i|\boldsymbol{\theta}^*) = q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_i)$, the acceptance criterion reduces to $u < \min\left(1, \frac{p(\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}_i)}\right)$. An improvement to the M-H algorithm is presented in Green (1995).

In the Gibbs sampler algorithm introduced by Geman and Geman (1984), the algorithm uses the full conditional distribution of the posterior distribution to generate a random sample of each parameter of interest in turn. This random sample is conditioned on the present values of the other parameters. The proposed move in the Markov chain is always accepted. In other words, the Gibbs sampling algorithm may be regarded as a special case of the M-H algorithm with an acceptance rate of one.

2.6   Concluding Remarks

In this section, a summary of major contributions and gaps in extreme value distribution of random fields and stochastic modelling of deterioration are presented in Table 2.2.

Table 2.2 Summary of major contributions and gaps in dependence modelling, stochastic deterioration modelling and extreme value distribution of random fields.

Major Contributions	Gaps
• Ditlevsen proposed a distribution for EV of one-dimensional standard Gaussian fields within a finite interval.	• The distribution has an unknown parameter G which limits its applicability. • No analytical solution for EV distribution of Gaussian random fields of higher dimensions.
• Asymptotic solutions from the classical EV and the outcrossing theories exist. • Double transformations of Gaussian fields to non-Gaussian fields have been demonstrated.	• Extensive literature review indicates that there is no analytical solution for the EV distribution of non-Gaussian random fields.
• Stochastic process models that capture temporal uncertainty in deterioration have been proposed.	• Spatial uncertainty modelling in deterioration receives little attention.
• Use of copulas to study dependence structures with respect to degradation modelling and reliability analysis have been demonstrated.	• Difficulty in parameter estimation due to complicated and analytically intractable models; actual applications of multivariate stochastic process model in civil engineering were rare

3. Extreme Value Distribution of Homogeneous Random Fields

In reliability problems, it is often required to calculate the distribution of maximum values of functions. When such problems are modelled with random fields, the EVs of random fields become important quantities in risk and reliability analyses. As a matter of fact, to assess overall risk during structural integrity assessment, the second order statistics of the stochastic field (mean, standard deviation and correlation length) are insufficient, and it is often desirable to obtain the full probability distribution of the extreme value. The extreme value (EV) theory provides useful probabilistic tools for handling the maximum values and their distribution.

This chapter presents the methodology and results of extreme value distribution of random fields. While the main focus of the research is on the EV distribution of two-dimensional random fields, we start off by evaluating EV distributions of one-dimensional random fields by comparing our simulation results with existing semi-analytical and analytical models. The simulation results are then extended to two-dimensional fields. For all the fields considered, the simulation study aims to answer the following questions:

(1) What distribution does the EV follow? How accurate is the asymptotic Gumbel distribution in the modelling of the EV? For fields with a quadratic exponential correlation model, how accurate are the Ditlevsen, approximate Ditlevsen distributions and the Poisson approximation?

(2) How does the statistics of the EV vary with the parameters of the random field, i.e., the size of the field (L), the correlation length (θ), correlation model, and for gamma fields, the shape parameter (α)? For lognormal random fields, the relationship between the EV statistics and parameters of the field are studied as well. In this study, all three correlation functions listed in Section 3.1 are investigated.

Put differently, the objectives of the study are threefold: 1) introduce a non-Gaussian random field model (i.e. copula-based gamma field) for the modelling of spatial uncertainty in deterioration; 2) use a spectrally based simulation algorithm to evaluate the

extreme value distribution of the one-dimensional Gaussian, lognormal and gamma fields; 3) evaluate the accuracy of existing asymptotic and other analytical or semi-analytical solutions and 4) extend the spectrally based simulation algorithm to evaluate the EV distribution of two-dimensional standard normal and gamma fields. Although the study is performed in the context of deterioration modelling, the results and methodology are nevertheless applicable to other areas where the extreme values of a stochastic field or process within a finite time interval are of practical significance.

The chapter is organized as follows. The basics of the copula-based gamma field are first introduced in terms of definition and characterization. The KL expansion in one dimension and multidimensional KL expansion are discussed. Furthermore, the Nystrom method that numerically solves the Fredholm integral equation of the second kind using the Gauss-Legendre quadrature is discussed. Then, the results from the one-dimensional standard normal, lognormal and gamma fields are discussed. Finally, the results of EV distribution of two-dimensional fields are presented.

3.1   Definition and Characterization

A gamma random field can be simply defined as a random field $X(s)$ of which the values taken at any n distinct locations $X(s_1), X(s_2), ..., X(s_n)$ follow a multivariate gamma distribution for any positive integer n. However, there are a number of multivariate gamma distributions that may be qualified for the definition of the gamma random field. Therefore, we need to present the formal definition of the gamma field and explain our choice of multivariate gamma distribution.

A random field $X(s)$ in a finite d-dimensional Euclid space Ω is called a homogeneous gamma field if the following conditions are satisfied:

(1) For any point $s \in \Omega$, the value of the field $X(s)$ is a random variable that follows a gamma distribution with shape α and scale β, i.e., $X(s) \sim Ga(\alpha, \beta)$.

(2) For any n points $s_1, s_2, ..., s_n \in \Omega$, the field values $X(s_1), X(s_2), ..., X(s_n)$ follow a multivariate gamma distribution that is defined as Eqs. (2.23) and (2.24) with $\alpha_i =$

α and $\beta_i = \beta$ for $i = 1, \dots, n$, and the z-correlation coefficient r_{ij} $(i, j = 1, \dots, n)$ defined by $r_{ij} = r(\boldsymbol{h}_{ij})$, where $r(\cdot)$ is a correlation function, and $\boldsymbol{h}_{ij}$ a certain distance measure between the two points s_i and s_j.

Therefore, the definition of the gamma field is very similar to that of a homogeneous Gaussian field, except that the marginal distributions are different and that the correlation in gamma field is not defined in the x-space but in the z-space. In Grigoriu's terminology, the gamma field is a non-Gaussian translation field (Grigoriu 1984).

The gamma distribution mentioned in the first condition is a two-parameter continuous probability distribution whose probability density function (PDF) is expressed as

$$g(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha} \Gamma(\alpha)} \qquad 3.1$$

for $x \geq 0$, where $\alpha > 0$ and $\beta > 0$ are the shape and scale parameters, respectively. The cumulative distribution function (CDF) is expressed as a ratio of two gamma functions

$$G(x; \alpha, \beta) = \frac{\Gamma(\alpha, x/\beta)}{\Gamma(\alpha)} \qquad 3.2$$

where $\Gamma(p, q) = \int_0^q x^{p-1} e^{-x} du$ is called the lower incomplete gamma function, and $\Gamma(p) = \Gamma(p, \infty)$ the complete gamma function, which is used in Eq. (3.1) as well.

This study adopts the Gaussian copula for construction of a multivariate gamma distribution because of its ease in concept and computation as well as its flexibility in dependence modelling. For detailed information about constructing multivariate distributions, refer to Joe (2001) and Kotz et al. (2000).

The correlation function is a very important second-order characteristic of a homogeneous random field. For one-dimensional field, commonly used correlation functions are exponential, Gaussian, spherical, and triangular (Figure 3.1).

- Exponential: $r(h) = e^{-3h/\theta}$

- Quadratic exponential: $r(h) = e^{-3(h/\theta)^2}$

- Triangular: $r(h) = 1 - h/\theta$ for $h \leq \theta$, or 0 otherwise
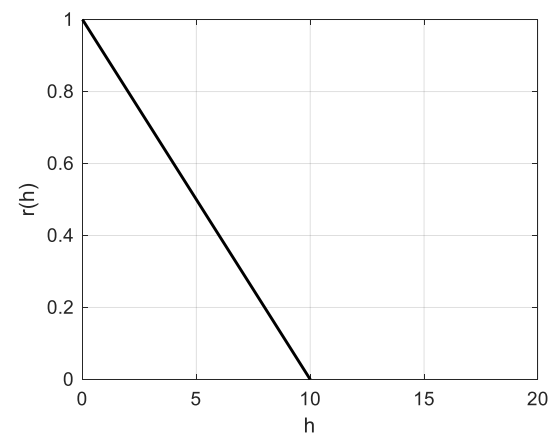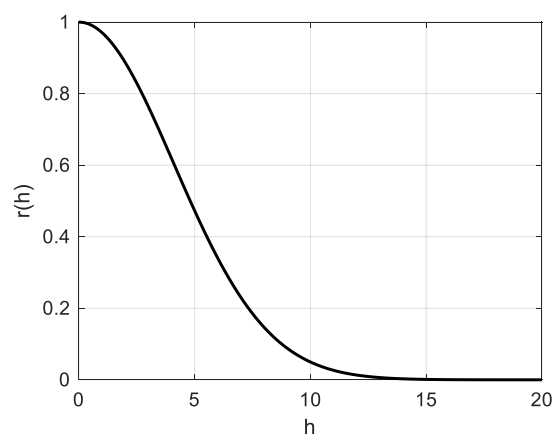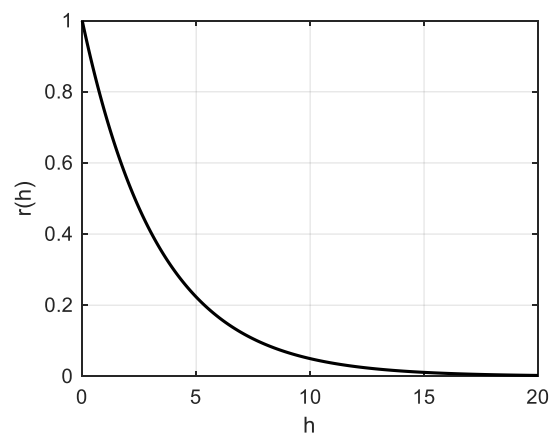
Figure 3.1 Commonly used correlation functions with θ=10 (Top: exponential; Middle: quadratic exponential; Bottom: triangular).

In these models, h represents the Euclidean distance between any two points in the line, and the parameter θ in the above correlation models is called correlation length. For multi-dimensional field, isotropy and separability are two important considerations in correlation modelling. VanMarcke (1983) discussed isotropic, ellipsoidal, separable, and direction-dependent correlation models for multi-dimensional fields. An isotropic model is very similar to a one-dimensional model as described above, except that $h = \|s_1 - s_2\|_2^{1/2}$ is defined the distance between two points. In contrast, an m-dimensional separable correlation model is expressed as $r(h_1, \dots, h_m) = \prod_{k=1}^{m} \rho_k(h_k)$, where $\rho_k(h_k)$ $(k = 1, \dots, m)$ represents the one-dimensional correlation model in the kth dimension. This study focuses on one-dimensional gamma field.

Altogether, the shape parameter α, the scale parameter β, and the correlation length θ define the 1-D homogeneous gamma field. In stochastic degradation modelling, the three parameters can be estimated from inspection data by using, e.g., maximum likelihood method and Bayesian method.

3.2   KL Expansion for Gaussian Fields

3.2.1 One-Dimension Gaussian Fields

The KL expansion is a widely used representation of stochastic processes and fields based on spectral decomposition of the autocovariance function (Ghanem & Spanos, 1991). Since the Gaussian fields considered in the study are exclusively zero mean and unit variance, the autocovariance function is the same as the correlation function. It has been shown that a stationary standard normal field $Z(s; \theta)$ defined on a finite one-dimensional domain Ω can be represented by the following sum of infinite series:

$$Z(s; \theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\theta) f_i(s) \qquad\qquad 3.3$$

where $\xi_i(\theta)$ are independent standard normal variables and λ_i and $f_i(s)$ are eigenvalues and eigenfunctions of the correlation function. These eigenvalues and eigenfunctions are obtained by solving the following Fredholm integral equation of the second kind:

$$\int_{\Omega} r(s_1, s_2) f_i(s_1) ds_1 = \lambda_i f_i(s_2) \qquad\qquad 3.4$$

where $r(s_1, s_2)$ is the correlation function, which for stationary fields is in the form of $r(h)$ in which $h = |s_1 - s_2|$.

The eigenvalues in Eq. (3.3) are often sorted in the nonincreasing order of magnitude. Since random fields in practice are often narrow banded, the KL expansion can often be truncated after a finite number of terms because the terms with very small eigenvalues can be safely neglected without affecting the accuracy of random field approximation. Hence, the random field $Z(s, \theta)$ is approximated as

$$Z(s; \theta) = \sum_{i=1}^{n} \sqrt{\lambda_i} \xi_i(\theta) f_i(s) \qquad\qquad 3.5$$

in which n represents the total number of K-L expansion terms.

Analytical solution of the integral equation (Eq. 3.4) exists for only a few functions such as exponential and triangular covariance functions as well as for the Brownian motion (Ghanem and Spanos, 1991; Wang, 2008). In view of this, numerical solutions are often employed and a number of methods have been developed. These methods are categorized into Galerkin method, collocation method, and Nystrom method; see Betz et al. (2014) for an overview and comparison of these methods used for solving the Fredholm integral in arbitrarily shaped domains. For one-dimensional random field, Melink and Korelc (2014) proposed a modified truncated exponential function to address the loss of positive definiteness within the framework of Galerkin method. Phoon et al. (2002) proposed a wavelet-Galerkin method that replaces polynomial and trigonometric bases with wavelet bases. Both methods aim to increase the numerical efficiency of the Fredholm integral solution in the KL expansion. In this paper, the Nystrom method is used to solve the linear Fredholm integral of the second kind for the eigenvalue problem. Delves and Mohamed (1997) have discussed the accuracy of this method qualitatively.

3.2.2 Extension to m-Dimensional Field

The extension of the one-dimensional KL expansion introduced in section 3.2.1 to a multivariate KL expansion is discussed in this section. A zero mean Gaussian random field $Z(s_1, \dots, s_m; \boldsymbol{\theta})$ defined on a finite m-dimensional domain Ω can be represented by the K-L expansion with infinite number of terms as

$$Z(s_1, \dots, s_m; \boldsymbol{\theta}) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \xi_k(\boldsymbol{\theta}) f_k(s_1, \dots, s_m) \qquad 3.6$$

where $\xi_k(\theta)$ are independent standard normal variables, λ_k and $f_k(s_1, \dots, s_m)$ are eigenvalues and eigenfunctions of the covariance function, and $\boldsymbol{\theta}$ is a vector of directional correlation lengths. Assuming separability of the autocovariance function in m dimensions, i.e.

$$C(\boldsymbol{s}, \boldsymbol{t}) = \prod_{i=1}^{m} C(s_i, t_i) \qquad 3.7$$

the eigenvalue for the m-dimensional Karhunen-Loeve expansion is the product of the corresponding one-dimensional eigenvalues (Sudret and Der Kiureghian, 2000).

$$\lambda_k = \prod_{j=1}^{m} \lambda_{i_j} \qquad 3.8$$

where λ_{i_j} are the component univariate eigenvalues. Similarly, the m-dimensional eigenfunction is the product of one-dimensional eigenfunctions.

$$f_k(s_1, \dots, s_m) = \prod_{j=1}^{m} f_{i_j}(s_j) \qquad 3.9$$

Similar to the 1-dimensional case, the eigenvalues are sorted in descending order of magnitude. For full details of simulation of two-dimensional random fields using the KL expansion, see Section 4.2.1.

3.3   Nystrom Method for the Fredholm Integral Equation

Using Gauss-Legendre quadrature of n integration points for the integral, Eq. (3.4) is approximated as

$$\sum_{j=1}^{n} w_j r(t, s_j) f_i(s_j) = \lambda f_i(t) \qquad 3.10$$

where w_j's are the quadrature weights and s_j's the integration points. Setting the above equation for $s = s_1, \ldots, s_n$ yields the following equation in matrix notations:

$$\mathbf{RWf} = \lambda \mathbf{f} \qquad 3.11$$

where $\mathbf{R} = \{r_{ij} = r(s_i, s_j)\}$, $\mathbf{W} = diag\{w_1, \ldots, w_n\}$, $\mathbf{f} = \{f_{ij} = f_i(s_j)\}$,   and $\lambda = diag\{\lambda_1, \ldots, \lambda_j\}$. Multiplying both sides of Eq. (3.11) by $\mathbf{W}^{1/2}$ yields

$$\mathbf{Ah} = \lambda \mathbf{h} \qquad 3.12$$

where $\mathbf{A} = \mathbf{W}^{1/2} \mathbf{R} \mathbf{W}^{1/2}$ and $\mathbf{h} = \mathbf{W}^{1/2} \mathbf{f}$.   Note that Eq. (3.12) is a typical eigenvalue problem. It can be readily shown that $\mathbf{A}$ is positive definite and therefore the eigenvalues are real positive values.   Solving the eigen problem for the eigenvalues λ and the corresponding eigenvectors $\mathbf{h}$, one can further obtain the discrete eigenvector at the integration points $\mathbf{f} = \mathbf{W}^{-1/2} \mathbf{h}$. Finally, the following Nystrom interpolation formula is used to obtain the value of the eigenfunction at any point s:

$$f_i(s) = \frac{1}{\lambda_i} \sum_{j=1}^{n} w_j r(s, s_j) f(s_j) \qquad 3.13$$

For an in-depth discussion of the Nystrom method with the Gauss-Legendre quadrature rule, see Atkinson (1997) and Press et al (2007).

The Nystrom method is coded in MATLAB. To validate it, we used the Nystrom method to calculate the eigenvalues of the exponential correlation model and compare the numerical results with the exact solutions. For illustration, Figure 3.2 shows one of the comparisons for $\theta = 1$ in log scale. Clearly the eigenvalues decrease faster than an exponential decay, which would be a straight line in the log scale. This gives us a good justification for the truncated approximation of the KL expansion. On the other hand,

59

although the relative difference seems increasing for higher order eigenvalues, the absolute values of those eigenvalues are already very small. Therefore, the numerical accuracy of the Nystrom method is deemed sufficient. See the appendix for more illustrations.
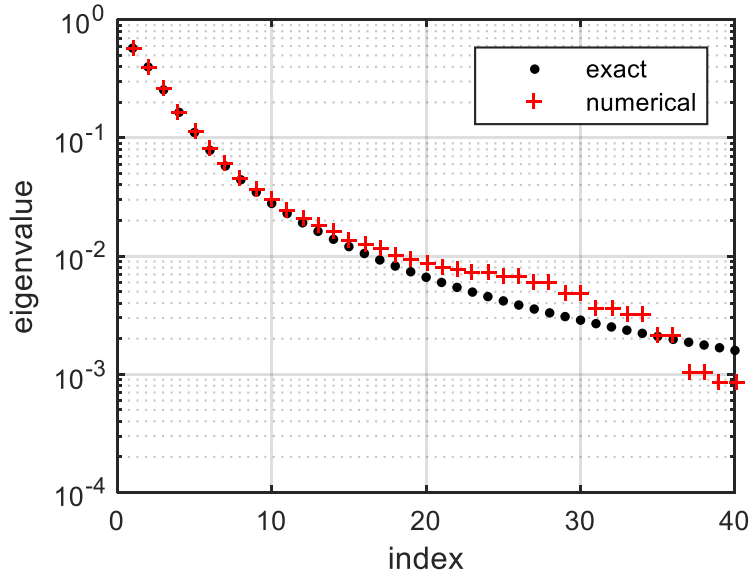
Figure 3.2 Comparison of the first 40 eigenvalues obtained from the Nystrom method with the exact solution for exponential correlation model (L=2,θ=1).

3.4   Simulation Algorithm

The simulation of the one-dimensional gamma field is based on memoryless transformation of the standard Gaussian field, which is generated by the KL expansion incorporated with the Nystrom method. As shown above, once the eigenvalues and eigen functions are numerically determined, the generation of the standard Gaussian field is a simple application of Eq. (3.5), which entails simply generation of n independent standard normal random variates. Note that the number of KL terms in Eq. (3.5) must be equal to the number of quadrature points in Eq. (3.10). However, much finer equal-distance mesh can be used in the interpolation of the eigenfunctions as shown in Eq. (3.13). After the standard Gaussian field is generated, the standard normal CDF of the Gaussian variate at each interpolation point is evaluated. Finally, the corresponding gamma field is

generated by taking the inverse of the gamma CDF with the known shape and scale parameters. The maximum value of the generated gamma field can then be evaluated among all interpolation points of the field. Repeat this process for a prescribed number of simulation runs. Finally, the statistics of the maximum values can be evaluated. This includes the calculation of the moments, empirical distribution, distribution fitting and various hypothesis testing.

Clearly, the number of KL terms (and therefore the integration points) n, the number of interpolation points (or grid size) m, and the number of simulation runs N_{sim} are three key parameters of the KL-based Monte Carlo simulation. The number of KL terms usually can be determined by the size of the smallest eigenvalue. In this study, n is chosen so that the truncated largest eigenvalue $\lambda_{n+1} \leq 10^{-15}$. The number of interpolation points also affects the accuracy of the extreme values. This value is determined by convergence test. For illustration, a gamma field with $L = 2$, $\alpha = 2, \beta = 1$, and quadratic exponential correlation model with $\theta = 1$ is simulated. Based upon the truncation criterion, the number of KL terms is $n = 22$. As shown in Figure 3.3, both the mean and standard deviation of the extreme value converges quickly with the number of interpolation points (m). As one would expect, the statistics exhibit greater fluctuation with the number of simulations. However, the figure shows that when N_{sim} is greater than 10^5, the statistics become stabilized. Based on this and a lot more similar convergence tests, the following study takes $m = 201$ and $N_{sim} = 10^6$.
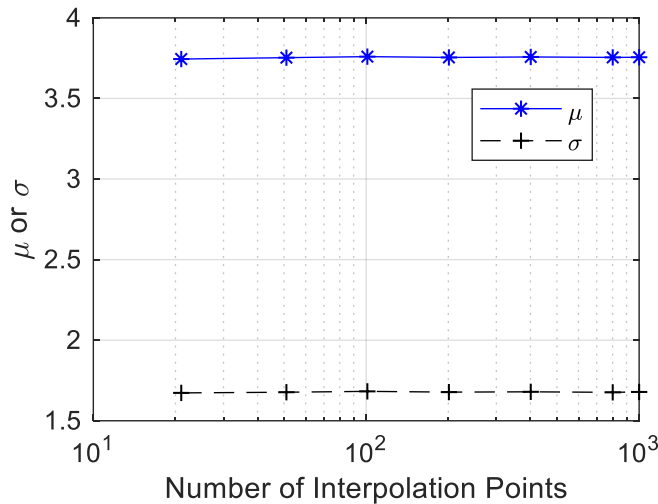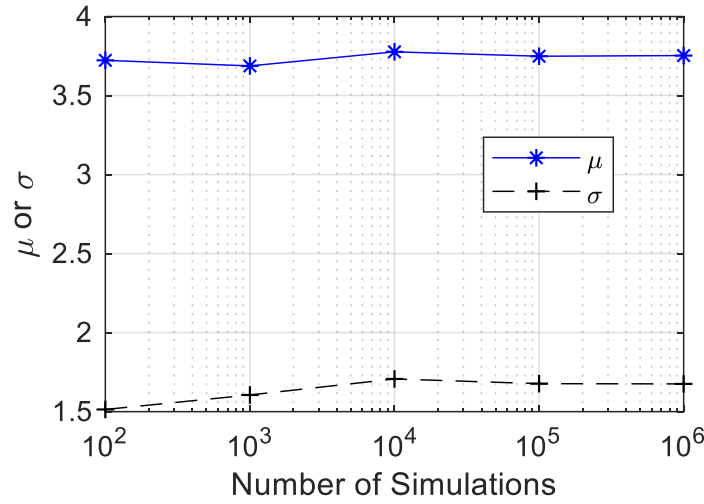
Figure 3.3 Convergence test for the number of interpolation points (top) and the number of simulations (bottom)

Using the KL expansion-based simulation algorithm presented above, the extreme values of the Gaussian, gamma and lognormal fields within a finite interval are studied.

Since both the gamma and lognormal fields are simply double transforms of the standard normal field, we discuss the key findings of the standard normal fields at first.

3.5   Standard Normal Fields

3.5.1   Distribution of the EV

As reviewed in Section 3, there are four possible distribution models that can be used to approximate the extreme values of a standard normal field with a quadratic exponential correlation model. They are: Gumbel, Ditlevsen, approximate Ditlevsen (without the G term), and the Poisson approximation. For the other two correlation models (triangular and exponential), only the Gumbel distribution is applicable because the two correlation models are not twice differentiable at the origin, a basic requirement for the calculation of the mean crossing rate. To evaluate the goodness of fit, these distributions are compared with the empirical distribution obtained from simulation. The results for the quadratic exponential model are discussed first.

Take $L = 2$ and $\theta = 1$. Figure *3.4* compares the PDF and the probability of exceedance function of the empirical (E), Gumbel (G), Ditlevsen (D), and approximate Ditlevsen (aD) distributions as well as the Poisson approximation (P). Clearly, the Ditlevsen distribution offers a very accurate result for the whole range of the distribution. Without the empirically estimated G term, even the approximate Ditlevsen distribution gives a surprisingly good fit. In contrast, the asymptotic Gumbel distribution fits poorly the empirical EV data from simulation. The Poisson approximation has worse performance when the threshold level is small, which is expected.

In the upper tail of the distributions, as shown in Figure *3.4*(b). For the probability of exceedance function, the poor fit of the asymptotic Gumbel is shown even more clearly. The Poisson approximation gives an amazingly close fit to the empirical distribution curve, up to the level beyond 10^{-4}. Note that the divergence in the empirical curve for $x > 4.5$ is caused by random error of Monte Carlo simulation (Recall $N_{sim} = 10^6$). If the number of simulation increases, continuously close fit can be expected for the Ditlevsen distribution, approximate Ditlevsen distribution, and Poisson approximation.
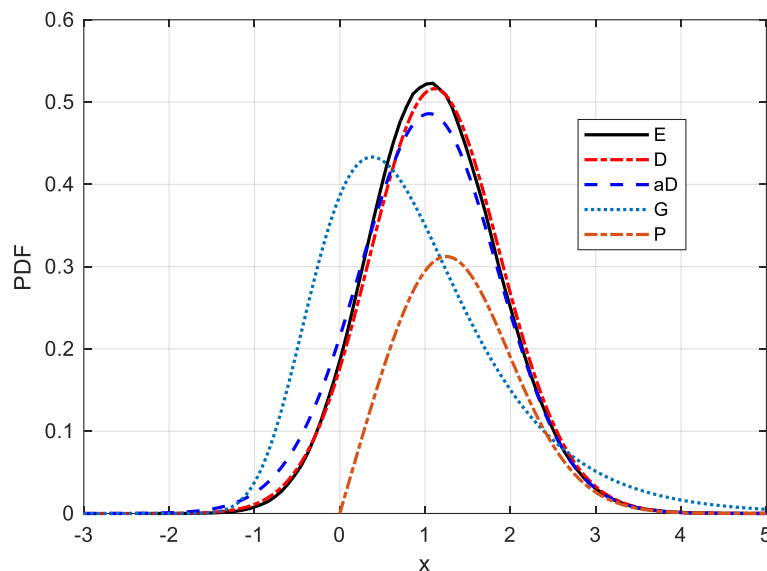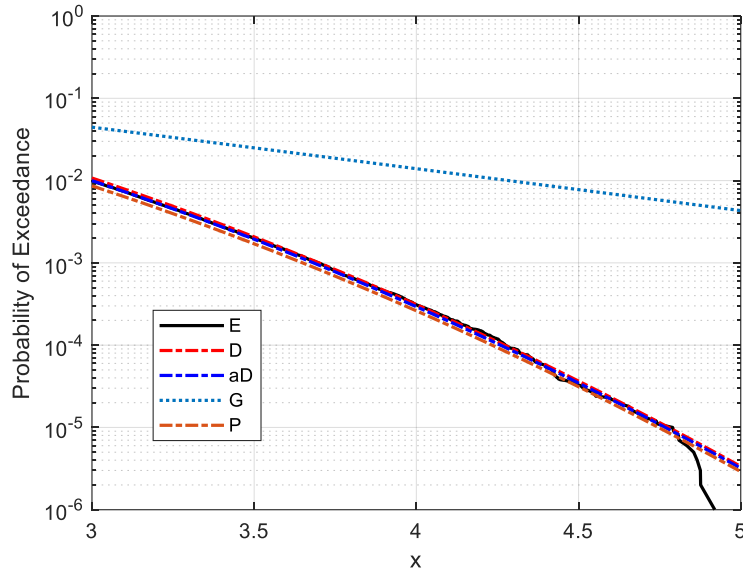
Figure 3.4 Comparison of the PDFs (top) and the probability of exceedance functions (bottom) of the empirical, asymptotic Gumbel, Ditlevsen and approximate Ditlevsen distributions for the standard normal field with L=2,θ=1 and $\rho(h) = exp\left(-3\left(\frac{h}{\theta}\right)^2\right)$.

Further study of other cases reveals the same trend of the comparison. That is, the Ditlevsen distribution provides an excellent fit for the whole range of x. The approximate Ditlevsen distribution is slightly worse than the exact version, yet still fits the empirical data very well in engineering applications. The Poisson approximation does not perform well for the left half of the distribution but provide an amazingly accurate prediction of the exceedance probability in the upper tail.

The overall good performance of the approximate Ditlevsen distribution is particularly interesting, because this distribution does not require simulation. To better understand this, we studied how the G term of the exact Ditlevsen distribution varies. First, it is found that the G value depends only on L/θ−as long as L/θ is kept constant, G remains the same. Second, the G value is a modest number. As shown in Figure 3.5(a), even when L/θ reaches 10, G is still less than 1.5. This verifies the earlier observation that although less accurate than the exact version, the approximate Ditlevsen model still provides a decent approximation to the empirical distribution in most cases.

To further improve the predictability, a power function is used to fit the $G \sim L/\theta$ curve, yielding

$$G = 0.3826 \left(\frac{L}{\theta}\right)^{0.6031} \qquad\qquad 3.14$$

for $0 \leq L/\theta \leq 10$. For this model fitting, the coefficient of determination (R-squared) is 0.9793 while the root mean square error (standard error) is 0.0594. Figure 3.5(b) shows how the difference between the empirical and fitted G terms varies with L/θ. Equation 3.14 is very useful for illustration. Instead of taking zero for G as suggested by Ditlevsen (2004), we suggest using the empirical formula shown above to obtain the G value, which is substituted into Eq. (2.28) to obtain a modified Ditlevsen (mD) distribution. To test this, Figure 3.5(c) shows the comparison of the modified Ditlevsen (mD) distribution with the empirical distribution for $L/\theta = 5$. Note that $L/\theta = 5$ represents one of the spots that has the poorest fit in Figure 3.5(a), yet the mD distribution still provides a very good fit.
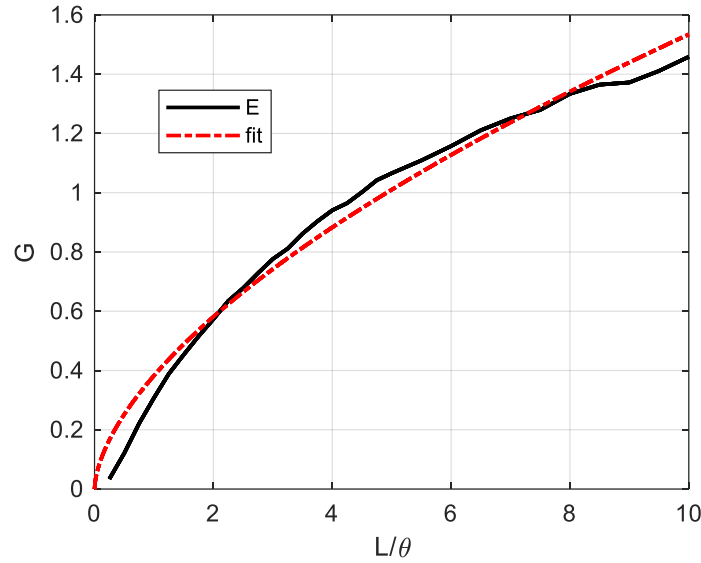
(a)

(b)

(c)
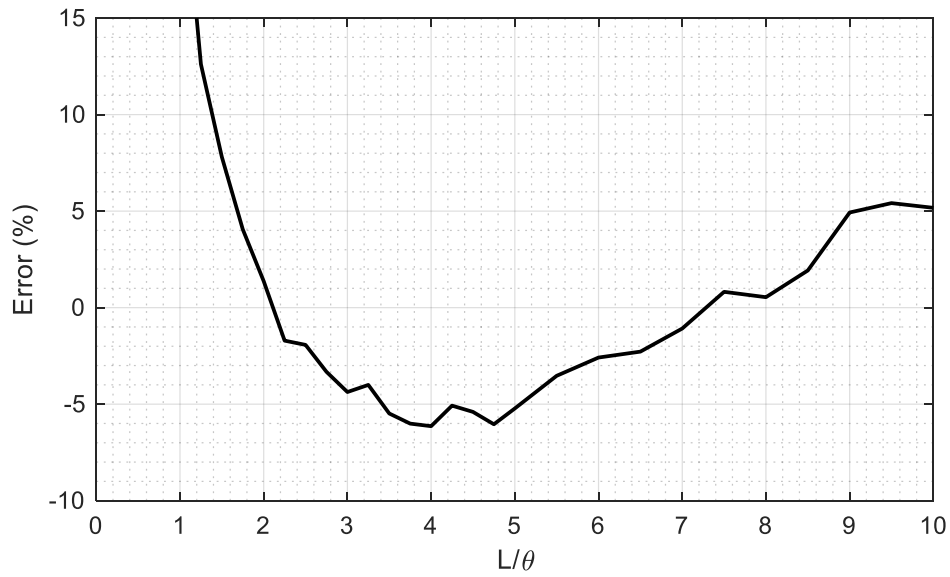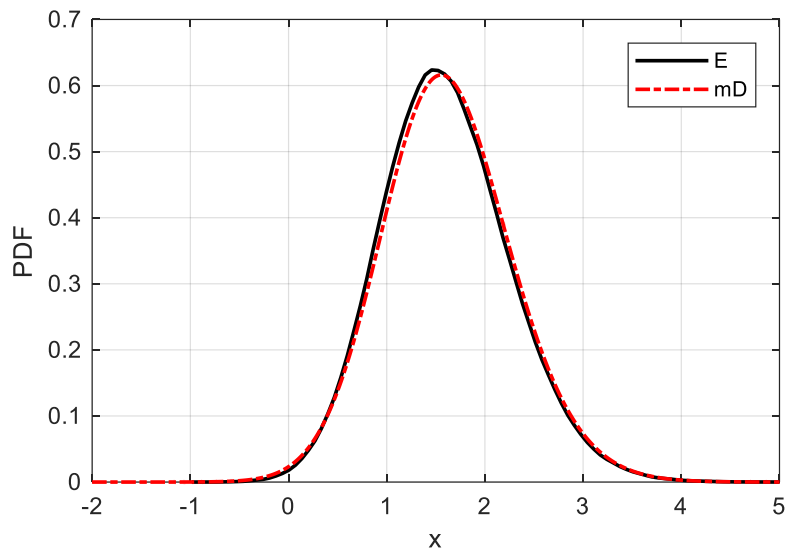
Figure 3.5 The G~L/θ relation and performance of the modified Ditlevsen distribution.

Considering the high computation cost in the estimation of G, we recommend that the modified Ditlevsen model be used in practice. Note, again, that the Ditlevsen, approximate Ditlevsen, and the Poisson approximate methods are only applicable to

mean-square differentiable random fields. If the random field is not mean-square differentiable, one must return to the KL expansion-based simulation.

For standard normal fields with either an exponential or a triangular correlation model, only the asymptotic Gumbel distribution can be used to benchmark the empirical distribution. However, simulation results have shown that in general, the asymptotic Gumbel model is not found to be a good fit for all three correlation models, although for some values of L/θ, the Gumbel distribution can be very close to the empirical CDF. Sensitivity analyses found that the performance of the asymptotic Gumbel distribution depends highly upon the size of the random field. This is because the two normalization parameters (a_L and b_L) of the Gumbel distribution both are influenced by L and θ differently, whereas the empirical distribution of the EV is in fact invariant for a fixed L/θ. As an extreme case, when $L \leq 1$ (regardless of θ), the asymptotic Gumbel model cannot be used at all because $a_L = \sqrt{2 \log L}$ returns a complex number.

For illustration, Figure *3.6* compares the Gumbel distribution with the empirical CDF at different sizes of the fields while L/θ is kept to be 5. Three correlation models – quadratic exponential, exponential and triangular models are considered. For all three models, the empirical distribution remains unchanged for the same L/θ, whereas the Gumbel distribution keep changing as the field size changes. The same trend is observed for other L/θ values as well. It is thus concluded that the asymptotic Gumbel distribution should NOT be used in the modelling of the extreme value of random field with a finite region, regardless of the correlation model.
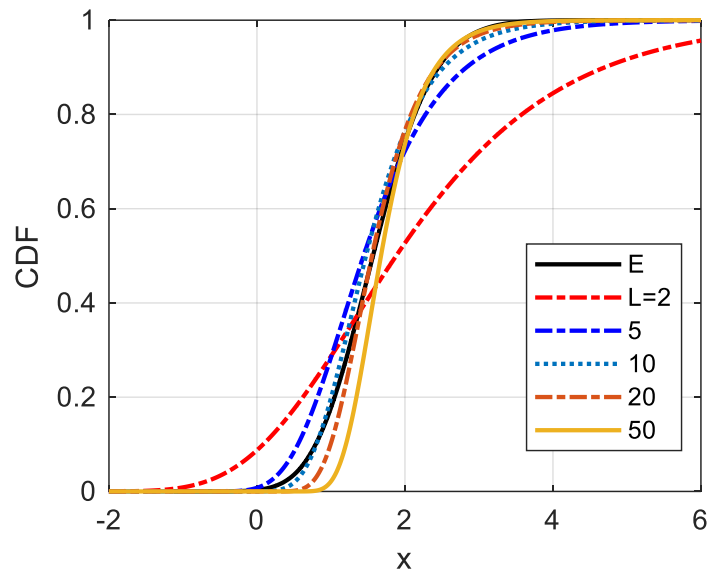
(a)
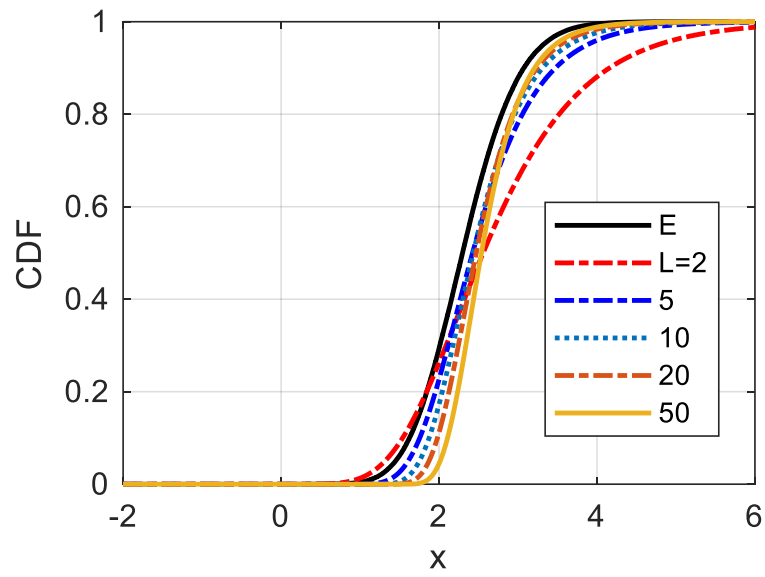
(b)
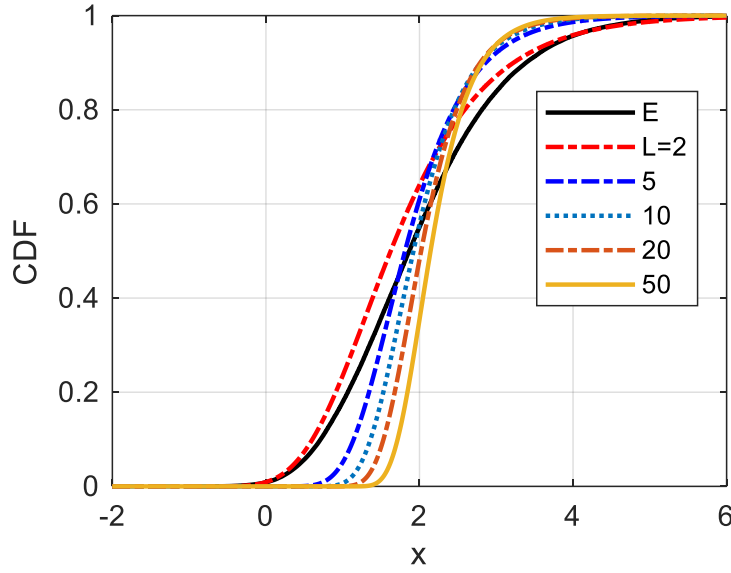
(c)

Figure 3.6 Comparison of empirical and asymptotic Gumbel distributions of standard normal fields (L/θ=5) for three different correlation models: (a) quadratic exponential; (b) exponential; (c) triangular.

3.5.2 Statistics of the EV

An important finding of the study is that the statistics of the EV of a finite stationary random field depend upon only L/θ for all three correlation models. For stationary standard normal fields with a quadratic exponential correlation function, simulation results show that the mean and standard deviation of the extreme values are a function of the ratio of the length of the random field to the correlation length L/θ. It is also established from simulation that for a constant L/θ, the probability or cumulative distribution of extreme values is invariant. An investigation of the reason behind this observation reveals that, while keeping L/θ constant, if one changes L by a factor, say c, the distances between the integration points will change by c. As a result, the quadrature weights will become $c\boldsymbol{W}$, and yet the correlation matrix of the integration points $\boldsymbol{R}$ will remain the same. Following the derivation in Section 3.4, it can be shown that all the eigenvalues will be multiplied by c, and the corresponding eigenfunctions will be scaled by $c^{-1/2}$. Taking the root of the eigenvalues and multiplying it with the corresponding eigenfunction, as required in Eq. (3.5), yields the

69

same KL expansion for the scaled stochastic field. For the exponential and triangular correlation models, Ghanem and Spanos (1991) provided the analytical expressions for the eigenvalues and eigenfunctions. It can be readily shown that the same scaling occurs in them as it does in the quadratic exponential model (see the appendix for details). Therefore, our study of the trend of the EV statistics was performed in terms of L/θ.

Figure *3.7* shows the trends of the mean and standard deviation of the extreme values along L/θ. The figure also compares the empirical results from the simulation with the analytical results derived from the asymptotic Gumbel distribution. Based upon the simulation, the mean extreme value increases as L/θ increases, as it should be. However, the standard deviation decreases with L/θ, except for the triangular model's empirical standard deviation. The increasing standard deviation for the triangular correlation model can be explained by the fact that any points beyond L/θ are independent of the points within.

Another interesting observation about the mean extreme values is that with the same parameters, different correlation models will cause very different mean values. Moreover, the ordering sequence among the correlation models changes with L/θ. For $L/\theta \leq 7$, the order of the mean extreme values is Exponential > Triangular > Quadratic Exponential. For $L/\theta > 7$, the order changes to Triangular > Exponential > Quadratic.

The asymptotic mean values derived from the Gumbel distribution demonstrate the same trends for the mean values for all three correlation models. As L/θ increases, the asymptotic mean diverges from the empirical mean for all cases, although the difference is the least in the quadratic exponential model. The reason for this surprising phenomenon has yet to be found. Note also that the asymptotic mean underestimates the empirical value of the triangular model whereas an opposite trend is observed in the other two correlation models.

According to the asymptotic EV theory discussed in Section 3.2, the asymptotic standard deviations of the EV are the same for all three correlation models. As shown in Figure *3.7*, the asymptotic result exhibits a better convergence to the corresponding empirical value for both the quadratic exponential and exponential models. However, the

decreasing asymptotic standard deviation is completely opposite to the increasing trend shown in the empirical curve.
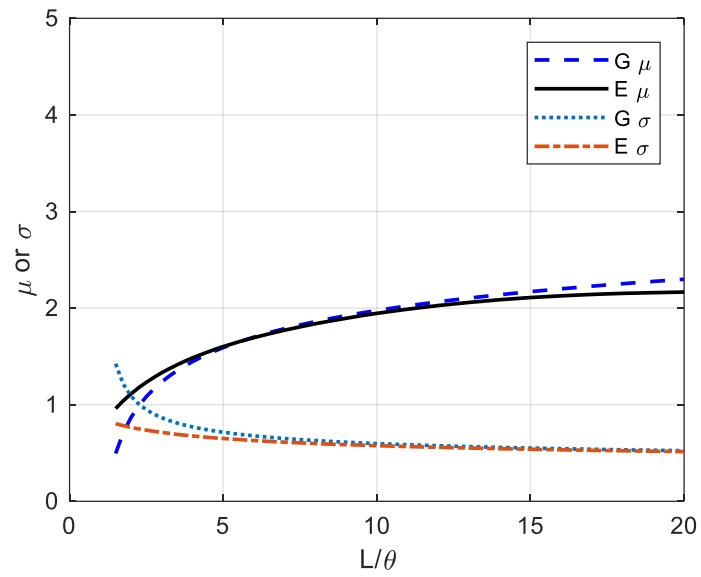
(a)
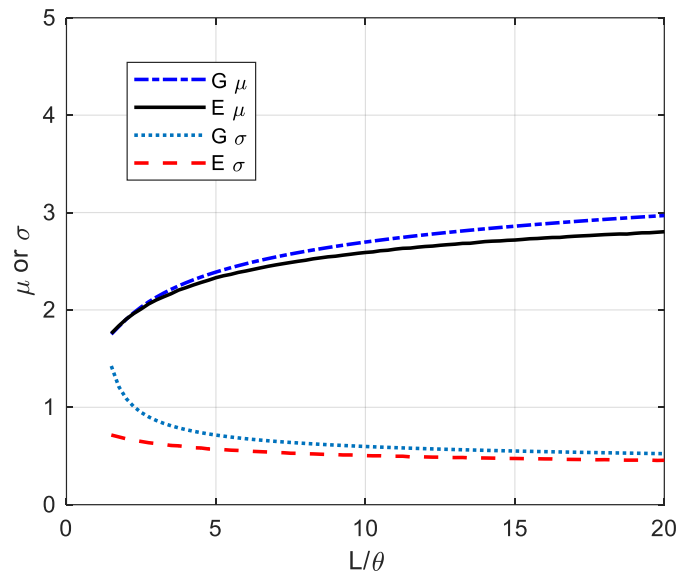
(b)
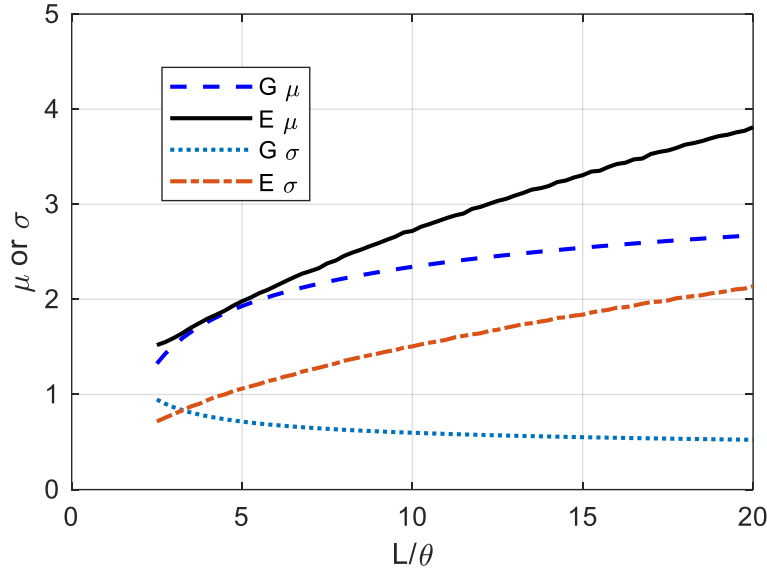
(c)

Figure 3.7 The statistics of the EV of a standard normal field with different correlation model: (a) quadratic exponential; (b) exponential; and (c) triangular.

3.6 Gamma Fields

Before the study, it is important to note that the scale parameter β of the gamma distribution has only a multiplicative effect on the extreme value of gamma fields. Therefore, a unit scale (i.e. $\beta = 1$) is used for all analyses of the subsequent discussions. Since the empirical EV distribution of the gamma field is simply the double transform of the EV distribution of the corresponding standard normal field, it is expected that major findings from the standard normal fields are observed in the gamma field. Therefore, the major interest of this part of study is to evaluate the impact of the shape parameter α on the statistics of the EV and the goodness of fit of the Ditlevsen, approximate Ditlevsen, asymptotic Gumbel, and Poisson approximation.

3.6.1 Distribution of the EV

Three cases of shape parameters ($\alpha = 0.5, 2, 10$) are studied for the quadratic exponential correlation model. In Figure 3.8, the simulation results show that both the transformed Ditlevsen, approximate Ditlevsen and modified Ditlevsen models generally produce good fits compared with the numerical simulation, although the exact and modified Ditlevsen distributions are slightly more accurate than the approximate version.
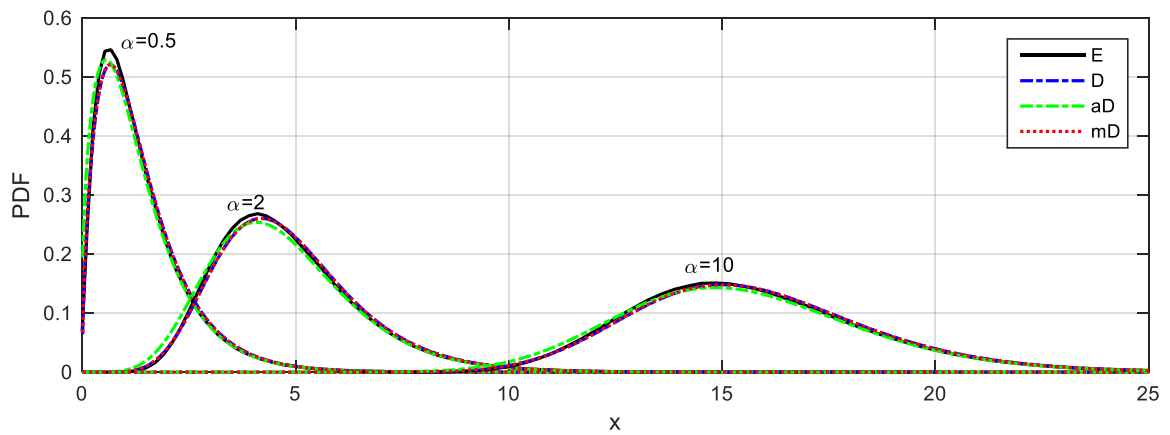
Figure 3.8 Comparison of the empirical, Ditlevsen, approximate Ditlevsen and modified Ditlevsen distributions for gamma field $L / \theta = 5, \rho(h) = exp\left(-3\left(h/\theta\right)^2\right)$
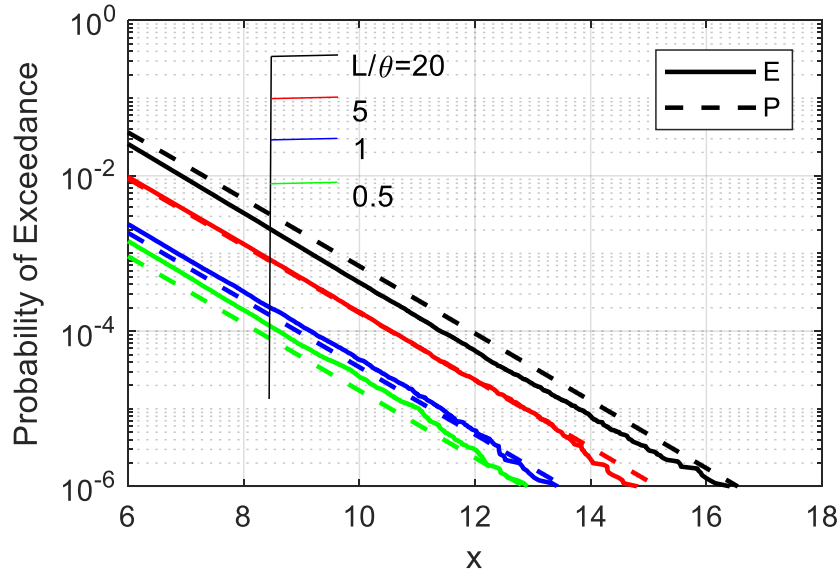
Figure 3.9 Performance of the Poisson approximation for gamma field with α=0.5, $\rho(h) = exp\left(-3\left(\frac{h}{\theta}\right)^2\right)$

Unlike in the standard normal field, the Poisson approximation does not perform well in the gamma field. As shown in Figure *3.9*, the Poisson approximation underestimates the probability of exceedance for $L/\theta \leq 1$, but overestimates the probability for high L/θ (e.g. $\frac{L}{\theta} \geq 20$).

Since the asymptotic Gumbel distribution has poor performance for standard normal fields, the model is also inaccurate for gamma fields in general. Numerous simulations have confirmed this postulate; for illustration purpose, only the case $\frac{L}{\theta} = 5 \,\&\, \alpha = 2$ is shown in Figure *3.10*.
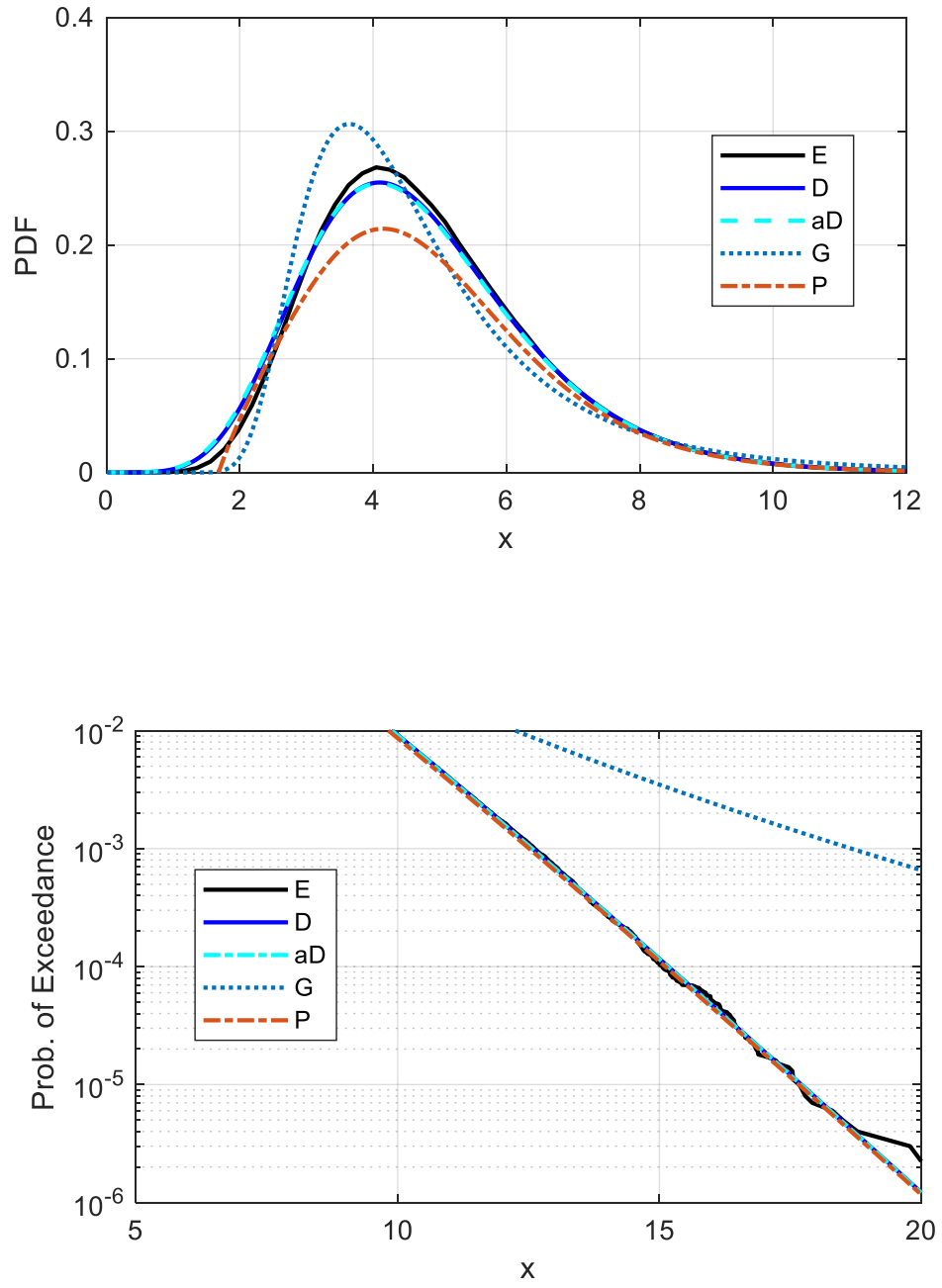
Figure 3.10 Comparison of the PDFs (top) and the probability of exceedance functions (bottom) of the empirical, Ditlevsen, approximate Ditlevsen, asymptotic Gumbel distributions and Poisson approximation for the Gamma field with $\frac{L}{\theta} = 5, \alpha = 2$ and $\rho(h) = exp\left(-3\left(\frac{h}{\theta}\right)^2\right)$.

3.6.2 Statistics of the EV

With a unit scale parameter, the mean and standard deviation of the extreme value are normalized by α and $\sqrt{\alpha}$, respectively. With respect to L/θ, the observation on extreme value distribution of the standard normal field is valid for gamma fields because the latter are obtained by monotonic transformation of the former. Figure *3.11* shows the trend of the normalized extreme value mean along L/θ for different shape parameters. Like the case in standard normal fields, for any given shape parameter α, as L/θ increases, the normalized mean of the extreme value increases. Although, for a given L/θ, the absolute value of the extreme value mean increases with α, the normalized mean decreases with α. Another explanation for Figure *3.11* is that a gamma distribution with a bigger shape parameter has a lighter right tail than one with a smaller shape parameter. This means there is a relatively higher probability of getting more extreme values from a gamma distribution with a small α (c.f. Section 2.1). In addition, increasing L/θ implies more randomness in the field so; the effective number of independent random variables in the field thereby increases.

The trend of the normalized standard deviation of the extreme value is slightly different from that of the normalized mean. As shown in Figure *3.11*, for a given α, the trend of the normalized standard deviation depends on the correlation model used for the gamma field. Moreover, there appears to be an anomalous trend for quadratic exponential correlation function at small shape parameters (Figure *3.11*(b)). The reason for this surprising trend is still unknown.

The ordering sequence of the mean value among the correlation models is the same as that of the standard normal fields. The quadratic exponential model has the least mean EV for the whole range of L/θ, whereas the exponential model has the greatest mean EV when $L/\theta \leq 7$, beyond which the quadratic model dominates.

(a)

(b)

(c)

(d)

78
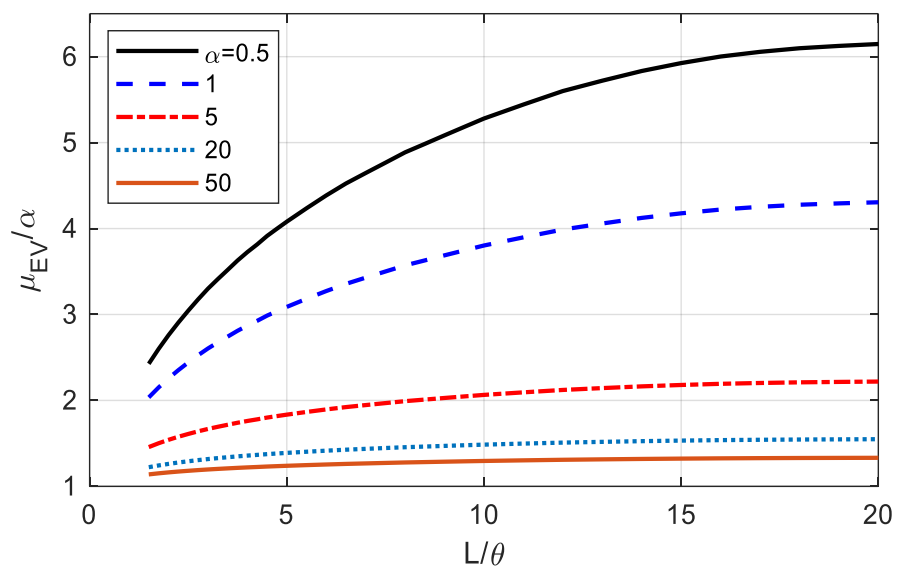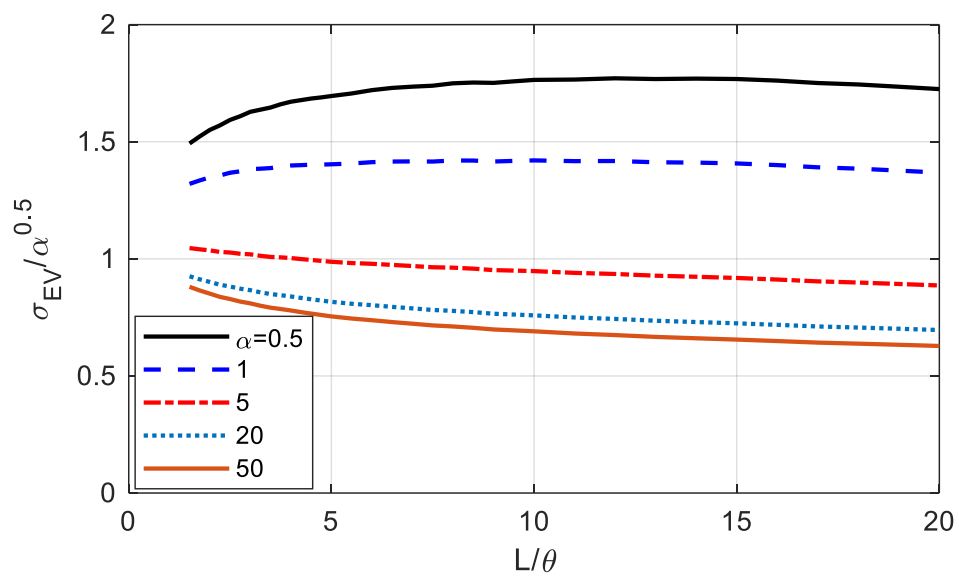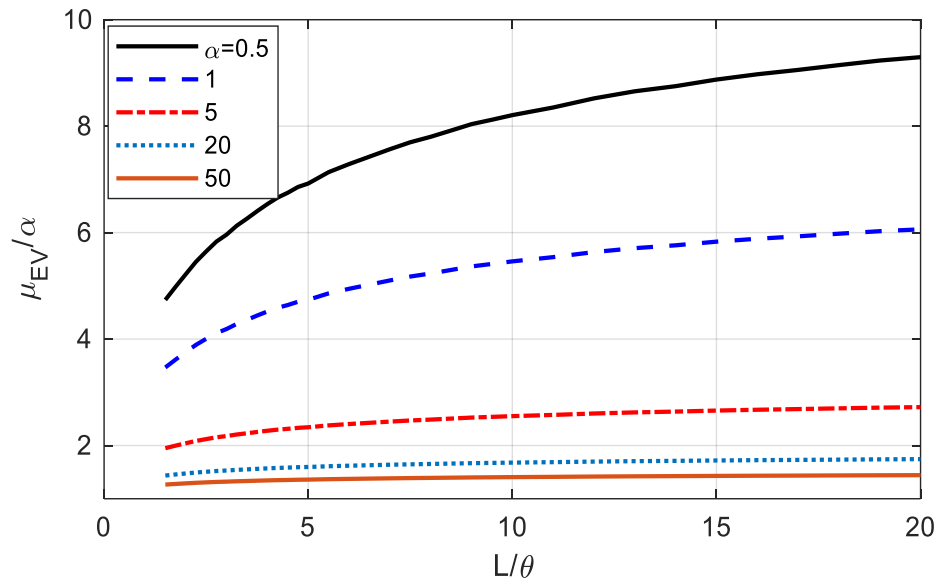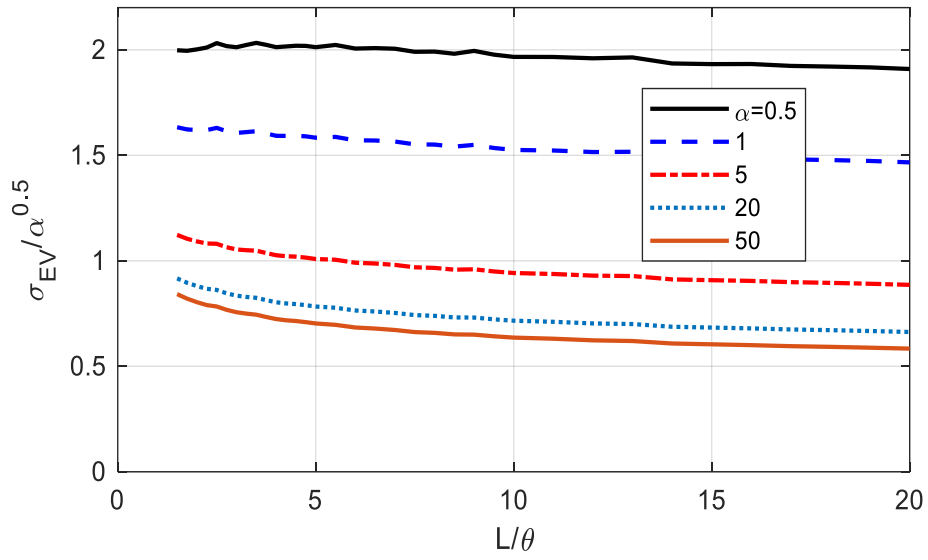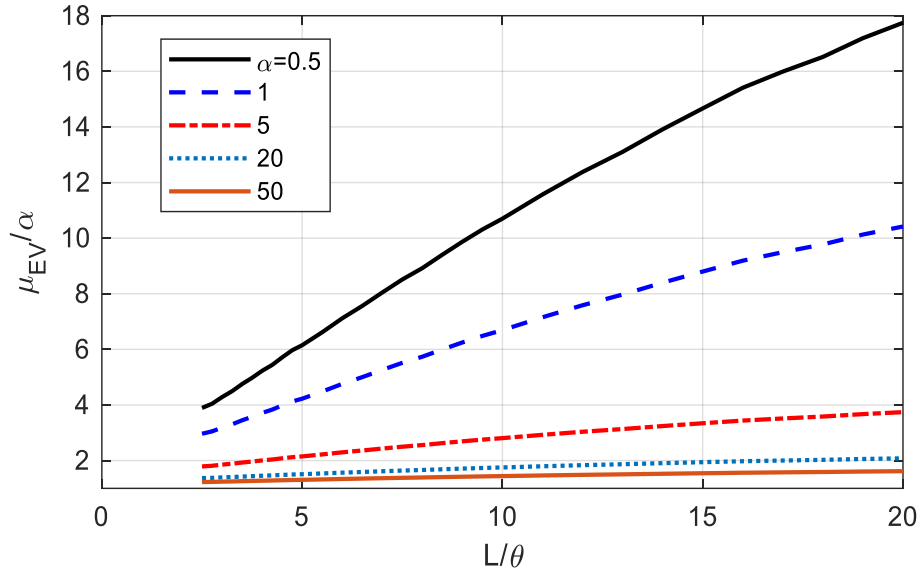
(e)
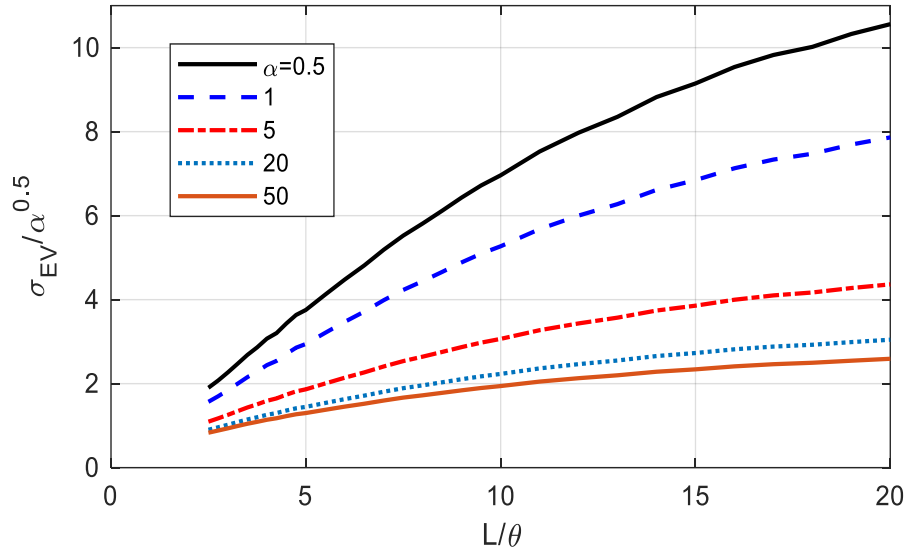
(f)

Figure 3.11 The EV statistics of gamma fields with three correlation models ((a)&(b)—squared exponential; (c)&(d)—exponential; (e)&(f)—triangular).

Alternatively, Figure *3.11* can be presented in a slightly different form by plotting normalized mean and standard deviation against shape parameter of the gamma field. This is shown in Figure *3.12*.
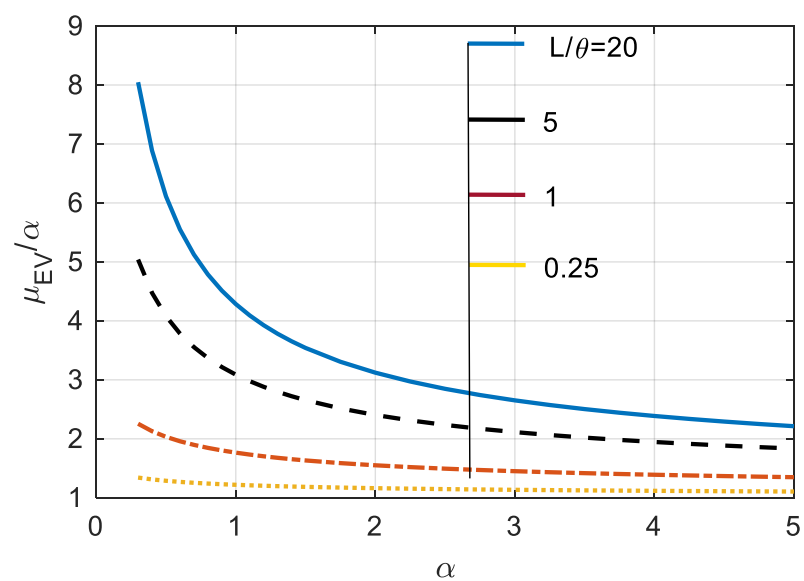
(a)
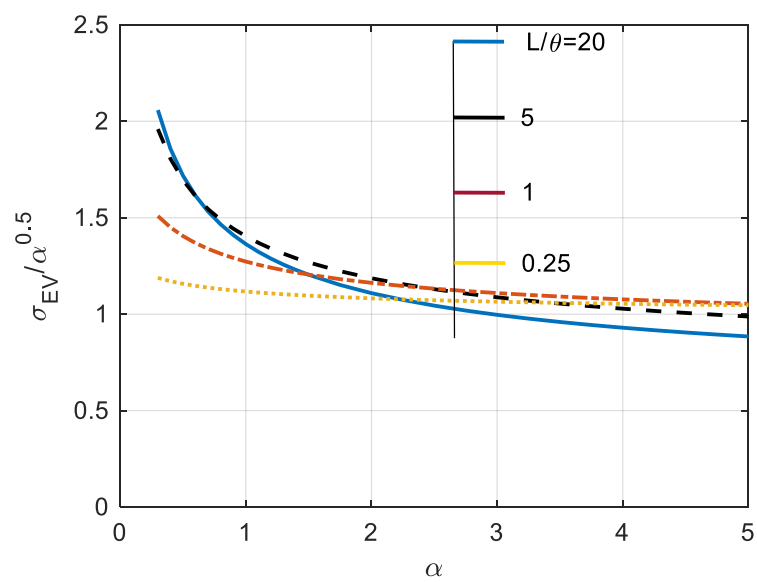
(b)
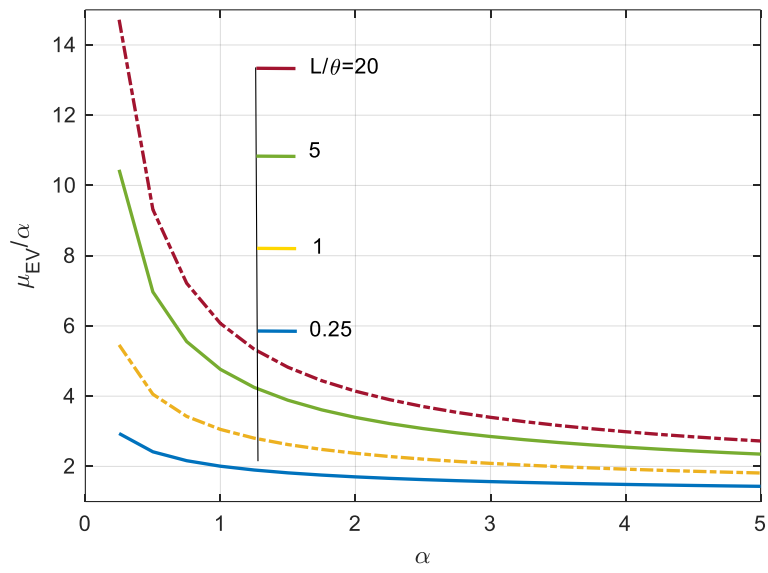
(c)
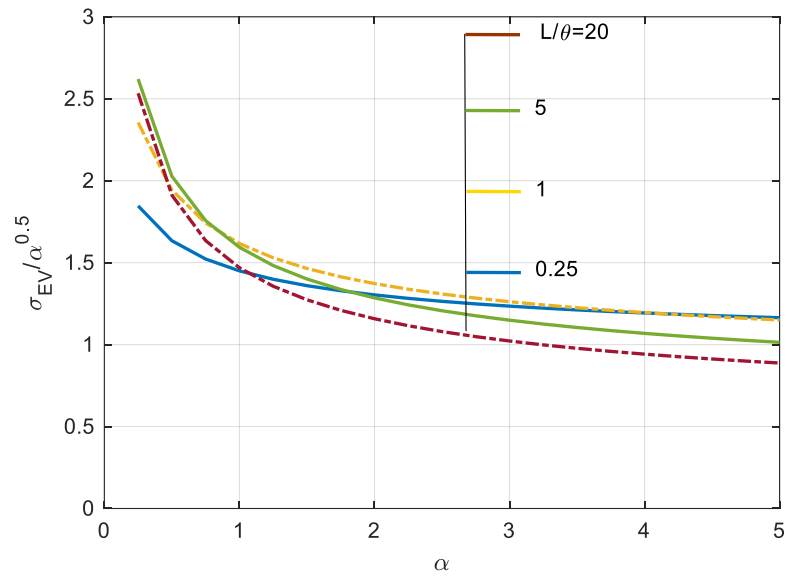
(d)
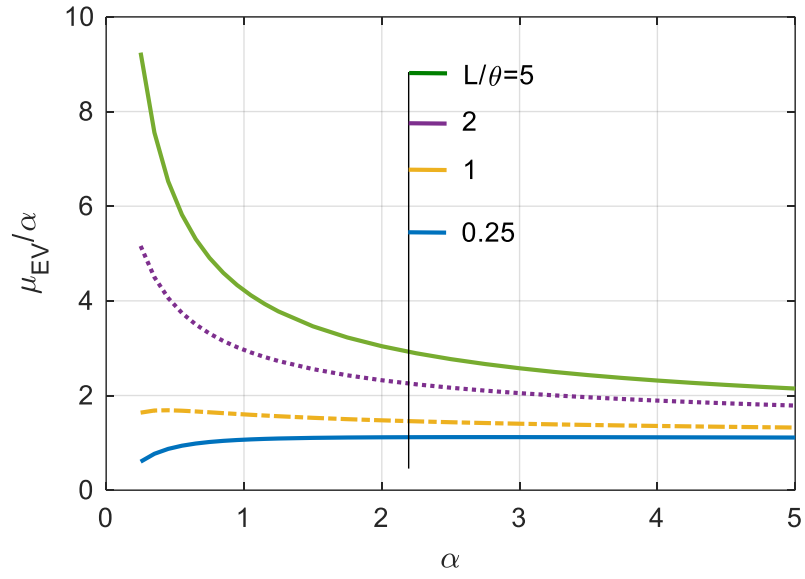
(e)
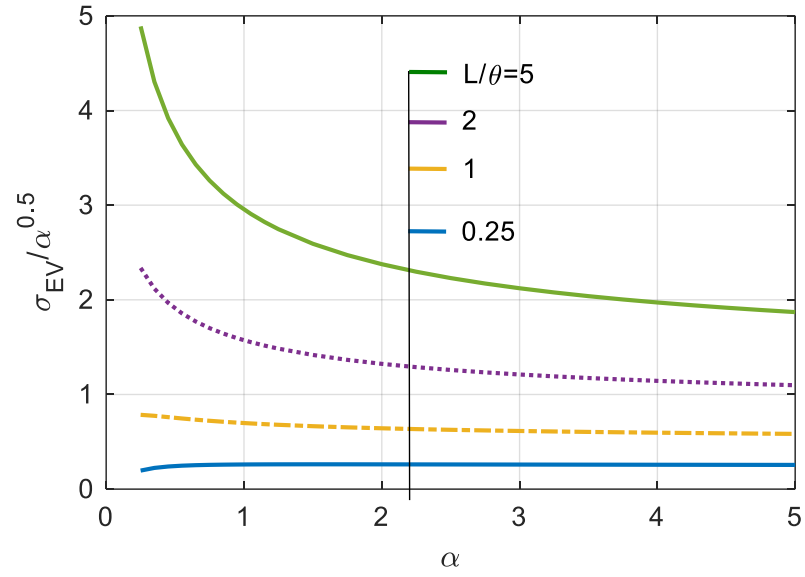
(f)

Figure 3.12. Normalized mean and standard deviation of the EV of gamma fields with different correlation functions ((a)&(b) Squared exponential (c)&(d) Exponential (e)&(f) Triangular).

3.7   Lognormal Fields

The lognormal random field has two parameters namely the mean μ and standard deviation σ of the associated Gaussian distribution. Similarly, the empirical EV distribution of the lognormal field is the double transform of the EV distribution of the corresponding standard normal field. As a result, major findings from the standard normal fields apply to the lognormal field. The effect of the mean μ and standard deviation σ on the statistics of the EV and the goodness of fit of the Ditlevsen, approximate Ditlevsen, modified Ditlevsen and Poisson approximation are presented in this section.

3.7.1   Distribution of the EV

Keeping $\frac{L}{\theta}$ and the mean μ constant, three cases of standard deviation ($\sigma = 0.3, 0.5, 0.7$) are presented for the quadratic exponential correlation model in Figure 3.13(a). In Figure 3.13(b), three cases of mean ($\mu = 0, 0.5, 2$) are presented while keeping the standard deviation constant. In Figure 3.13, the simulation results show that both the exact and modified Ditlevsen distributions are slightly more accurate than the approximate Ditlevsen.
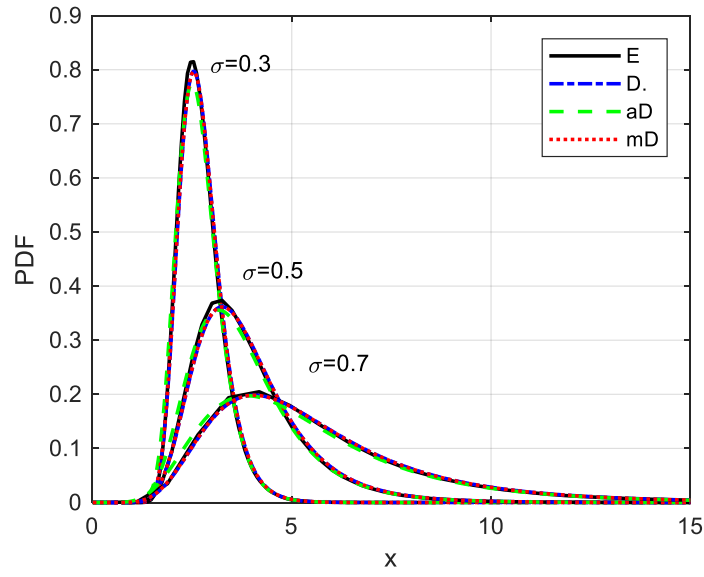
(a)
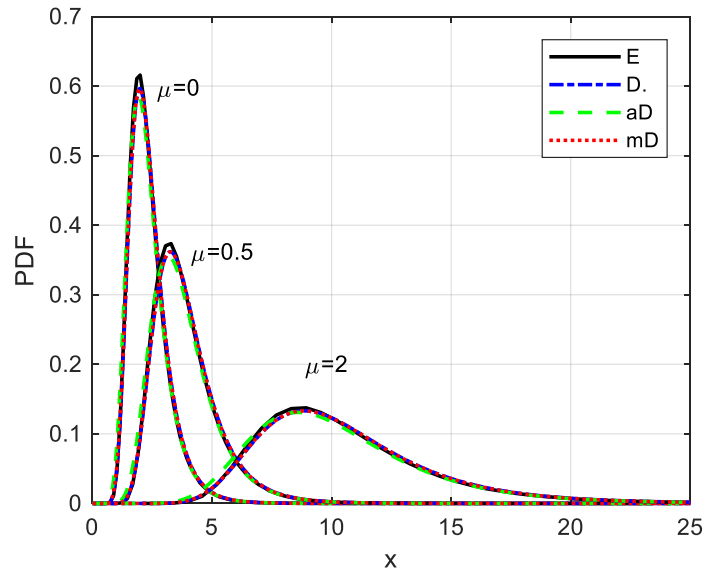
(b)

Figure 3.13. Comparison of the empirical, Ditlevsen, approximate Ditlevsen and modified Ditlevsen distributions for lognormal field ($L/\theta = 5$, $\rho(h) = e^{-3(h/\theta)^2}$). (a) $\mu = 0.5$ (top). (b) $\sigma = 0.5$ (bottom).
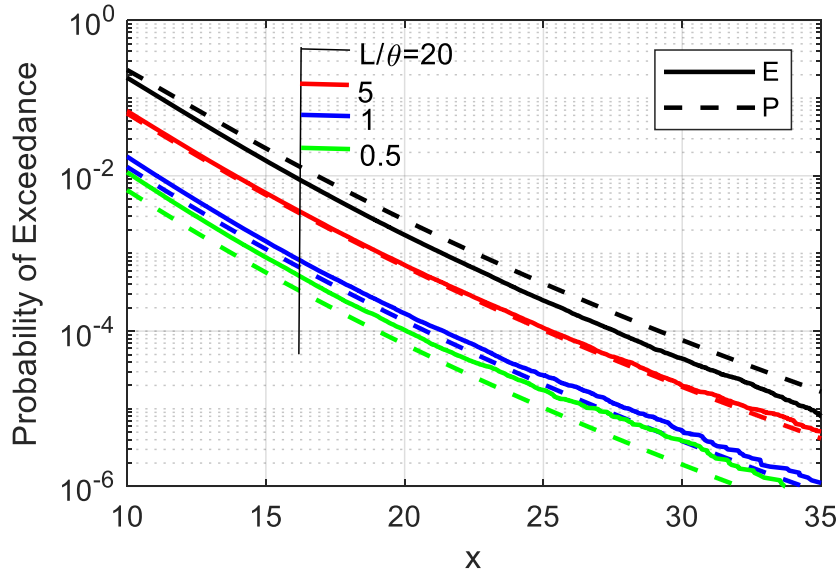
Figure 3.14. Performance of the Poisson approximation for lognormal field with $\mu = 1, \sigma = 0.5$, $\rho(h) = e^{-3\left(\frac{h}{\theta}\right)^2}$

As shown in Figure 3.14, the Poisson approximation underestimates the probability of exceedance for low L/θ, but overestimates the probability for high L/θ.

3.7.2 Statistics of the EV

The mean and standard deviation of the lognormal random variable are functions of the mean μ and standard deviation σ of the associated Gaussian distribution. For a lognormal random variable, its mean is $\mu_{LN} = \exp\left(\mu + \frac{\sigma^2}{2}\right)$ and variance given as $\sigma_{LN}^2 = [\exp(\sigma^2) - 1]\exp(2\mu + \sigma^2)$. Therefore, the extreme value statistics obtained from simulation are normalized by μ_{LN} and σ_{LN}. Figure 3.15 shows the trend of the normalized extreme value mean and standard deviation along L/θ for different standard deviations of the associated normal distribution. Like the case of gamma fields, for any given standard deviation σ, as L/θ increases, the normalized mean of the extreme value increases. The ordering sequence of the mean value among the correlation models remains unchanged. The quadratic exponential model has the least mean EV for the whole range of L/θ, followed by the exponential model and lastly the triangular model.

The trend of the normalized standard deviation of the extreme value is different from that of the normalized mean. As shown in Figure 3.15, for a given σ, the trend of the normalized standard deviation depends on the correlation model used for the lognormal field. Similar observations are made in the case of gamma fields.
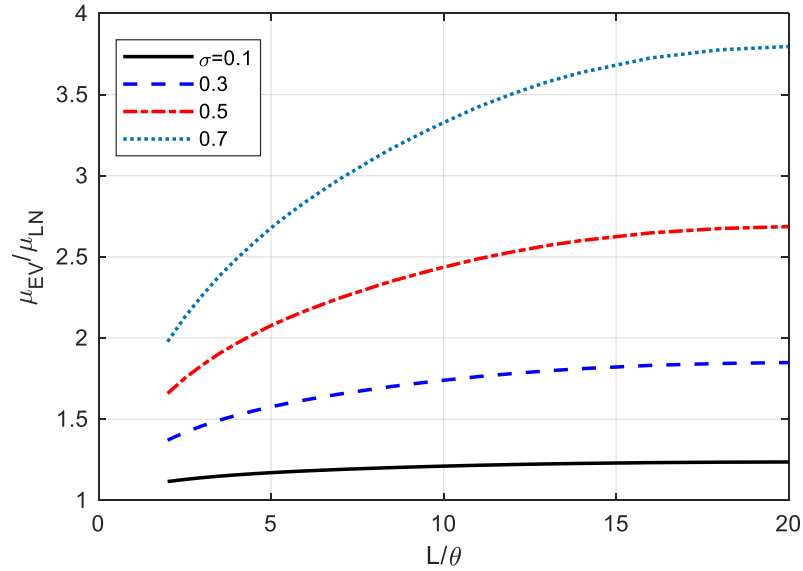
(a)
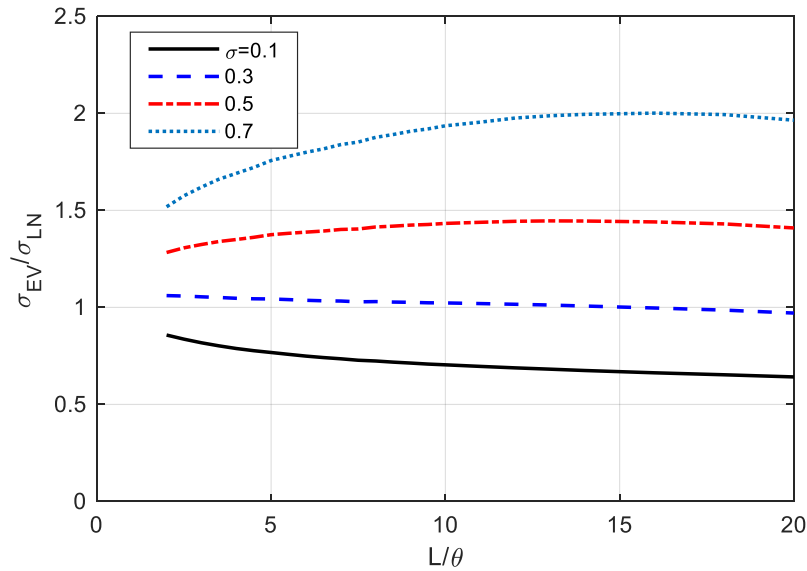
(b)
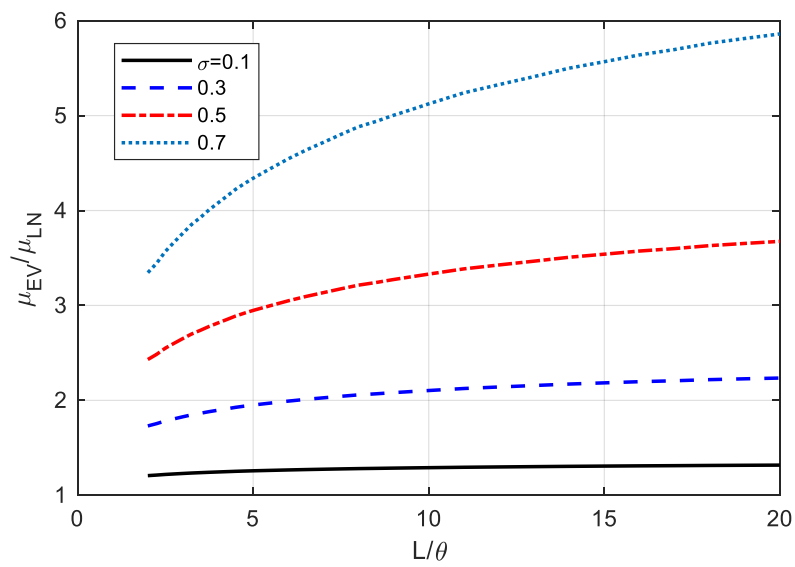
(c)
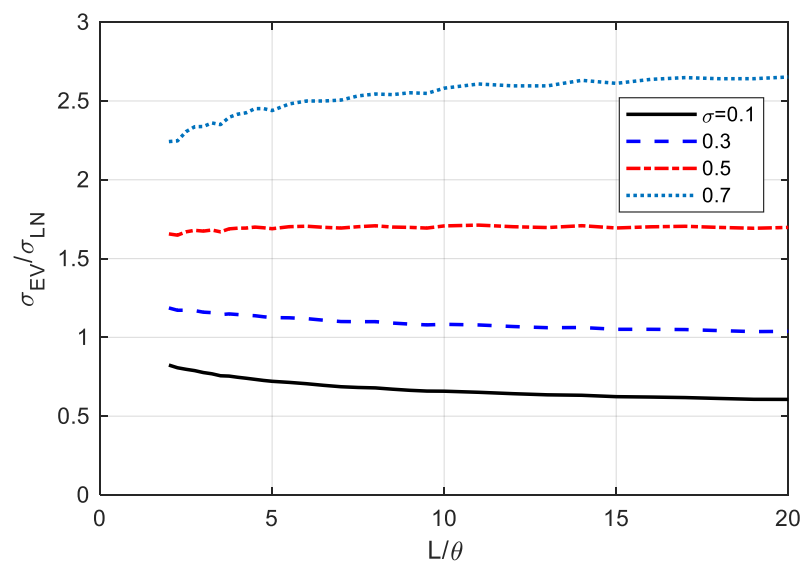
(d)
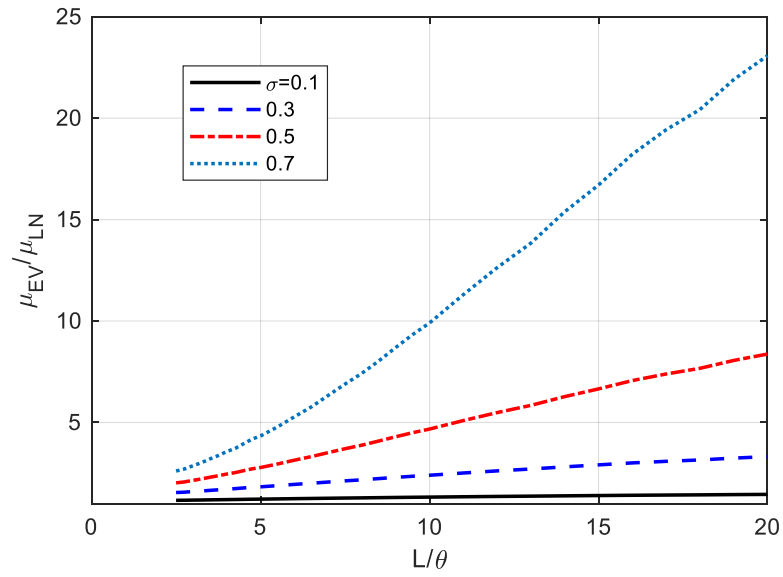
(e)
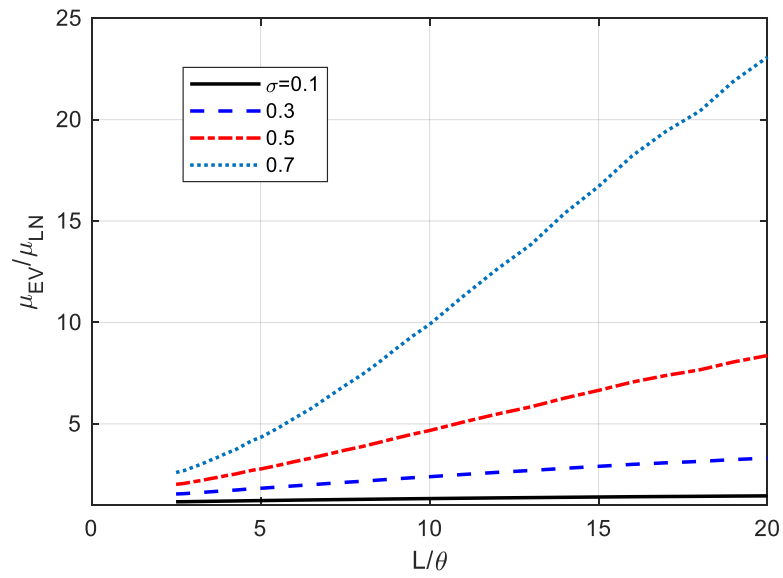
(f)

Figure 3.15. The EV statistics of lognormal fields μ = 0.5 with three correlation models ((a)&(b)—quadratic exponential; (c)&(d)—exponential; (e)&(f)—triangular).

This study is confined to be one-dimensional gamma fields. The EV of two- and three-dimensional gamma fields, and other practical questions related to the proposed gamma field (e.g., parameter estimation, impact on inspection sampling) will be discussed in subsequent sections.

3.8   Two-dimensional Random Fields

Using the same KL method, we present the simulation procedure and results of statistics of the EV of two-dimensional random fields. The two-dimensional Gaussian and gamma fields are simulated as follows.

1. Choose a pair of L_x/θ_x and L_x/θ_y, where L_x and L_y are the dimensions of the random field.
2. Use the Nystrom method described in section 3.4 to determine the eigenvalues and eigenvectors in both horizontal and vertical directions.
3. Plug these eigenvalues and eigenvectors into the two-dimensional K-L expansion (eqn. 3.15) to generate a realization of a 2-D Gaussian field

$$\xi(\boldsymbol{s}, \boldsymbol{t}) = \sum_{j=1}^{n} \sum_{i=1}^{n} \sqrt{\lambda_i \lambda_j} \zeta_{ij}(\omega) f_i(\boldsymbol{s}) f_j(\boldsymbol{t}) \qquad 3.15$$

where the number of terms $n = 43$.

3. Repeat step 3 until enough realizations of the 2-D Gaussian field are generated.
4. Obtain the maximum value of each Gaussian field realization.
5. Calculate the mean and standard deviation of the field.
6. Repeat steps 1 to 5 for another pair of L_x/θ_x and L_x/θ_y. The range of both ratios is chosen as 2 to 20.

To obtain the two-dimensional gamma fields in Figure 3.18, the maximum values got in step 4 are translated to gamma fields using a scale parameter $\beta = 1$ and four shape parameters $\alpha = 0.5, 1, 5$ and 20.

As noted in the previous sections on one-dimensional standard normal, gamma and lognormal random fields, the statistics of the EV of two-dimensional random fields depends upon L/θ_x and L/θ_y when the fields are separable. *Figure 3.16* show the trend of the mean of extreme values along L/θ_x and L/θ_y for homogeneous standard normal fields with a quadratic exponential correlation function. The mean extreme values increase as L/θ_x and L/θ_y increase, as expected. In Figure 3.17, the normalized standard deviation decreases as L/θ_x and L/θ_y increase.

For two-dimensional gamma fields with a quadratic exponential correlation function, the mean of the EV is normalized by α. With respect to L/θ_x and L/θ_y, Figure 3.18 shows the trend of the normalized extreme value mean for different shape parameters. Similar to one-dimensional gamma fields, the normalized mean of the extreme value increases as L/θ_x and L/θ_y increase, for any given shape parameter α. As shown in Figure 3.19, for a given α, the trend of the normalized standard deviation decreases as L/θ_x and L/θ_y increase.
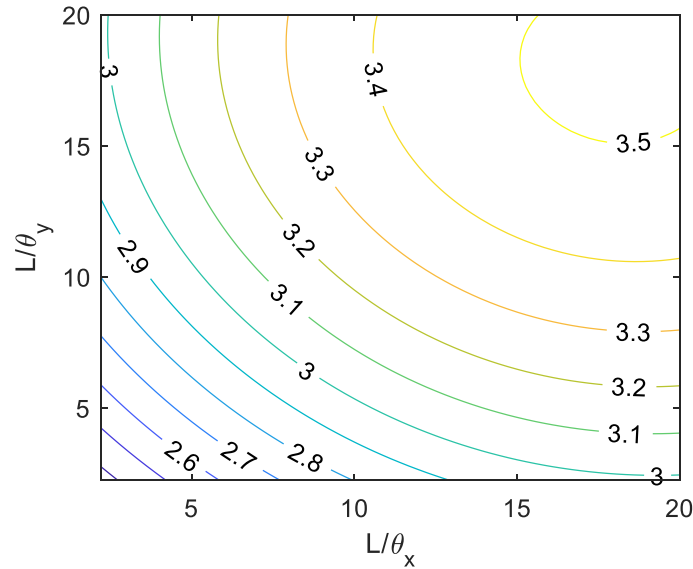
Figure 3.16. The EV mean of standard normal fields with a quadratic exponential correlation model.
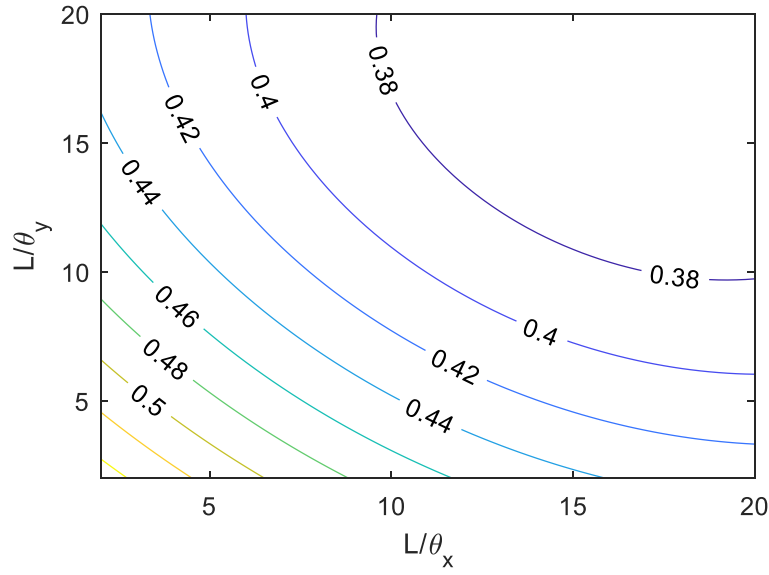
Figure 3.17. The EV standard deviation of standard normal fields with a quadratic exponential correlation model.
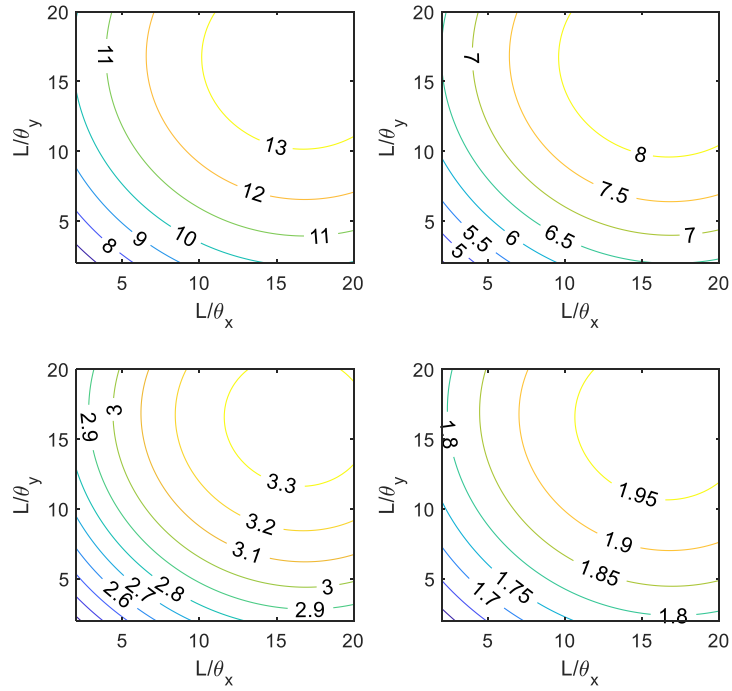
Figure 3.18. The EV mean of homogeneous gamma fields with a quadratic exponential correlation model (Top left: $\alpha = 0.5$, Top right: $\alpha = 1$, Bottom left: $\alpha = 5$, Bottom right: $\alpha = 20$).
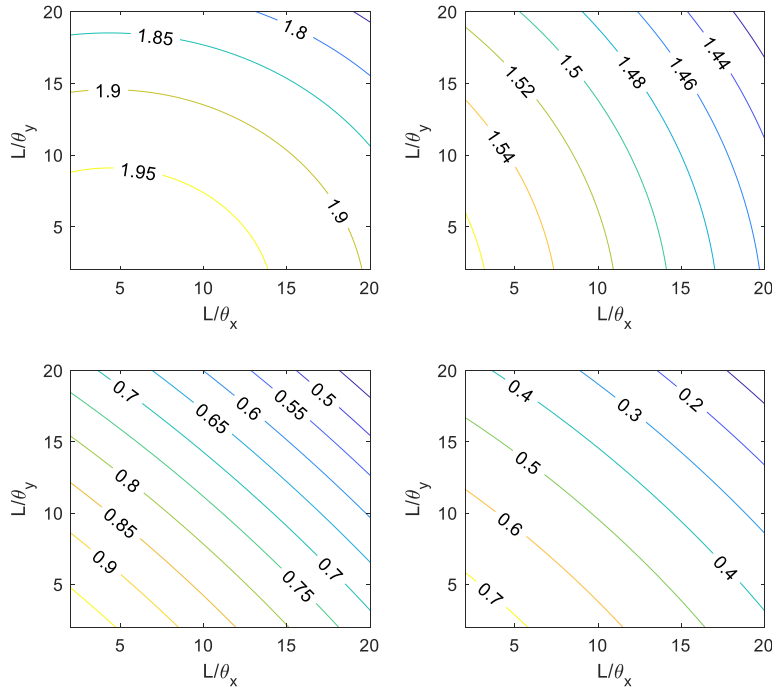
Figure 3.19. The EV standard deviation of homogeneous gamma fields with a quadratic exponential correlation model (Top left: $\alpha = 0.5$, Top right: $\alpha = 1$, Bottom left: $\alpha = 5$, Bottom right: $\alpha = 20$).

3.9   Conclusions

Based on the study on extreme values of one-dimensional random fields, a number of conclusions are drawn:

- Preliminary studies have indicated that simple random simulation based on fine mesh will not work due to convergence problem. Hence spectrally based simulation techniques must be employed. The proposed KL expansion-based simulation method incorporating the Nystrom quadrature technique can be effectively used to study the EV of homogeneous random fields.

- The exact Ditlevsen distribution, with the G term calibrated with simulation data, fully characterizes the EV of the standard normal field, and with a further double transform, characterizes the EV of the gamma and lognormal fields as well. With an empirically calibrated G term, as shown in Eq. (3.14), the modified Ditlevsen distribution can predict more accurately than the approximate Ditlevsen with zero G for both types of fields.

- In the upper tail of the distribution, the Poisson approximation also performs surprisingly well for standard normal fields. However, the performance deteriorates for gamma and lognormal fields.

- The asymptotic Gumbel distribution is generally not recommended for either standard normal, lognormal or gamma fields.

- The statistics of the EV were found to be a function of size-to-correlation length ratio for standard normal fields, and of the shape parameter for gamma fields and of the standard deviation of the underlying Gaussian field for lognormal fields as well.

- Among the three correlation models studied, the quadratic exponential model provides the least mean EV for both standard normal and gamma fields. For large fields ($L/\theta > 7$), the triangular model includes the greatest mean EV, and the exponential model dominates otherwise. For lognormal random fields, the exponential model has the least mean EV for the whole range of L/θ, followed by the exponential model and lastly the triangular model.

- For separable two-dimensional random fields, the statistics of the EV were found to be a function of size-to-correlation length ratios along horizontal and vertical directions for standard normal fields, and of the shape parameter for gamma fields.

4. Case Study of a Nuclear Feeder Pipe

As discussed in Chapter 1, nuclear feeder pipes experience wall thinning due to flow-accelerated corrosion (FAC) over their service life. The FAC rate of a feeder varies from location to location and over time. Models that capture temporal uncertainty of this degradation mechanism have been proposed in the literature. However, such models do not take spatial uncertainty into consideration. Therefore, a more sophisticated probabilistic model that considers these temporal and spatial uncertainties is required for efficient management of the feeders. We propose a more advanced gamma random field model to capture the spatial uncertainty with a view to answering the following questions:

1. How are missing ultrasonic scan data estimated? How are different neighbouring scan patches of wall thicknesses aligned?
2. How are the parameters of the gamma field estimated? How can the field be simulated, and its EV distribution obtained?

From a practical point of view, it is necessary to consider uncertainties in any proposed model. The gamma random field model possesses aleatory uncertainty. In addition to the aleatory uncertainty, there is epistemic uncertainty because limited amount of scan data set is used for parameter estimation. The chapter is organized as follows. We start by verifying an assumption regarding the underlying distribution of scan data. Parameter estimation of one-dimensional gamma fields and evaluation of their extreme value distribution based on the modified Ditlevsen distribution are then presented. This is followed by the method of "repairing and stitching" to address the issue of missing and unobserved scan data. Basics of the two-dimensional gamma field are presented in terms of its simulation and extreme value distribution. The parameter uncertainty is considered by applying the asymptotic normality property of the maximum likelihood estimator. Finally, the effect of missing data on extreme value distribution is investigated.

4.1   Validation of Assumptions

There are 2 common types of ultrasonic scanning probes used in scanning CANDU feeder pipes. These typically consists of 6 or 14 probes. In this study, only the data from 14-probes are used. The 14-probe technology allows scanning along the axial direction. To cover the full circumference, a feeder pipe is typically scanned four runs, each covering one of the four zones namely the extrados, right cheek, intrados and left cheek of the pipe. The scan data were drawn from wall thickness measurements of a feeder pipe from an anonymous power generation company (Figure 4.1). The pipe has a nominal thickness of 5.54 mm and an outside diameter of 60 mm, which corresponds to a circumference of 188.5 mm.

In Figure 4.1, it is obvious that the wall thickness at any location must be a non-negative number. For the reason, the amount of deterioration or wall thickness loss at any location equally takes a non-negative value. The wall thickness loss is defined as the difference between the average initial wall thickness and current wall thickness. Even though the initial wall thickness is a random variable, a fixed value of 6mm is assumed because the feeder was not scanned before it was put into operation.

Before we process the scan data, it is reasonable to try out a number of non-negative long-tailed random variables as a possible probability distribution of the data. Preliminary plots of the extrados and left cheek data suggest bimodal distributions. This can be explained by the fact both scan patches contain significant regions of local thinning as well as regions with less corrosion. As the focus is on the distribution of maximum values, the upper part of the mixed distribution is used for distribution fitting of both the extrados and left cheek scan data. Eight candidate probability distributions are considered by fitting each distribution to the histogram of raw deterioration data in Figure 4.2. From Figure 4.2, gamma, lognormal and log-logistic distributions provide decent fits to the data, with long right tails. Hence the empirical distribution of the deterioration data is assumed to follow a gamma distribution.
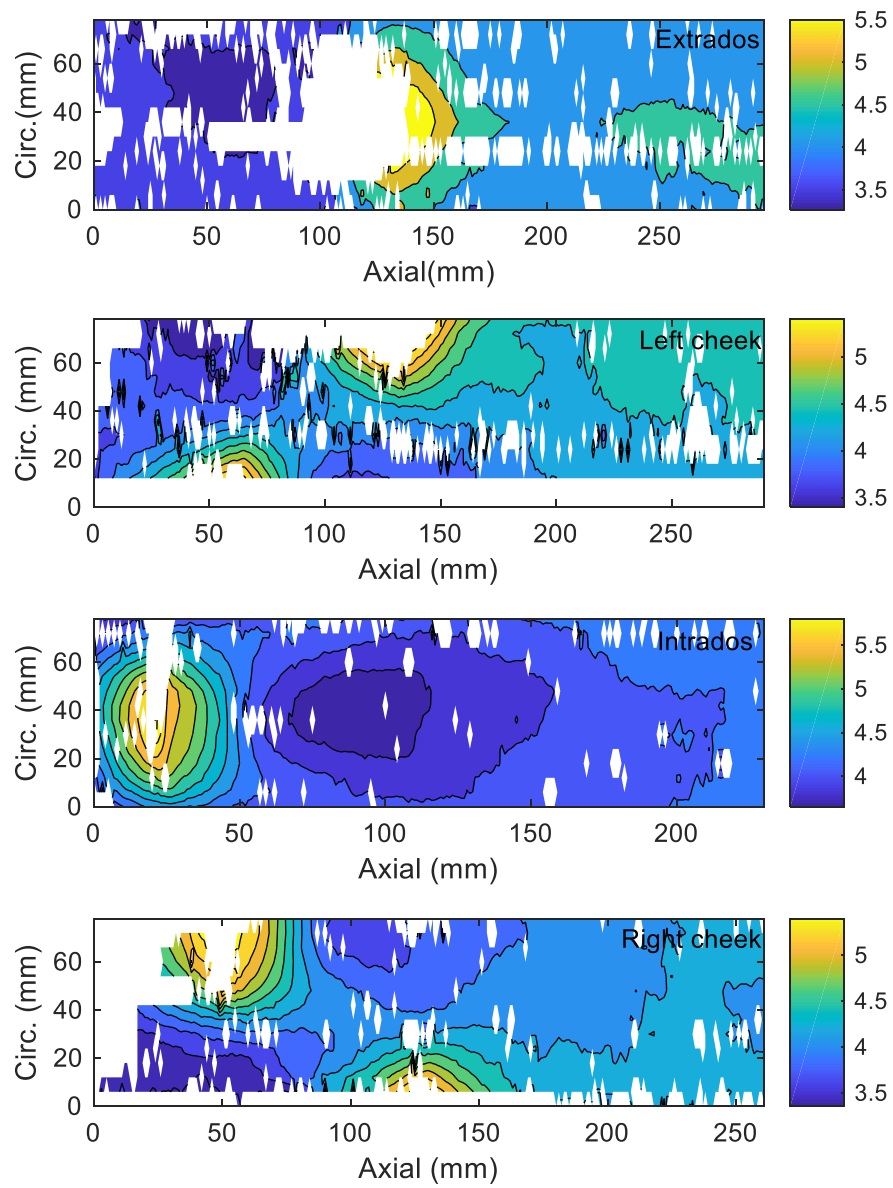
Figure 4.1 Ultrasonic scan plots showing missing data.
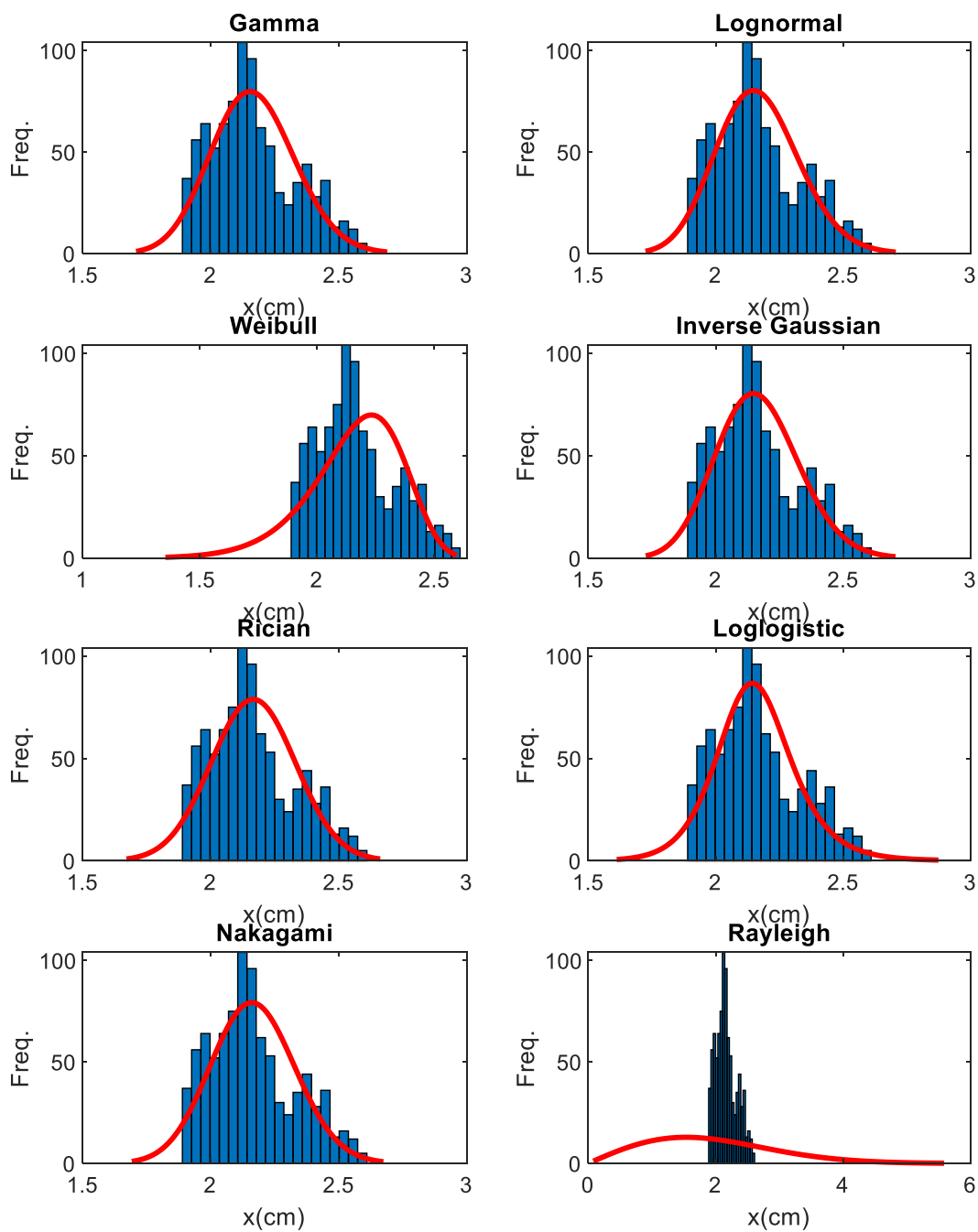
(a)
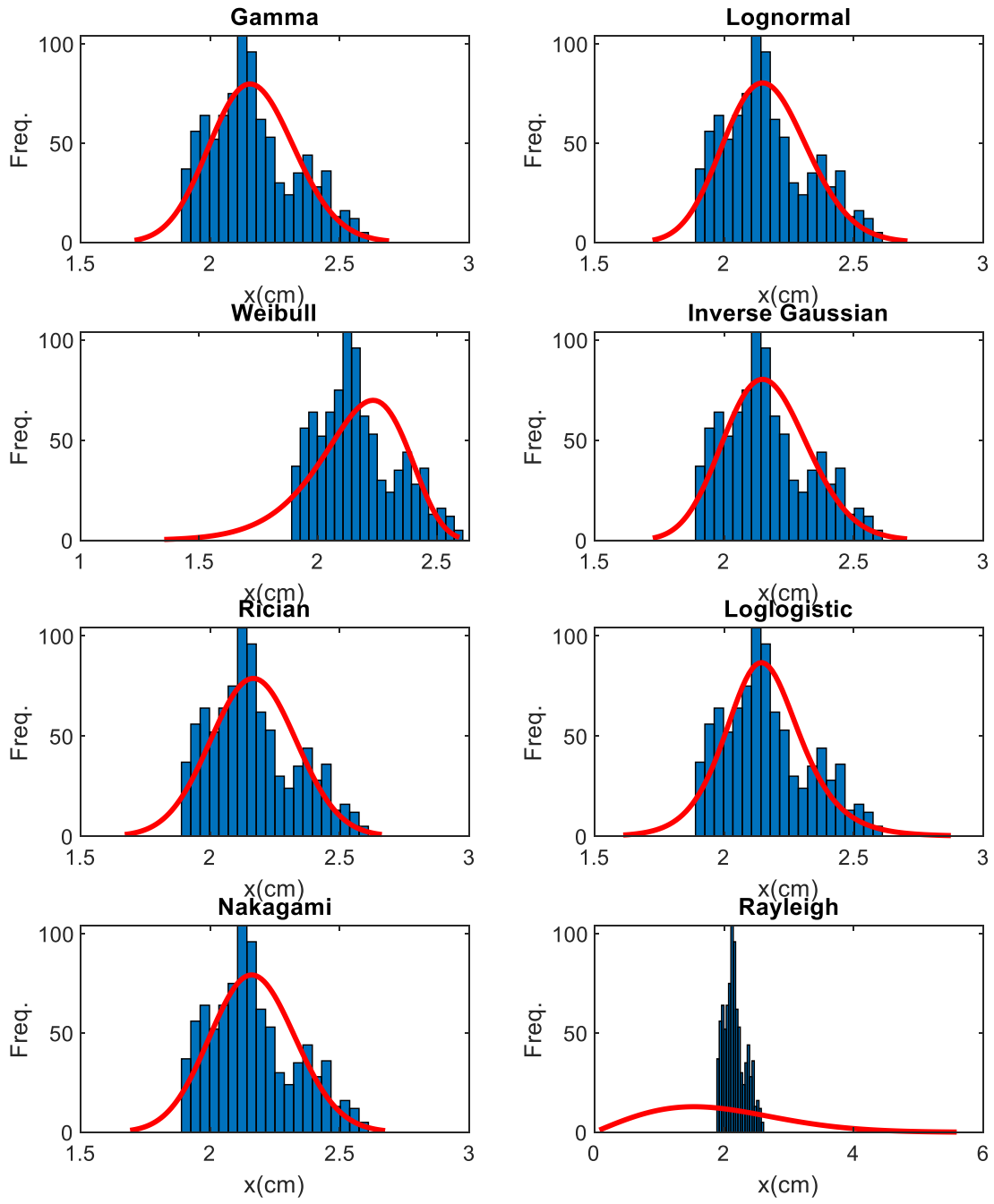
(b)
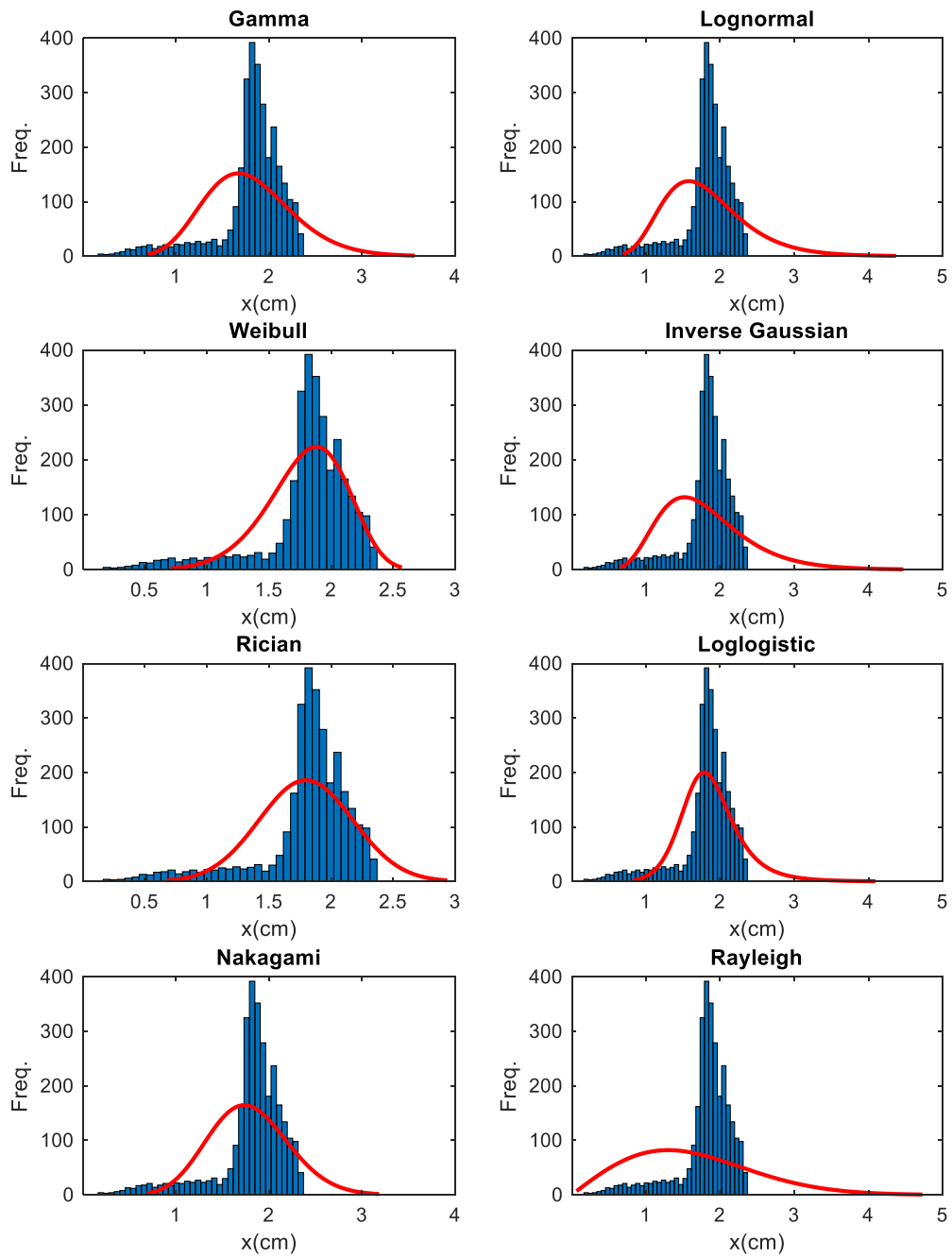
(c)
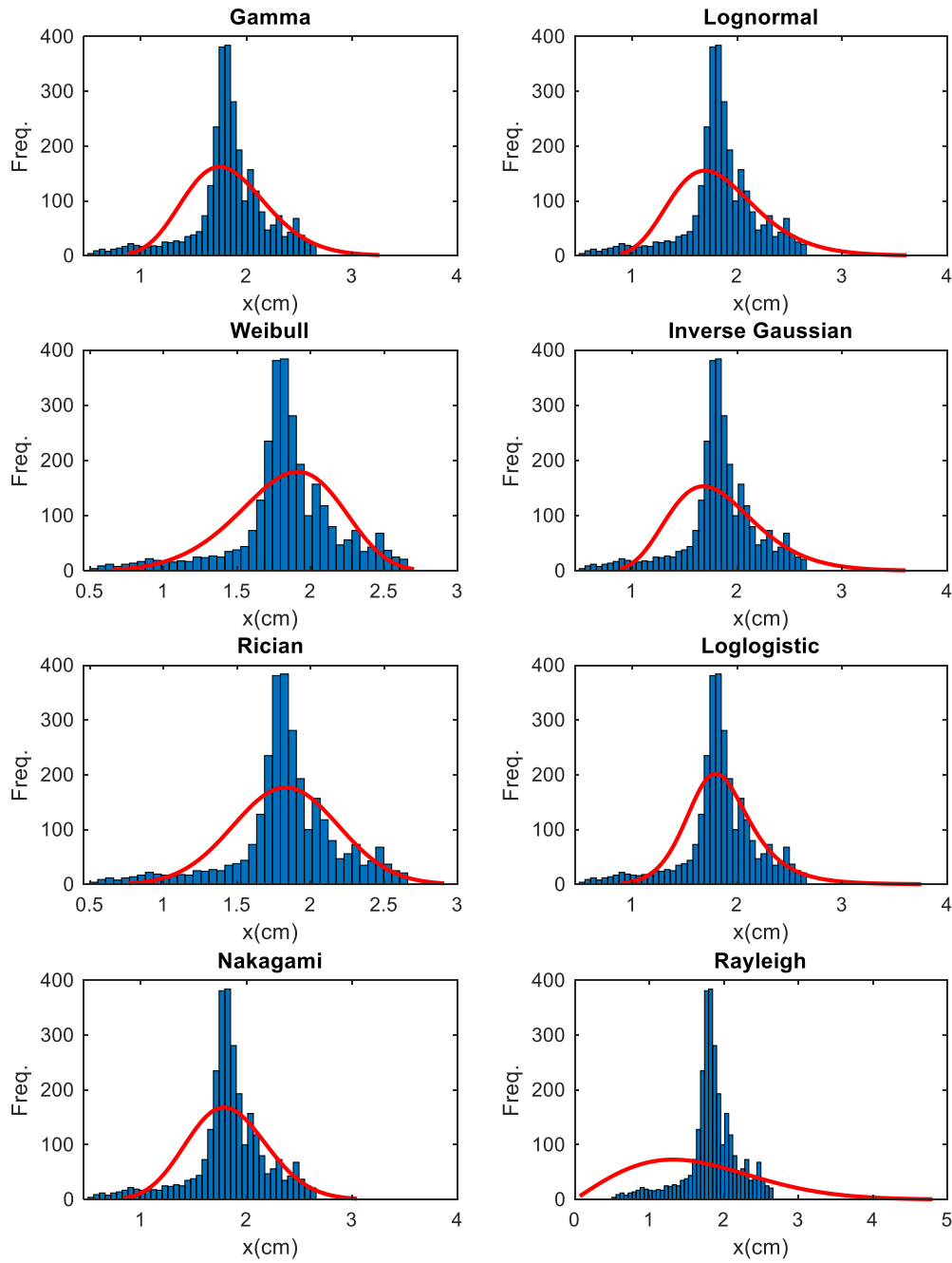
(d)

Figure 4.2 Histograms of deterioration data (wall thickness loss) with distribution fits (a-extrados; b-left cheek; c-intrados; d-right cheek.

In this chapter, the Gaussian copula is assumed to capture the spatial variability in the deterioration data. The choice of the Gaussian copula is based on the ease of construction of its PDF. However, the best practice is to fit a number of copula functions to the deterioration data and then use criteria such as Akaike information criterion (AIC) and Bayesian information criterion (BIC), which are both based on the likelihood function, to decide the best copula model.

4.2 One-dimensional Gamma Fields

The raw data for the extrados scan (see Figure 4.1) is used to demonstrate the applicability of the modified Ditlevsen distribution because the scan contains the minimum wall thickness. Measurements obtained from each probe is assumed to be a one-dimensional gamma field. Essentially, there are 14 fields. Starting from 0 in the circumferential direction, the probes are numbered from 1 through 14. One of the terms in the likelihood function is the correlation matrix. For each random field, the maximum number of data points is 297. If all these data points were used, the size of the correlation matrix would be 297 by 297. To reduce computational time and round-off error during iterations, each field is divided into a block with 6 data points. The maximum deterioration in each block is chosen as the representative of the block. In case there is no data in a block, the block is disregarded.

The pipe is assumed to have an average initial wall thickness of 6mm. Therefore, the amounts of degradation at every location is calculated by subtracting the wall thickness obtained from the combined random field from the initial average wall thickness of the feeder, w_o taken as 6mm.

The quadratic exponential correlation model is used to describe the correlation between any two points for each of the 14 gamma fields. Using the likelihood function of equation 2.11,

$$L(\alpha, \beta, \theta \,|x_i, \dots, x_n) \;=\; \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{R}^{-1}\mathbf{z}\right) \prod_{i=1}^{n} \frac{g(x_i;\alpha_i,\beta_i)}{\phi(z_i)} \qquad 4.1$$

the estimated parameters of the fields are shown in

Table 4.1.

Table 4.1 Estimated parameters of one-dimension gamma fields from extrados scan.

Probe No.	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
1	51.8	0.0361	16.92
2	40.24	0.0445	18.93
3	25.91	0.0672	17.99
4	15.13	0.1156	18.08
5	6.954	0.2489	22.96
6	5.304	0.2758	23.12
7	5.762	0.2938	22.02
8	2.807	0.568	24.61
9	9.926	0.184	18.16
10	7.066	0.2596	18.74
11	20.66	0.0926	19.87
12	33.58	0.0593	20.69
13	44.85	0.0445	17.69
14	75.69	0.0264	19.61

Using the estimates in Table 4.1, the extreme value distributions can be obtained by first calculating a G term based on Eq. (3.14) and then plugging the G into Eq. (2.18) for each one-dimensional field. The results of this procedure are presented in Figure 4.3 together with the observed maximum thickness loss of each probe line (indicated on the horizontal axis).

(a)

(b)

103
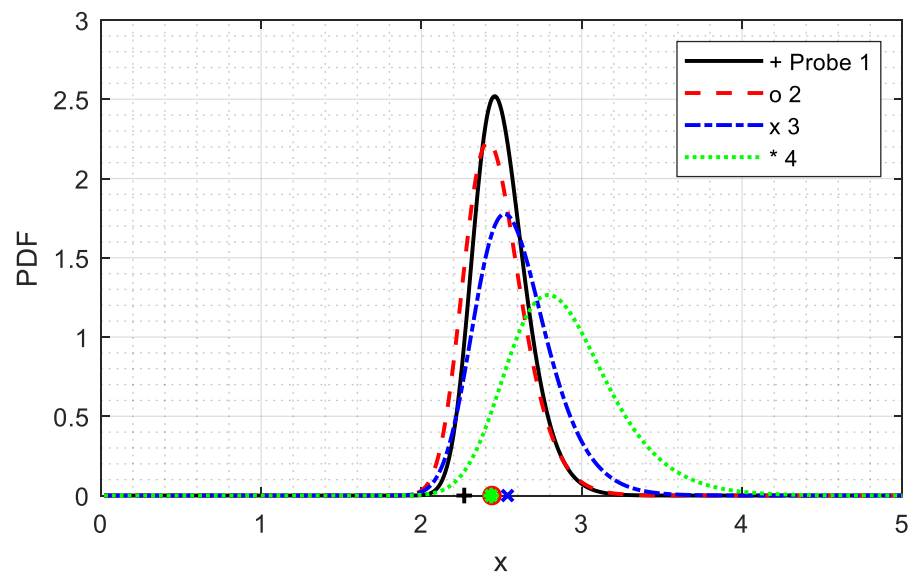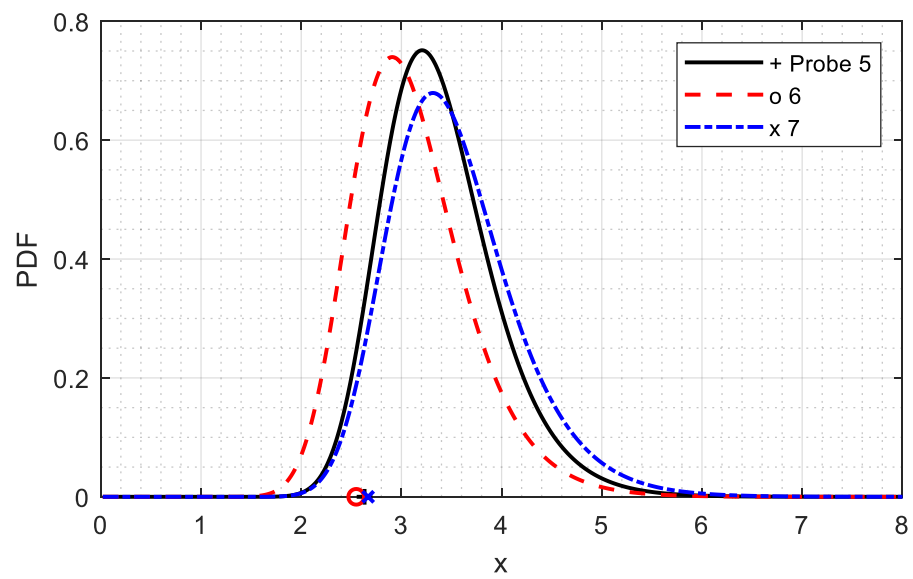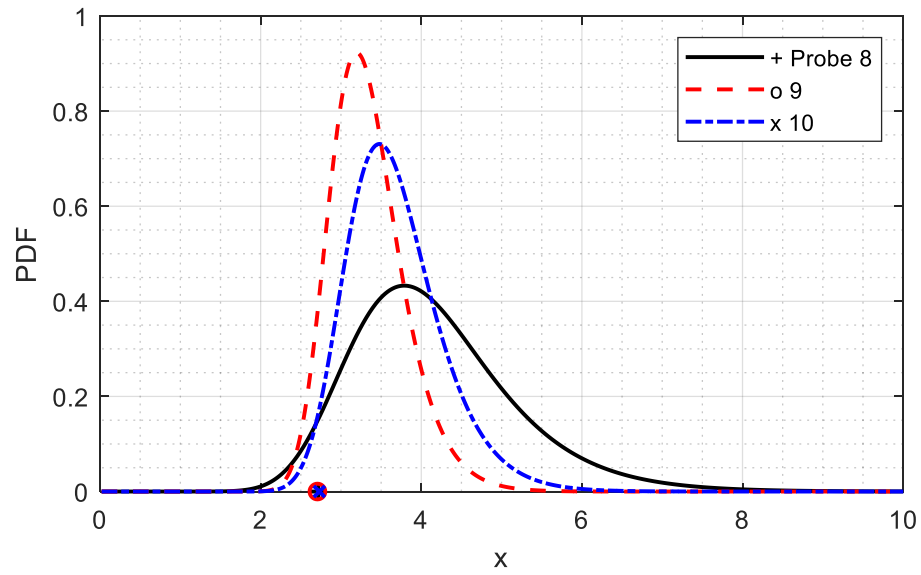
(c)
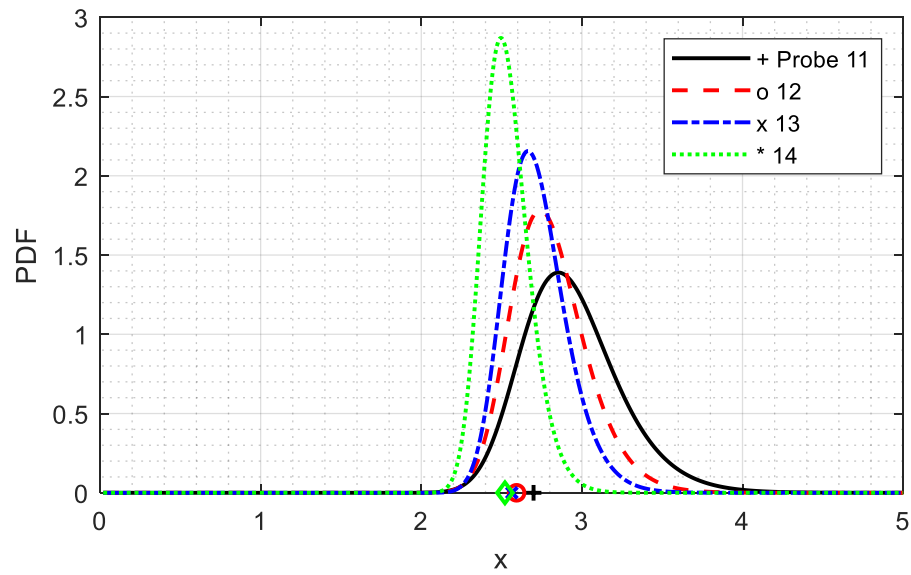
(d)

Figure 4.3 Extreme value distributions of 1-D gamma fields.

Figure 4.3(a) shows that the variability of the EV distributions increases from probe 1 through probe 4 (about the first 20mm in the circumferential direction). This observation

can be attributed to the amounts of data captured by the probes. Similar trend is observed in the EV distributions of probe 11 through 14 i.e. about the last 20mm in the circumferential direction (Figure 4.3(d)). In Figure 4.3(b)&(c), the EV distributions of intermediate probes i.e. probes 5 through 10, have the greatest variability because that region has much more significant missing data in terms of amount and importance. By importance means the region contains the minimum observed wall thickness. This also explains why their EV distributions have greater means, if calculated.

4.3   Repairing and Stitching

Suppose we are interested in using all the four scan patches together for EV analysis, two major challenges arise in order to construct a full wall thickness profile from the scan data. First, interpreting missing data present in each scan and second, aligning the different ultrasonic scans. Missing data includes wall thickness at the unscanned locations in addition to data at the loose contact positions. Figure 4.1 shows graphically the extent of missing data due to loss of contact between the probe and the pipe surface. Quantitatively, the rates of missing data are 20.95%, 24.74%, 6.18% and 19.9% in the extrados, left cheek, intrados and right cheek, respectively. The scan spacing of the data in the axial direction is 1 mm while the circumferential spacing (i.e., the distance between two neighboring probes) is approximately 6 mm. This means that the best alignment accuracy in axial and circumferential directions are 1mm and 6 mm, respectively. However, a finer grid of data e.g. 1mm x 1mm is required because the scanning positions at the overlapping zones do not necessarily match.

Although a number of inspections are typically carried out along the service life of feeder pipes, with each inspection outage generating a snapshot of the random field, predicting the time trend of wall thinning remains a third challenging issue. This issue is, however, not covered in this research. Meanwhile, the aligning methodology used in this section for different scan patches can be equally applied to the alignment of the scan data obtained during different inspection outages. Wall thinning in nuclear pipes is a nonstationary nonhomogeneous random field. It is nonhomogeneous because evidences

support that the wall thinning tends to localize at the bends and grayloc positions. It is also nonstationary because there is a relatively clear trend of wall thinning along time.

To construct a full circumferential wall thinning profile, two major tasks are involved. We call them 'repairing' and 'stitching,' respectively.

4.3.1 Repairing

The repairing task refers to the estimation of the missing and unobserved data – we simply fix the 'holes' in Figure 4.1. For this, a moving window kriging is used to account for nonhomogeneity in the spatial data. Within each kriging window the random field is assumed to be homogeneous. Therefore, in this study, the kriging window size is determined based on the correlation length of the semivariogram function defined in Eq (4.3).

In spatial statistics, the semivariogram function $\gamma(h)$ is a very important concept defined as $\gamma(h) = 0.5var[W(s+h) - W(s)]$, and empirically estimated as

$$\gamma(h) = \frac{1}{2N} \sum_{i=1}^{N} \{w(s_i) - w(s_{i+h})\}^2$$

$$4.2$$

where h is the distance between a pair of data, N is the number of paired data. A theoretical model is then fitted to the empirical semivariogram to obtain the correlation length. To account for possible anisotropy, the semivariograms along axial and circumferential directions are calculated and fitted, respectively. The correlation lengths of the two directions θ_x and θ_y are then used to determine the side lengths of the rectangular kriging window. The moving kriging window is then centred at the location where the missing data is to be estimated.

Twice the semivariogram term $2\gamma(h)$ is known as the variogram function. For a homogeneous random field, it is readily shown that $2\gamma(h) = \sigma^2(1 - \rho(h)) = \sigma^2 - C(h)$, where $\rho(h)$ is the auto-correlation function. A number of theoretical semivariogram models are available, refer to Cressie (1993) for more details. Common semivariogram

models in the literature include the Gaussian (i.e. quadratic exponential) and exponential models shown below:

$$\gamma(h) = s\left(1 - e^{-3(h/\theta)^2}\right) \qquad 4.3$$

$$\gamma(h) = s\left(1 - e^{-3h/\theta}\right) \qquad 4.4$$

where s is the sill and θ the correlation length. Based on the equations above, the correlation length is defined as the spatial distance between two locations over which the correlation approximately equals to 0.05. This definition is adopted for kriging purpose as observations beyond this distance becomes little significance in kriging.

Kriging is an effective tool for estimating missing or unobserved data. It is basically a prediction method that involves a linear combination or weighted average of neighboring measurements. The ordinary kriging predictor, which is used in this case study, belongs to a class of best linear unbiased predictor or estimator (Isaaks and Srivastava 1989; Stein 1999). The predictor is best because it minimizes the variance of the estimator. The kriging predictor $\widehat{W}(s_0)$ is expressed as

$$\widehat{W}(s_0) = \sum_{i=1}^{n} \lambda_i W(s_i) \qquad 4.5$$

where λ_i, s_0 and s_i are kriging weights, the location of the point where an estimate is required and the locations with observed attributes or values, respectively. The estimator is unbiased if

$$E\{\widehat{W}(s_0) - W(s_0)\} = 0 \qquad 4.6$$

where $W(s_0)$ represents the true value of the random field at position s_0; whereas the minimum variance dictates the determination of the weights λ_i's so that

$$\sigma_E^2 = \min_{\lambda_i} \ var\{\widehat{W}(s_0) - W(s_0)\} \qquad 4.7$$

For a homogeneous random field, the ordinary kriging weights λ_i sum to 1. In this case, it can be readily shown that the optimal weights can be calculated by using the following equation:

$$\lambda = \Gamma^{-1}\left(\gamma_0 - \frac{\mathbf{1}'\Gamma^{-1}\gamma_0}{\mathbf{1}'\Gamma^{-1}\mathbf{1}}\right) \qquad\qquad 4.8$$

where $\lambda = (\lambda_1, \dots, \lambda_n)'$; $\Gamma = [\gamma(s_i, s_j)]$, $i, j = 1, \dots, n$, is the semivariogram matrix for the observations; $\gamma_0 = (\gamma(s_0, s_i))$ is the semivariogram vector, each element being the semivariogram function value for s_0 and s_i, and $\mathbf{1} = (1, \dots, 1)'$ is the unit vector of n elements. Details of derivation of Eq. (4.8) can be found in Cressie (1993).

The empirical semivariograms of the four patches in both axial and circumferential directions are estimated and illustrated in Figure 4.4. The exponential and Gaussian models are used to fit the empirical semivariograms, and the Gaussian model (Eq.4.3) is found to outperform the exponential model for all eight cases. The blue solid lines in Figure 4.4 represent the fitted model.
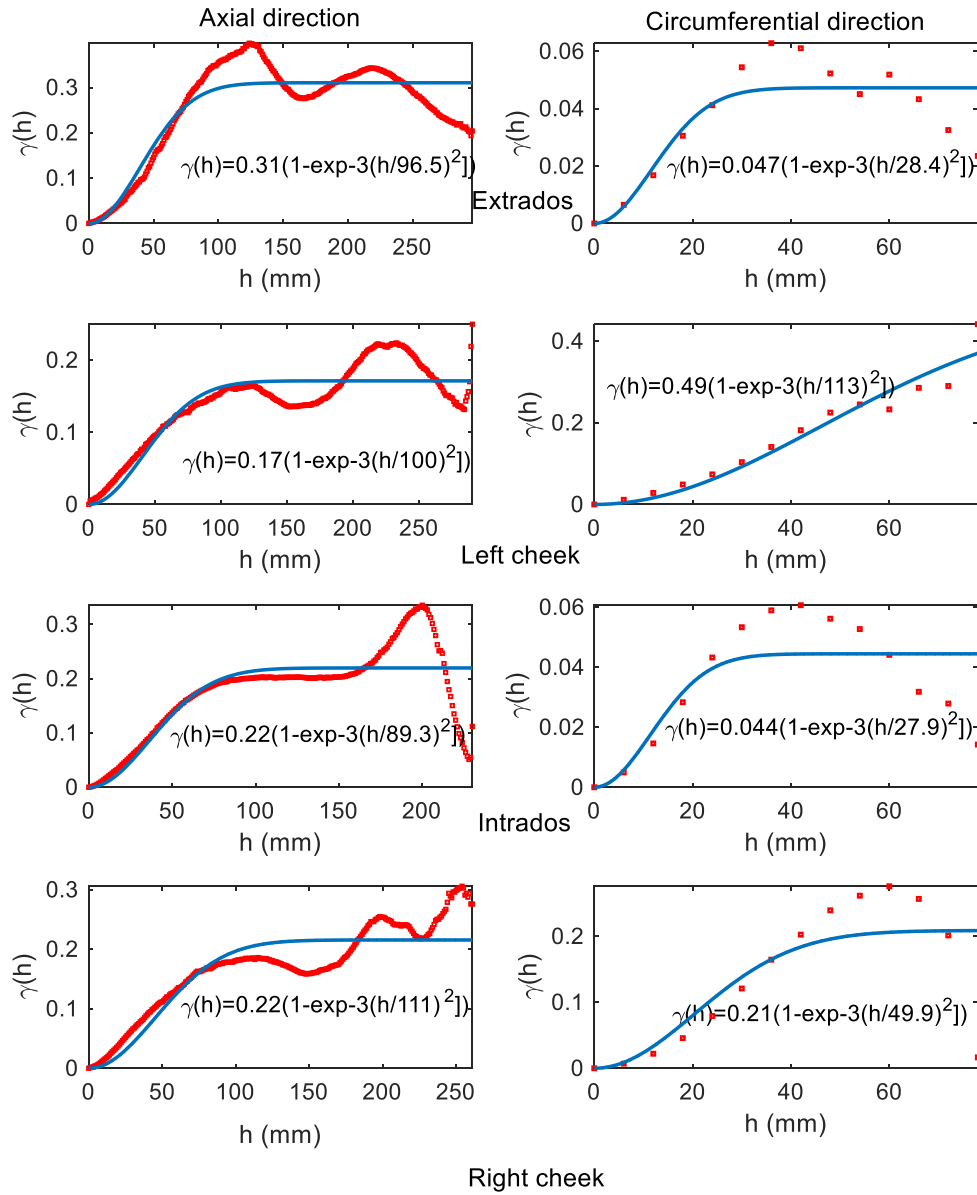
Figure 4.4 Empirical and fitted semivariograms.

The sine wave pattern is consistently revealed in the axial semivariograms. Theoretically, the sine wave usually indicates some sort of periodic pattern in the data. In this study, the axial direction undergoes two bends, and the wall thickness does show some regular pattern when entering and leaving a bend. Therefore, one does not need to worry about this. It is also interesting to see that the correlation lengths of the four patches are fairly

close. This shows certain homogeneity around circumference. By contrast, the circumferential correlation length exhibits greater variation, particularly between extrados/intrados and the cheeks. This is also reasonable, because the flow in terms of turbulence characteristics at the intrados or extrados and that at the cheeks are very different. Therefore, the wall thinning characteristics are also different.

Within each kriging window, the raw data are used to re-estimate the semivariogram functions, which are further used in Eq. (4.8) to estimate the optimal weights. Thus, a kriging window of 100 mm (axial) x 30 mm (circumferential) is used for the extrados and intrados. For the two cheeks, the window size is selected to be 100 mm (axial) x 78 mm (circumferential). In the end, the spacing was reduced to 1 mm apart in both directions after kriging.
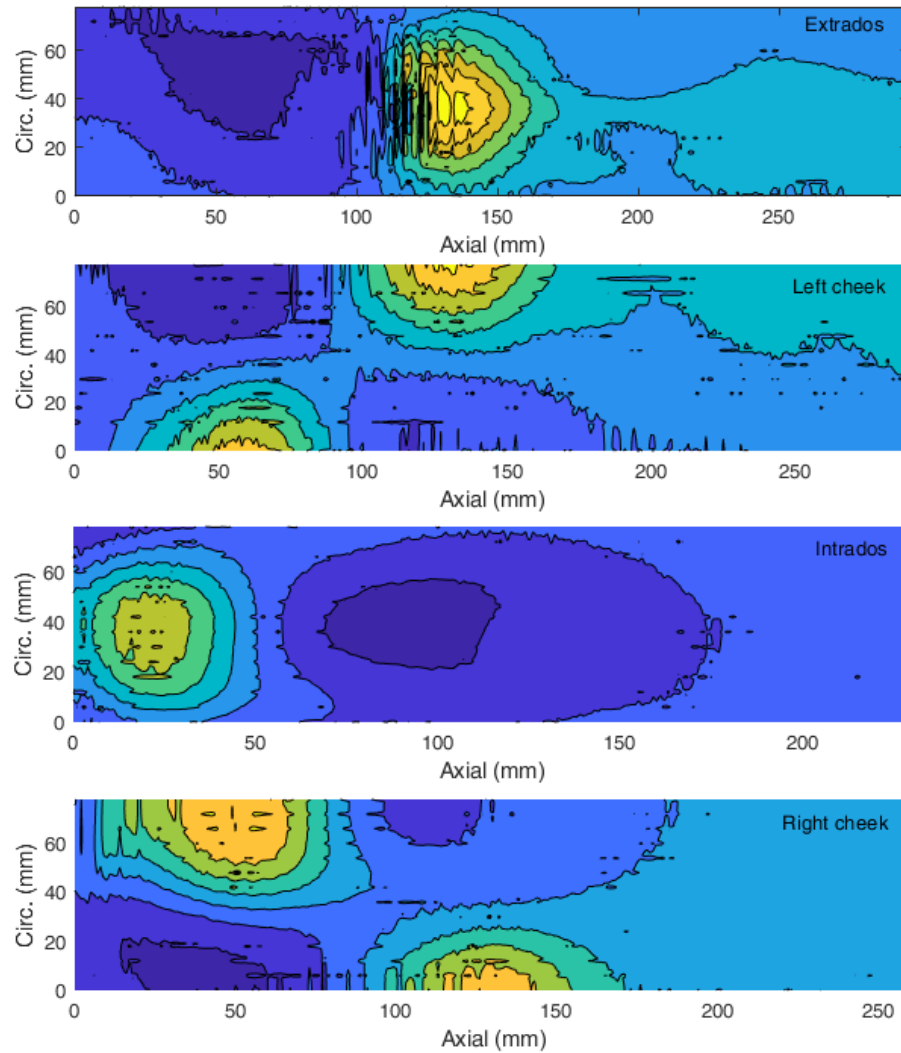
Figure 4.5 Random fields of the four scan patches.

4.3.2 Stitching

After each of the patches are 'repaired,' they need to be stitched together to make a full 360° wall thickness profile along the whole pipe length. This stitching involves both circumferential and axial alignments. In view of these alignments, a two-dimensional cross correlation matrix is computed for each pair of neighboring patches, and the matching point corresponds to the coordinates with the highest positive cross correlation coefficient. The results of the two-dimensional correlation are shown in Figure 4.6. The

111

location with the highest correlation is chosen as the matching point for any two neighbouring scan patches. These points are denoted with the red crosses.
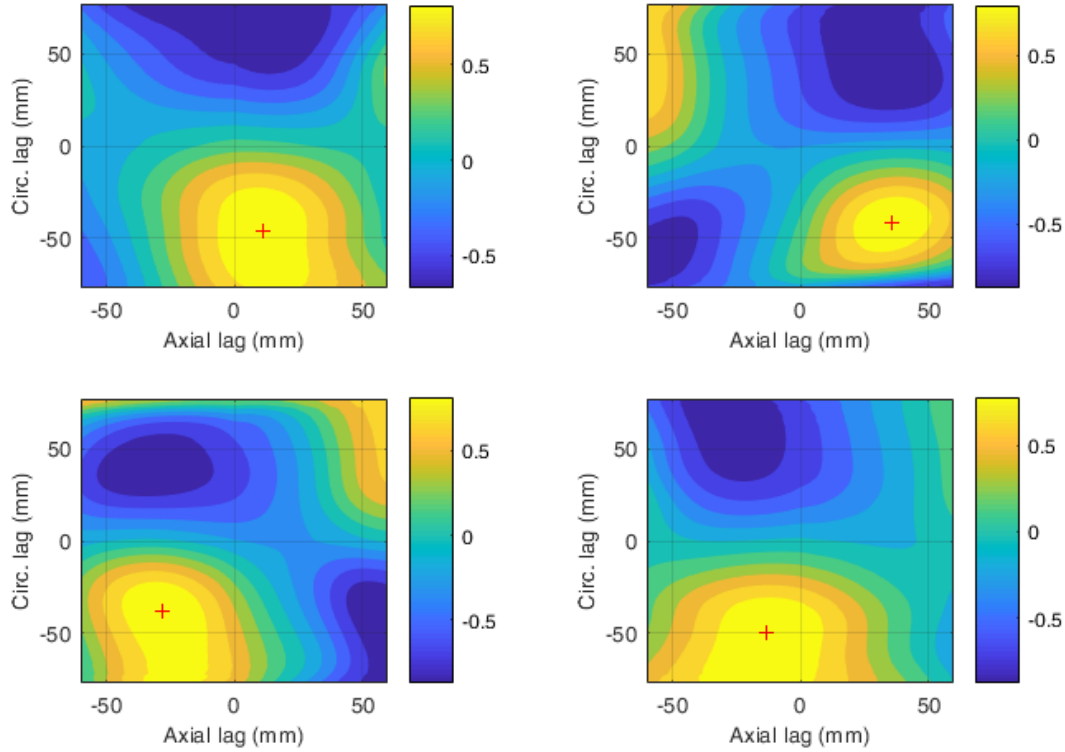
Figure 4.6 Cross correlation plots for the four pairs of scan patches (Top left: Extrados-Left cheek; Top right: Left cheek-Intrados; Bottom left: Intrados-Right cheek; Bottom right: Right cheek-Extrados).

Based on the obtained matching points in Figure 4.6, the four scan patches in Figure 4.5 are stitched together to obtain a full characterization of the random field. For any overlapped position, the minimum value is chosen. The contour plot of the combined random field is shown in Figure 4.7. This plot starts with extrados at the bottom, followed by right cheek and intrados, and finally the left cheek at the top. The two bends are clearly shown, one centered around 60 mm in axial direction, and the other centered around 140 mm. Also it can be seen that the extrados at the first bend becomes intrados (thicker) at the second bend. The globally minimum wall thickness remains unchanged from the raw scan data, which is 3.269 mm.
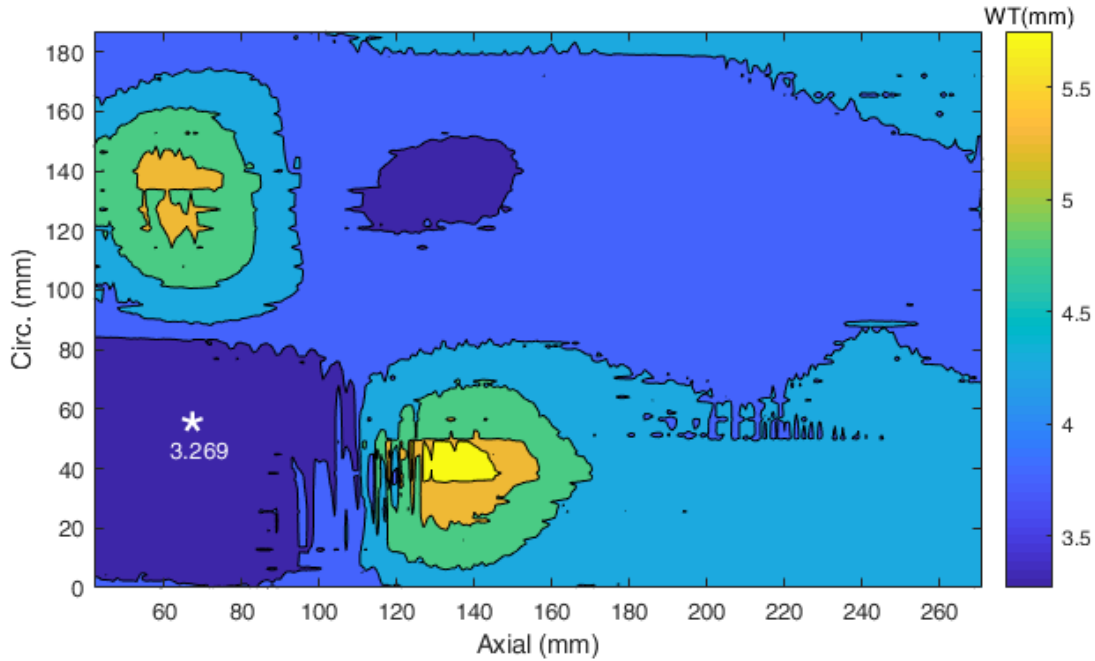
Figure 4.7 Contour plot of combined random field.

4.4 Two-dimensional Gamma Field

One of the underlying Kriging assumptions is that the predicted values have a Gaussian distribution, however, the empirical distribution of the deterioration data is skewed (Figure 4.2). So, realizations at locations in the gamma field are assumed to follow gamma distribution. The combined random field shown in Figure 4.7 now serves as the input data for the parameter estimation of the gamma field. This is coded in MATLAB. In the combined random field, the number of data points is 43 240. If all these data points were used, the size of the correlation matrix would be 43240 by 43240. At this size, the correlation matrix becomes close to singular during iterations and results may be inaccurate. In order to avoid this issue, the combined random field is divided into smaller blocks of data points. The maximum deterioration in each block is chosen as the representative of the block. This approach drastically reduces the size of the correlation matrix in the likelihood function. Theoretically, the quality of the estimates depends on the number of blocks used during parameter estimation. In other words, as the number of blocks increases, the estimates obtained get closer to the true values. The block size is chosen as 8.5mm by 10mm. After deducting the wall thickness at every location from the

initial wall thickness, the maximum deterioration is 2.731mm which occurs at the location with the observed minimum wall thickness.

We assume that the random field is fully separable. The two-dimensional quadratic exponential correlation model is used to describe correlation between any two points as

$$\rho(h_x, h_y) = \exp\left(-3\left(\left(\frac{h_x}{\theta_x}\right)^2 + \left(\frac{h_y}{\theta_y}\right)^2\right)\right) \qquad 4.9$$

where h_x and h_y are horizontal and vertical distances respectively, θ_x and θ_y are the correlation lengths in both directions. Using the likelihood function of equation 2.11 (section 2.1.2),

$$L(\alpha, \beta, \theta_x, \theta_y | x_i, \dots, x_n) = g_n(x) =$$

$$= \frac{1}{(2\pi)^{n/2}|R|^{1/2}} \exp\left(-\frac{1}{2} z^T R^{-1} z\right) \prod_{i=1}^{n} \frac{g(x_i; \alpha_i, \beta_i)}{\phi(z_i)} \qquad 4.10$$

the estimated parameters of the two-dimensional gamma field are found to be $\hat{\alpha} = 45.87$, $\hat{\beta} = 0.0414$, $\widehat{\theta_x} = 23.52$ and $\widehat{\theta_y} = 19.29$.

4.4.1 Simulation

The simulation of two-dimensional gamma fields using the KL expansion is described below.

1. Using $\widehat{\theta_x}$ and $\widehat{\theta_y}$ obtained in the previous section, apply the Nystrom method described in section 3.4 to determine the eigenvalue and eigenvectors in both horizontal and vertical directions.
2. Plug these eigenvalues and eigenvectors into the two-dimensional K-L expansion (eqn. 4.11) to generate a realization of a 2-D Gaussian field

$$\xi(s, t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \sqrt{\lambda_i \lambda_j} \zeta_{ij}(\omega) f_i(s) f_j(t) \qquad 4.11$$

where the number of terms $n = 43$.

3. Repeat step 2 until 100 000 realizations of the 2-D Gaussian field are generated.

4. Obtain the maximum value of each Gaussian field realization.

5. Transform the maximum values of the Gaussian field to 2D gamma field maxima by using $\hat{\alpha}$ and $\widehat{\beta}$ obtained in Section 4.3.

6. The extreme value distribution is obtained from the translated maximum values and shown in Figure 4.8.

4.4.2 Extreme Value Distribution

The extreme value distribution, shown in Figure 4.8, has a mean and coefficient of variation (COV) of 2.953mm and 5.20%, respectively. From the observed scan data, the maximum wall thickness loss is 2.731mm which corresponds to the 5th percentile of the extreme value distribution. As expected, the EV distribution of the gamma field has a long right tail typical of gamma distributions.
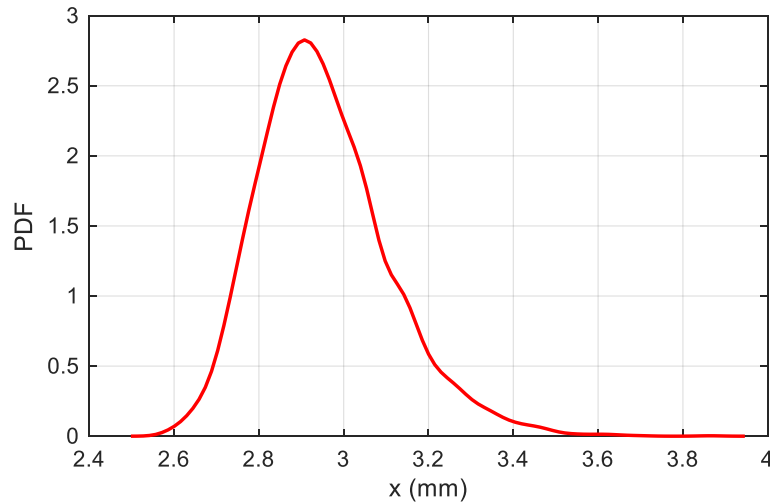
Figure 4.8 Extreme value distribution of the 2D gamma field (wall thickness loss)

4.4.3 Extreme Value Distribution with Parameter Uncertainty

In section 4.3, the parameters of the two-dimensional gamma field are estimated via the maximum likelihood method. Then, these point estimates, together with the 2-D K-L

expansion, are used to evaluate the extreme value distribution of the degradation in the feeder pipe. This distribution captures the aleatory uncertainty in the model. In this case study, however, epistemic uncertainty, which includes the uncertainty in the model parameters, has not been accounted for.

To characterize the epistemic uncertainty in the model parameters, the general theory of maximum likelihood is applied. The theory says that the distribution of an estimate $\hat{\theta}$ is normal with mean θ and variance equal to the inverse of the information matrix, as the sample size goes to infinity. The negative of the Hessian matrix, which is essentially a square matrix of the negative of the second-order partial derivatives of the objective function logarithm with respect to the decision variables evaluated at the maximum likelihood estimates, was obtained from MATLAB's *fmincon*. The information matrix, I is equal to this negative of the Hessian matrix.

$$I = \begin{bmatrix} 2.239 & 2.414e3 & 0.1413 & 0.21 \\ & 2.724e6 & -356.6 & -326.8 \\ & & 6.631 & 4.885 \\ Sym. & & & 8.053 \end{bmatrix}$$

Using the inverse of the information matrix above, together with the mean of maximum likelihood estimator, 5750 sets of parameters are generated from the multivariate normal distribution. Their distributions are shown in Figure 4.9. Each set of parameters is then used to generate realizations of two-dimensional gamma fields by applying the method described in section 4.3.1. The result of the extreme value distributions is then compared with Figure 4.8. See Figure 4.10 below.
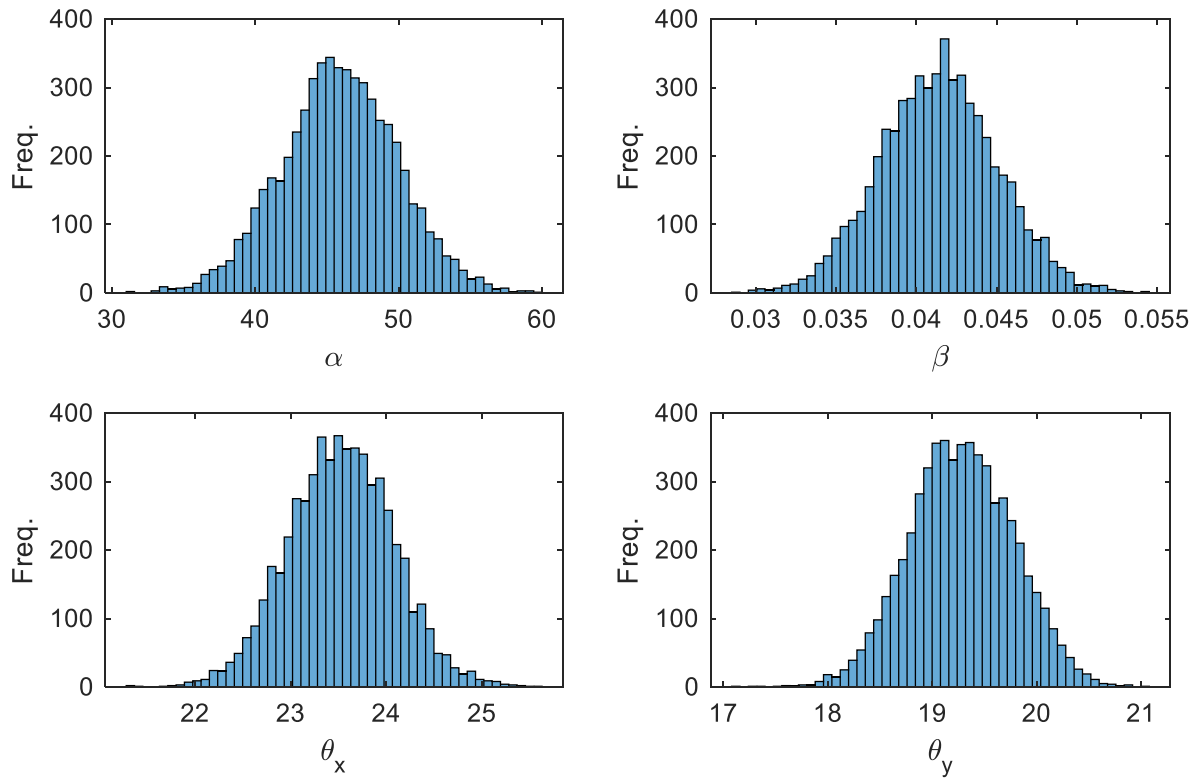
Figure 4.9. Distribution of parameters of random fields
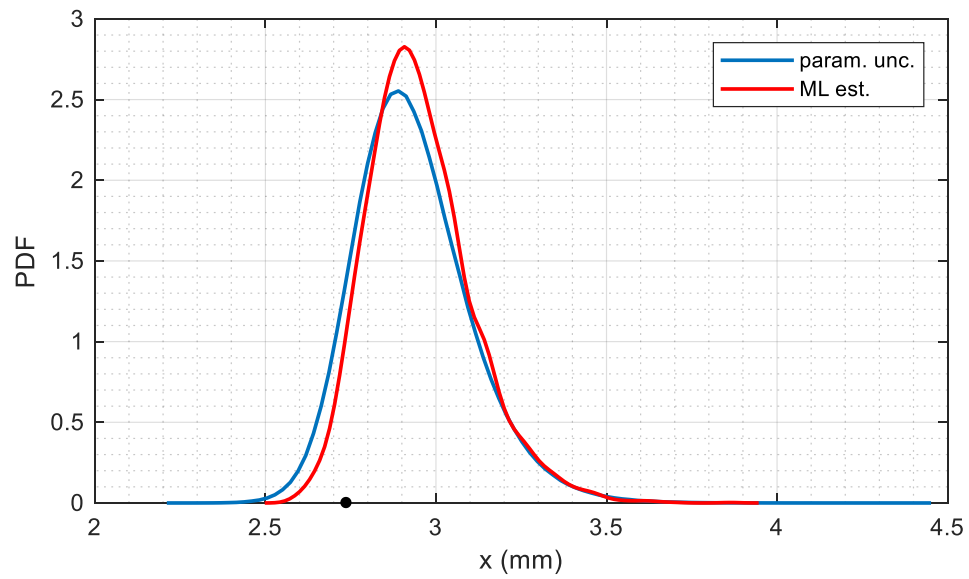
Figure 4.10 Extreme value distribution of 2D gamma fields (wall thickness loss) considering parameter uncertainty (Observed maximum thickness shown with a black dot).

The mean and COV of the extreme value distribution with parameter uncertainty are 2.932mm and 5.75%, respectively. The COV is slightly greater than the COV of the extreme value distribution obtained from the ML estimates. The extreme value distribution with parameter uncertainty has longer tails as expected because introducing parameter uncertainty into the distribution increases the variability.

Every structure, including nuclear components, is designed and built with a target reliability. The acceptance criterion for feeders is that the predicted wall thickness shall not be less than the design minimum wall thickness based on the N285.4 -14 Clause 13.2.5.2. In this case, this thickness is 2.54 mm which corresponds to 3.46 mm wall thickness loss. Hence the probability of failure, defined as $F = \mathrm{P}\{X \geq w_o - 2.54\}$, is equal to 0.0059. This probability of failure is one of the components of risk defined as the product of failure probability and consequence of failure. It is commonplace in engineering to carry out risk reduction which, in this case, may be achieved by either finding ways to reduce this feeder's probability or consequence of failure or both.

4.5   Effects of Missing Data

To evaluate the effect of missing data on the EV distribution, the EV distribution of one-dimensional random fields from the raw data of extrados scan is compared with that of the kriged data. The size of data blocks for kriged fields is fixed at 6mm for the sake of consistency. Based on the procedure described in section 4.1, *Table 4.2* presents the estimated parameters of random fields covering regions with significant missing data and localised wall thinning. Furthermore, the modified Ditlevsen distribution was used to evaluate corresponding EV distributions.

Table 4.2 Estimated parameters of one-dimension gamma fields from extrados scan after kriging.

Probe No.	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\theta}$
3	25.95	0.0671	18.54
4	16.28	0.1066	14.6
5	14.17	0.1259	16.96
6	7.48	0.2246	15.49
7	6.14	0.2686	17.39
8	7.84	0.2191	18.74
9	9.29	0.1955	14.33
10	15.2	0.1233	16.12
11	19.74	0.0979	17.44
12	33.63	0.0589	18.34

For probes with less missing data such as probes 3, 4, 11 and 12, Figure 4.11 shows that the EV distributions of random fields based on only raw data are similar to the distributions obtained using both raw and kriged data. In other words, missing data has little effects on EV distribution. However, the opposite is the case for EV distributions obtained from regions with more missing data coupled with localised thinning. For example, the EV distribution obtained from raw data for probes 5, 8 and 10 is way different to the EV distribution after kriging. Specifically, the EV distributions from raw data exhibits greater variability that those obtained from combined raw and kriged data due to extensive missing data and proximity to the region of localised thinning including the observed minimum wall thickness. Conversely, the EV distributions of random fields using raw and kriged data have lower variabilities because of additional information obtained from kriging i.e. using linear combinations of measurements from neighbouring probes.
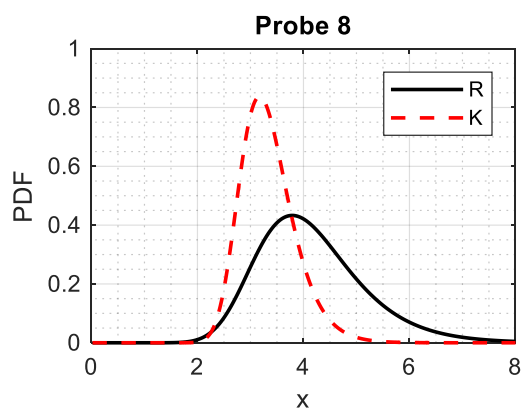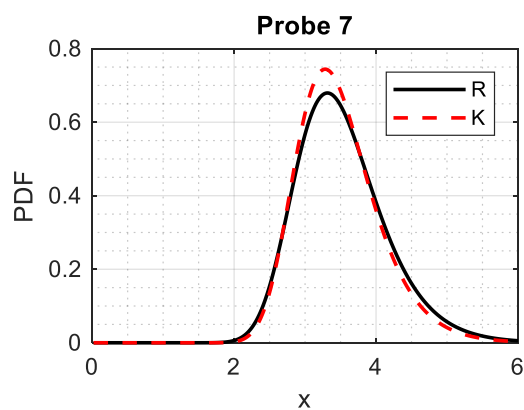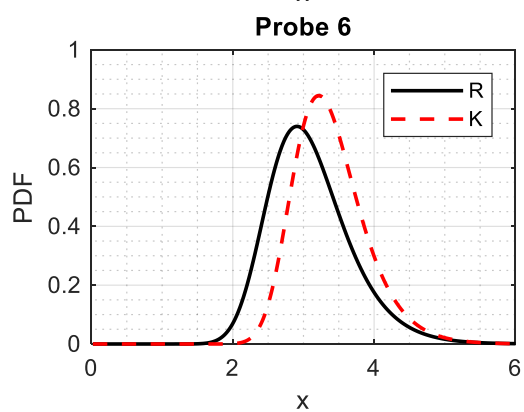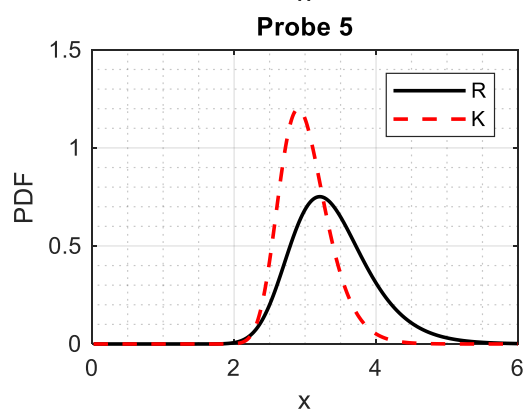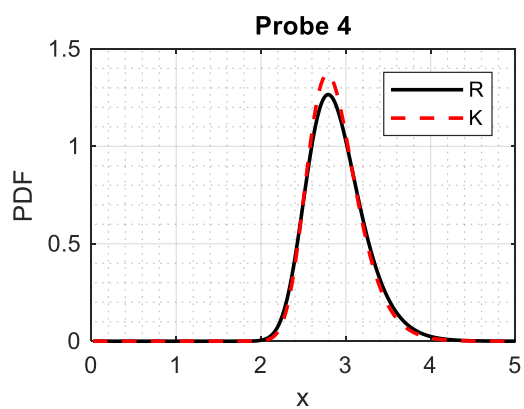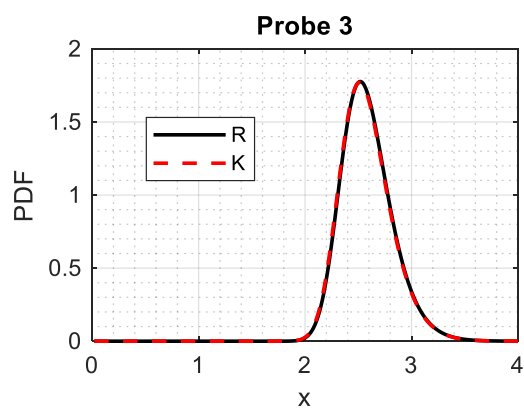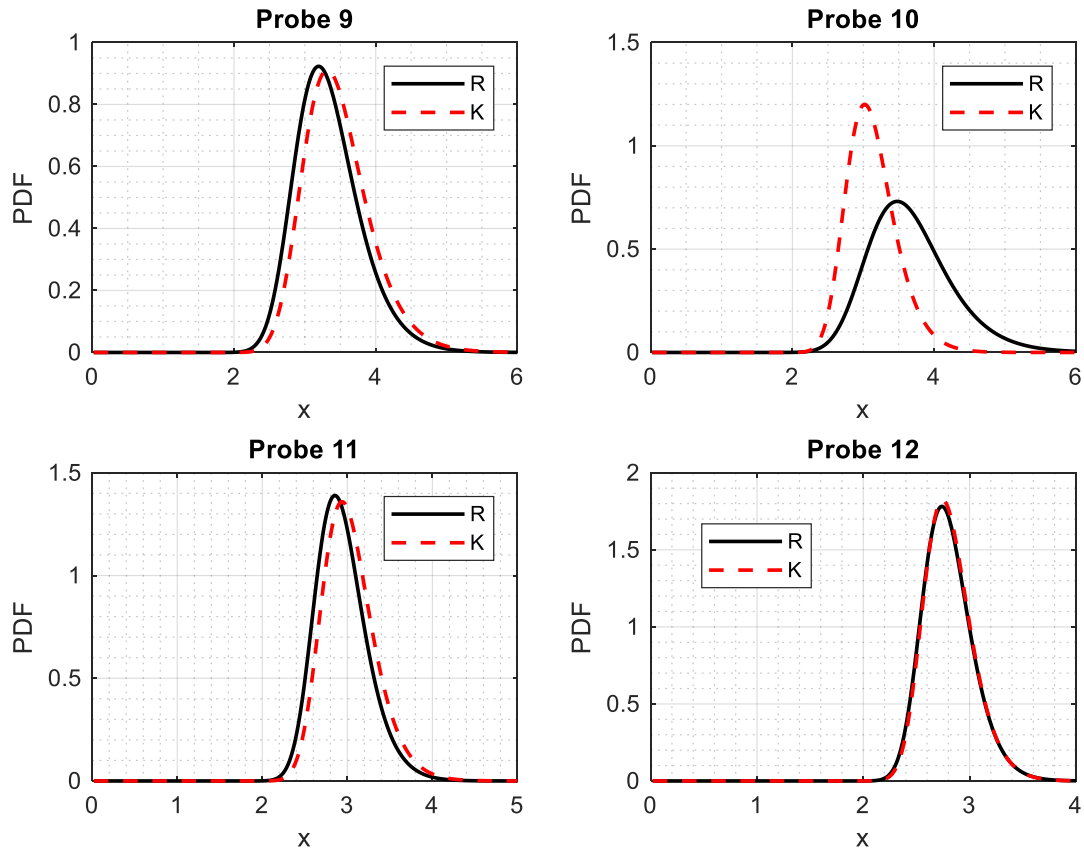
Figure 4.11 Comparison of EV distribution of 1D gamma fields (R=raw data vs K=kriged+raw data).

4.6   Conclusions

Based on the case study of the nuclear pipe, the following conclusions are drawn:

- The variability of EV distributions of one-dimensional gamma fields increases with increasing amount of missing data in the extrados scan.
- The minimum wall thickness of the nuclear feeder pipe after kriging-based 'repairing and stitching' remains the same as that observed in the scan data. Moreover, the 'repairing and stitching' technique also provides a useful tool for future spatiotemporal degradation modelling.

- The combined wall profile of a feeder pipe can be treated as a two-dimensional gamma field and its parameters estimated by the maximum likelihood method.

- The extreme value distribution obtained for the gamma field has a mean and coefficient of variation (COV) of 2.953mm and 5.20%, respectively. Furthermore, the observed maximum deterioration (i.e. minimum wall thickness) corresponds to the 5th percentile of the extreme value distribution.

- With the introduction of the parameter uncertainty in the extreme value distribution of the two-dimensional gamma field, the resulting distribution has much longer tails than the distribution obtained from the point estimates of the parameters.

- The higher the rate of missing data, the greater the variability of the EV distribution of the one-dimensional gamma fields.

5. Multivariate Gamma Process for Dependent Degradation

Over their service life, flexible pavements undergo several forms of deterioration which include cracking, potholes, depressions, cracking, rutting, shoving, etc. These forms of deterioration are known to occur together in pavements and often dependent due to shared underlying causes such as traffic load, construction and material defects and environmental conditions. Modelling each degradation mode separately might not give an accurate prediction of the remaining lifetime and possibly affect inspection and maintenance decisions regarding the pavement. Therefore, there is a need to capture the dependence and model the degradation modes together.

The purpose of this chapter is to present a model suitable for modelling competing degradation phenomena. The multivariate gamma process described in this chapter is a variation of the copula-based gamma field described in Section 3.1. The formal definition, properties and simulation are presented in Section 5.1. Also presented in Section 5.1 is the statistical estimation of the model parameters using the maximum likelihood method. This is followFed by the discussion of the first passage time and remaining lifetime prediction. A case study of multiple dependent deterioration mechanisms in a flexible pavement section is presented to illustrate the applicability of the model in terms of remaining lifetime prediction and condition-based maintenance decisions.

5.1 Multivariate Gamma Process

5.1.1 Definition and Key Properties

An n-dimensional multivariate gamma process $X(t) = \{X_1(t), \dots, X_n(t)\}$ with $t \geq 0$ satisfies the following conditions:

(1) $X_j(0) = 0$ for all $j = 1, \dots, n$.

(2) For any time $t \geq 0$, $X_j(t)$ is a nonstationary gamma process with increments that follow a gamma distribution with shape $\alpha_j(t)$ and scale β_j, i.e., $\Delta X_j(t) \sim Ga\left(\alpha_j(t), \beta_j\right)$.

(3) For any time interval $0 \le t_1 < t_2$, the increments $X_j(t_2) - X_j(t_1)$ follow a multivariate gamma distribution that is defined as Eqs. (2.10) and (2.11) with α_j and β_j for $j = 1, \ldots, n$, and the correlation coefficient is defined as correlation between two stochastic processes.

Again, the gamma distribution mentioned in the second condition is a two-parameter continuous probability distribution whose probability density function (PDF) and cumulative distribution function (CDF) are defined in Eqs. 3.1 & 3.2. The shape parameter is assumed to follow a power law $\alpha_j(t) = a_j t^{c_j}$ for $j = 1, \ldots, n$ and $a_j, c_j > 0$. When $0 < c < 1$, the rate of increase of the shape parameter decreases with time. On the other hand, when $c > 1$, the rate of increase of the shape parameter increases with time. In both scenarios, the stochastic process is non-stationary. However, the stochastic process is said to be stationary when the shape parameter is linear with time i.e. $c = 1$.

In the context of degradation modelling, we consider $X_j(t)$ to be a degradation process that represents the cumulative amount of deterioration observed in a component. Although the multivariate gamma process technically starts at zero as per its formal definition (1st condition), it can be used to model either the cumulative increment or the cumulative decrement of a particular performance degradation phenomenon that does not necessarily start at zero. In other words, degradation phenomena may have a non-zero starting point which must be included as part of the degradation model.

5.1.2 Simulation

Sample paths of a multivariate gamma process can be simulated by the procedure described in this section. The procedure involves generating multivariate Gaussian variates and then transforming them to multivariate random variates with gamma-distributed marginals via copula.

Suppose we are interested in simulating a multivariate gamma process whose dimension is four over a specified planning horizon. A sample path of the multivariate gamma process can be simulated by the procedure described below.

1. Generate random variates from the multivariate normal distribution z of dimension n, with zero mean and a positive definite correlation matrix.

2. Repeat step 1 until the number of sets of variates generated equals the number of increments required for the simulated degradation paths.

3. Perform a double transformation of the variates in steps 1 and 2. The transformation involves calculating the standard normal CDF u_j at each value of z and setting $x_j = G_j^{-1}(u_j)$ where $j = 1, \dots, 4$ and G_j^{-1} is the inverse univariate gamma cumulative distribution function with shape and scale parameters $\alpha_j \Delta t$ and β_j respectively.

Compute the cumulative $X_j(t)$ for each realization. For illustration, see Figure 5.1. Figure 5.2 and Figure 5.3 show the correlation of increments in the standard normal and gamma spaces, respectively. A key assumption of the multivariate gamma process is that the increments of the individual gamma processes are coupled by the Gaussian copula. The copulas in Figure 5.2 are elliptical because the Gaussian copula belongs to the elliptical family.
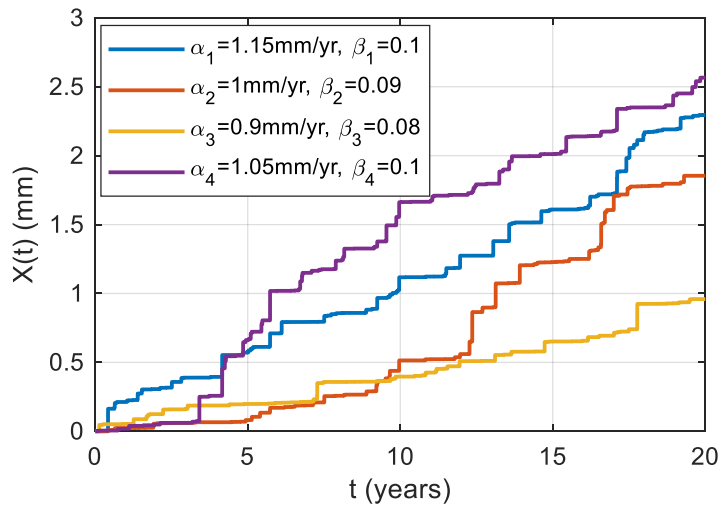
Figure 5.1 Simulated degradation path for a multivariate gamma process ($\Delta t = 2, c_j = 1$ for $j = 1, \dots, 4$; $\{\rho_{12} = 0.7, \rho_{13} = 0.5, \rho_{14} = 0.7, \rho_{23} = 0.4, \rho_{24} = 0.6, \rho_{34} = 0.5\}$).
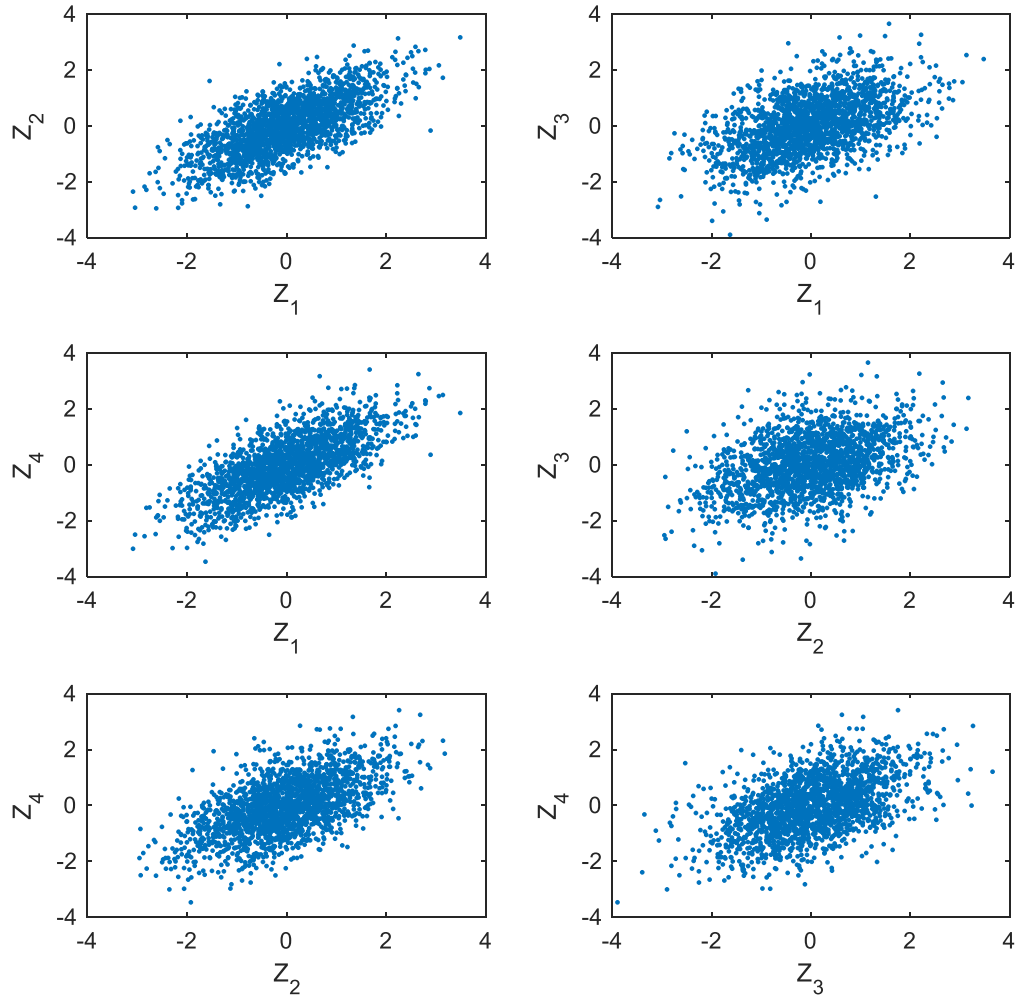
Figure 5.2 *Correlation of increments for the same time interval in standard normal space. 2000 realizations. ($\{\rho_{12} = 0.7, \rho_{13} = 0.5, \rho_{14} = 0.7, \rho_{23} = 0.4, \rho_{24} = 0.6, \rho_{34} = 0.5\}$).*
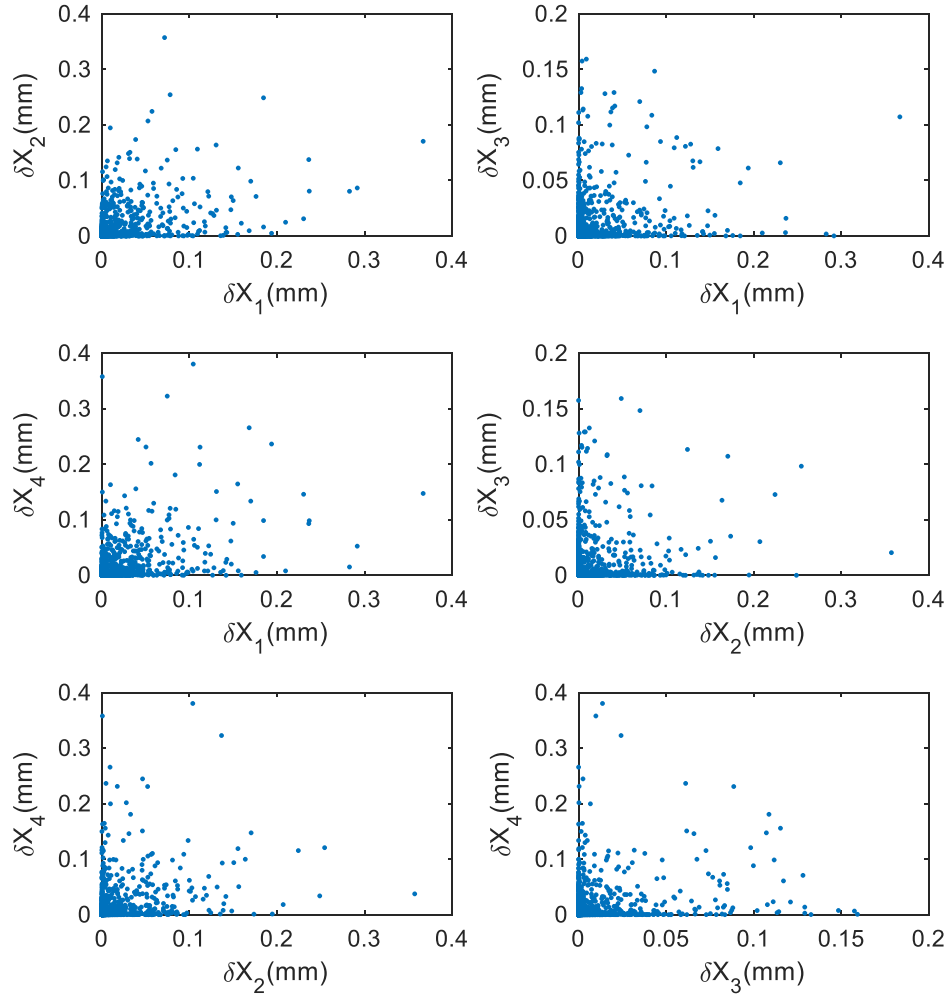
Figure 5.3 *Correlation of increments for the same time interval in gamma space. 2000 realizations. ($\Delta t = 0.1, c_j = 1$ for $j = 1, ..., 4$; {$\rho_{12} = 0.7, \rho_{13} = 0.5, \rho_{14} = 0.7, \rho_{23} = 0.4, \rho_{24} = 0.6, \rho_{34} = 0.5$}).*

5.1.3 Parameter estimation

With a properly structured degradation dataset, the model parameters can be estimated by utilizing the independent increments property. Suppose there exists datasets from m inspection outages of a component experiencing n number of competing degradation phenomena. It is also assumed that all degradation phenomena have common inspection

times $t_0, t_1, \ldots, t_m$, where t_0 is the time the component was put into service. Considering the initial state of the component x_{0j}, where $j = 1, \ldots, n$, there will be m increments for each degradation phenomenon. An increment is defined as $\Delta x_{ij} = x_{ij} - x_{i-1,j}; 1 \leq i \leq m$ for a fixed j. In other words, the degradation data are $X_1(t) = [x_{01}, x_{11}, \ldots x_{m1}], \ldots, X_n(t) = [x_{0n}, x_{1n}, \ldots x_{mn}]$ while the increments are $\Delta X_1(t) = [\Delta x_{11}, \Delta x_{21}, \ldots \Delta x_{m1}], \ldots, \Delta X_n(t) = [\Delta x_{1n}, \Delta x_{2n}, \ldots \Delta x_{mn}]$. For any two consecutive inspection outages, the joint probability density function (PDF) of the multivariate gamma distribution is expressed as

$$g_n(\boldsymbol{\Delta x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{z}^T \mathbf{R}^{-1} \mathbf{z}\right) \prod_{j=1}^{n} \frac{g(\Delta x_{ij}; a, \beta_j, c_j)}{\phi(z_{ij})} \tag{5.1}$$

Consequently, the likelihood function for the joint distribution function is the product of independent multivariate gamma densities of the increments

$$L\left(\boldsymbol{a_j}, \boldsymbol{\beta_j}, \boldsymbol{c_j} | \Delta x_{11}, \ldots, \Delta x_{mn}\right) = \prod_{i=1}^{m} g_n\left(\boldsymbol{\Delta x_{ij}}, \boldsymbol{a_j}, \boldsymbol{\beta_j}, \boldsymbol{c_j}\right) \tag{5.2}$$

The maximum likelihood estimates of $\boldsymbol{a}, \boldsymbol{\beta}$ and $\boldsymbol{c}$ are obtained by numerically maximizing the likelihood function. This is equivalent to computing the first partial derivatives of the likelihood function with respect to each of the parameters of the multivariate gamma process

$$\frac{\partial L}{\partial a_1}, \ldots, \frac{\partial L}{\partial a_n}, \frac{\partial L}{\partial \beta_1}, \ldots, \frac{\partial L}{\partial \beta_n}, \frac{\partial L}{\partial c_1}, \ldots, \frac{\partial L}{\partial c_n} = 0 \tag{5.3}$$

It is always mathematically convenient to take the logarithm of the likelihood function during parameter estimation. The parameter estimation was done in MATLAB using *fmincon*. To ensure that the correlation matrix $\boldsymbol{R}$ remained positive definite at every iteration during the parameter estimation, the Cholesky decomposition of $\boldsymbol{R}$ was used in the likelihood function. After the solution converged, the correlation matrix was reassembled i.e. $\boldsymbol{R} = \boldsymbol{L}\boldsymbol{L}^T$, where $\boldsymbol{L}$ is a lower triangular matrix.

5.2 First Passage Time and Remaining Life Prediction

Generally, failure is said to occur in engineering when degradation exceeds the threshold specified in the code(s). In pavement engineering, however, the definition of failure is dependent on what really matters to the planner. For example, a pavement experiencing multiple degradation phenomena may be said to fail when any individual degradation process $X_j(t)$ reaches its critical threshold ζ_j. In other words, each degradation process determines the failure of the component. Mathematically, the probability of failure is defined as

$$F(t) = P(T \leq t) = 1 - P(X_1(t) < \zeta_1, \dots, X_n(t) < \zeta_n) \qquad 5.4$$

The other extreme is when failure is defined as when degradation phenomena all reach their respective thresholds. Mathematically,

$$F(t) = P(T \leq t) = P(X_1(t) \geq \zeta_1, \dots, X_n(t) \geq \zeta_n) \qquad 5.5$$

Alternatively, a pavement subjected to multiple degradation processes may be said to have failed when a process reaches its failure threshold, two specific processes both reach their thresholds, any two processes both reach their thresholds or any combination thereof.

As the state of each degradation process can be observed, the probability density function takes into account this information. Suppose the degradation processes are last observed at surviving time s, then at future time t the probability of a degradation increment of $\zeta_j - X_j(s)$ is an updated pdf $f_{X_j(t)-X_j(s)}$. To estimate the remaining lifetime distribution, growth of each process over time has to be predicted based on the updated pdf. So, future degradation process is given as

$$X_j(t) = X_j(s) + \Delta X_j(t - s) \qquad 5.6$$

where $\Delta X_j(t - s)$ is the addition of all future increments up to time t when failure occurs.

5.3   Case Study: Highway Pavement

Flexible pavements experience multiple degradation over time as a result of normal wear and tear caused by traffic load and climatic conditions. Other contributing factors to pavement material breakdown are construction failure and prolonged exposure to atmospheric substances such as rain and sunlight. Examples of common degradation phenomena in pavement include cracking and rutting. In this case study, three measures of pavement degradation are considered. These are rut depth, International Roughness Index (IRI) and Distress Manifestation Index (DMI). MTO uses the Automatic Road Analyzer (ARAN®) to monitor the pavement quality. This includes measuring IRI and rut depths and collecting information on other forms of pavement defects which are subsequently aggregated under the DMI. A little background on the performance indices in flexible pavements is presented in the following paragraphs.

Rutting is a permanent deformation along the wheel path on the road surface and increases over time. A newly constructed road, for instance, has a zero rut depth. The ARAN® uses Laser Crack Management System to detect and measure rut depths on freeways and over long distances. In order to obtain a representative rut depth for a pavement section, the mean of rut depths taken at regular intervals is calculated.

The IRI is a dimensionless measure of road roughness in the sense that its unit is mm/m or m/km. IRI increases over time until there is an intervention in the form of a maintenance action. Ideally, a newly-built road is expected to have a zero IRI, but this is hardly the case. A rough unpaved road can have an IRI as high as 10 m/km (MTO 2012).

DMI refers to a subjective weighted sum of all distresses and is a measure of overall service damage for the road section. Theoretically, DMI values range from 0 (worst condition) to 10 (excellent condition). The weight, severity and density of each distress are taken into account to calculate the DMI, which decreases over time for any road pavement until there is an intervention. MTO (2012) lists 15 distress types which include raveling, flushing, shoving, rutting, distortion, wheel track and different forms of cracking. For asphalt concrete pavement, the DMI is calculated as shown in Eq. 5.7 (MTO 2012).

$$DMI = 10 \left(208 - \sum_{k}^{n} w_k (S_k + D_k) \right) / 208 \qquad\qquad 5.7$$

where S_k, D_k, and w_k represent the severity rate, density rate and weight factor of distress k and n is the number of distresses in the pavement.

Having described how degradation data are obtained in a pavement, the case study is now presented. For the case study, the assumptions are:

i.   The road section is subjected to multiple degradation processes $X_j(t)$, where $j = 1, 2, 3$ and these processes are assumed to be dependent. Each $\{X_j(t), t \geq 0\}$ is a nonstationary gamma process with shape and scale parameters $\alpha_j(t)$ and β_j respectively.

ii. Contrary to the first condition in the definition of the multivariate gamma process, degradation phenomena do not necessarily start from zero, so $X_j(t) = x_{0j} \pm Ga(\alpha_j(t), \beta_j)$ where x_0 is the initial measure of the degradation.

Table 5.1 presents the degradation data for a section of a flexible pavement road. The table shows the measurements of DMI, rut depths and IRI covering a 7-year period. Measurements were taken on a yearly basis. To incorporate the DMI values in the increasing gamma process, the absolute values of the changes are used in the parameter estimation. Measurement error in the observed data is not accounted for in the model.

Table 5.1 Degradation data for a road section.

Year	DMI	IRI (m/km)	Rut depth (mm)
2005	9.49	1.12	3.49
2006	9.03	1.21	4.56
2007	8.73	1.29	4.85
2008	8.54	1.35	5.44
2009	7.83	1.44	5.76
2010	7.51	1.54	5.99
2011	7.02	1.68	6.61

The procedure described in Section 5.1.3 is used to estimate the parameters of the multivariate gamma process model. The estimated parameters are shown in Table 5.2.

Table 5.2 Estimated parameters of the multivariate gamma process

Parameter	DMI	IRI	Rut depth
α	5.89	15.1	12.8
c	1.01	1.10	0.65
β	0.07	0.01	0.08

Correlation coefficients	DMI	IRI	Rut depth
DMI	1	0.59	0.08
IRI		1	0.32
Rut depth	*sym.*		1

Table 5.2 reveals that the power term, $\hat{c}$ of the shape parameters is less than 1 for rut depth. This confirms the initial assumption of nonstationarity i.e. the mean rates of the rutting phenomenon is not linear with time. However, the mean rates of increase of the DMI and IRI are close to 1. The correlation coefficients shown as part of Table 5.2 are nonlinear correlations among decrements in DMI as well as increments in IRI and rut depth. From the table, DMI and IRI are modestly correlated (i.e. 0.59). A probable explanation for this is that most of the distresses aggregated under DMI are also responsible for roughness in the pavement. However, there is a very low correlation between rut depth and DMI (0.08) because wheel track rutting is just one of the distresses in DMI but has nothing to do with the other fourteen.

To study the effect of modelling the stochastic processes as dependent as against individual univariate stochastic processes, the parameters of individual nonstationary gamma processes are estimated by numerically maximizing the likelihood function

$$L\left(a,\beta,c \mid \Delta x_{1j},\dots,\Delta x_{mj}\right) = \prod_{i=1}^{m} \frac{\Delta x_{ij}^{\left(a(t_i^c - t_{i-1}^c)-1\right)} \exp\left(-\Delta x_{ij}/\beta\right)}{\beta^{\left(a(t_i^c - t_{i-1}^c)\right)} \Gamma\left(a(t_i^c - t_{i-1}^c)\right)} \qquad 5.8$$

The parameters in Table 5.2 are compared with corresponding parameters of individual nonstationary gamma processes shown in Table 5.3. Both tables reveal that shape parameters for the multivariate gamma process model are greater than the shape parameters from corresponding individual gamma process models. Meanwhile, the tables suggest that the scale parameters and the power term in both multivariate and individual gamma process models are comparable.

Table 5.3 Estimated parameters of nonstationary gamma processes

Parameter	DMI	IRI	Rut depth
α	5.87	14.7	13.2
c	1.02	1.12	0.64
β	0.07	0.01	0.08

5.4 Remaining Lifetime Prediction

Monte Carlo simulation is used to generate one million sample paths and failure probability evaluated by dividing the number of times $X_j(t)$ exceeds ζ_j by the total number of simulation runs. The simulation uses the parameters shown in Table 5.2 in the previous section together with failure thresholds for freeways in Table 5.4. These are the thresholds that need to be satisfied for Ontario freeway pavement.

Table 5.4 Failure thresholds of degradation phenomena in freeways

Phenomenon	ζ
DMI	6
IRI	1.9m/km (MTO 2012)
Rut depth	12mm (Yuan et. al. 2017)
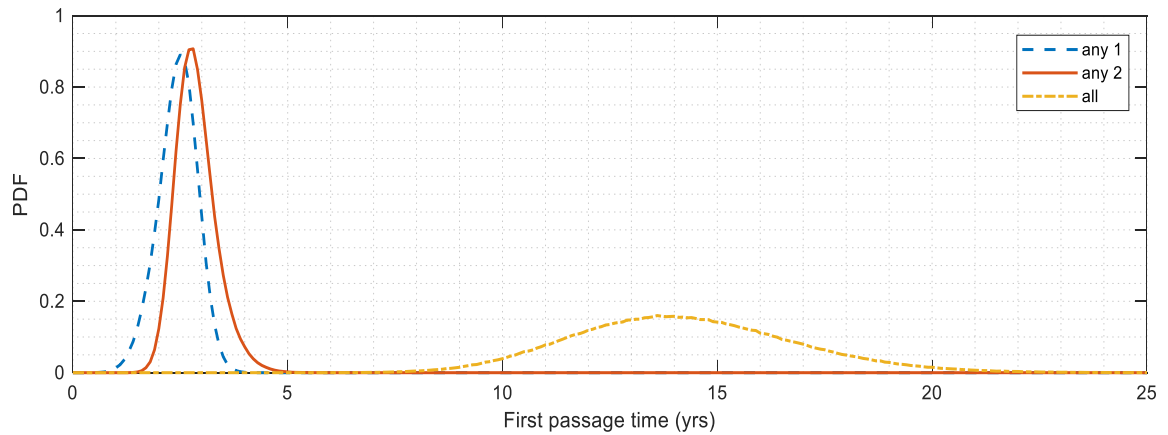
Figure 5.4 Remaining lifetime distribution for three scenarios.

Figure 5.4 shows the remaining lifetime distribution based on current state of the pavement section. Three scenarios are considered namely when failure is defined as any degradation phenomenon reaching its failure threshold, any 2 phenomena both reaching their thresholds and all three reaching their thresholds. Figure 5.4 reveals that as the definition of failure is relaxed, the distribution of remaining lifetime shifts to the right as expected. Furthermore, the mean of the distribution for each scenario estimated numerically is shown in Table 5.5. The means are estimated to be 2.44, 2.87 and 14.2 years, respectively.
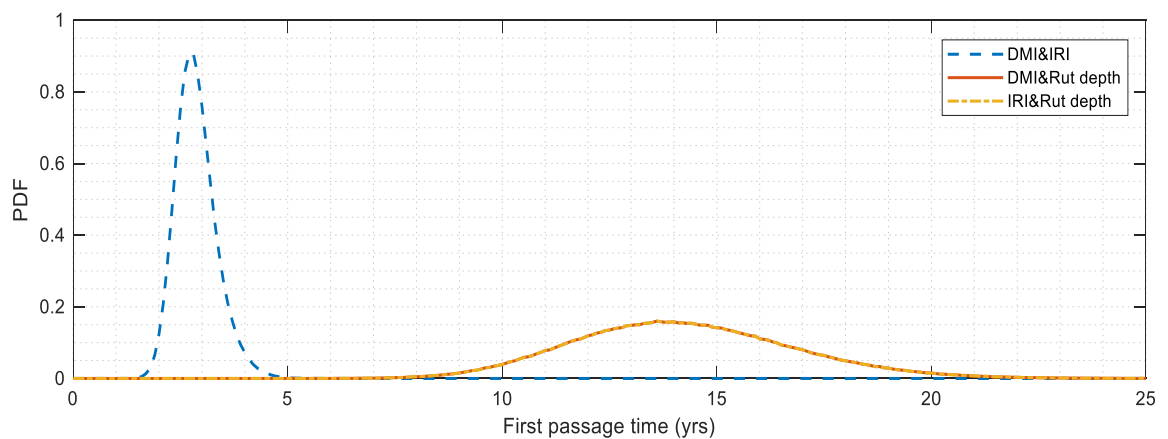
Figure 5.5 Remaining lifetime distribution for specific pairs of degradation phenomena.

Figure 5.5 shows another three scenarios as well. These are when failure is defined as when DMI and IRI both reach their failure thresholds, DMI and rut depth both reach their thresholds and IRI and rut depth both reach their thresholds. Numerical estimation of the expectations of these distributions are presented in Table 5.5. For the first three failure criteria in Table 5.5, the magnitude of the remaining lifetime means depends on how strict the definition of what constitutes failure in the pavement section is. In all three cases, the remaining lifetime distributions obtained depend largely on which degradation phenomenon dominates. In general, the lifetime distributions and their means are a function of dominance of a degradation phenomenon over another (first six failure criteria only) as well as the degradation parameters and how far the failure thresholds are from the most recent condition of the pavement.
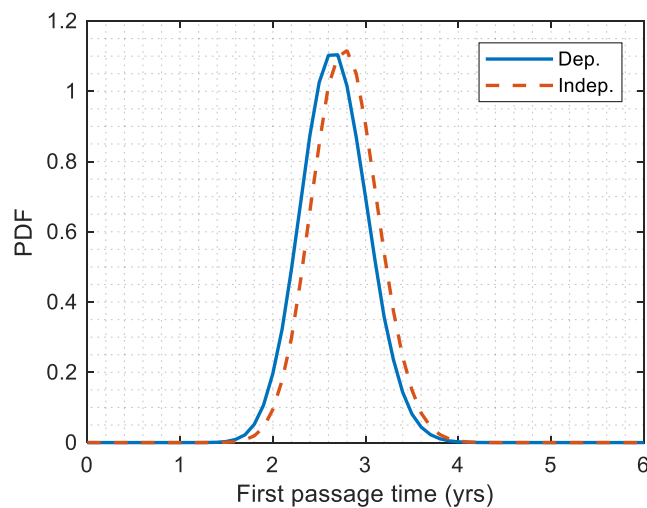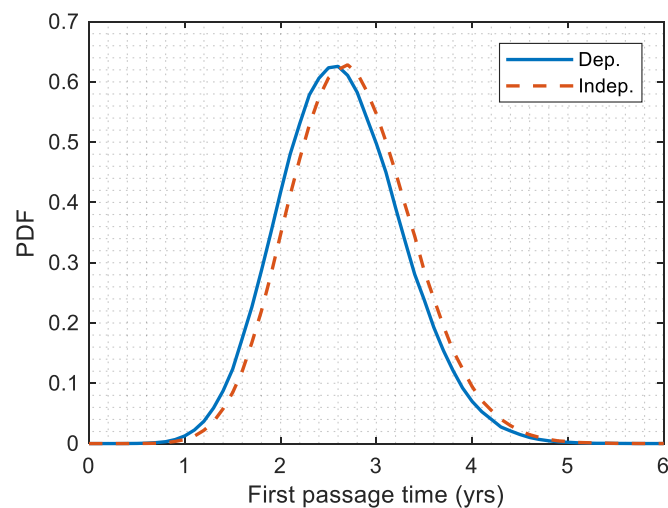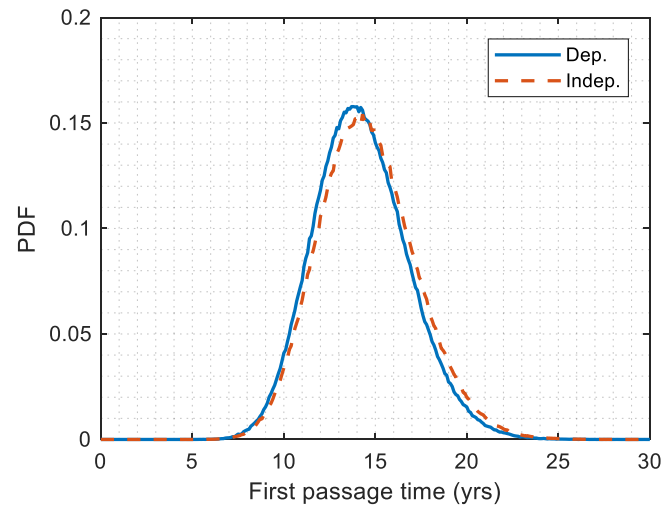
Figure 5.6 Comparison of remaining lifetime distribution when only one phenomenon matters: Dependent vs Independent. Top: DMI; Middle: IRI; Bottom: Rut depth.

In Figure 5.6, it is assumed that failure occurs when a specific stochastic process exceeds its failure threshold. Using Monte Carlo simulations, three scenarios are presented in the figure while corresponding means of the distributions are shown in Table 5.5. Comparisons are made between the remaining lifetime distributions obtained from the multivariate gamma process and univariate gamma processes. The results show that modelling dependent gamma processes as univariate gamma processes may slightly overestimate (Figure 5.6) the lifetime distributions.

Table 5.5 Remaining lifetime means based on different failure criteria (multivariate gamma process)

Scenario	*any* 1	*any* 2	*all*	DMI & IRI	DMI & Rut depth	IRI & Rut depth	DMI	IRI	*Rut depth*
Mean (years)	2.44	2.87	14.2	2.87	14.2	14.2	2.64	2.67	14.2

5.5   Condition-Based Maintenance Decisions

5.5.1   Inspection Optimization

In this section, we present what optimal decisions should be taken based on the multivariate gamma process model proposed in the previous sections. Based on the data used in the model, periodic inspections are being done on an annual basis. Suppose the failure cost C_F, preventive maintenance cost C_R and inspection cost C_I are known, is it economically justifiable to continue to carry out annual inspections of the pavement? This research question is formulated as an optimization problem with an overall objective of finding the inspection time that minimizes the total expected cost over the entire service life of the pavement.

For the condition-based optimization, each degradation grows until when inspection is carried out. At inspection, only the cost of inspection is incurred if the pavement deterioration is below the preventive maintenance threshold. However, if the amount of deterioration is greater or equal to the preventive maintenance threshold but less than the failure threshold, preventive maintenance is done. Once deterioration reaches the failure threshold, the pavement must be immediately inspected and fixed. In other words, failures are self-revealing. In reality, both types of thresholds are set by the local/regional authority in charge of transportation infrastructure.

In practical terms, the preventive maintenance thresholds can be compared to serviceability limit levels. When deterioration in pavement reaches these thresholds, preventive maintenance actions are required to extend the pavement service life and to forestall much more expensive failure maintenance due to extensive deterioration in the pavement. Examples of preventive maintenance actions include applying overlays and seals of different kinds. When the deterioration levels in the pavement are way past the preventive maintenance thresholds, failure maintenance actions in the form of patching, pothole repair, etc. need to be carried out.

A number of assumptions are made in this model. First, it is assumed that inspections are perfect. This implies, there are negligible measurement errors and inspections reveal the

true state of the pavement. Second, we assume every future preventive or failure maintenance returns the pavement conditions to perfect conditions. In other words, there is perfect preventive and failure maintenance. Third, the times spent to carry out preventive and failure maintenance are assumed to be negligible, hence there are no downtimes. See Figure *5.7* for illustration.
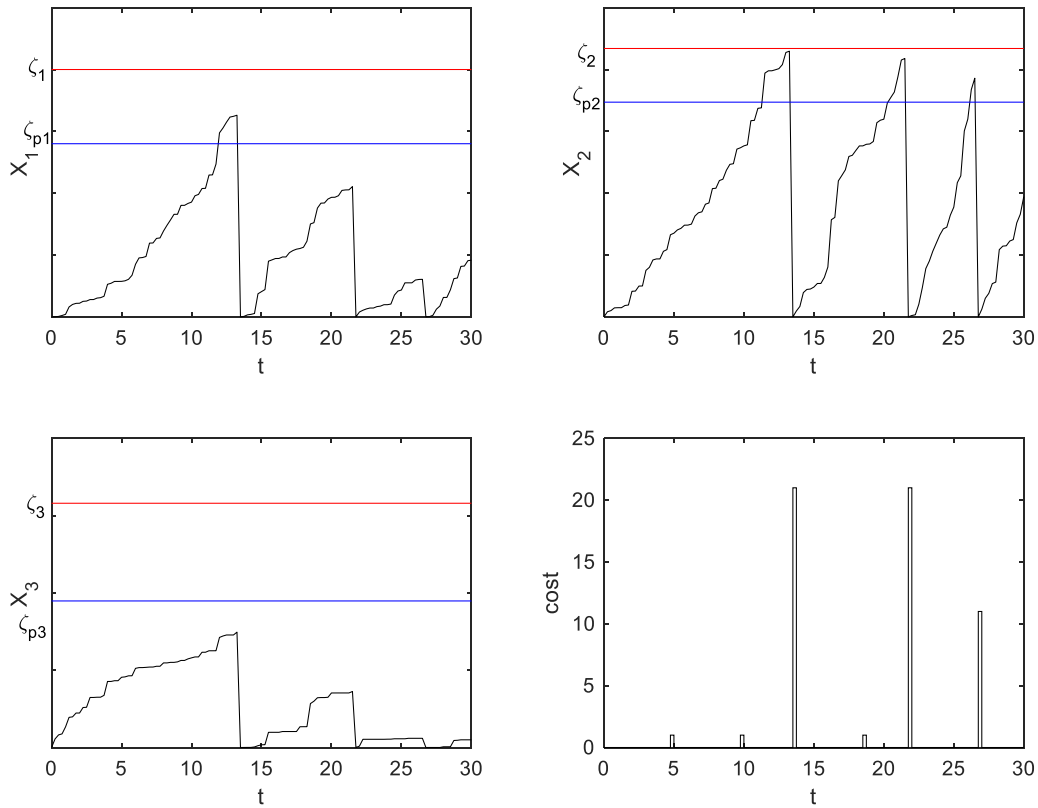

Figure 5.7 Realization of deterioration processes and associated costs when failure is defined any process reaching its failure threshold. 5-year inspection interval; time step – 0.25. Top left: DMI; top right: RI; bottom left: Rut depth; bottom right: costs incurred.

According to Park (1988), the total mean cost over the service life is given as

138

$$E[C] = \sum_{n=1}^{\infty} \left((nC_I + C_P)P\{preventive\ maintenance\ at\ nt_I\} \right.$$
$$\left. + [(n-1)C_I + C_F]P\{unit\ fails\ in\ [(n-1)t_I, nt_I]\} \right)$$

5.9

where $P\{\}$ is the probability of an event. The inspection cost C_I is defined as the cost of obtaining deterioration data. The failure cost C_F is defined as the cost of fixing the pavement when deterioration exceeds the failure threshold ζ. The preventive maintenance cost C_P is the cost of fixing the pavement whenever inspection shows that deterioration exceeds the preventive maintenance threshold ζ_{Pj}. In this case study, preventive maintenance thresholds are $\zeta_{Pj} = k_j\zeta_j$, where k_j are arbitrarily taken as $k_1 = 0.7, k_2 = 0.8, k_3 = 0.6$. In order to illustrate the optimization of the expected cost, we assume the following cost parameters: $C_I = 1$, $C_P = 10$ and $C_F = 20$. The optimization procedure is based on a 30-year planning horizon and 100 000 Monte Carlo simulation runs.

Figure 5.8 compares the expected cost and 95th percentile of the cost when failure is defined as Eq. (5.4). The figure shows that the expected cost shows some form of undulation. This is because the expected cost depends on whether an interval is more associated with preventive or failure replacement. As a matter of fact, the mean occurrence of preventive maintenance per planning horizon decreases with inspection interval up to 6 years (Figure 5.9). However, the mean occurrence of preventive maintenance increases and reaches a maximum at an inspection interval of 8 years and then decreases to zero at much higher inspection intervals. Regarding the occurrence of failure maintenance, the mean rate increases with inspection interval, peaking at 6 years and then falls to local minimum at 8 years. Afterwards, the occurrence rate increases to a maximum of 2.7. Generally, the inspection cost decreases as the inspection interval increases (Figure 5.10). As a result, the cost of inspection relative to the other costs is small so the inspection cost has little effect in determining the shape of the curve for higher inspection intervals. The combination of these trends shown in Figure 5.9 explains why the expected cost function has multiple peaks and troughs. The optimal inspection

interval and corresponding (expected and 95th percentile) costs are 8 years, 39.6 and
54.0 units, respectively.
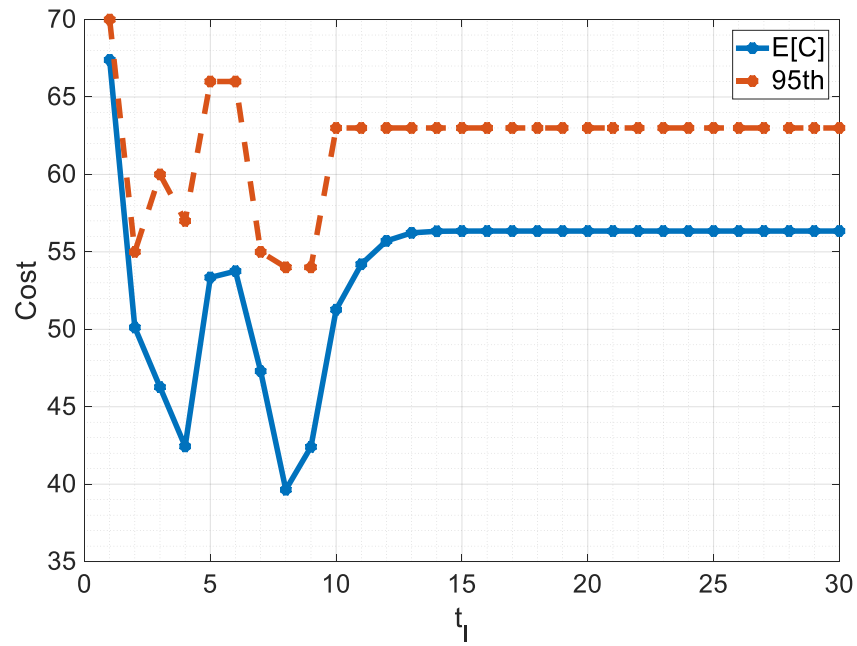
Figure 5.8. Expected and 95th percentile costs for when failure means any phenomenon
reaches its threshold.
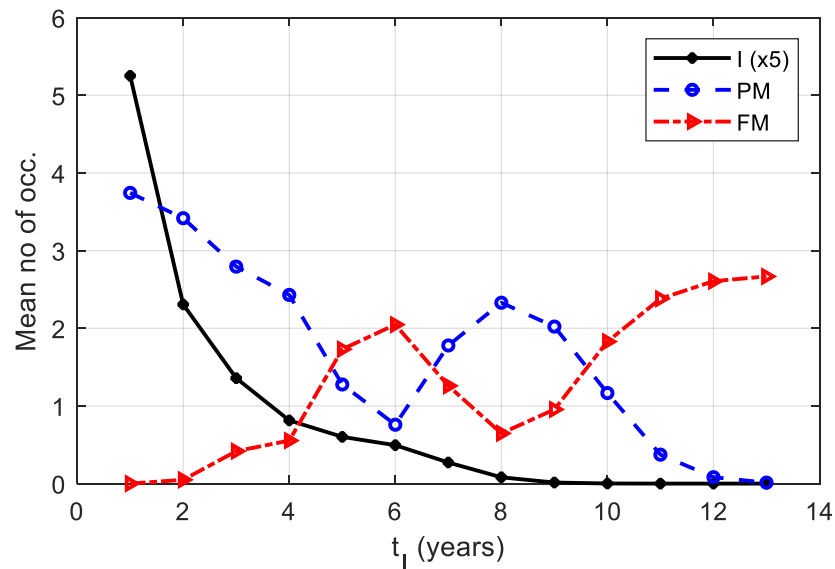
Figure 5.9. Mean occurrence rates of inspections, preventive and failure maintenance actions
against inspection interval.
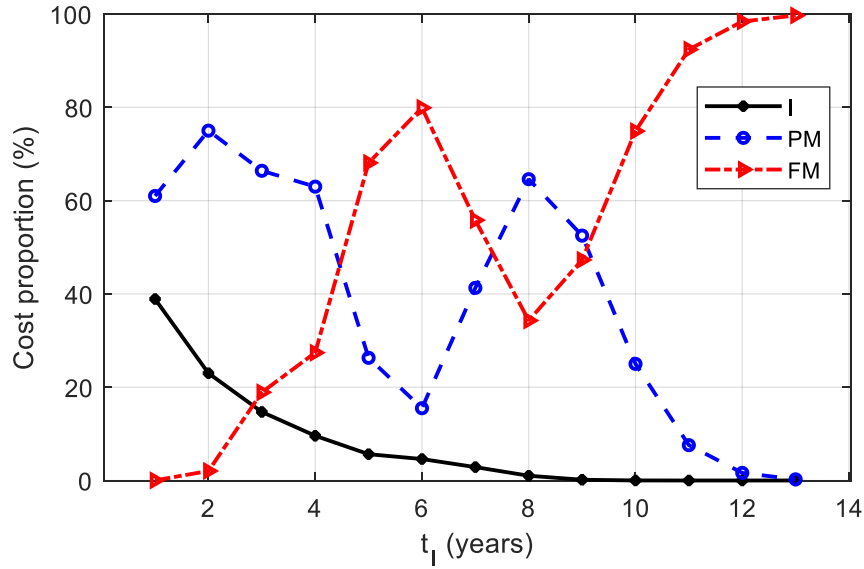
Figure 5.10. Cost proportions of inspections, preventive and failure maintenance actions against inspection interval.

In summary, based on the expected cost criterion, pavement inspections should be done every 8 years. This result is due to hypothetical cost parameters used in the optimization. It is important to note that the optimum inspection interval obtained is way different from the current MTO practice whereby pavement inspections are carried out annually partly due to regulatory and safety reasons.

Changing some of the cost parameters could result in a lower optimum inspection interval. To illustrate how the cost parameters influence expected and 95th percentile costs and optimal inspection interval, the optimization is repeated with a different failure maintenance cost C_F while all other parameters are kept constant. Figure 5.11 shows that when $C_F = 50$, the optimal inspection interval is 2 years with an expected cost of 51.6 units.
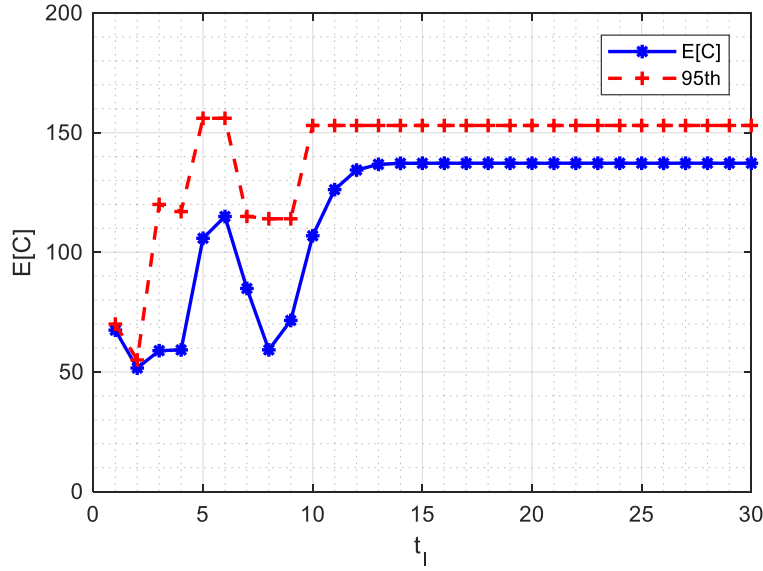
Figure 5.11. Expected and 95th percentile costs for when failure means any phenomenon reaches its threshold ($C_I = 1$, $C_P = 10$, $C_F = 50$).

5.5.2 Multidimensional Optimization

In the previous section, the only decision variable in the optimization is inspection interval. Preventive maintenance thresholds k_j are parameters. The optimization problem can be formulated in a different way by treating inspection interval t_I and preventive maintenance thresholds $k_j, j = 1,2,3$ as decision variables with the objective to minimize expected cost $E[C]$. For the optimization, k_j are continuous variables with lower and upper bounds of 0 and 1, respectively while t_I is an integer decision variable. All other assumptions in section 5.5.1 hold. The cost parameters used are $C_I = 1$, $C_P = 10$, $C_F = 50$.

Using a genetic algorithm solver and 10,000 Monte Carlo simulations, the minimized expected cost is found to be 49.8 units while the optimal inspection interval is 2 years with optimal thresholds $k_1 = 0.72$, $k_2 = 0.54$, $k_3 = 0.37$.

5.6 Conclusions and Summary

Based on the study on multivariate gamma process for multiple degradation, these conclusions are made:

- The proposed multivariate gamma process can be effectively used to model dependent degradation phenomena in civil infrastructure such as a flexible pavement.

- The parameters of the multivariate gamma process were used to generate realizations of future degradation paths which are subsequently used to evaluate the remaining lifetime distribution based on a number of failure scenarios.

- The remaining lifetime distribution of any failure scenario depends on the current state and parameters of the degradation phenomenon/phenomena involved, failure threshold(s) and the degradation phenomenon which dominates the failure.

- The mean remaining lifetime of the pavement increases as the failure definition is relaxed.

- Using the results from the multivariate gamma process modelling as input, the condition-based inspection and maintenance optimization reveals that the optimal inspection interval depends on the relative values of cost parameters used in the optimization.

6.   Summary, Conclusions and Recommendations

6.1   Summary and Conclusions

The dissertation presents advanced dependence models for the modelling of spatial and temporal uncertainties in deterioration of civil and nuclear infrastructural components and systems and examines the models' effects on decision making. The thesis answers a few research questions such as how to model spatial uncertainty in degradation, what is the extreme value distribution of random fields and what are the effects of inherent spatial (aleatory) and parameter (epistemic) uncertainty on the extreme value distribution of one- and two-dimensional random fields. The thesis also addresses the question of modelling multivariate dependent degradation phenomena as well as lifetime distribution and condition-based maintenance optimization of such components and systems. Moreover, the proposed models have been applied to deterioration in a feeder pipe and flexible pavement.

Regarding EV distribution of random fields, Ditlevsen had proposed a distribution for one-dimensional standard Gaussian fields within a finite interval with an unknown parameter G. Meanwhile, EV distribution of non-Gaussian fields within a finite region remained unknown. The thesis proposes a Monte Carlo simulation-based methodology to calibrate the parameter G. Moreover, simulation based on mesh refining does not work for EV distribution of homogeneous random fields due to convergence issues. Therefore, a spectrally based simulation technique based on the KL expansion is employed.

To evaluate EV distributions of non-Gaussian random fields in a finite region, use of double transformations of the Gaussian field to any translation fields of interest is demonstrated. Based on the empirically calibrated G, a modified Ditlevsen distribution is proposed to evaluate EV distribution of Gaussian and non-Gaussian fields (with known parameters) without the need for simulation. Random fields studied include Gaussian, lognormal and gamma fields. The EV statistics are a function of size-to-correlation length ratio L/θ for all the random fields studied. The exact Ditlevsen distribution and modified Ditlevsen, which uses an empirical G term, are found to make more accurate prediction of EV distribution of one-dimensional random fields than the approximate Ditlevsen with

zero G. The Poisson approximation performs well in the upper tail of the EV distributions. Regarding the correlation models, the triangular model generally gives greater EV mean than exponential and quadratic exponential models for large random fields.

The copula-based gamma field proposed in this thesis takes into consideration spatial uncertainty in degradation. Following the 'repairing and stitching' procedure, the combined wall profile of the feeder pipe was treated as a two-dimensional gamma field. After parameter estimation, the EV distribution of the maximum wall thickness loss is obtained based on the KL expansion and double transformation of the Gaussian field to gamma field. With the introduction of parameter uncertainty, the resulting EV distribution has a much longer tail than the distribution obtained from the point estimates of the parameters.

The thesis proposes a multivariate gamma process model, which is a variation of the gamma field model, for dependent degradation phenomena in civil infrastructure. It was illustrated with a case study of multiple degradation in a highway pavement section. The thesis shows that modelling degradation phenomena as independent stochastic processes overestimates the remaining lifetime prediction, when compared to the multivariate gamma process model. Generally, the remaining lifetime distribution of any failure scenario depends on the current degradation state, parameters of the degradation phenomena involved, failure thresholds and the degradation phenomenon which dominates the failure. In the context of condition-based maintenance, the optimal inspection interval that minimizes the expected cost over a planning horizon depends largely on the cost parameters.

6.2   Recommendations for Future Study

In this thesis, a gamma field model for capturing spatial uncertainty has been proposed. However, the underlying assumption that the Gaussian copula is adequate to capture the spatial variability in the wall thinning data is not checked in the current study. More research work should be done to address this issue by trying out some other copula functions. Furthermore, more work should be done to extend the 'repairing and stitching'

technique to scan data from multiple inspection outages and to develop a spatiotemporal degradation model. Homogeneity is another strong assumption used in the extreme value analysis. Extension of the results to nonhomogeneous fields requires further research.

By treating the feeder wall profile as a random field, the number of locations is infinite. Therefore, it is practically impossible to inspect every location in the field. Essentially, there is epistemic uncertainty which may be reduced by collecting more data. Thus, the quantification of such uncertainty and its effect on risk-based decisions are important concepts in degradation modelling worth exploring. Therefore, research needs to be done to determine how much inspection data is needed to optimally assess the fitness for service (FFS) of nuclear feeders. This may be otherwise referred to as value of information studies. Another area of future investigation is developing a stochastic finite element-based approach for reliability analysis and probabilistic integrity assessment. This is sequel to the random field-based approach to deterioration modelling proposed in this thesis.

Chapter 5 of the thesis discusses the multivariate gamma process model for competing degradation. In its current form, the model is section-specific and does not account for measurement error. The proposed model can be greatly improved by incorporating measurement error. Furthermore, a more sophisticated mechanistic-empirical version of the multivariate gamma process is worth exploring. Another possible area of future investigation is upscaling the section-specific deterioration model to the pavement network level.

Appendices

Exponential and Triangular Covariance Functions - Scaling

Consider a one-dimensional Gaussian process $Z(\zeta)$ defined on $[-1/2, 1/2]$ with an exponential covariance function given as $C(s_1, s_2) = e^{-c|s_1 - s_2|}$, where $c = 3/\theta$ and $\zeta = s/L$. As shown in Ghanem & Spanos (1991), the analytical solution of the Fredholm integral of the second kind for the exponential covariance function is a pair of transformed transcendental equations:

$$1.5\eta - \Omega \tan(\Omega) = 0; \quad \Omega + 1.5\eta \tan(\Omega) = 0$$

where $\eta = \frac{L}{\theta}$ and $\Omega = \frac{\omega L}{2}$

The eigenfunctions of the two equations are

$$f_n(\zeta) = \frac{\cos(2\Omega_n \zeta)}{\frac{\sqrt{L}}{2}\sqrt{1 + \frac{\sin(2\Omega_n)}{2\Omega_n}}}; \quad f_n^*(\zeta) = \frac{\sin(2\Omega_n^* \zeta)}{\frac{\sqrt{L}}{2}\sqrt{1 - \frac{\sin(2\Omega_n^*)}{2\Omega_n^*}}}$$

and the corresponding eigenvalues are

$$\lambda_n = \frac{1.5\eta L}{\Omega_n^2 + 2.25\eta^2}; \quad \lambda_n^* = \frac{1.5\eta L}{\Omega_n^{*2} + 2.25\eta^2}$$

Finally, the truncated KL expansion for the Gaussian field is a sum of odd and even terms and is given as

$$Z(\zeta, \theta) = \sum_{i=1}^{n}\left[\xi_i \sqrt{\lambda_i} f_i(\zeta) + \xi_i^* \sqrt{\lambda_i^*} f_i^*(\zeta)\right]$$

where ξ and ξ^* are independent standard normal variates. Therefore, the KL expansion is independent of the absolute value of the random field size L.

For the triangular covariance function in section 2.3 defined on the domain $[0, 1]$, the analytical solution is a pair of equations $\Omega_n = \frac{n\pi}{L}$ and $\Omega\left(\frac{2}{\eta} - 1\right)\tan\Omega = 1$ for odd and even n, respectively. The eigenfunctions are

$$f_n(\zeta) = \frac{\cos(2\Omega_n\zeta) + \tan(\Omega_n)\sin(2\Omega_n\zeta)}{\frac{\sqrt{L}}{2}\sqrt{2 + (\tan^2(\Omega_n) - 1)(1 - \frac{\sin(4\Omega_n)}{4\Omega_n}) + \frac{\sin^2(2\Omega_n)}{\Omega_n}\tan(\Omega_n)}}; \quad f_n^*(\zeta) = \frac{\cos(\omega_n^* s)}{\frac{\sqrt{L}}{2}\sqrt{1 + \frac{\sin(4\Omega_n^*)}{2\Omega_n^*}}}$$

for odd and even n, respectively, and the eigenvalues are $\lambda_n = \frac{\eta L}{2\Omega_n^2}$; $\quad \lambda_n^* = \frac{\eta L}{2\Omega_n^{*2}}$. These

results are well established in the literature of stochastic field.

Additional Plots: Comparison of analytical and numerical eigenvalues.
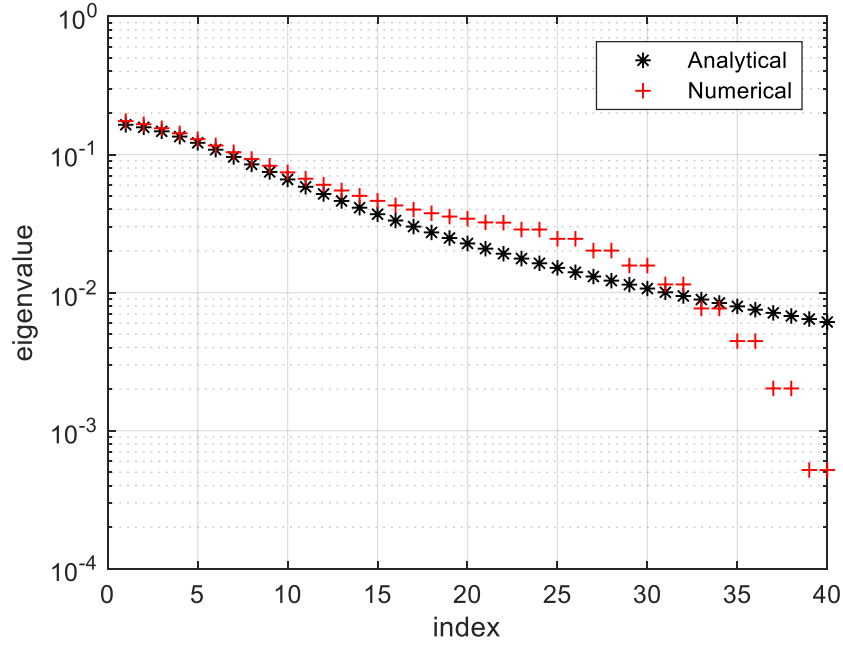
(a)
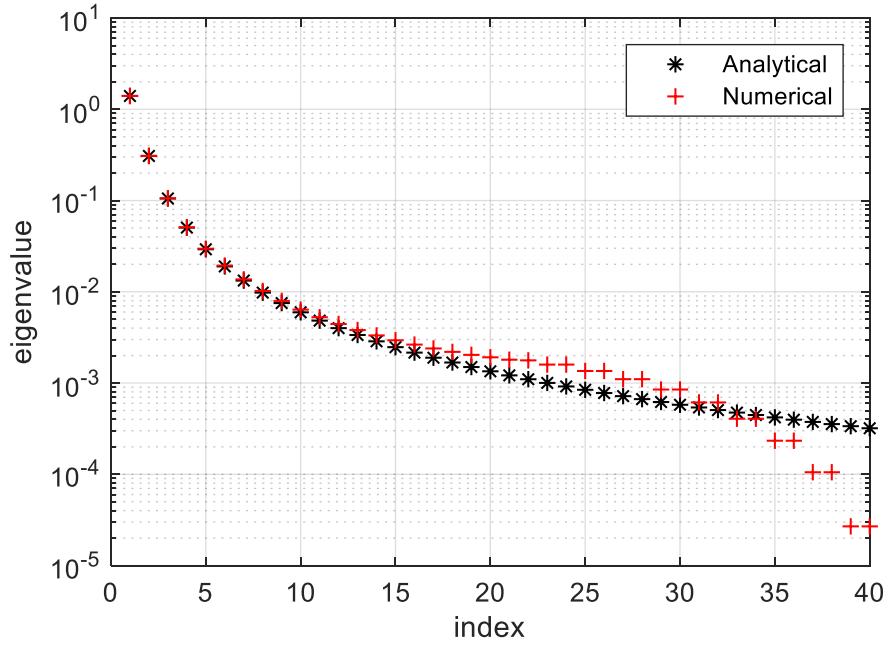
(b)
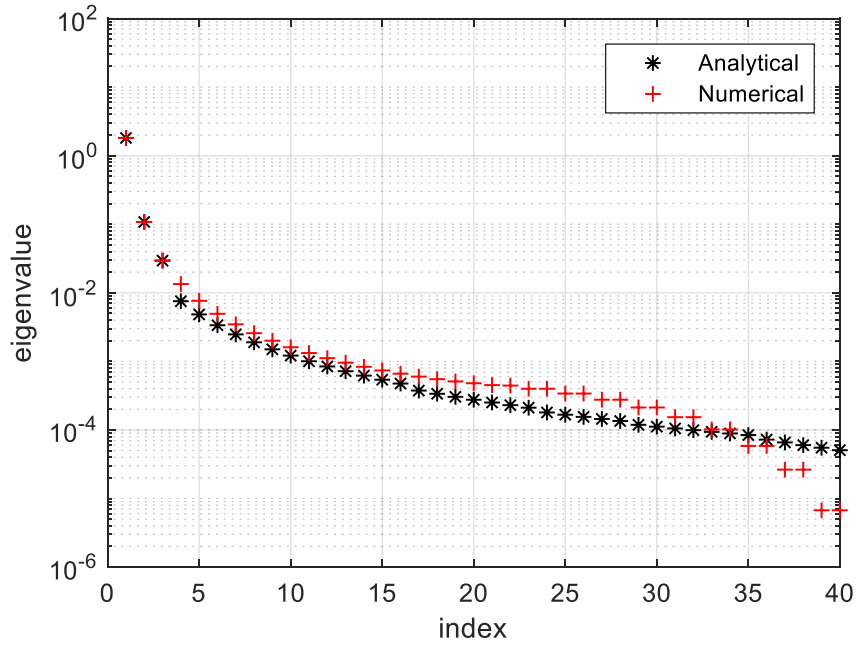
(c)

Figure. Comparison of first 40 eigenvalues obtained from exact analytical solution and numerical solution (Nystrom method). Exponential correlation function $\rho(h) = \exp\left(-\frac{3h}{\theta}\right)$, θ ((a) $\theta = 0.25$, (b) $\theta = 5$, (c) $\theta = 20$). Domain [-1, 1].

Estimator

This section demonstrates few of the properties of maximum likelihood estimate. To test the accuracy of the code used to estimate the parameters in section 4.2, true values of the parameters of a 2D gamma field are specified as $\alpha = 10$, $\beta = 0.5$, $\theta_x = 40$ and $\theta_y = 30$ while the random field size was 230 by 188. The estimation was done a total of 5000 times and their biases calculated as -0.1588, 0.0023, 0.0077 and 0.0053, respectively.
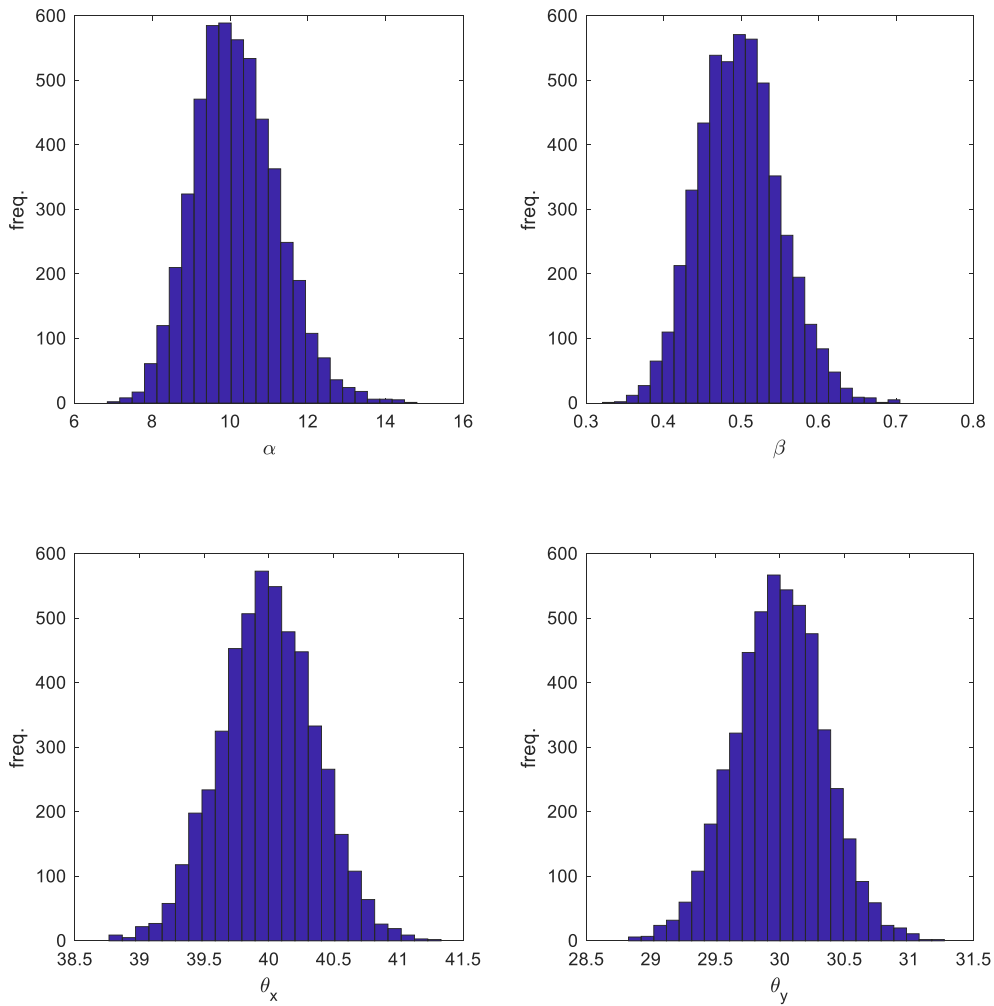
Validation of two-dimensional KL expansion

The accuracy of the extreme value distribution of one-dimensional random fields based on the Karhunen-Loeve expansion has been demonstrated by comparing the distributions with Ditlevsen and approximate Ditlevsen distributions in Section 3.6 – 3.8. As the Ditlevsen distribution does not exist in two or higher dimensions, this section aims to validate extreme values of two-dimensional random fields, based on the two-dimensional KL expansion, by comparing the results with the already validated extreme values of one-dimensional random fields.

In the figure below, the parameters of the one-dimension gamma field are $\alpha = 30$, $\beta = 0.1$ and $\theta_x = 40$. Same parameters are retained for 2D gamma fields of 200 x150, while the correlation length in the other direction θ_y is gradually increased from 20 to 1000. Using the KL expansions in one and two dimensions to generate 100 000 realizations of gamma fields for each scenario, the maximum values are recorded and their mean taken. The results show that at very high values of θ_y, extreme values of two-dimensional fields approach that of the one-dimensional field.

$$Z(s; \omega) = \sum_{i=1}^{n} \sqrt{\lambda_i} \zeta_i(\theta) f_i(s)$$

$$Z(\boldsymbol{s}, \boldsymbol{t}) = \sum_{j=1}^{n} \sum_{i=1}^{n} \sqrt{\lambda_i \lambda_j} \zeta_{ij}(\theta) f_i(\boldsymbol{s}) f_j(\boldsymbol{t})$$
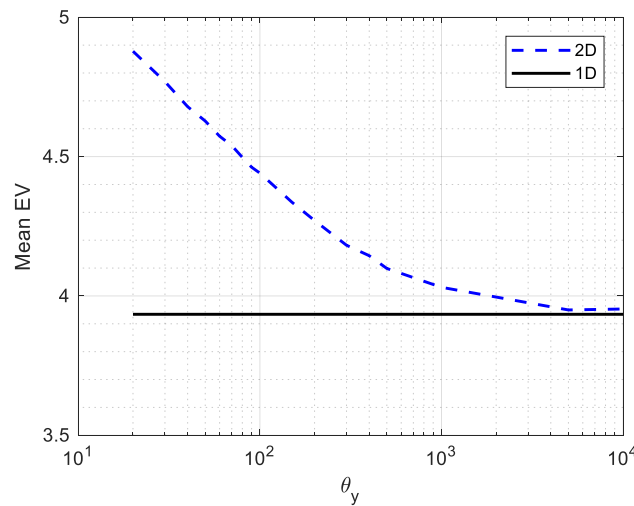
References

Adegbola, A., Yuan, X.-X., Wang, M. (2015). Random Field Modelling for the Prediction of Wall Thickness of Nuclear Pipes Considering Data Misalignment, *12th International Conference on Applications of Statistics and Probability in Civil Engineering*, July 12-15, 2015, Vancouver, BC, Canada

Adler, R., Taylor, J. (2007). Random fields and geometry. Springer, New York, NY, USA.

Ahmad, R., Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. Computers and Industrial Engineering, 63: 135-149.

Alaswad, S., Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliability Engineering and System Safety, 157: 54-63.

Aryai, V., Mahmoodian, M. (2017). Spatial-temporal reliability analysis of corroding cast iron water pipes. Engineering Failure Analysis, 82: 179-189.

Atkinson, K. (1997). The numerical solution of integral equations of the second kind. Cambridge University Press, Cambridge, UK.

Azuri, G., Jódar, J., Carrera, J., Gupta, H. (2013). Stochastic simulation of nonstationary rainfall fields, accounting for seasonality and atmospheric circulation pattern evolution. International Association for Mathematical Geosciences, 45, 621–645.

Balakrishnan, N., Lai, C. (2009). Continuous bivariate distributions, 2nd ed. Springer.

Betz, W., Papaioannou, I., Straub, D. (2014). Numerical methods for the discretization of random fields by means of the Karhunen-Loeve expansion. Comput. Methods Appl. Mech. Engrg., 271: 109 – 129.

Caballé, N. C., Castro, I. T., Pérez, C. J., Lanza-Gutiérrez, J. M. (2015). A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. Reliability Engineering and System Safety, 134: 98-109.

Canadian Infrastructure Report Card (Organization). (2016). Informing the future: The Canadian infrastructure report card. Retrieved from canadainfrastructure.ca/ /downloads/Canadian_Infrastructure_Report_2016.pdf.

Canadian Standards Association (2005). Periodic inspection of CANDU nuclear power plant components, N285.4-05.

Castro, I. T., Caballé, N. C., Pérez, C. J. (2015). A condition-based maintenance for a system subject to multiple degradation processes and external shocks. International Journal of Systems Science, 46(9): 1692-1704

Ching, J., Lin, C. (2014). Probability distribution for mobilized shear strengths of saturated undrained clays modeled by 2-D stationary Gaussian random field - A 1-D stochastic process view. Journal of Mechanics, 30(3): 229.

Chung, H.S (2010). A review of CANDU feeder wall thinning. Nuclear Engineering and Technology, 42(5): 568-575.

Cressie, N. A. C. (1993). Statistics for Spatial Data. Revised edition. John Wiley & Sons, Hoboken, New Jersey.

de Haan, L.; A. Ferreira (2006). Extreme Value Theory: An Introduction. New York, Springer.

Delves, L., Mohamed, J. (1997). Computational methods for integral equations. Cambridge University Press, Cambridge, UK.

Denis, A., Cremoux, F. (2002). Using the entropy of curves to segment a time or spatial series. Mathematical Geology, 34(8), 899-914.

Ditlevsen, O. (1966). Extremes of realizations of continuous time stationary stochastic processes on closed intervals. Journal of Mathematical Analysis and Applications, 14: 463 – 474.

Ditlevsen, O. (2004). Extremes of random fields over arbitrary domains with application to concrete rupture stresses. Probabilistic Engineering Mechanics, 19(4), 373-384.

Duan, X., Kozluk, M., Li, M. (2009). Comprehensive integrity assessment of carbon steel feeder pipes/elbows subject to wall thinning. Proceedings of Pressure Vessels and Piping Division Conference, Prague, Czech Republic.

Duan, F., Wang, G., & Wang, H. (2018). Inverse gaussian process models for bivariate degradation analysis: A bayesian perspective. *Communications in Statistics - Simulation and Computation, 47*(1), 166-186.

Elsayed, E., Liao, H. (2004). A geometric Brownian motion model for field degradation data. International Journal of Materials and Product Technology, 20(1-2):51-72.

Geman, S., Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6(6): 721–741.

Genest, C., Rivest, L. (1993). Statistical inference procedures for bivariate Archimedean copulas. J. Am. Stat. Assoc., 88:1034–1043.

Gill, V., Lawson, J. (2013). Where the rubber meets the road: How much motorists pay for road infrastructure. The Conference Board of Canada. Retrieved from https://www.conferenceboard.ca/temp/f4442f03-45e6-483c-bf07-b8dbf7eaa07b/14-083_MotoristCostRecovOnt_RPT.pdf.

Giorgio, M., Guida, M., Pulcini, G. (2011). An age- and state-dependent Markov model for degradation processes. IIE Transactions, 43(9): 621 – 632.

Guida, M., Pulcini, G. (2011). A continuous-state Markov model for age- and state-dependent degradation processes. Structural Safety 33(6): 354–366.

Guillaumot, V., Durango-Cohen, P., Madanat, S. (2003). Adaptive optimization of infrastructure maintenance and inspection decisions under performance model uncertainty. Journal of Infrastructure Systems, 9(4): 133-139.

Green, P.J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika, 82: 711-732.

Grigoriu, M. (1984). Crossings of non-Gaussian translation processes. Journal of Engineering Mechanics, 110(4), 610-620.

Haas, T. (1990). Kriging and automated variogram modeling within a moving window. Atmospheric Environment, 24(7): 1759-1769.

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1): 97-109.

Hao, H., & Su, C. (2014). Bivariate nonlinear diffusion degradation process modeling via copula and MCMC. *Mathematical Problems in Engineering, 2014*, 1-11.

Harris, P., Charlton, M., Fotheringham, A. (2010). Moving window kriging with geographically weighted variograms. Stohastic Environmental Research and Risk Assessment, 24(8): 1193-1209.

Hu, Y., Li, H., Shi, P., Chai, Z., Wang, K., Xie, X., & Chen, Z. (2018). A prediction method for the real-time remaining useful life of wind turbine bearings based on the wiener process. *Renewable Energy, 127*: 452-460.

Isaaks, E., Srivastava, R (1989). Applied Geostatistics. Oxford University Press. *New York.*

Jin, J., Awad, R. (2011). Fitness for service assessment of degraded CANDU feeder piping-Canadian regulatory expectations. Journal of Nuclear Engineering and Design, 241(3):644-647.

Jin, Y., Mukherjee, A. (2014). Markov chain applications in modelling facility condition deterioration. Int. J. Critical Infrastructures, 10(2): 93-112.

Joe, H. (2001). Multivariate models and dependence concepts. Monographs on Statistics and Applied Probability, 73. Chapman & Hall/CRC.

Khorasgani, H., Biswas, G., & Sankararaman, S. (2016). Methodologies for system-level remaining useful life prediction. *Reliability Engineering and System Safety, 154*, 8-18.

Kibble, W. (1941). A Two-Variate Gamma Type Distribution. Sankhyā: The Indian Journal of Statistics, 5(2): 137-150.

Kotz S., Balakrishnan N., Johnson N. (2000). Continuous multivariate distributions, Vol. 1: Models and Applications. John Wiley & Sons.

Lawless, J., Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Analysis, 10:213-227.

Le Son, K., Fouladirad, M., Barros, A., Levrat, E., & Lung, B. (2013). Remaining useful life estimation based on stochastic deterioration models: A comparative study. Reliability Engineering and System Safety, 112: 165-175.

Leadbetter, M., Lindgren, G., Rootzen, H. (1983). Extremes and related properties of random sequences and processes. Springer-Verlag, New York, NY.

Lee, Y. S., Lee, S. H., Hwang, K. M. (2016). Cause analysis of flow accelerated corrosion and erosion-corrosion cases in Korea nuclear power plants. Corrosion Science and Technology, 15(4): 182-188.

Li, X., Xue, P. (2014). Multivariate storage degradation modeling based on copula function. Advances in Mechanical Engineering, 2014:1-12.

Liu, Z., Ma, X., Yang, J., Zhao, Y. (2014). Reliability modeling for systems with multiple degradation processes using inverse Gaussian process and copulas. Mathematical Problems in Engineering.

Melink, T., Korelc, J. (2014). Stability of Karhunen-Loeve expansion for the simulation of Gaussian stochastic fields using Galerkin scheme. Probabilistic Engineering Mechanics, 37: 7-15.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. (1953). Equation of state calculations by fast computing machines. J. Chem. Phys., 21(6): 1087-1092.

Ministry of Transportation Ontario (Organization), (2013). Pavement design and rehabilitation manual. 2nd ed. Materials Engineering and Research Office, Ministry of Transportation Ontario. Retrieved from http://www.bv.transports.gouv.qc.ca/mono/1165561.pdf.

Moran, P. (1969). Statistical inference with bivariate gamma distributions. Biometrika, 56(3): 627-634.

Mori, N., Yasuda, T. (2002). A weakly non-Gaussian model of wave height distribution for random wave train. Ocean Engineering, 29(10): 1219-1231.

MTO (2012). Ontario's Default Parameters for AASHTOWare Pavement ME Design. Interim Report. Ministry of Transportation Ontario, Pavement and Foundations Section, Materials Engineering and Research Office, Downsview, Ontario, Canada.

Na, U. J., Kwon, S. J., Ray Chaudhuri, S., Shinozuka, M. (2012). Stochastic model for service life prediction of RC structures exposed to carbonation using random field simulation. KSCE Journal of Civil Engineering, 16(1): 133-143.

Na, U. J., Ray Chaudhuri, S., Shinozuka, M. (2009). Effects of spatial variation of soil properties on seismic performance of port structures. Soil Dynamics and Earthquake Engineering, 29(3), 537-545.

Nelsen, R. (2006). An Introduction to Copulas, 2nd ed. Springer series in statistics, New York, NY.

Nystad, B., Gola, G., Hulsund, J. (2012). Lifetime models for remaining useful life estimation with randomly distributed failure thresholds. First European conference of the prognostics and health management society, vol. 3.

Onoufriou, T., Frangopol, D. (2002). Reliability-based inspection optimization of complex structures: a brief retrospective. Computers and Structures, 80(12): 1133–1144.

Oumouni, M., Schoefs, F., Castanier, B. (2019). Modeling time and spatial variability of degradation through gamma processes for structural reliability assessment. Structural Safety, 76: 162-173.

Pan, Z., Balakrishnan, N. (2011). Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. Reliability Engineering and System Safety, 96(8): 949–957.

Pan, Z., Balakrishnan, N., Sun, Q., & Zhou, J. (2013). Bivariate degradation analysis of products based on wiener processes and copulas. *Journal of Statistical Computation and Simulation, 83*(7), 1316-1329.

Pan, Z., Feng, J., & Sun, Q. (2016). Lifetime distribution and associated inference of systems with multiple degradation measurements based on gamma processes. Eksploatacja I Niezawodnosc - Maintenance and Reliability 18(2): 307-313.

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

Park, C., Padgett, W. (2005). Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime data analysis, 11(4): 511 – 527.

Park, K.S. (1988). Optimal continuous-wear limit replacement under periodic inspections. IEEE Transactions on Reliability, 37(1): 97-102.

Peng, C. (2015). Inverse gaussian processes with random effects and explanatory variables for degradation data. *Technometrics, 57*(1): 100-111.

Peng, L., Stewart, M. G. (2014). Spatial-time dependent reliability analysis of corrosion damage to RC structures with climate change. Magazine of Concrete Research, 66(22): 1154-1169.

Peng, W., Li, Y., Mi, J., Yu, L., & Huang, H. (2016). Reliability of complex systems under dynamic conditions: A bayesian multivariate degradation perspective. *Reliability Engineering and System Safety, 153*: 75-87.

Peng, W., Li, Y., Yang, Y., Huang, H., Zuo, M. (2014). Inverse Gaussian process models for degradation analysis: A Bayesian perspective. Reliability Engineering & System Safety, 130: 175 – 189.

Peng, W., Li, Y., Yang, Y., Zhu, S., & Huang, H. (2016). Bivariate analysis of incomplete degradation observations based on inverse gaussian processes and copulas. *IEEE Transactions on Reliability, 65*(2): 624-639.

Possan, E., Andradec, J. (2014). Markov chains and reliability analysis for reinforced concrete structure service life. Materials Research, 17(3): 593-602.

Phoon, K., Huang, S., Quek, S. (2002). Implementation of Karhunen-Loeve expansion for simulation using a wavelet-Galerkin scheme. Probabilistic Engineering Mechanics, 17: 293-303.

Podofillini, L., Zio, E., Vatn, J. (2006). Risk-informed optimisation of railway tracks inspection and maintenance procedures. Reliability Engineering and System Safety, 91(1): 20–35.

Press, W., Teukolsky, S., Vetterling, W., Flannery, B. (2007). Numerical recipes. The art of scientific computing. Cambridge University Press, 3 ed., Cambridge, UK.

Rackwitz, R., Joanni, A. (2009). Risk acceptance and maintenance optimization of aging civil engineering infrastructures. Structural Safety, 31(3): 251–259.

Rangel-Ramírez, J., Sørensen, J. (2012). Risk-based inspection planning optimisation of offshore wind turbines. Structure and Infrastructure Engineering, 8(5): 473-481.

Rice, S.O. (1945). Mathematical analysis of random noise. The Bell System Technical Journal, 24(1): 46 - 156.

Rodríguez-Picón, L., Flores-Ochoa, V., Méndez-González, L., Rodríguez-Medina, M. (2017). Bivariate degradation modelling with marginal heterogeneous stochastic processes. Journal of Statistical Computation and Simulation, 87(11): 2207.

Rodríguez-Picón, L. A., Flores-Ochoa, V. H., Méndez-González, L. C., & Rodríguez-Medina, M. A. (2017). Bivariate degradation modelling with marginal heterogeneous stochastic processes. *Journal of Statistical Computation and Simulation, 87*(11): 2207-2226.

Sari, J. J., Newby, M. M., Brombacher, A. A., & Tang, L. (2009). Bivariate constant stress degradation model: LED lighting system reliability estimation with two-stage modelling. *Quality and Reliability Engineering International, 25*(8), 1067-1084.

Shafei, B., Alipour, A. (2015). Application of large-scale non-Gaussian stochastic fields for the study of corrosion-induced structural deterioration. Engineering Structures, 88: 262-276.

Sharabah, A., Setunge, S., Zeephongsekul, P. (2006). Use of Markov chain for deterioration modeling and risk management of infrastructure assets. International Conference on Information and Automation, 384 - 389

Shemehsavar S. (2014). A bivariate gamma model for a latent degradation process. Communications in Statistics – Theory and Methods, 43(9): 1924-1938.

Sklar, A. (1973). Random Variables, Joint Distribution Functions and Copulas. Kybernetika 9: 449–460.

Slade, J., Gendron, T. (2005). Flow accelerated corrosion and cracking of carbon steel piping in primary water – Operating experience at the Point Lepreau Generating Station. Proceedings of 12th International Conference on Environmental Degradation of Materials in Nuclear Power System – Water Reactors, Salt Lake City, 773-784.

Stein, M. (1999). Interpolation of Spatial Data: some theory for kriging. Springer, New York.

Stewart, M. G., Mullard, J. A., Drake, B. J., Al-Harthy, A. S. (2007). Utility of spatially variable damage performance indicators for improved safety and maintenance decisions of deteriorating infrastructure. Civil Engineering and Environmental Systems, 24(2): 149–163.

Sudret, B. (2008). Probabilistic models for the extent of damage in degrading reinforced concrete structures. Reliability Engineering and System Safety, 93(3): 410-422.

Sudret, B., Der Kiureghian, A. (2000). Stochastic finite element methods and reliability : a state-of-the-art report. Department of civil and environmental engineering, University of California, Berkeley.

Tsai, C., Tseng, S., Balakrishnan, N. (2012). Optimal design for degradation tests based on gamma processes with random effects. IEEE Transactions on Reliability, 61(2): 604 – 613.

Tseng, S.T., Tang, J., & Ku, L.H. (2003). Determination of optimal burn-in parameters and residual life for highly reliable products. Naval Research Logistics 50(1): 1-14.

van Noortwijk, J. (2007). A survey of the application of gamma processes in maintenance. Reliability Engineering and System Safety, 94:2-21.

Vanmarcke, E. (2010). Random fields analysis and synthesis. Rev. and expanded new ed. World Scientific, New Jersey.

Wang, L. (2008). Karhunen-Loeve expansions and their applications. PhD dissertation, the London School of Economics and Political Science.

Wang, X., Balakrishnan, N., Guo, B., Jiang, P. (2015). Residual life estimation based on bivariate non-stationary gamma degradation process. Statistical Computation and Simulation, 85(2): 405-421.

Wang, X., Xu, D. (2010). An inverse Gaussian process model for degradation data. Technometrics, 52(2): 188 – 197.

Wei, Q., Xu, D. (2014). Remaining useful life estimation based on gamma process considered with measurement error. 10th International Conference on Reliability, Maintainability and Safety. IEEE: 645 - 649.

Whitmore, G. A., Crowder, M. J., and Lawless, J. F. (1998). Failure inference from a marker process based on a bivariate Wiener model. *Lifetime Data Analysis*, 4: 229-251.

Webster, L. F. (1997). The Wiley dictionary of civil engineering and construction. John Wiley & Sons.

Wu, S.H., Ou, C.Y., Ching, J.Y., Juang, C.H. (2012). Reliability-based design for basal heave stability of deep excavations in spatially varying soils. Journal of Geotechnical and Geoenvironmental Engineering, 138(5), 594-603.

Xu, A., Shen, L., Wang, B., & Tang, Y. (2018). On modeling bivariate wiener degradation process. *IEEE Transactions on Reliability, 67*(3), 897-906.

Xu, D., Wei, Q., Elsayed, E. A., Chen, Y., & Kang, R. (2017). Multivariate degradation modeling of smart electricity meter with multiple performance characteristics via vine copulas: Multivariate degradation modeling of SEM via vine copulas. *Quality and Reliability Engineering International, 33*(4): 803-821.

Xu, J., Xu, D., Xing, M., Wei, Q., Qin, Y., Chen, Y., & Kang, R. (2018). Failure behavior modeling and reliability estimation of product based on vine-copula and accelerated degradation data. *Mechanical Systems and Signal Processing, 113*: 50-64.

Ye, Z., Chen, N. (2014). The inverse Gaussian process as a degradation model. Technometrics, 56(3):302-311.

Ye, Z., Chen, L., Tang, L., Xie, M. (2014). Accelerated degradation test planning using the inverse Gaussian process. IEEE Transactions on Reliability, 63(2).

Yu, S.Z. (2009). Hidden semi-Markov models. Artificial Intelligence 174 (2), 215-243.

Yu, Y., Ma, L., Gu, Y., Zhou, Y. (2008). Confidence interval of lifetime distribution using bootstrap method. Third World Congress on Engineering Asset Management and Intelligent Maintenance Systems Conference, Beijing, China.

Yuan, X.-X. (2007). Stochastic modeling of deterioration in nuclear power plant components. PhD dissertation, University of Waterloo.

Yuan, X.-X., Lee, W., Li, N. (2017). Ontario's local calibration of the MEPDG distress and performance models for roads: A summary. Innovations in Pavement, Management, Engineering and Technology Session of the 2017 Conference of the Transportation Association of Canada. St. John's, NL.

Yuan, X.-X., Pandey, M. D. (2009). A nonlinear mixed-effects model for degradation data obtained from in-service inspection. *Reliability Engineering & System Safety*, 94: 509-519.

Zhang, L., & Singh, V. P. (2006). Bivariate flood frequency analysis using the copula method. Journal of Hydrologic Engineering, 11(2), 150-164.

Zhang, S. (2014). Development of probabilistic corrosion growth models with applications in integrity management of pipelines. PhD Dissertation, the University of Western Ontario.

Zhang, S., Zhou, W. (2014). An Efficient Methodology for the Reliability Analysis of Corroding Pipelines. Pressure Vessel Technology, 136(4): 041701-7.

Zhao, L., Yun, G. J. (2018). Probabilistic service life of reinforced concrete structures with randomly distributed corrosion-induced cracking. Structure and Infrastructure Engineering, 14(3): 333-347.

Zhou, W., Hong, H., Zhang, S. (2012). Impact of dependent stochastic defect growth on system reliability of corroding pipelines. International Journal of Pressure Vessels and Piping, 96-97:68-77.

Zhou, Z.J., Hu, C.H., Xu, D.L., Chen, M.Y., & Zhou, D.H. (2010). A model for real-time failure prognosis based on hidden Markov model and belief rule base. European Journal of Operational Research 207: 269-283.