Maintenance Policies to Guarantee Optimal Performance of Stochastically Deteriorating Multi-Component Systems

by

Colin Thomas Barker

A thesis submitted for the degree of
Doctor of Philosophy
in the subject of

Applied Probability and Statistics

Centre for Risk Management, Reliability and Maintenance
School of Engineering and Mathematical Sciences
City University
London

October 2006

Contents

List of Tables

List of Figures

Acknowledgments

Many people have contributed in their own way to the completion of this thesis. It is with great pleasure and a lot of respect that I seise the opportunity to thank them now.

First of all, I wish to express my sincere gratitude to my supervisor Professor Martin Newby, without whom this work could certainly not have been achieved. His invaluable comments, availability and kindness have undoubtedly motivated me throughout my research. His presence both as a supervisor but also as a friend will not be forgotten.

I am deeply grateful to my girlfriend Lidwine Therrey, whose choice and willingness to accompany me to London will always stay in my heart. Her constant encouragements, joy and love throughout the years definitely played a precious role in the realization of this work.

My deepest gratitude and thanks to my brother Julien Barker, my mother Danie Granier, my father Ken Barker, my grandparents Jeanette et Fabien Granier and my friend Robin Smyth for their love and happiness that accompanies me through my life. My entire studies could certainly not have been done without their constant support.

I am sincerely grateful to Mathieu Jurgawczynski for contributing to my interest in mathematics and motivating me to undertake a Ph.D.. His presence and musical enthusiasm surely take an important place in my life.

The time spent on my research could not have been appreciated as such without the friendly environment that surrounded me. I particularly want to thank Michelle Hickey, Daniel Nankoo, Panagiotis Rentzos, Ibrahim Elgayar, Evangelos Sagianos, Christos Papanagnou, Giannis Kiskiras, Dimitrios Polyzos and Christos Mermigis for all the moments and conversations I was privileged to share with them.

Sans oublier tous mes amis de France et d'Angleterre pour leur précieuse présence au fil des ans.

"I grant the powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgment".

Thesis supervisor
Martin Newby

Author
Colin Thomas Barker

Maintenance Policies to Guarantee Optimal Performance of Stochastically Deteriorating Multi-Component Systems

Abstract

Guaranteeing a prescribed level of reliability for a complex multi-component system is the problem considered in this thesis. For this, optimal maintenance and inspection policies are derived, taking into account the different deteriorations the components in the system are subject to. These degradations are described with the use of continuous stochastic processes in time and are assumed not to be directly observable. Maintenance decisions are based on a performance measure defined by a functional acting on the system state process. The transient property of the performance measure enables a modified control limit rule, based on last exit times, to be considered. For this a critical level is defined and the probability of never returning below it is used in the decision making approach. A condition based maintenance policy is investigated with the use of a bijective function of the performance measure's value, that determines the required amount of repair. Both periodic and non-periodic inspections are studied. The non-periodic approach is handled with the use of an inspection scheduling function which assigns the amount of time between two consecutive inspections to the value of the performance measure at inspection. Two main types of models are proposed. The first type considers one threshold and focuses in guaranteeing a prescribed level of reliability for systems where crossing of a critical boundary does not cause immediate failure but will require action later. Examples include physical infrastructures such as roads. Models of the second type take failure of the system into account with the incorporation of a second threshold. Examples include aeroplanes where safety regulations imply regular inspections and repairs. Occurrence of unfrequent catastrophic failures must however be considered.

Notations

Stochastic processes

- $B_t^{(i)}$, $i \in \{1, 2, \ldots, n\}$: standard Brownian motions
- $W_t^{(i)}$, $i \in \{1, 2, \ldots, n\}$: Brownian motions with drift and volatility terms describing the state of deterioration of the un-maintained component C_i
- μ_i, $i \in \{1, 2, \ldots, n\}$: drift term for $W_t^{(i)}$
- σ_i, $i \in \{1, 2, \ldots, n\}$: volatility term for $W_t^{(i)}$
- $\mathbf{W}_t$: n-dimensional Wiener process describing the state of the un-maintained system
- R_t : stochastic process describing a performance measure of the considered un-maintained system (refers to a Bessel process with drift)
- R_t^* : stochastic process describing a performance measure of the considered maintained system
- $G_{\mathcal{F}}^x$: first hitting time of threshold $\mathcal{F}$ by the process R_t given that it initially started from state $R_0 = x$
- H_ξ^x : last exit time from the interval $[0, \xi)$ for the process R_t given that it initially started from state $R_0 = x$
- f_τ^x : transition density function for R_t from state x after a time interval of length τ
- h_ξ^x : probability density function for H_ξ^x
- $g_{\mathcal{F}}^x$: probability density function for $G_{\mathcal{F}}^x$

- $\widetilde{g_{\mathcal{F}}^x}$: Laplace transform of $g_{\mathcal{F}}^x$

Cost models

General notation

- ξ : critical threshold for the one-threshold models, repair threshold for the two-threshold models
- $\mathcal{F}$: failure threshold in the two-threshold models
- C_i : cost of inspection of the system
- C_f : cost of failure of the system
- C_{rep} : cost of repair of the system
- C_r : cost of repair function
- d : maintenance function
- k : parameter describing the amount of maintenance undergone on the system
- x_i, y_j, $i, j \in \{1, 2, \ldots, N\}$: points used for the Gauss-Legendre rule
- w_i, $i \in \{1, 2, \ldots, N\}$: weights of the Gauss-Legendre rule
- Π : inspection policy

Periodic inspections

- τ : period of inspection
- τ^* : optimal period of inspection resulting in a minimum expected cost of maintenance per unit time

- *One-threshold model*
  - $V_\tau(x)$: cost of inspection and maintenance per cycle with period of inspection τ, given that the value of the considered critical threshold is $\xi - x$
  - $v_\tau(x)$: expected cost of inspection and maintenance per cycle with period of inspection τ, given that the value of the considered critical threshold is $\xi - x$
  - $L_\tau(x)$: length of a cycle with period of inspection τ, given that the value of the considered critical threshold is $\xi - x$
  - $l_\tau(x)$: expected length of a cycle with period of inspection τ, given that the value of the considered critical threshold is $\xi - x$
  - $C_\tau(x)$: expected cost of maintenance per unit time, with period of inspection τ, given that the value of the considered critical threshold is $\xi - x$
- *Two-threshold model*
  - V_τ^x: cost of inspection and maintenance per cycle with period of inspection τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$
  - v_τ^x: expected cost of inspection and maintenance per cycle with period of inspection τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$
  - L_τ^x: length of a cycle with period of inspection τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$
  - l_τ^x: expected length of a cycle with period of inspection τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$
  - C_τ^x: expected cost of maintenance per unit time, with period of inspection τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$

Non periodic inspections

- τ_i, $i \in \mathbb{N}$: times at which the system is inspected
- τ_i^+, $i \in \mathbb{N}$: times after maintenance of the system
- τ_i^*, $i \in \mathbb{N}$: times at which the system is replaced (renewal points)

- T_i, $i \in \mathbb{N}$: amount of time between inspections at times τ_{i-1} and τ_i
- m: inspection scheduling function determining the amount of time until the next inspection
- a, b: parameters of the inspection scheduling function m
- *One-threshold model*
  - $V_{\xi-x}$: total cost of maintenance, given that the value of the considered critical threshold is $\xi - x$
  - $v_{\xi-x}$: expected total cost of maintenance, given that the value of the considered critical threshold is $\xi - x$
- *Two-threshold model*
  - V^x: total cost of maintenance, given that at inspection time (prior to any maintenance action) $R_\tau = x$
  - v^x: expected total cost of maintenance, given that at inspection time (prior to any maintenance action) $R_\tau = x$

Mathematical notation

- $\mathbf{1}_{\{.\}}$: indicator function
- $\mathbb{P}$, $\mathbb{E}$: probability and expectation symbols
- I_ν: modified Bessel function of the first kind
- K_ν: modified Bessel function of the second kind

Other

- C_i, $i \in \{1, 2, \ldots, n\}$: components of the system
- K_i, $i \in \{1, 2, \ldots, p\}$: classes in which the components in the system fall into

to my family:
Julien, Danie, Ken,
Jeannette, Fabien,
Peggy, Stan
& Betty.

Chapter 1

Introduction

Motivations

Reliability plays a key role in all technological systems. It determines the success of investment in plant and machinery, and the level of customer satisfaction. When tackled with efficiency, it has been shown that considerable improvements in both system safety and maintenance expenses may be achieved. However, the increasing complexity of the systems often requires adaptiveness of the techniques considered. In the case of a highly reliable system for instance, its actual failure may only rarely be observed making models based on systems' lifetime distributions inappropriate. Focusing on the actual system's deterioration has proved to be an efficient way to handle the matter, see for instance [23], [46], [55], [62], [86]. The deterioration through time is usually described with the use of an appropriate stochastic process and interests are based on the time at which the process first reaches a particular value, called a critical threshold. These times usually define times at which the system must be maintained (repaired or replaced) and such a maintenance policy is often referred to as a "condition based maintenance policy with a control limit rule". Such a method usually involves one stochastic process describing the deterioration of the whole system and taking into account the degradations of all the components present does not seem to have been considered yet. The study of multi-component systems usually involves the use of structure functions, [12]. The approach consists in deriving a system's reliability from its components' reliability and requires knowledge on the structure of the system. This appealing technique has its limit when the size of the system becomes too large, leading to extremely complex or even untractable expressions for the structure functions. Hence, the study of

multi-component systems usually refers to series/parallel systems, [14], [78] [83], [93]. The apparent lack of models considering both condition based maintenance policies and complex multi-component systems set the initial motivation for the research proposed in the thesis. A complex multi-component system refers to a system consisting of several types of components, where each component endures its own type of degradation. The subjacent idea is to propose a condition based maintenance policy for a system, taking into consideration the different deteriorations. For this, the system's state is described by a multi-variate process made out of processes describing the different components' degradations. Whereas many proposed models are based on the assumption that the components' degradations are directly observable, this is often not the case in real life situations. Information on the system's state is generally gained at inspection times by evaluating a performance measure. Common examples include modern cars, these use extensive condition monitoring and built in testing: the On-Board Diagnostics system (OBD). When a car is serviced, a repair technician has access to the state of health information for various vehicle sub-systems with the use of the OBD. If the usage of the car has been exceptionally heavy a repair is undertaken.
'... *all major functions of an automobile are controlled, and monitored, by computers on their own networks. If something goes wrong with the car, the computer will know and record a fault code long before a light comes on on the dashboard (the automotive industry calls that a MIL, or Malfunction Indicator Light) alerting us that something is wrong. Some of those problems are simple and will be taken care of next time you go in for a service. Others are not and, if unchecked, can damage or destroy the engine*', [17].
An approach proposed in the literature to deal with this kind of situation, is the use of a functional $\mathcal{A}$ acting on the state process, [38], [60], [61], [82]. This new process represents the observable performance measure and decisions related to the system's maintenance are made on the basis of its value at inspection time. This technique has also been considered as an alternative to deal with non-monotone processes by looking at the maximum of the process, [61]. Rather than forcing the performance measure to have monotone trajectories, other important characteristics of the chosen performance process are exploited here. The Bessel process is the process considered to describe the system's performance measure and corresponds, under certain assumptions stated in Chapter 3, to the Euclidean norm of a multi-variate Wiener process. As mentioned above, the general control limit rule focuses on first hitting times of particular thresholds

by a process, generally monotone. The reason for this being that first hitting times are stopping times: knowledge of the process' past history only is required to be informed of their occurrence with certainty. The fact that Bessel processes have non-monotone trajectories motivated the elaboration of a different control limit rule. It is the transient property of such processes that is exploited and rather than concentrating on first hitting times only, the method proposes the use of last exit times from intervals. Last exit times are non-stopping times as knowledge of their occurrence requires knowledge of the future history of the process: the process can cross a threshold and return below it in the future (this may be seen as various intensities of use of the system). The extra complexity added to the models with the use of non-stopping times implies working with the probability of occurrence of such events.

A more general approach towards system maintenance was also wished to be included in the proposed models. Indeed, a considerable amount of models only consider perfect repairs of the system, meaning that after a repair the system is as good as a new one. Many examples in life show that this assumption is not always valid, increasing the will to consider a wider range of repairs. The fact that the Euclidean norm of a multi-variate Wiener process corresponds to a Bessel process only when it starts from the origin forces repairs to be treated differently. When considered, imperfect repairs are usually modelled by lowering the state of the process found at inspection, [26], [61]: the amount by which the state of the process is decreased represents the amount of repair undergone on the system. In the proposed models, repairs are tackled with the use of a repair function, which determines the amount by which the threshold(s) must be lowered. The process describing the performance measure is then re-set to zero.

Particular attention is paid to the considered inspection policies. Both cases of periodic and non-periodic inspections are treated. The periodic case consists in determining the optimal period of inspection resulting in a minimum expected cost per unit time. The non-periodic approach is inspired by the one proposed in [34]. The concept is based on the realistic assumption that the amount of time between two consecutive inspections depends on the system's state: the worse the system gets, the more frequently it should be inspected. This is put into mathematical form by defining a function (called the inspection scheduling function) that determines the next time of inspection from the value of the process describing the system's state at inspection time. In this thesis, it is the amount by which the threshold is decreased (given by the repair function), or the value of the performance measure, that determines the inspection times. Extensions

are provided with the study of different inspection scheduling functions with various convexity properties, enabling cases such as infant mortality to be considered.
Two types of models are derived, with both periodic and non-periodic inspection policies investigated in each case. The first type considers one threshold only and focuses in guaranteeing a prescribed level of reliability. Concrete examples of systems this model may apply to are roads: regular maintenance actions are usually undertaken to keep a certain level of reliability and even if they have not failed, reconstructions are usually planned after a certain amount of time for safety issues. For the second type of models, a second threshold is added in order to incorporate failure of the system. Repair is undertaken on the system to guarantee a fixed level of reliability but failure of the system is now plausible. Examples of systems thought of are airplanes, where regular inspections and repairs are usually undertaken to maintain a certain level of safety but unfrequent failures of such systems are to be taken into account.

The Chapters

The purpose of Chapter 2 is to provide a description of the problem considered in the thesis. The set up for the investigated system's performance measure is defined and so are the different types of control limit rules explored. The fundamental notions used when considering maintenance policies are stated: these concern the inspection policies, the maintenance actions and the type of process usually considered. A brief literature review is given to illustrate some of the important maintenance policies investigated in the past.
Chapter 3 focuses on the processes considered in the derived models. The chapter begins with a description of the properties required for the process describing the degradation of the components, the desired continuity property implies the consideration of Brownian motions. The stochastic process describing the system's performance measure is then introduced. The process considered is the Bessel process with drift: justifications for this choice and important properties are provided.
The methodology adopted constitutes the content of Chapter 4. The consideration of last exit times necessitates a discussion on stopping and non-stopping times. The chapter then concentrates on the approach chosen to deal with repairs and inspections. Repairs are modelled with the use of a maintenance function d that takes into con-

sideration the state of the performance measure at inspection time to determine the amount of maintenance to be undertaken on the system. Non-periodic inspections are dealt with with the use of an inspection scheduling function m that deduces the next inspection time from the state of the process describing the system's performance.

Chapter 5 concentrates on models with one threshold, for both periodic and non-periodic inspection policies. The features of the models, expressions for the expected costs and descriptions of the methods used to obtain the solutions are provided. The chapter ends with an analysis of the numerical results obtained.

Chapter 6 constitutes an extension of Chapter 5. The models proposed consider a second threshold that is introduced in order to incorporate instant failures of the system. This requires the numerical computation of the probability density function for the first time the process hits a threshold. The two types of inspection policies are also considered and so are particular cases of maintenance strategies. Explorations of the numerical results computed are also provided.

Chapter 2

Problem statement

2.1 Introduction

The purpose of the present chapter is to provide a concise description of the problem investigated in the thesis. For this, the fundamental concepts encountered in the theory of reliability and maintenance need to be described. A considerable amount of problems in this field concentrate on the state of an item. Under the reasonable assumption that most systems deteriorate through time (possibly in different ways), one can distinguish between two states:

1. *the working state*: the system's intended tasks or performance are ensured,

2. *the failure state*: the system does not perform its required function.

When no maintenance is undertaken on the system, its state will usually gradually evolve from the working state to the failure state. However, this evolution can be modified by considering maintenance policies and it must be noted that different maintenance policies on a given system have different repercussions on its state. Rather than concentrating only on the state of deterioration of a system, our interest also lies in its performance. We assume that the considered system's task may be ensured with different levels of efficiency and propose to determine an optimal maintenance policy that guarantees a prescribed level of performance.

In the glossary of maintenance terms in Terotechnology, [1], maintenance is defined as *'the combination of all technical and associated administrative actions intended to retain an item or a system in, or restore it to, a state in which it can perform its required function'*. According to Dekker, [27], its objectives can be summarized under four headings:

(i) ensuring system function (availability, efficiency and product quality);

(ii) ensuring system life;

(iii) ensuring safety;

(iv) ensuring human well-being.

A maintenance policy on the other hand is understood to be a combination of an inspection policy and one or more maintenance actions that act on the actual system's state of deterioration. The maintenance policies in the thesis aim at ensuring the system's efficiency and life. A similar approach may be adopted in the case where ensuring safety is of interest.

This chapter is organized as follows: section 2.2 explains how the performance of the system is evaluated and in particular how it is related to the deterioration of all the components in the system, bringing a multi-dimensional aspect to the problem. Major concepts related to inspection policies and maintenance actions are then described in section 2.3. This section also provides a description of the stochastic process' characteristics usually exploited when deriving cost expressions for the considered maintenance policies. Both inspection policies and maintenance actions are summarized in figure 2.3 and examples of different models considered in the literature are provided in section 2.4.

2.2 The system's performance measure

2.2.1 The set up

The main concern of the research is to derive optimal maintenance policies that guarantee a prescribed level of performance for a system. Since we extend the usual single component case to a case where the number of parts in the system is not restricted, it is a major issue that the proposed tool defining the item's performance takes the number of components present in the system into consideration. The idea is to take into consideration the degradation of all the components, denoted by $C_1, C_2, \ldots, C_n$, and to let the performance measure depend on these degradations, which seems to be a reasonable assumption. For this we assume that the state of each component is described by a continuous stochastic process in time, $W_t^{(i)}$ for component C_i say. The system's state can therefore be described by the multi-dimensional stochastic process

$$\mathbf{W}_t = \left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)}\right)$$

The performance of the system is then assumed to be described by another stochastic process being defined by a functional acting on $\mathbf{W}_t$:

$$R_t = \mathcal{A}\left(\mathbf{W}_t\right)$$

In such a way, the system's performance directly depends on the components' states. We further make the assumption that $\mathbf{W}_t$ is not observable, which is generally true in the case of large systems. However, at inspection times, the performance measure, R_t, is known.

The general set up required to deal with the performance of the system is the following. The performance of the system is described by values taken in the interval $\mathbb{R}^+$ and therefore the functional is chosen accordingly so that it also belongs to the interval (this is explained later in Chapter 3). The interval $\mathbb{R}^+$ is called the state space of the process R_t. When $R_t = 0$ the system's performance is assumed to be maximal (this corresponds to the case of a new system), and as the value for R_t increases the performance is understood to decrease. This set up seemed to be the most convenient, since it allows the performance of the system to get as low as possible (when R_t tends to $+\infty$) and only deals with positive values describing the performance. It is however possible to redefine the problem by setting a maximum value M and considering the process $\widetilde{R}_t = M - R_t$ on the interval $[0, M)$, setting $\widetilde{R}_t = M$ to denote a maximal performance and $\widetilde{R}_t = 0$ to denote the lowest performance the system may attain. The convenience of this being that the process $\widetilde{R}_t$ now decreases when the performance decreases. We shall however stick with R_t rather than $\widetilde{R}_t$.

2.2.2 Control limit rule

Since we wish to guarantee a certain level of performance for the system, we must somehow define under which conditions the performance criteria are met. To deal with this, we introduce a value $\xi \in \mathbb{R}^+$ called a critical threshold and partition the state space in two intervals

$$\mathbb{R}^+ = [0, \xi) \cup [\xi, +\infty)$$

In the general case, a control limit rule is considered. A process, usually monotone, describing the system's state of deterioration is considered and the approach consists in determining the first time at which the process hits the threshold, called the first hitting time:

$$G_{\xi}^{x} = \inf_{t \in \mathbb{R}^{+}} \{X_t = \xi | X_0 = x\}$$

Such a time defines a failure time for the system and appropriate maintenance is undertaken on the system.
We propose an extension to the general control limit rule in the sense that the process of interest now defines the performance of the system and is not assumed to be monotone anymore. Indeed, the performance of a system may fluctuate with time due to the system's usage at particular moments. Thus, rather than considering the first time at which the process hits the critical threshold, we are now interested in the time after which the process definitely stays above the critical threshold called the last exit time:

$$H_{\xi}^{x} = \sup_{t \in \mathbb{R}^{+}} \{X_t \leq \xi | X_0 = x\}$$

We note that for some non-monotone process this time may not exist (as illustrated in figure 2.1): its existence with the chosen process is explained in Chapter 3.

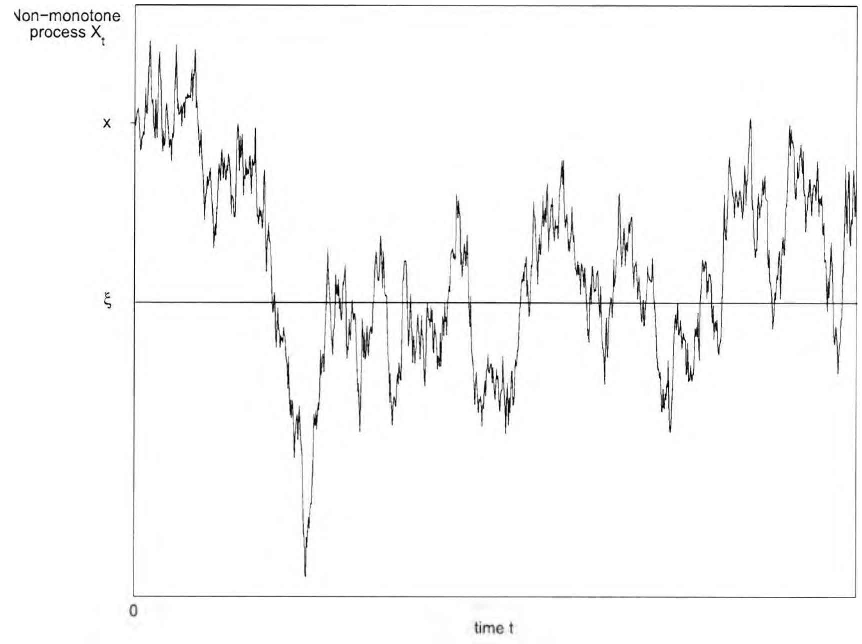


Figure 2.1: Example of a non-monotone process, where H_{ξ}^{x} does not exist.

The reason for considering such a time is the following: the non-monotonicity of the process implies that it may leave the interval $[0, \xi)$ to return to it later. Hence the performance of the system varies from being low for a certain amount of time (*i.e.* when it belongs to the interval $[\xi, +\infty)$) to being acceptable (i.e. when it belongs to the interval $[0, \xi)$). The value for the threshold ξ is chosen accordingly so that this type of scenario is accepted. However, due to the property of the chosen process (involving the transience property, see Chapter 3), it will eventually definitely escape from the interval $[0, \xi)$ meaning that the performance will never reach an acceptable level anymore. After such a time, the system is assumed to be aging in such a way that it needs to be replaced. The proposed models improve in complexity in the way that a second threshold $\mathcal{F}$ is then added. This new threshold allows to incorporate failures in the model. This is done by assuming that the new threshold satisfies $\mathcal{F} > \xi$ and by considering the first hitting time of $\mathcal{F}$:

$$G^x_{\mathcal{F}} = \inf_{t \in \mathbb{R}^+} \{X_t = \mathcal{F} | X_0 = x\}$$

With this new set up consisting of two thresholds ξ and $\mathcal{F}$, the state space is now partitioned as

$$\mathbb{R}^+ = [0, \xi) \cup [\xi, \mathcal{F}) \cup [\mathcal{F}, +\infty)$$

If at inspection times the process has definitely escaped from the interval $[0, \xi)$, it needs to be repaired. Moreover, when the process hits the critical threshold $\mathcal{F}$ the system is assumed to be in a failed state and needs to be replaced with a new one.

We note here the analogy of the approach with a sequential probability test. Such a test is used when it is desired to test that a sample $X_1, X_2, \ldots$ came from one distribution against the possibility that it came from another, *i.e* it is wished to test $\mathcal{H}_0 : X_i \sim f_0(.)$ versus $\mathcal{H}_1 : X_i \sim f_1(.)$ say. The values $\lambda_1, \lambda_2, \ldots$, where

$$\lambda_m = \frac{\prod_{i=1}^m f_0(x_i)}{\prod_{i=1}^m f_1(x_i)},$$

are computed sequentially and decisions are made corresponding to the value obtained for λ_m: if $\lambda_m \leq k_0$ reject $\mathcal{H}_0$, if $\lambda_m \geq k_1$ accept $\mathcal{H}_0$ and if $k_0 < \lambda_m < k_1$ compute λ_{m+1}, $m \in \mathbb{N}$, for fixed k_0 and k_1 satisfying $0 < k_0 < k_1$. The strategy opted in our model is similar in the sense that at inspection time decisions on the maintenance action to carry out depend on the probability of occurrence of the last exit time: repair, replace or do nothing and test at the next inspection.

2.3 Maintenance policies: fundamental concepts

2.3.1 Inspection policies

Ensuring the good functioning of a system requires accurate knowledge of its state of deterioration. In many cases errors in measurement arise, [26]. These are not taken into account here and make no difference to the models derived later. Indeed, this type of error is usually incorporated in the models with the use of a normally distributed random variable. The process describing the components' deteriorations being a Brownian motion with drift (see Chapter 3), errors in measurements may be taken into account with a change in the drift and volatility terms of the process: the sum of two normally distributed random variables is a normally distributed random variable.

In some cases, the state of degradation of a system is continuously observable: one may think of a light bulb for example. In most cases however, the state of the system (or subsystem) may only be known after an inspection. Hence knowledge of the state of deterioration may only be available at particular times and may require dismantling of the system. An optimal maintenance policy strongly depends on a appropriate inspection policy. Moreover, failing to determine such a policy may not only result in suboptimal maintenance policies but also in failures which may have catastrophic consequences, *e.g.* nuclear power plants.

In this section we propose an overview of the most considered inspection policies in the field of reliability and maintenance. We distinguish between instantaneous and non-instantaneous inspections, perfect and imperfect inspections, periodic and non-periodic inspections. This does not constitute an exhaustive list since combinations of the proposed policies may be considered.

No inspection

Under such an inspection policy, the system is never inspected and usually runs until it fails. The 'no inspection' policy is often associated with cases of technical and economic obsolescence: a repair on the system induces a higher cost than a replacement. Typical examples include light bulbs, first price DVD or CD players, where replacement of such items with new ones is usually more appropriate than a repair.

Instantaneous/Non-instantaneous inspections

The time required for an inspection may vary from system to system. It is often the case that the inspection time can be neglected when compared to the system life. However, for large systems it is sometimes the case that inspections require considerable amount of time and the associated cost may depend on the time required to inspect it. To deal with that matter, we define the concept of instantaneous inspection.

(i) *Instantaneous inspections:* This type of inspections refers to inspections whose time length may be neglected compared to the system's life time or to the time between any two consecutive inspections. If an inspection is planned at time $t = \tau$ say, the system's state is assumed to be known instantaneously. The time at which this state is known is usually denoted by $t = \tau^+$ (sometimes $t = \tau$). A fixed cost of inspection C_i is usually associated to such an inspection.

(ii) *Non-instantaneous inspections:* Here the inspection times are not neglected. This is often the case when an 'in-depth' inspection is considered or when the system is large and inspections require a major dismantling. The time required to inspect it needs not be neglected anymore and if the system is inspected at time $t = \tau_1$, its state of deterioration is known at time $t = \tau_2$, with $\tau_1 < \tau_2$. The resulting cost of inspection is usually a function of the time required to inspect the system, $C_i(\tau_2 - \tau_1)$.

Perfect/Imperfect inspections

It is essential to know up to what extend the information gained after inspection can be trusted. For this the concepts of perfect and imperfect inspections is introduced.

(i) *Perfect inspections:* After being inspected, the true state of degradation of the system is assumed to be known.

(ii) *Imperfect inspections:* The system state of degradation is not fully available. This may come from several reasons, the main ones being that the system was only partially inspected (leaving some parts uninspected) or that errors in assessing the state of the different parts occurred. In this last case we need to distinguish between two types of errors:

- *False positive:* the inspection incorrectly reports a positive result, i.e. the system is incorrectly reported to be in a satisfactory state.
- *False negative:* the inspection incorrectly reports a negative result, i.e. the system is incorrectly reported to be in an unsatisfactory state.

Periodic/Non-Periodic inspections

Another important feature of inspections is the time at which they occur. There may be conditions on such times due to the type of the considered item. For instance, in the aerospace industry one can reasonably assume that inspections of airplanes require appropriate tools and appropriate engineers, forcing inspections to take place in pre-determined bases at particular times. Hence, one may think of pre-determined equally spaced inspection times. However other type of systems may be inspected at any desired time, allowing the inspection planner greater freedom in the planning. This leads us to the definition of periodic and non-periodic inspections. For this, let $\Pi = \{\tau_1, \ldots, \tau_n\}$ denote an inspection policy over a finite time, where each of the τ_i denotes an inspection time (over an infinite time interval, the notation $\Pi = \{\tau_1, \ldots, \tau_n, \ldots\}$ is used).

(i) *Periodic inspection policy:* The amount of time between any two consecutive inspections is constant. This constant is called the period of inspection. Let τ be the period of inspection, the corresponding inspection policy is:

$$\Pi = \{\tau_1, \ldots, \tau_n\}, \text{ with } \tau_i = i \times \tau, \ \forall\, i \in \{1, \ldots, n\}$$

(ii) *Non-Periodic inspection policy:* The amount of time between any two consecutive inspections may be different. This amount may be fixed and pre-determined or may be random, it may depend on the system state for instance. The corresponding inspection policy is:

$$\Pi = \{\tau_1, \ldots, \tau_n\}, \text{ where } \tau_{i+1} - \tau_i \neq \tau_{j+1} - \tau_j$$

for at least one $j \in \{1, \ldots, n-1\}$, with $i \in \{1, \ldots, n-1\}$, $i \neq j$.

Remark 2.3.1. The *continuous condition monitoring policy* is briefly mentioned here. This strategy assumes that the system's state is known continuously. In this special case, a cost c_{ccm} per unit time is usually included. Note that this policy may be regarded as a limit case of the periodic inspection policy, with $\tau \rightarrow 0$.

2.3.2 Maintenance actions

As mentioned in the previous subsection, information on the state of degradation is obtained by performing inspections of different types. This new available information can then be taken into account to perform appropriate maintenance actions. A good inspection policy combined with poor maintenance actions is useless. It is then of prime importance to elaborate a concise maintenance strategy. This may involve making decisions based on the state of the system known after an inspection and taking actions. However other systematic maintenance strategies may consider actions without taking into account the system's state of deterioration.
This subsection aims at defining the important classes maintenance actions fall into.

Instantaneous/Non-instantaneous maintenance actions

Such maintenance actions are defined in a similar way as for instantaneous/non-instantaneous inspections. In the instantaneous case, the time length of a maintenance action is neglected compared to the system's life time and a fixed cost is usually associated. The non-instantaneous case takes the time needed to perform a maintenance action into account and the associated cost may depend on that required time.

Preventive and Corrective maintenance

Maintenance actions are principally considered in order to be able to deal with failure of the system. Such actions may be considered prior a failure or after.

(i) *Preventive maintenance:* Preventive maintenance includes all maintenance actions that are planned before observing failure of the system. Such a strategy aims at improving the system's state. We distinguish between the following three cases:

- the preventive maintenance is undertaken at pre-determined time intervals and no actions are considered before that time: this constitutes a systematic approach,
- the time at which the preventive maintenance is undertaken is chosen on the basis of the estimated remaining lifetime of the system,
- preventive maintenance is undertaken on the basis of the known state of degradation of the system or its performance.

(ii) *Corrective maintenance:* Corrective maintenance includes all maintenance actions undertaken after failure. Most models usually associate a higher cost with a corrective maintenance than with a preventive maintenance. An example of a maintenance policy based on corrective maintenance is the *no repair* strategy. The strategy consists in letting the system run until it fails, no maintenance actions are considered. This maintenance action is usually combined with the no inspection policy. Such a maintenance strategy can be undertaken only in particular cases where failure does not involve catastrophic consequences and where safety criteria are not an issue (e.g. a light bulb at home). This strategy is also considered in cases where the cost of repair of the failed item is less than the inspection cost: *i.e.* the cost induced by letting the system fail is lower than the inspecting cost.

Repairs and replacements

Whether preventive or corrective maintenance are considered, it is essential, when possible, to specify the degree of maintenance undertaken on the system. This ranges from doing nothing (no maintenance) to a complete replacement. Maintenance actions consist of repairs and replacement.

(i) *Repair:* This includes all maintenance actions that change the system state of degradation to a new state. The new state considered often corresponds to a lower state of degradation, assuming that repairs improve the system state. However, one must note that inappropriate repairs can lead to particular cases where the system state gets worse; such repairs are not considered in this thesis. By perfect repair it is understood a repair that resets the system state to the state of a new system: such a repair is said to change the system state to an 'as good as new' state. By imperfect repair it is understood any other kind of repair than a perfect repair. Imperfect repairs include the special case of minimal repair: such a repair leaves the system's state to its state prior to inspection, the system is said to be in an 'as good as old' state.

Excluding the cases of minimal and perfect repairs where the new state of degradation is fully known, the degree of repair might be known or only partially known. We distinguish between deterministic repairs and random repairs:

- *deterministic repair:* the state of degradation of the system after a repair is known with certainty:

$$X_{\tau^+} = K$$

with

$$0 \leq K \leq X_\tau,$$

where X_τ is the state of the system at inspection time, X_{τ^+} is the state of the system after the maintenance action (assumed instantaneous here with the notation) and K is a constant. Hence, the state of the system K, after a repair, is improved.

- *random repair:* the state of degradation of the system after a repair is not fully known. The amount of repair undertaken is usually modelled with the use of a random variable Θ whose realizations belong to a known sample space Ω. In the case where improvement of the system's state is certain but the amount is random, one has

$$\Omega \subseteq [0, X_\tau]$$

If the random repair on the system does not necessarily improve its state, *i.e.* possibility of failed repair or damage, then

$$\Omega \subseteq [0, \infty)$$

(ii) *Replacement:* Parts of the system or even the whole system is replaced with a new one. This has the same effect as a perfect repair. Examples of replacement strategy include the block replacement policy: the considered component is replaced on failure and at periodic times.

2.3.3 Regenerative phenomena

When dealing with maintenance policies for an item, it is of prime importance to model its deterioration accurately. This can be done by looking at the physics of the actual degradation and selecting the corresponding appropriate characteristics that the process, responsible for the description of this degradation, must satisfy. This is the matter of interest discussed in Chapter 3, where the choice of an appropriate stochastic process representing the performance measure of the un-maintained system is discussed. However, maintenance policies and expressions for the costs of maintenance of the system are made using a process representing the maintained system's state of deterioration. This process is derived from the initial one, representing the performance of the un-maintained system, by considering the effects of maintenance at appropriate times. Let R_t and R_t^* denote the processes modelling the un-maintained and maintained system respectively. The usual framework models repairs on the system by changing the value of the stochastic process R_τ^* at inspection time τ to a lower value at time τ^+ (to denote instantaneous repair), satisfying:

$$0 < R_{\tau^+}^* < R_\tau^*$$

hence assuming that the repair improves the performance of the system. At particular times though, the system's performance gets sufficiently low that replacement of the whole system is considered. Assume this happens at times $(\tau_i^*)_{i \in \mathbb{N}}$, hence:

$$R_{\tau_i^{*+}}^* = 0$$

i.e. the process restarts from its initial state at time $t = 0$. This property leads us to the following definitions:

Definition 2.3.2. (Regenerative process). Consider a process $(X_t)_{t \geq 0}$ having the property that there exist points at which the process (probabilistically) restarts itself. That is, suppose that with probability one, there exists a time T_1, such that the continuation of the process beyond T_1 is a probabilistic replica of the whole process starting at 0 (note that this implies the existence of further times $T_2, T_3, \ldots$, having the same property as T_1). Such a process is known as a regenerative process.

Definition 2.3.3. (Renewal process). Consider a counting process for which the inter-arrival times are independent and identically distributed with an arbitrary distribution. Such a counting process is called a renewal process. Formally, let $(X_k)_{k \in \mathbb{N}}$ be a sequence of nonnegative independent random variables with a common distribution F. To avoid trivialities, suppose that $\mathbb{P}[X_k = 0] < 1$. Let

$$S_n = \sum_{i=1}^{n} X_i, \; n \geq 1$$

with

$$S_0 = 0.$$

Define

$$N_t = \sup\{n : S_n \leq t\}$$

Then, the process $(N_t)_{t\geq 0}$ is a renewal process.

Remark 2.3.4. We shall say that a *cycle* is completed every time a renewal occurs.

We note that generalizations to the renewal process may be considered:

- *Alternating renewal process:* it is sometimes the case that renewal processes have intervals of different types, characterized by different probability density functions. For instance, suppose that $(X_k)_{k\in\mathbb{N}}$ and $(Y_k)_{k\in\mathbb{N}}$ are two independent sequences of random variables, each sequence consisting of mutually independent and identically distributed random variables. Let X_k be a type 1 interval ending with a type 1 event and Y_k a type 2 interval ending with a type 2 event. Suppose further that random variables are alternatively taken from the two series, the resulting process is illustrated in figure 2.2. The $(2n-1)$th event is a type 1 event and occurs after time $X_1+\ldots+X_n+Y_1+\ldots+Y_{n-1}$ and the $2n$th event is a type 2 event occurring after time $X_1+\ldots+X_n+Y_1+\ldots+Y_n$. This system is called an alternating renewal process. In a reliability context, one may think of $(X_k)_{k\in\mathbb{N}}$ as representing the lengths of operations of a system and $(Y_k)_{k\in\mathbb{N}}$ the lengths of repairs of the system. Models dealing with such scenarios are studied in [5].

- *Modified renewal process:* consider the case where the first random variable X_1 in definition 2.3.3 has a distribution G that is not identical to F. The process is then referred to as a modified renewal process (or delayed renewal process). A typical example of such a scenario being when the considered component at time $t=0$ is not new. After failure it is replaced by a new one and so on: the distribution for the first failure-time is therefore not identical to the remaining ones. In the special case where $G(x)=\displaystyle\int_0^x \frac{1-F(u)}{\mathbb{E}(X_i)}\,du$, the process is called an equilibrium (or stationary) renewal process, [25].

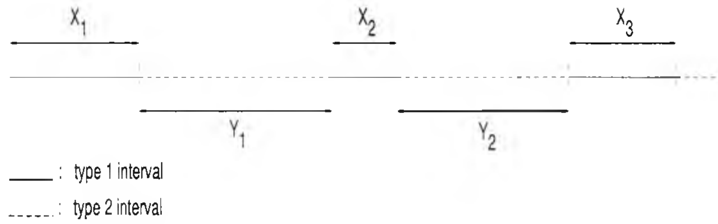


Figure 2.2: Alternating renewal process.

From the above definitions, one may deduce that R_t^* is a regenerative process and the sequence of regeneration points $(\tau_i^*)_{i\in\mathbb{N}}$ defines a renewal process. These properties of the investigated process are then exploited to derive expressions for the costs of maintenance.
In the special case where inspections occur periodically, the renewal reward theorem, [76], is usually used to calculate the expected long run cost per unit time.

Theorem 2.3.5. *Consider a renewal process with interarrival times $X_1, X_2, \ldots$ Suppose further that a reward Y_n is earned at the time of the nth renewal. Y_n may (and will usually) depend on X_n (the length of the renewal interval), but suppose that the pair (X_n, Y_n), $n = 1, 2, \ldots$ are independent and identically distributed. Let*

$$Y_t = \sum_{n=1}^{N_t} Y_n$$

denote the total reward earned by time t.
If $\mathbb{E}[|Y_n|]$ and $\mathbb{E}[X_n]$ are finite, then

(i) with probability 1,

$$\frac{Y_t}{t} \to \frac{\mathbb{E}[Y]}{\mathbb{E}[X]} \quad as\ t \to +\infty$$

(ii) $\dfrac{\mathbb{E}[Y_t]}{t} \to \dfrac{\mathbb{E}[Y]}{\mathbb{E}[X]}$ *as* $t \to +\infty$

i.e. the renewal reward theorem states that the expected long-run return per unit time is just the expected return earned during a cycle divided by the expected length of a cycle.
Thus, when dealing with periodic inspections, the aim consists in deriving an expression for the expected cost of maintenance over a cycle and the expected length of a cycle by considering the different possible scenarios. Once these are obtained, the expected long run cost of maintenance per unit time may be computed with the use of theorem 2.3.5. In the more general case where inspections are non-periodic, the renewal reward theorem may also be considered. However, rather than calculating the expected cost per unit time over an infinite time horizon, the expected total cost over a finite time horizon is usually evaluated: such an expression is also obtained with the use of renewal points and provides more insight since it allows to price a maintenance strategy for a pre-defined project length.

2.4 Models investigated

2.4.1 Maintenance policies considered in the literature

Many maintenance policies have been considered in the literature. These consist in combinations of the different concepts introduced in sections 2.3.1 and 2.3.2. To summarize, figure 2.3 states the different inspection policies and maintenance actions encountered in the two previous sections.

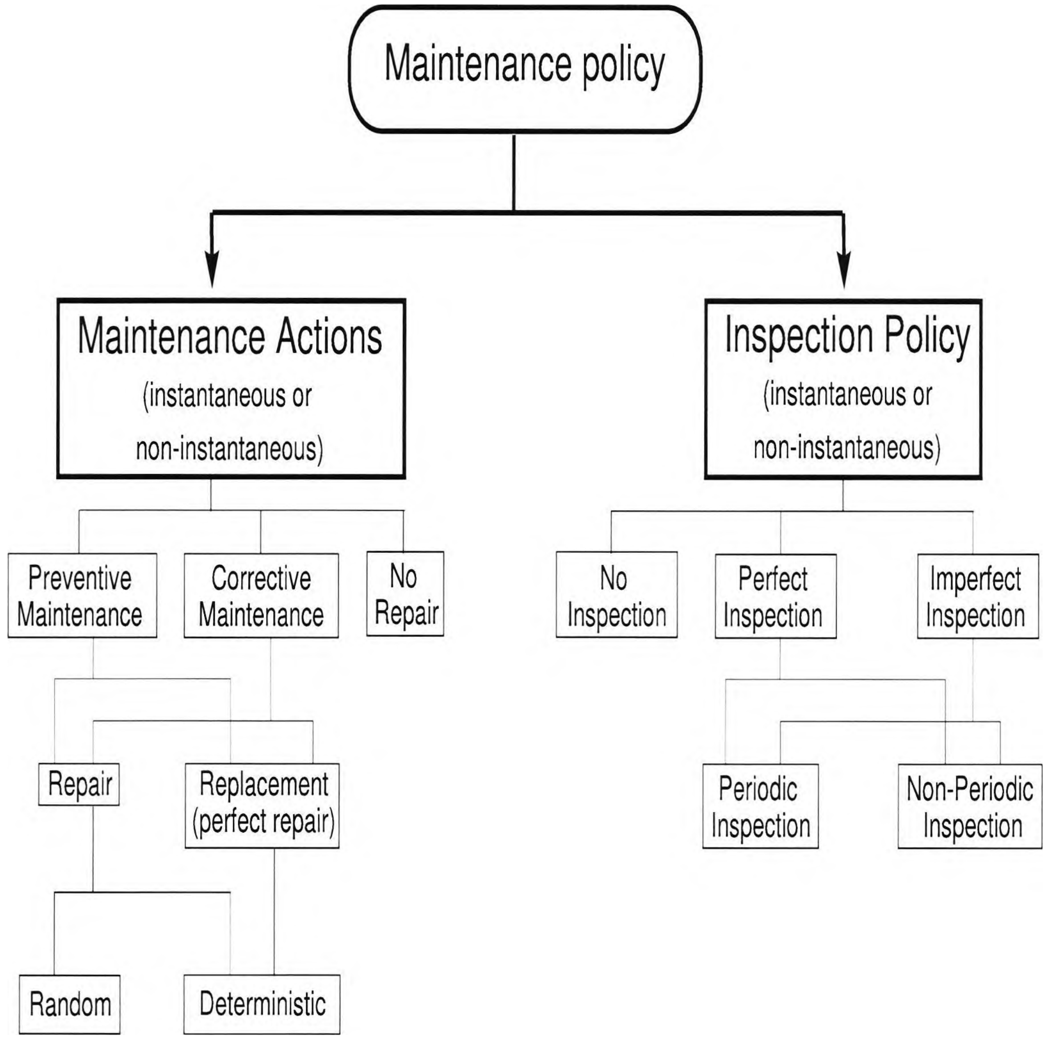


Figure 2.3: Maintenance Policies.

To start with, we mention the extension to the above inspection policies proposed by Scarf and Deara, [77], [78]. The models proposed consist in determining an optimal maintenance policy for a two-component system (a clutch system used in a bus fleet), taking into account both failure and economic dependence. In order to deal with that kind of scenario, various age-based replacement policies and block replacement policies (such as simple block replacement and modified block replacement) were investigated. Age-based replacement policies replace a component on failure or at age T, whichever is sooner. As mentioned in subsection 2.3.2, a simple block replacement policy replaces a component on failure (corrective maintenance) and at time $t = k\Delta,\ k = 1, 2, \ldots$. A modified block replacement policy replaces a component on failure and at time $t = k\Delta,\ k = 1, 2, \ldots$, if the component's age $\tau > b$ for some suitable threshold b. Due to the presence of economic interaction, opportunistic replacement are also introduced: at an opportunity (e.g. an inspection time), a component is replaced if its age $\tau > b'$ for some $b' < b$. These block replacement strategies are also combined to give more complex maintenance policies and results show that the modified block replacement policies provide lower maintenance costs. However, as mentioned in the article, the implementation of such policies requires more level of maintenance reporting and control to gain successful results. Thus the economic gains from using a more complex policy have to be weighted up against the additional investment required to implement them. Block replacement models have also been studied by Aven and Dekker. In [4], a decision variable $T > 0$, which affects the times of maintenance renewing the system, is introduced. Such a decision variable can be the time at which the system is replaced preventively or a critical time after which the first suitable moment is awaited to renew the system. The aim consists in determining the value for T that would minimize a long term average cost expression $g(T)$ of the form:

$$g(T) = \frac{c + \int_0^T m(t)\, h(t)\, dt}{d + \int_0^T h(t)\, dt},$$

with c and d representing the cost and the time of a preventive replacement respectively. The generality of the proposed framework enables to develop a number of different models (depending on the chosen expression for the functions m and h) such as:

(i) the standard minimal repair model: m is the cost of a minimal repair times the rate of occurrence of failures, $h(t) = 1$ implying that a cycle always has length T and $d = 0$ meaning that preventive replacements take no time;

(ii) the age replacement model: m is the failure replacement costs times the hazard rate, h is the survival function and $d = 0$;

(iii) a block replacement model: m corresponds to the expected failure replacement costs times the renewal density, $h(t) = 1$ and $d = 0$.

As far as imperfect inspections are concerned, we mention the work form Barros, [13]. The author investigates the effects of the quality of inspections on maintenance policies. The considered maintenance actions consist in replacing the components in the system (two components in parallel) preventively or correctively. The complexity of the model is introduced by assuming that the components' state of degradation is only partially known. This kind of assumption enables problems such as delay in assessing failure of a component, false-alarm and no-detection to be considered. The situation considered here is the no-detection one: a probability p that a component fails and failure is not detected is introduced. One may think of such a scenario in the case where the inspection tool fails before failure of one of the components in the system. The proposed method utilizes the stochastic process $\mathbf{1}_{\{T<t\}}$ (where T represents the life time of the system), whose expression is given as a smooth semi-martingale:

$$\mathbf{1}_{\{T<t\}} = \int_0^t \mathbf{1}_{\{T>s\}} \Lambda_s ds + M_t,$$

where

$$\mathbf{1}_{\{T>s\}} \Lambda_s = \lim_{\delta s \to 0} \frac{1}{\delta s} \mathbb{P}\left[s < T < s + \delta s \mid \mathcal{G}_s\right]$$

is called the failure rate process, $\mathcal{G}_s$ is a sub-σ-algebra and M_t is a $\mathcal{G}$ martingale with the properties $\mathbb{E}\left[M_{t+u} | \mathcal{G}_t\right] = M_t$, $\forall\, u > 0$ and $\mathbb{E}\left[M_t | \mathcal{G}_0\right] = \mathbb{E}\left[M_0\right] = 0$. This process uses the available information on the system's components given by the supervision, taking into account that the information given is subject to errors.

An important contribution to the case of periodic and non-periodic inspections with both deterministic and random maintenance can be found in the work of Newby and Dagg, [26], [61]. The proposed models aim to determine an optimal inspection and maintenance policy with the use of a maintenance cost as a measure of policy. Maintenance decisions are taken on the basis of the value of a bivariate process $\{(X_t, M_t), 0 \leq t < \infty\}$, where M_t represents the maximum of a non-monotone process X_t (a Wiener process with drift). In the periodic case, the problem consists in finding the optimal inspection period τ^* determined from

$$\tau^* = \underset{\tau \in \mathbb{R}^+}{\operatorname{argmin}} \{\mathcal{C}(\tau)\},$$

where $\mathcal{C}(\tau)$ denotes the average maintenance cost per unit time. Different models, such as the no inspection policy, the continuous condition monitoring and the age

replacement model are considered. A discounted cost criterion is also introduced and the case of non-periodic inspections is dealt via dynamic programming.
Another way of handling the complexity of non-periodic inspections is proposed by Grall *et al.* in [34]. At inspection times, decisions on both the type of maintenance action and the next inspection time are being taken. The next inspection time is chosen with the use of a function m, called the inspection scheduling function. It is a decreasing function of the state of the system after an instantaneous maintenance action. This framework assumes the sensible argument that the worse the system is, the more frequently it needs to be inspected. Both preventive and corrective maintenance are included in the model by partitioning the state space, in which the state of the system evolves, into three intervals: $[0, L)$, $[L, M)$, $[M, \infty)$. Upon inspection, if the state of the system is found in the first interval the system is left unchanged, if it is found in the second region a preventive perfect maintenance is performed bringing the state to an 'as good as new' state and if the state is greater than M a corrective maintenance is performed (M corresponding to the failure threshold). The optimal policy, which minimizes a long run expected cost, is then determined by choosing the appropriate inspection scheduling function and threshold L.

2.4.2 Covariates

An important factor acting on the way a system deteriorates is the environment in which it evolves. Hostile environments (e.g. humidity, frost, dryness, high pressure, etc) may result in accelerated degradations. Moreover, a system's degradation is certainly affected by its operating history: an old item which has already experienced some kind of failure and/or repairs will not deteriorate in the same manner as a new one. It is possible to include these ideas by introducing variables in the model representing factors that affect the deterioration. Such factors that affect the reliability of the system are called covariates.
One of the most important models is the proportional hazard rate proposed by Cox, [24]. This model treats the presence of covariates in the case of perfect repairs. If we let T denote the lifetime of an item, its hazard rate is given by:

$$\lambda(t) = \frac{f(t)}{R(t)},$$

with f and R being the density and reliability function of the random variable T respectively. The covariates are introduced in the model by considering a hazard rate of the form:

$$\lambda(t;z) = \lambda_0(t) . \psi(z;\beta),$$

where λ_0 is the baseline hazard rate (system's hazard rate under the assumption of no covariates), z is a row vector with each of the z_i representing covariates associated to the system, β a column vector with each of the β_i being unknown parameters of the model defining the effect of the covariates and ψ a function of the covariates. A common expression for ψ is the exponential function, resulting in the following expression for the hazard rate:

$$\lambda(t;z) = \lambda_0(t) . e^{(\beta_1 z_1 + \ldots + \beta_n z_n)}$$

Numerous extensions to this model have been proposed, among these Kumar *et al.* investigated the issue of time-dependent covariates, [41], [42], [43]. Percy and Alkali, [67], developed the Generalized Reduction of Intensity Model, which considers intensity functions instead of hazard rates and allows the case of minimal repair to be considered.

2.5 Summary

Most items experience deterioration through time and thus evolve from a working state towards a failure state. For both economic reasons and safety issues, it is of interest not only to prevent such failures but also to guarantee a certain level of performance for the system. This is done by considering maintenance policies. A maintenance policy consists in an inspection policy and some maintenance actions. An appropriate choice of maintenance policy may result in considerable savings and/or increase of the system's safety.

In this chapter, we first introduced a way to relate a performance measure of a system with the deterioration of its components. The possible characteristics of an inspection policy and the different types of maintenance actions usually considered were then defined. Part of the chapter also concentrates on the stochastic process' properties required to calculate the maintenance costs. Some examples of maintenance strategies considered in the literature were given to illustrate a few of the many possible combinations between inspection policies and maintenance actions. Eventually, we ended this chapter mentioning that factors such as the environment in which the system evolves or its operating history may affect the system's deterioration: such factors are called covariates.

Chapter 3

The processes

3.1 Introduction

The present chapter focuses on the choice of appropriate stochastic processes for the description of the components' deteriorations on the one hand, and for the description of a performance measure of the whole system on the other hand. The chosen process describing the system's performance measure must somehow take into consideration the state of degradation of a complex multi-component system. For this we wish to take into account the degradation of each component in the system and have a summary description of these at any desired fixed time in the future. This summary description can be seen as a performance measure that is used to make decisions on the type of maintenance to be undertaken on the system. For this, we assume that the considered system S consists of n components. The n chosen processes modelling the degradation of the components in the system are then grouped to form an n-dimensional process. The only available information on the system's deterioration through time is given by the performance measure, which is described by an appropriate functional acting on the n-dimensional stochastic process. This concept of applying a functional to an underlying process describing the system's state has been investigated in the past by considering bivariate processes, [26], [60], [61]. The approach consists in beginning with the underlying process X_t and work with a bivariate process (X_t, Y_t), where Y_t is a performance metric. The process Y_t may be constructed by applying a functional to the basic process, $Y_t = \mathcal{A}(X_t)$. The advantages of the approach are that decisions can be based on X_t, Y_t or even on the pair (X_t, Y_t). The associated process Y_t can incorporate the process history or some other important aspect of the process. The system is inspected to determine its state (X_t, Y_t) and a repair or replacement is chosen on the basis of the system state at the inspection. Many examples are available in risk analysis for engineering projects. Fatigue crack growth in pressure vessels and in aircraft structures has been described by Sobczyk and Spencer [82] and Newby [57], [58]; optimal inspection and maintenance policies for degradation processes are studied by Newby and Dagg [61]. Similar approaches are used in epidemiology: Jewell and Kalbfleisch [38] use marker processes in the study of CD4 counts in HIV infected patients; Betensky [15] used Wiener process models in designing sequential tests for differences in treatment effect of drugs. All of these mentioned examples construct a decision making process through using a stochastic process and using an associated process, usually a transform of the underlying process, as a decision variable. Natural examples of functionals of an underlying process X_t are:

(i) the maximum process, $Y_t = \sup_{0\leqslant s\leqslant t} X_s$;

(ii) the Euclidean norm of a multivariate process $Y_t = \|X_t\|_2$;

(iii) an accumulation process $Y_t = \int_{0\leqslant s\leqslant t} X_s ds$;

(iv) a usage measure $Y_t = \int_{0\leqslant s\leqslant t} |X_s|\, ds$;

(v) errors in measurement $Y_t = u(X_t, \varepsilon)$ where ε is a noise term;

(vi) covariate processes where a distribution $F(X_t \mid Y_t)$ describes the dependence of X_t on covariate Y_t.

When the underlying process X_t is a Wiener process (i) was extensively studied by Dagg and Newby, see [26], (ii) is a Bessel process (if the Wiener process starts at the origin), and (iii) is the Kolmogorov diffusion [53].
The first section of the chapter focuses on modelling the components' degradation: attention is paid to the required properties the chosen stochastic process must satisfy. For this the class of Markov and Lévy processes is introduced. The second section deals with the actual stochastic process representing the performance measure of the system: relevant properties of this process for the considered models are then stated.

3.2 Modelling degradation

The present section introduces the processes used to model the state of degradation of the considered components and system. Justifications for these particular choices are given. This is of prime importance for the rest of the thesis since all the models derived in Chapters 5 and 6 clearly depend on the chosen stochastic processes.
To start with, Lévy processes are defined and their relevant properties to the problem stated. An argument based on the desired continuity property for the chosen process leads to Brownian motions with drift: justifications for this choice of Lévy process as the underlying process modelling the state of degradation of a component are given.

3.2.1 Lévy processes

In order to find an appropriate maintenance strategy for a considered system (by appropriate we understand one which minimizes a maintenance cost function, or one which maximizes some kind of safety of the system), decisions must be made under uncertainty. This uncertainty mainly comes from the way the system deteriorates through time and our main concern is to be able to define a time at which the system is either considered as unsafe or as failed. Two strategies have been opted in the past. The first being based on the lifetime distribution of the system through time, and the second one based on the physics of failure and the characteristics of the operating environment. The disadvantage of the lifetime distribution are that the only information available is whether the system is functioning or not. In order to represent ageing, failure rates are considered. However, failure rates cannot be observed or measured for a particular component. Hence, for engineering structure it is generally more interesting to base the modelling of degradation on the physics of failure, [81]. It is therefore recommended to model the deterioration of the system with the use of a time-dependent stochastic process. Moreover, according to Barlow and Proschan in [11], deterioration is usually assumed to be a Markov process. A Markov process is defined as follows

Definition 3.2.1. A process $(X_t)_{t\geq 0}$ with values in $\mathbb{R}^d$ is called a Markov process if

$$P\left[X_t \leq x | X_{t_0} = x_0, X_{t_1} = x_1, \ldots, X_{t_n} = x_n\right] = P\left[X_t \leq x | X_{t_n} = x_n\right] \tag{3.1}$$

whenever $t_0 < t_1 < \ldots < t_n < t$.

Hence a process is Markov if the conditional distribution of the future given the present and the past, is independent of the past. We will see in Chapter 4 that the Markov property of the process enables the extension of perfect repair models to more general repairs. An extensive amount of work in the field of maintenance and reliability has been done using time-dependent Markov stochastic processes, the most commonly used being the compound Poisson process, the gamma process and the Brownian motion: definitions of these processes are now given.

Definition 3.2.2. (Compound Process, Compound Poisson process). Given a counting process $(N_t)_{t\geq 0}$ and a sequence of independent random variables $(X_i)_{i\in\mathbb{N}^*}$, which are also independent of N_t, the random variable

$$S_t = \sum_{i=1}^{N_t} X_i$$

is called a compound process.
If $(N_t)_{t\geq 0}$ is a Poisson process and $(X_i)_{i\in\mathbb{N}^*}$ is a sequence of independent and identically distributed random variables also independent of N_t, S_t is called a compound Poisson process.

Definition 3.2.3. (Gamma Process). The gamma process with shape function $v\,(t) > 0$ (non-decreasing, right continuous, real valued function for $t \geq 0$ with $v\,(0) \equiv 0$) and scale parameter $u > 0$ is a continuous-time stochastic process $(X_t)_{t\geq 0}$ with the following properties

(i) $X_0 = 0$ with probability one;

(ii) $X_\tau - X_t$ has a gamma distribution $\Gamma\left(v\,(\tau) - v\,(t)\,, u\right),\ \forall\ \tau > t \geq 0$;

(iii) X_t has independent increment.

Remark 3.2.4. The gamma process can be regarded as a compound Poisson process of gamma-distributed increments in which the Poisson rate tends to infinity and increment size tend to zero in proportion, [44].

Definition 3.2.5. (Brownian motion). Let $(\Omega, F, \mathbb{P})$ be a given probability triple. A real-valued stochastic process $(B_t)_{t\in\mathbb{R}^+}$ is a Brownian motion if it has the properties:

(i) $B_0\,(\omega) = 0,\ \forall\ \omega \in \Omega$;

(ii) the map $t \rightarrow B_t(\omega)$ is a continuous function of $t \in \mathbb{R}^+$, $\forall\, \omega$;

(iii) $\forall\, t, h \geq 0$, $B_{t+h} - B_t$ is independent of $\{B_u : 0 \leq u \leq t\}$, and has a Normal distribution with mean 0 and variance h.

Among the many publications using the gamma process to model deterioration, we refer to [3, 26, 34, 55, 84, 85, 87, 88, 89, 86] and as for the one based on Brownian motion we refer to [6, 7, 8, 9, 10, 26, 45, 47, 61, 62, 63, 90].
The reason why such Markov processes (the gamma process and the Brownian motion) have been and still are extensively used being that they belong to a smaller class of time-dependent stochastic process known as Lévy processes.

Definition 3.2.6. (Lévy process) A processes $(X_t)_{t \geq 0}$ with values in $\mathbb{R}^d$ is called a Lévy process (or process with stationary independent increments) if it has the properties

(i) for almost all w, $t \rightarrow X_t(w)$ is right continuous on $[0, +\infty)$, with left limits on $(0, +\infty)$;

(ii) for $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables $Y_j := X_{t_j} - X_{t_{j-1}}$ $(j = 1, \ldots, n)$ are independent;

(iii) the law of $X_{t+h} - X_t$ depends on h, but not on t (stationary increments).

Remark 3.2.7. We note here that Lévy processes may have non-continuous paths. Properties *(ii)* and *(iii)* imply that Lévy processes have independent and stationary increments.

To understand why Lévy processes constitute good candidates for deterioration modelling we must first make sure how deterioration and failure of a system are viewed. These are commonly seen as the result of accumulation of shocks, wear and tear through the lifetime of the system. Shocks on the system are usually dealt with shock models: these assume that the considered system is exposed to shocks at random times, with each shock causing a random accumulating amount of damage. The main interest is the time at which the accumulated amount of damage exceeds a fixed amount. According to van Noortwijk, [84], [85], [88], [89], in many cases where the system is subject to shocks the order in which these accumulating shocks occur is immaterial, hence suggesting that the random deteriorations in equal interval of time are exchangeable random variables. Moreover this also suggests that the amount of deterioration in a fixed interval is independent of the starting time of that interval, implying the stationarity

property of the process. Exchangeable and stationary increments are similar to the stronger properties of stationary and independent increments of Lévy processes, [20]. The restriction to stationary increments is outweighed by the analytical advantages of using Lévy processes, [59]. We note that the stationary property of the increments assumes that the deterioration does not take into account the age of the system, which may seem inaccurate for certain types of systems. The use of Lévy processes has also been considered in the degradation-threshold-shock models proposed by Lehmann, [45], [46]: these consist in a combination of degradation threshold model (control limit rule) with a shock model.
The following result about Lévy processes plays an important role in the choice of process made to describe the deterioration of the components:

Theorem 3.2.8. *(Lévy -Khintchine Formula) A function ϕ is the characteristic function of an infinitely divisible distribution if and only if it has the form*

$$\phi(u) = \exp\{\Psi(u)\},\ u \in \mathbb{R}$$

where

$$\Psi(u) = iau - \frac{1}{2}\sigma^2 u^2 - \int \left(1 - e^{iux} + iux\mathbf{1}_{\{|x|<1\}}\right)\mu(dx)$$

for some real $a, \sigma \geq 0$ and Lévy measure μ.

Moreover, the distribution of a Lévy process is infinitely divisible and consequently its characteristic function is given by theorem 3.2.8. These results allows the following construction of Lévy processes:

Theorem 3.2.9. *Let $(X_t)_{t\geq 0}$ be a Lévy process, then*

$$X_t = X_t^{(1)} + X_t^{(2)} + X_t^{(3)}$$

where the $X_t^{(i)}$ are independent Lévy processes: $X_t^{(1)}$ is Gaussian, $X_t^{(2)}$ is a compound Poisson process with jumps of modulus greater or equal to 1 and $X_t^{(3)}$ is a compensated countably infinite sum of jumps of modulus strictly less then 1 independent of $X_t^{(2)}$.

See [16] or [74] for more details. Hence Lévy processes may be expressed as the sum of a Brownian motion with a drift and two jump processes. If continuous paths are required, the appropriate choice of Lévy process is given by the following theorem:

Theorem 3.2.10. *(Lévy). If $(X_t)_{t \geq 0}$ is a continuous Lévy process in $\mathbb{R}^d$ then X_t is expressible in the form*

$$X_t = \sigma B_t + \mu t \tag{3.2}$$

where B_t is a Brownian motion in $\mathbb{R}^d$, $\mu \in \mathbb{R}^d$ and σ a $d \times d$ matrix.

In other words, the only continuous Lévy processes are Brownian motions with drift and therefore the choice of continuity property implies non-monotonicity: Lévy processes cannot be monotone and continuous at the same time. We note that a Brownian motion with drift can appear monotone by choosing a large drift and a small volatility. If the monotonicity property of the process is required the Brownian motion part needs to be abandoned in order to deal with the jump processes only: the compound Poisson process $X_t^{(2)}$ is usually considered, with the gamma process as a particular case (see remark 3.2.4).

3.2.2 Brownian motion with drift describing the degradation of a component

Taking into account the previous subsection, the process chosen to represent the evolution of the system's deterioration will belong to the class of Lévy processes. However, it must be specified whether the chosen process will be continuous or not. Keeping in mind the various applications of such processes in the field of reliability and maintenance, the choice must be made between the gamma process and the Brownian motion with drift. Moreover, the required tools to solve the considered models, such as the transition density functions and probability density functions of time to reach a certain threshold, are available for both of the processes, [74]. Whereas the Brownian motion is a continuous process, it is not monotone and can be negative. On the other hand, the gamma process is not continuous (since it is a jump process) but has positive jumps. Hence it is a monotone increasing positive process. As far as deterioration of a system through time is concerned, the positivity and monotonicity of the gamma process seem to be relevant properties. Indeed, it implies that the system's state through time does not improve by itself and the state of the system does not get better than when it is new (*i.e.* when it is equal to zero). These are the main reasons why a gamma process may be preferred to a Brownian motion. However, in our particular case the continuity of the Brownian motion is given preference to the appealing monotonicity of the gamma process. In the case of a process with continuous paths, the first hitting and last exit time may be determined with greater accuracy. It must be taken into consideration, however, that the deterioration of a component cannot improve by itself through time and the way the non-monotonicity of the chosen process is interpreted is as follows. It is assumed that the chosen process not only models the natural deterioration of the considered component but also takes into account the minor repairs undertaken on the component. These types of small repairs are neglected in the maintenance actions considered due to the resulting small change in the state of the process but are somehow incorporated in the model through this choice of non-monotone process. For example, one may think of the maintenance of a road, which may include rebuilding parts of the pavement or parts of the way as major maintenance actions. However, filling in small holes constitutes some sort of minor maintenance for the road that can be neglected compared to more important actions. These are the type of minor maintenance actions that are assumed to be included in the process and therefore justify its non-monotonicity. The effect of such minor repairs on the state of the component may be handled by choosing an appropriate value for the volatility term associated to the Brownian motion in comparison with the value for the drift term. The approach usually consists in choosing a large value for the drift term compared to the value of the volatility term, preventing the process from being negative. We note that the use of non-monotone processes to model crack and growth was justified by Sobczyk [82]. Moreover, Whitmore [90] used Wiener processes to model degradation. In the following section we explain under which criteria decisions are made: rather than being based on the considered processes, they depend on the value of a functional acting on an multi-dimensional Brownian motion with drift. This new process is also non-monotone and this is handled with the use of last exit times, as will be shown.

So far, we have chosen to model the degradation of each component in the system with the use of a Brownian motion with a drift and a volatility term, referred from now on as a Wiener process. The framework is as follows: let S denote the system of interest and assume that S consists of n components, $C_i, i \in \{1, \ldots, n\}$, each of which experiences its own way of deteriorating through time. The n deteriorations are assumed to be independent, *i.e.* the deterioration of any component in the system has no influence on the deterioration of the $n-1$ remaining ones. The proposed model takes into account the different n deterioration processes. Each component undergoes a deterioration described by a Wiener process. For $i \in \{1, \ldots, n\}$, $W_t^{(i)}$ denotes the

Wiener process describing the state of degradation of component C_i, where

$$\begin{aligned} W_t^{(i)} &= \mu_i t + \sigma B_t^{(i)} \\ W_0^{(i)} &= 0 \end{aligned} \tag{3.3}$$

with $\mu_i, \sigma \in \mathbb{R}^+$.

The above Wiener processes have different drift terms (the μ_i's) but the volatility terms (σ) are identical and each component is assumed to be new at time $t = 0$. The independence of the degradations is modelled by considering n independent Brownian motions $B_t^{(i)}$'s.

The next step consists in considering the following n-dimensional Wiener process:

$$\begin{aligned} \mathbf{W}_t &= \left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)}\right) \\ &= \underline{\mu} t + \sigma \mathbf{B}_t \\ \mathbf{W}_0 &= \underline{0} \end{aligned} \tag{3.4}$$

with

$$\underline{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \mathbf{B}_t = \begin{pmatrix} B_t^{(1)} \\ \vdots \\ B_t^{(n)} \end{pmatrix} \tag{3.5}$$

Two strong assumptions are made: the first being that the degradation of each of the components can be described by a Wiener process with drift and the second being the consideration of a similar volatility term for all the considered Wiener processes. Justifications for the first assumption have already been given and the assumption on the volatility terms is purely made for simplifications purposes and tractability of the models. The initial thoughts were to consider different drift terms and volatility terms for each of the considered Wiener process, making the model more general: each component would experience its own way of deteriorating through time. Difficulties immediately arose when dealing with this scenario. Indeed, in the case of possibly different volatility terms σ_i, equation (3.4) needs to be re-written as:

$$\begin{aligned} \mathbf{W}_t &= \left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)}\right) \\ &= \underline{\mu} t + \Sigma \mathbf{B}_t \\ \mathbf{W}_0 &= \underline{0} \end{aligned} \tag{3.6}$$

with $\underline{\mu}$, $\mathbf{B}_t$ defined as above and

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & \dots & & 0 \\ 0 & \sigma_2 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 0 & \sigma_n \end{pmatrix} \tag{3.7}$$

This corresponds to a multi-dimensional Wiener process with a diagonal volatility matrix $\mathbf{\Sigma}$ (since the Brownian motion are assumed to be independent). As explained in the following section, maintenance decisions for the models are made with the use of a performance metric, which is a functional of the considered n-dimensional process. The complexity arises when performing the desired functional (the Euclidean norm) on that kind of process. This is explained in the final chapter, where possible methods (such as considering the distribution of the trace of a Wishart matrix) to solve that issue are presented as part of extensions to the thesis.

A simplifying way of dealing with this case of different volatility terms is to group components whose corresponding Wiener process have the same variance into subsystems and look at these subsystems individually. Consider the same system S, where the degradation of C_i is now modelled by:

$$\begin{aligned} W_t^{(i)} &= \mu_i t + \sigma_i B_t^{(i)}, \\ W_0^{(i)} &= 0 \end{aligned} \tag{3.8}$$

for $1 \leq i \leq n,\ \mu_i,\ \sigma_i \in \mathbb{R}^+$.

We then perform a grouping of the components into different classes. Classes of components differ according to the volatility σ_i chosen to model their degradation. If $K_1, \ldots, K_p$ denote the p different classes which the degradation models of the components fall into, we classify components in the following way:

$$C_i \in K_j \Leftrightarrow W_t^{(i)} = \mu_i t + \sigma_j B_t^{(i)} \tag{3.9}$$

Clearly $1 \leq p \leq N$ and $K_i \cap K_j = \emptyset$, for $i \neq j,\ 1 \leq i, j \leq p$. Note that only the value σ_j for the volatility term, and not the value μ_i of the drift, is taken into account for the classification of the components. A justification for such a classification is the presence of identical components in almost all large systems, hence whose degradation may possibly be modelled with the use of an identical volatility term.

Let's now focus on a particular class of components, K_i say. Assume that K_i contains

n_i components. Relabelling them, assume that these n_i components are $C_1, C_2, \ldots, C_{n_i}$. The corresponding degradation processes considered are:

$$\begin{aligned} W_t^{(1)} &= \mu_1 t + \sigma_i B_t^{(1)} \\ W_t^{(2)} &= \mu_2 t + \sigma_i B_t^{(2)} \\ &\vdots \\ W_t^{(n_i)} &= \mu_{n_i} t + \sigma_i B_t^{(n_i)} \end{aligned} \tag{3.10}$$

With initial values:

$$W_0^{(j)} = 0, \text{ for } 1 \leq j \leq n_i$$

Hence:

$$W_t^{(j)} \sim N_1(\mu_j t, \sigma_i^2 t), \ \forall\, 1 \leq j \leq n_i$$

Let:

$$W_t^{K_i} = \left(W_0^{(1)}, W_0^{(2)}, \ldots, W_0^{(n_i)}\right)$$

where the superscript K_i refers to the considered class of components.
Clearly:

$$W_t^{K_i} = \underline{\mu}^{K_i} t + \sigma_i \mathbf{B}_t^{K_i}, \ W_0^{K_i} = \underline{0} \tag{3.11}$$

with

$$\underline{\mu}^{K_i} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_{n_i} \end{pmatrix}, \ \mathbf{B}_t^{K_i} = \begin{pmatrix} B_t^{(1)} \\ \vdots \\ B_t^{(n_i)} \end{pmatrix} \tag{3.12}$$

3.3 The performance metric

The process in hand that describes the state of the system is an n-dimensional Wiener process. However, this does not really reflect the overall amount of deterioration. We are not interested in making maintenance decisions based on the individual state of deterioration of each component but we wish to be able to deal with situations such as: all the components are in a satisfactory state but the overall performance of the system is sufficiently low so that the system is considered as unsafe. Moreover, the only available information on the system's state at inspection time is often given by evaluating a performance measure, the true state of degradation of the components is not known. For this a functional is applied to the multi-dimensional process (recall the examples given in section 3.1) that encapsulates the overall degradation of the system.

The functional considered in this thesis is the Euclidean norm and rather than looking at the bivariate process $(X_t, \|X_t\|_2)$, decision are made using the functional $\|X_t\|_2$ itself. We shall define Bessel processes and more particularly Bessel processes with drift and state some of their properties relevant for the rest of the thesis. We then explain how such processes are used as a performance measure in our models.

3.3.1 Properties of the Bessel process

This section aims at defining the Bessel process. Some important definitions and properties are stated. An extensive amount of work on such processes has been published and among the many contributions we refer the reader to the work of Shiga and Watanabe [80], Revuz and Yor [73], Pitman and Yor [70], Going-Jaeschke and Yor [31] and Dufresne [29]. The field of applications of the Bessel process is quite wide and ranges through finance (option pricing) [21], health related issues [15] and queuing theory [48]. What we propose here is the use of such processes in the field of reliability and maintenance, something that as far as we know does not seem to have been considered in the past.

Bessel process

A natural and convenient way to define the Bessel process is to define it as the square root of another process, namely the square of a Bessel process.

Definition 3.3.1. For every $n \geq 0$ and $x_0 \geq 0$ the unique strong solution to the equation

$$X_t = x_0 + nt + 2\int_0^t \sqrt{|X_s|}dB_s$$

is called the square of an n-dimensional Bessel process started at x_0 and is denoted by $BesQ^n_{x_0}$.

Definition 3.3.2. The square root of $BesQ^n_{x_0^2}$, $n \geq 0$, $x_0 \geq 0$ is called the Bessel process of dimension n started at x_0 and is denoted by $Bes^n_{x_0}$.

For $n > 1$, a $Bes^n_{x_0}$ process X_t satisfies $E\left[\int_0^t \left(\frac{ds}{X_s}\right)\right] < \infty$ and is the solution to the equation

$$X_t = x_0 + \frac{n-1}{2}\int_0^t \frac{ds}{X_s} + B_t, \; n > 1 \tag{3.13}$$

We will use the number $\nu \equiv \frac{n}{2} - 1$, called the index of the process, and write $Bes_{x_0}(\nu)$ instead of $Bes^n_{x_0}$.
The process X_t is called the n-dimensional Bessel process because its generator is the Bessel differential operator

$$Af(x) = \frac{1}{2}f''(x) + \frac{n-1}{2x}f'(x),$$

see [65].

Remark 3.3.3. The Bessel process can also be defined for $n \leq 1$, [31]. This case is less simple and will not be treated here since it is not relevant to our work.

The reason why Bessel processes caught our attention is that when the dimension n is equal to an integer greater or equal to two and $x_0 = 0$ (the process starts at 0), this process is the radial part (Euclidean norm) of an n-dimensional Brownian motion, *i.e.* let $B_t^{(1)}, \ldots, B_t^{(n)}$ be n mutually independent one-dimensional Brownian motions and $\mathbf{B}_t = \left(B_t^{(1)}, \ldots, B_t^{(n)}\right)$, then

$$\begin{aligned} X_t &= \|\mathbf{B}_t\|_2 \\ &= \sqrt{\sum_{j=1}^{n} \left(B_t^{(j)}\right)^2} \end{aligned} \tag{3.14}$$

defines a Bessel process with index $\nu \equiv \frac{n}{2} - 1$, [73], [80]. From now on we shall only consider Bessel processes that correspond to the Euclidean norm of a multi-dimensional Brownian motion, hence assuming that $n \in \mathbb{N}, n \geq 2$ and $x_0 = 0$.
Another important result dealing with the dimension of the process is given in the following theorem.

Theorem 3.3.4. *For a Bessel process of dimension $n \in \mathbb{N}, n \geq 1$, the point 0 is*

(i) a reflecting boundary if $n = 1$,

(ii) an entrance boundary if $n \geq 2$.

Equivalently, for $n \geq 2$ the process will never reach 0 for $t > 0$, and for $n = 1$ reaches zero almost surely, [37] and [80].
Before stating any other relevant properties of Bessel processes we first need to give a few definitions.

Definition 3.3.5. (Filtration). By a filtration $\{\mathcal{G}_t : t \in \mathbb{R}^+\}$ on the probability triple $(\Omega, \mathcal{G}, \mathbb{P})$, we mean an increasing family of sub-σ-algebras of $\mathcal{G}$:

$$\text{for } 0 \leq s \leq t,\ \mathcal{G}_s \subseteq \mathcal{G}_t \subseteq \mathcal{G}_\infty := \sigma\left(\bigcup_{u\in\mathbb{R}^+} \mathcal{G}_u\right) \subseteq \mathcal{G}$$

The setup $(\Omega, \mathcal{G}, \mathbb{P}, \{\mathcal{G}_t : t \in \mathbb{R}^+\})$ is then called a filtered space. A σ-algebra $\mathcal{G}_t$ being the collection of events which may occur before or at time t, *i.e.* the set of possible pasts up to time t.

Let $(\Omega, \mathcal{G}, \mathbb{P}, \{\mathcal{G}_t, t \in \mathbb{R}^+\})$ be a filtered space. A $\{\mathcal{G}_t\}$-stopping time may be defined as

Definition 3.3.6. (Stopping time). A map $T : \Omega \to [0, +\infty]$ is called a $\{\mathcal{G}_t\}$-stopping time if

$$\{T \leq t\} := \{\omega : T(\omega) \leq t\} \in \mathcal{G}_t,\ \forall\, t \leq \infty$$

where $\mathcal{G}_t$ is a sub-σ-algebras of $\mathcal{G}$.

In other words, it should be possible to decide whether or not $T \leq t$ has occurred on the basis of the knowledge of $\mathcal{G}_t$, the history up to time t.
Bessel processes are Markov processes (they even enjoy the strong Markov property, [73]) with continuous paths in $\mathbb{R}^+$, [73], and clearly inherit the non-monotonicity property from the Brownian motions. Moreover, for $n \geq 3$ such processes are transient. More precisely (with $\nu = \frac{n}{2} - 1,\ n \in \mathbb{N}$):

Theorem 3.3.7. *Let X_t be a $Bes_{x_0}(\nu)$ with $\nu > 0$. Then:*

$$\lim_{t\to\infty} X_t = +\infty$$

almost surely.

We now introduce two processes associated to the Bessel process , which are going to play a major role in the models developed later. Namely the first hitting time and the last exit time

Definition 3.3.8. Let X_t be a $Bes_0(\nu)$, with $\nu > 0$ and $\xi \in (0, +\infty)$.

1. By first hitting time of the threshold ξ for the process X_t starting from x_0, we mean the process:

$$G_\xi^{x_0} = \inf_{t\in\mathbb{R}^+} \{X_t = \xi | X_0 = x_0\}$$

2. By last exit time from the interval $[0, \xi)$ for the process X_t starting from x_0, we mean the process

$$H_{\xi}^{x_0} = \sup_{t \in \mathbb{R}^+} \{X_t \leq \xi | X_0 = x_0\}$$

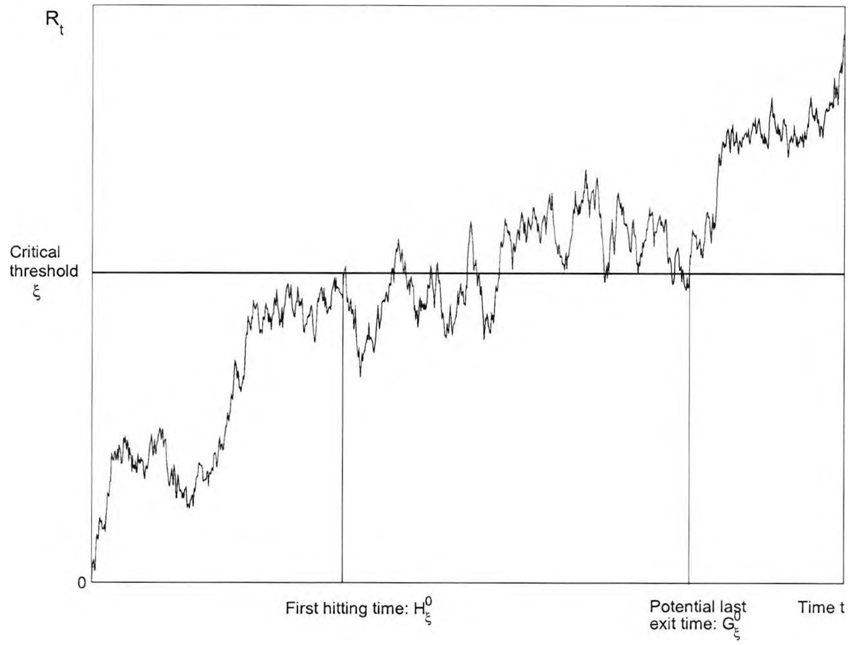


Figure 3.1: First hitting time and potential last exit time for a Bessel process.

Figure 3.1 shows both up and down-crossings of the critical threshold ξ by a Bessel process, with the first hitting time and a potential last exit time.

Remark 3.3.9.

(i) The existence of both the last exit time and first hitting time for a Bessel process is justified by theorem 3.3.7.

(ii) Note that the first hitting time is a stopping time, whereas the last exit time for a Bessel process is not a stopping time.

(iii) As mentioned earlier, Bessel processes are not increasing processes. However they are processes that 'tend' to increase. By this it is meant that for a fixed threshold value ξ, there exists time T_ξ after which the process will always be greater than ξ, *i.e*:

$$\forall\, t \geq T_\xi,\ X_t \geq \xi$$

The transition density function for Bessel processes and the probability density functions for both the first hitting time and last exit time are known, [39], [70]. A part of Patie's thesis, [66], concentrates on the first time a Bessel process hits a curve (whose equation is given by $f \in C\left(\mathbb{R}_0^+, \mathbb{R}^+\right)$): an explicit expression in the case $f(t) = a + bt, t \geq 0,\ a > 0,\ b \in \mathbb{R}$ is given. Shao and Yin derived an expression for the joint distribution of the future infinimum and its location for a transient Bessel process, [79]. Other interesting and useful properties, such as the inversion property and the additivity property of Bessel processes were derived by Shiga and Watanabe, [80]. To avoid any redundancy in the thesis, these are not given in this section but can be found in the referenced literature. We prefer to state these properties and give the analytic expressions for the transition and probability density functions in the more general case of a Bessel process with drift. The reason for this being that these are the type of processes that are considered in the rest of the thesis.

Bessel process with drift

In this section and the rest of the thesis, R_t will denote a Bessel process with drift. By Bessel process with index $\nu \geq 0$ and drift $\mu \geq 0$, we mean a diffusion with values in $\mathbb{R}^+$ and generator:

$$A = \frac{1}{2}\frac{d^2}{dx^2} + \left[\frac{2\nu+1}{2x} + \frac{\partial}{\partial x}\log\psi_\nu\left(\mu, x\right)\right]\frac{d}{dx}$$

where

$$\begin{aligned} \psi_\nu\left(\mu, x\right) &= \left(\frac{\mu x}{2}\right)^{-\nu}\Gamma\left(\nu+1\right)I_\nu\left(\mu x\right), \text{ if } \mu x > 0 \\ &= 1, \text{ if } \mu x = 0 \end{aligned} \tag{3.15}$$

We use the notation $Bes_{x_0}\left(\nu, \mu\right)$ to denote such a process starting at x_0.

Remark 3.3.10. Note that $Bes_{x_0}\left(\nu, 0\right)$ corresponds to $Bes_{x_0}\left(\nu\right)$.

The terminology 'drift' is being used since in the particular case where $x_0 = 0$ and $n \in \mathbb{N}$, $Bes_0\left(\nu, \mu\right)$ corresponds to the radial part (Euclidean norm) of an n-dimensional Brownian motion with drift, [69], [75], where $\nu = \frac{n}{2} - 1$. More precisely, let $\mathbf{W}_t$ be an n-dimensional Brownian motion ($n \geq 2,\ n \in \mathbb{N}$) with a drift $\underline{\mu} = (\mu^{(1)}, \ldots, \mu^{(n)})$:

$$\begin{aligned} \mathbf{W}_t &= \underline{\mu} t + \mathbf{B}_t \\ &= \begin{pmatrix} \mu^{(1)} \\ \vdots \\ \mu^{(n)} \end{pmatrix} t + \begin{pmatrix} B_t^{(1)} \\ \vdots \\ B_t^{(n)} \end{pmatrix}, \end{aligned} \tag{3.16}$$

where the $B_t^{(i)}$ are independent Brownian motions. Then

$$R_t = \|\mathbf{W}_t\|_2 = \sqrt{\sum_{j=1}^{n}\left(\mu_i t + B_t^{(j)}\right)^2} \tag{3.17}$$

is a $Bes_0(\nu, \mu)$, with

$$\mu = \|\underline{\mu}\|_2 = \sqrt{\sum_{j=1}^{n}\mu_i^2} \tag{3.18}$$

As mentioned earlier, we will be working with that particular kind of Bessel process only, *i.e.* making the assumptions $x_0 = 0$ and $n \in \mathbb{N}$, $n \geq 2$.

Bessel processes with drift have continuous non-monotone paths in $\mathbb{R}^+$. They are Markov processes (and also enjoy the strong Markov property), result of theorem 3.3.4 still apply here and so does the transience property, [70].

We now state two relevant properties for $Bes_0(\nu, \mu)$. The first one dealing with the sum of such processes and the second with the inversion of time.

Theorem 3.3.11. *Let $(X_t)_{t\geq 0}$ be a $Bes_0(\nu_1, \mu_1)$, and let $(Y_t)_{t\geq 0}$ be an independent $Bes_0(\nu_2, \mu_2)$, where $\nu_1, \nu_2 > 1$ and $\mu_1, \mu_2 \geq 0$. Then the process*

$$\left(\sqrt{X_t^2 + Y_t^2},\ t > 0\right)$$

is $Bes_0\left(\nu_1 + \nu_2 + 1, \sqrt{\mu_1^2 + \mu_2^2}\right)$

Theorem 3.3.12. *(Watanabe). For all $\nu > -1$, μ_1, $\mu_2 \geq 0$ a process $(R_t,\ t > 0)$ is a $Bes_{x_0}(\nu, \mu)$ if and only if $\left(tR_{\frac{1}{t}},\ t > 0\right)$ is $Bes_\mu(\nu, x_0)$.*

Both of the above theorems were first derived by Shiga and Watanabe in the case of a $Bes_{x_0}(\nu)$, [80]. Pitman and Yor proposed these extensions to the case of Bessel processes with drift in [70].

The expressions for the transition density functions of such processes are known and are given now. For this let $p_t^{\nu,\mu}(x, y), t > 0$, denote the transition density from state x to state y for a Bessel process with drift

- For $x, y > 0$:

$$p_t^{\nu,\mu}(x, y) = \frac{y}{t}\frac{I_\nu(\mu y)}{I_\nu(\mu x)} I_\nu\left(\frac{xy}{t}\right) e^{-\left(\frac{x^2+y^2+\mu^2 t^2}{2t}\right)}$$

- For $x = 0$ and $y > 0$:

$$p_t^{\nu,\mu}(0, y) = \left(\frac{1}{\mu}\right)^{\nu} \left(\frac{y}{t}\right)^{\nu+1} I_\nu(\mu y) e^{-\left(\frac{y^2+\mu^2 t^2}{2t}\right)}$$

where I_ν denotes the modified Bessel function of the first kind. The second expression (for $x = 0$) is derived from the first one using the series expansion for the modified Bessel functions of the first kind, we refer to Appendix A for details.

In a similar way that we defined the last exit time for a $Bes_0(\nu)$ we may define the last exit time for a $Bes_0(\nu, \mu)$. Using the same notation, the expressions for the probability density functions of $H_\xi^{x_0}$ are:

- For $\mu > 0, x_0 > 0, \nu \geq 0$

$$\mathbb{P}^{\nu,\mu}\left[H_\xi^{x_0} \in dt\right] = \frac{I_\nu(\frac{x_0 \xi}{t}) e^{-\left(\frac{x_0^2+\xi^2+\mu^2 t^2}{2t}\right)}}{2t I_\nu(\mu x_0) K_\nu(\mu \xi)} .dt \tag{3.19}$$

- For $\mu > 0, x_0 = 0, \nu \geq 0$

$$\mathbb{P}^{\nu,\mu}\left[H_\xi^{0} \in dt\right] = \frac{\xi^\nu e^{-\left(\frac{\xi^2+\mu^2 t^2}{2t}\right)}}{2t (\mu t)^\nu K_\nu(\mu \xi)} .dt \tag{3.20}$$

where K_ν denotes the modified Bessel function of the second kind. Explanations on how the above expressions are derived can be found in [69] and [70]. Expressions for the density functions of the first time a Bessel process with drift hits a line $[y = \alpha t]$, with slope $\alpha > 0$, are given in [70]. However, as far as we know, no analytical expression for the density functions of the first time a Bessel process with drift hits a fixed threshold is given in the literature. Since some of the models considered later require the expression for such a density functions, several attempts were made to derive such an expression but did not produce any convincing results. Nevertheless, in [70] Pitman and Yor (and Yin in [91]) managed to derive an expression for the Laplace transform of such a density function. Let L_ξ denote the first time $Bes_0(\nu, \mu)$ hits the value ξ, it was derived that:

$$\mathbb{E}_0[e^{-\frac{1}{2}\beta^2 L_\xi}] = \left(\frac{\sqrt{\beta^2+\mu^2}}{\mu}\right)^{\nu} \frac{I_\nu(\mu \xi)}{I_\nu\left(\xi\sqrt{\beta^2+\mu^2}\right)}, \quad \nu \geq 0 \tag{3.21}$$

The expression in the case of $x_0 > 0$ can also be found in the referenced papers: it is not given here since it is not relevant to the model developed in this thesis.
Symbolic computations for the inverse of the Laplace transform given by equation (3.21) did not produce any coherent results because of singularities, due to the presence of the modified Bessel function $I_\nu\left(\xi\sqrt{\beta^2+\mu^2}\right)$ at the denominator. The same problem arose when considering rational approximations, such as the Padé approximation. Since we were unable to find an expression for the density function of the first hitting time and were in possession of its Laplace transform only, a second thought was to solve the proposed cost models in the Laplace domain rather than in the time domain. If the expression for the density function of the first hitting time is only known in the Laplace domain, expressions for the Laplace transform of the transition density function for a Bessel process with drift can be computed, and so can the Laplace transform of the last hitting time, [70]. However, this change of domain required the computation of convolutions in the Laplace domain (due to the presence of products of functions of time in the time domain), adding to the model some extra complexity. This idea was dropped and the way the problem was handled was to consider a numerical inversion of this Laplace transform. This was done using the EULER method proposed by Abate and Whitt in [2]. Simulations for the first hitting times were performed and compared to the results obtained with the numerical approximation and the conclusions drawn were more than satisfactory. Description of the method, explanation on how it was used and comparisons with simulation results can be found in Appendix B.
We end this section on Bessel processes by stating a theorem relating a Brownian motion, its maximum and a three dimensional Bessel process, [75].

Theorem 3.3.13. *Let $(B_t, t > 0)$ be a Brownian motion on the line with drift μ and $B_0 = 0$. Let*

$$\begin{aligned} M_t &= \max_{0\leq s\leq t} B_s, \\ Y_t &= 2M_t - B_t \end{aligned} \tag{3.22}$$

Then the process $(Y_t, t \geq 0)$ is a time homogeneous diffusion identical in law to the radial part of a three dimensional Brownian motion with drift of magnitude $|\mu|$, started at the origin.

Figure 3.2 illustrates the three considered processes of theorem 3.3.13. The special case of theorem 3.3.13 with no drift was first established by Pitman, [68].

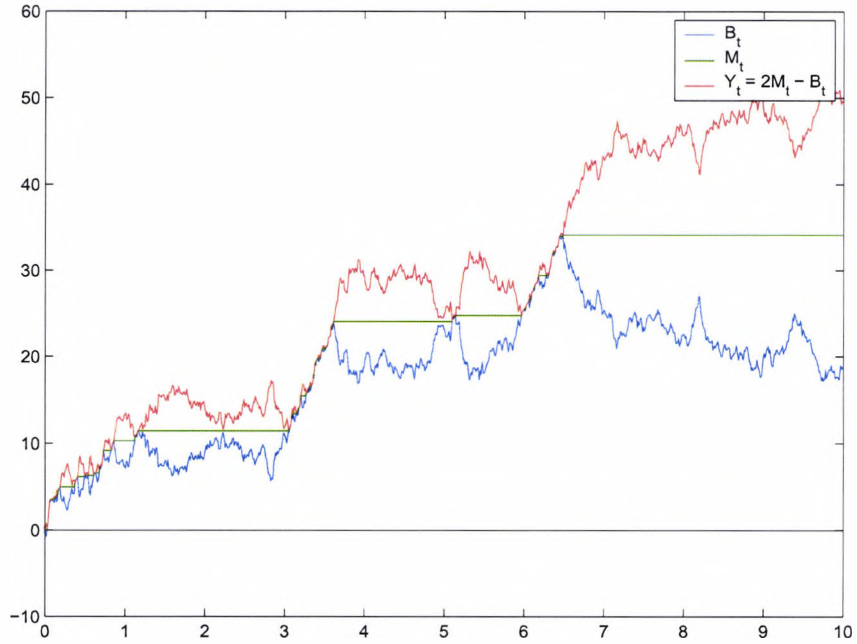


Figure 3.2: The processes considered in theorem 3.3.13.

Remark 3.3.14. The reader should not confuse the process of interest, the Bessel process with drift , with what is called the Bessel process with *naive* drift: the latter was introduced in [92] and corresponds to a Bessel process with an added drift (and not the radial part of an $\mathbb{R}^n$ Brownian motion with drift).

3.3.2 The Bessel process as a performance measure

Recall from section 3.2.2 that the following n-dimensional Wiener process was considered:

$$\begin{aligned} \mathbf{W}_t &= \left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)}\right) \\ &= \underline{\mu} t + \sigma \mathbf{B}_t \\ \mathbf{W}_0 &= \underline{0} \end{aligned}$$

where $W_t^{(i)}$ is used to model the deterioration of component C_i. As mentioned, the degradation process used to describe a performance measure of the system is a functional on the underlying n-dimensional process $\mathbf{W}_t$. The chosen functional is the radial part R_t of $\mathbf{W}_t$:

$$R_t = \|\mathbf{W}_t\|_2 = \sqrt{\sum_{i=1}^{n}(W_t^{(i)})^2} \tag{3.23}$$

R_t is the radial part of a drifting Brownian motion with volatility term σ starting at the origin: hence using the results of section 3.3.1, it corresponds to the process $\sigma Bes_0(\nu, \mu)$: σ times a Bessel process starting at the origin with index ν and drift μ, [73, 75], where:

$$\nu = \frac{n}{2} - 1, \qquad \mu = \sqrt{\sum_{i=1}^{N}\mu_i^2} \tag{3.24}$$

Figure 3.3 illustrates the sample paths for three Wiener processes with drifts and the corresponding Bessel process with drift.

Remark 3.3.15.

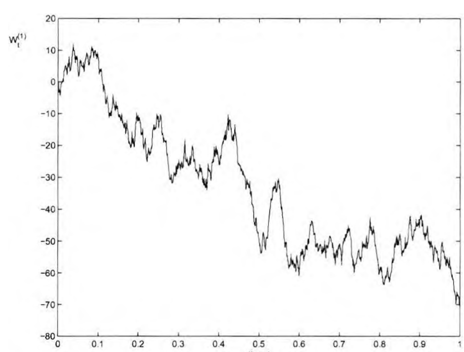
(i) Note that $\forall\ \sigma \neq 0$

$$\begin{aligned} P\left[R_t < A\right] &= P\left[\|\mathbf{W}_t\|_2 < A\right] \\ &= P\left[\left\|\sigma(\frac{1}{\sigma}\mu t + \mathbf{B}_t)\right\|_2 < A\right] \\ &= P\left[\|\sigma\|_2 . \left\|\frac{1}{\sigma}\mu t + \mathbf{B}_t\right\|_2 < A\right] \\ &= P\left[\left\|\frac{1}{\sigma}\mu t + \mathbf{B}_t\right\|_2 < \frac{A}{\|\sigma\|_2}\right] \end{aligned} \tag{3.25}$$

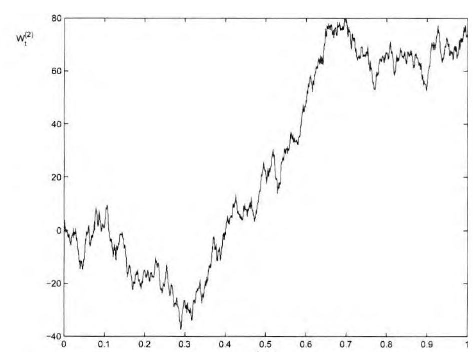
Therefore, without loss of generality, we may assume that $\sigma = 1$ for the rest of the thesis.

(ii) We emphasis on the fact that the radial part of a Brownian motion with drift starting from x_0 corresponds to a Bessel process with drift $Bes_{x_0}(\nu, \mu)$ only if $x_0 = 0$.

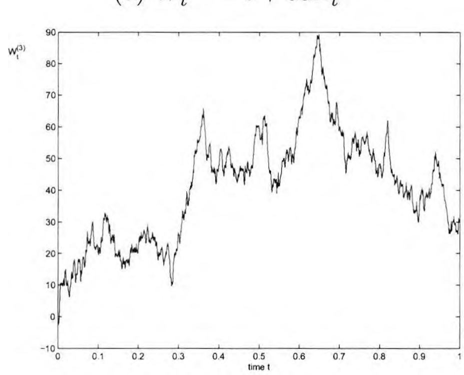
Remark 3.3.15(ii) will play a major role in the next chapter where expressions for the expected costs of maintenance are derived. Indeed, we will see that, rather than considering a recursive argument on the state of the system after maintenance, the



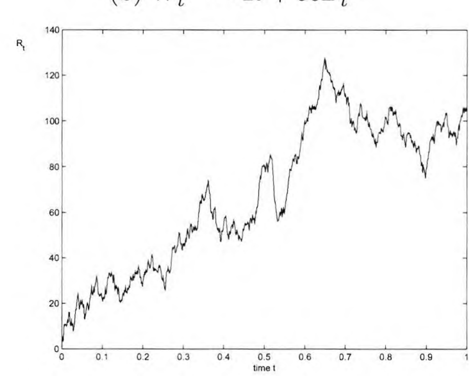
(a) $W_t^{(1)} = t + 50B_t^{(1)}$



(b) $W_t^{(2)} = 2t + 50B_t^{(2)}$



(c) $W_t^{(3)} = 3t + 50B_t^{(3)}$



(d) $R_t = 50Bes_0\left(0.5, \sqrt{\frac{14}{2500}}\right)$

Figure 3.3: Wiener processes with drift and corresponding Bessel process with drift R_t.

models consider a recursive argument on the critical threshold levels.

The reasons why the Bessel process with drift was chosen as a performance measure of the system are the following. First of all we decided to deal with Wiener processes to describe the state of deterioration of the components in the system (the reasons are explained in section 3.2.2), hence the functional had to be a functional acting on Wiener processes. The desired functional needed to preserve the continuity property satisfied by the Wiener processes, as this was the main reason for choosing them as our underlying deterioration processes. Moreover, the state space of the process needs to be bounded below in order to define a point (the lower bound) as a starting point for the process (corresponding to the state of a new system). This is the case for the Bessel process with drift since it has path in $\mathbb{R}^+$. Hence the performance measure's value of a new system corresponds to the process being equal to 0 and as this value increases, the system's performance is assumed to decrease. Besides, for $t > 0$, $\nu \geq 0$ the point 0 is never reached (3.3.4) meaning for the model that the system can never be as good as new unless appropriate maintenance is undertaken. The Markov property was also required in order to consider more general repair models than a simple perfect repair model. We note that it is not a monotone process: this is not a major issue since performance measures of a system need not be monotone and may fluctuate with respect to the particular usage of the system. Non-monotone trajectories usually result when system usage is followed. For example, gas turbines have a maximum power output rating and users are advised not to exceed a fixed percentage of the maximum in normal operation. Use above the advised output indicates extra maintenance at the next service action. Built in test equipment records this type of information and reports it to a diagnostic computer.

The models derived in Chapters 5 and 6 use a control limit policy, based on one or two thresholds (a replacement threshold and a failure one). However, these differ from the usual control limit rules, as explained in Chapter 4, due to the consideration of non-stopping times but also due to the multi-dimensionality of the underlying process $\mathbf{W}_t$ describing the deterioration of all the components in the system. To explain this, we consider the case of a two dimensional Bessel process illustrated in figure 3.4. Let X_t and Y_t be the two Wiener processes with volatility term set to be equal to one and R_t the corresponding Bessel process. Figure 3.4 shows the two planes in which both of the Wiener processes evolve and clearly illustrates the control limit rule considered here. As decisions for the maintenance of the system are entirely based on the Bessel process , our interest lies in times at which the process escapes from a cylinder, with base centered at the origin and radius equal to the critical threshold ξ. For clarity reasons, the illustrating figure shows the first hitting time of the cylinder rather than the last exit time. The angle at which the cylinder is hit is irrelevant here. If more than one threshold is considered, the different non-overlapping areas between cylinders may be seen as different levels of performance of the system and we may define the cylinder with base of greatest radius as representing the critical failure threshold. Figure 3.5 illustrates the fact that even if the individual Wiener processes have not reached a critical value, the functional (the Bessel process) may itself exceed a critical value. Recall that this is the reason why we decided to base our decisions using this functional rather than the n-dimensional Wiener process itself.

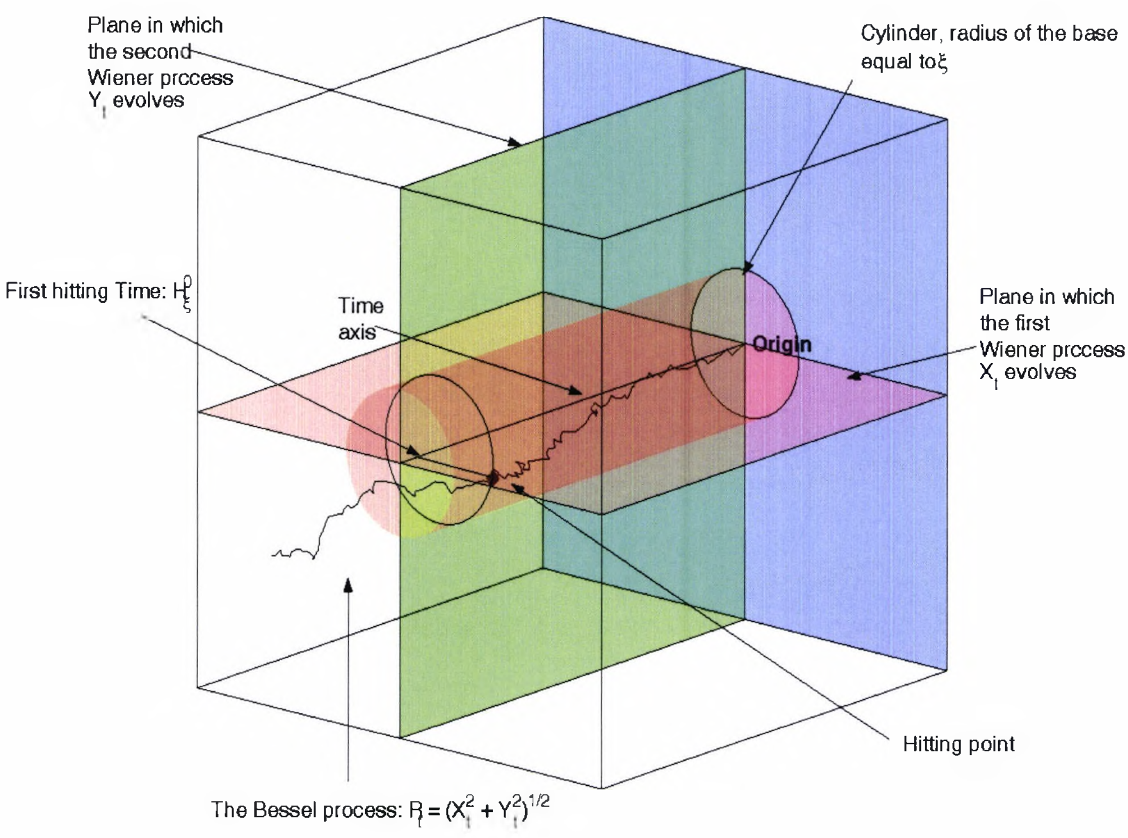


Figure 3.4: Two-dimensional Bessel process and first hitting time of a cylinder.

The interesting fact about theorem 3.3.11 is its application in the case where a new set of components is added to the considered system and this can be seen as an upgrade

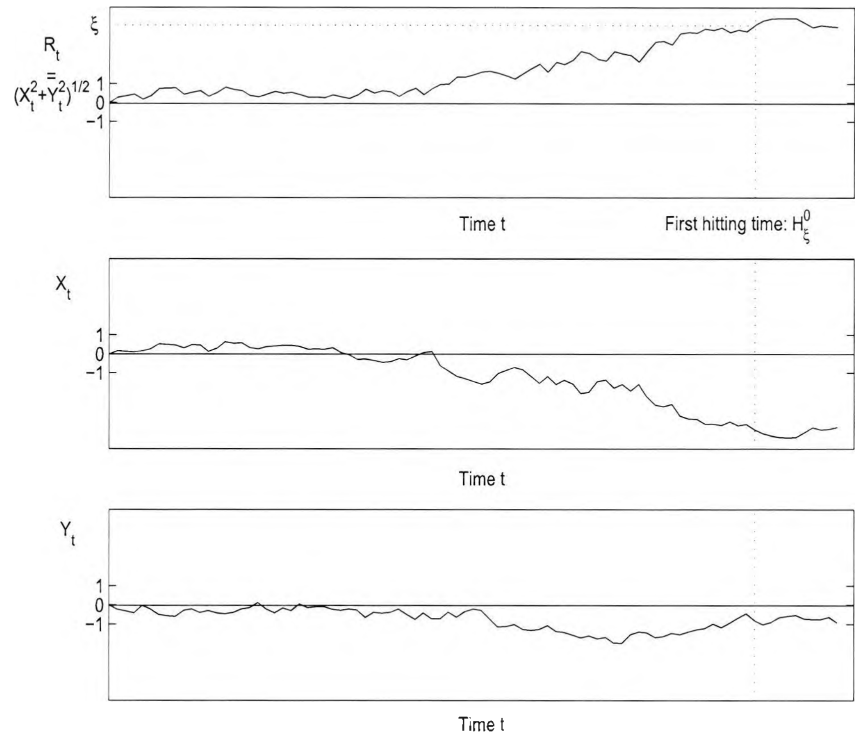


Figure 3.5: First hitting time for a 2-dimensional Bessel process and corresponding position of the two underlying Wiener processes.

for the system. The new components' deteriorations however must be modelled with the same common volatility σ, but are allowed different drifts (note that one may have $\mu_1 \neq \mu_2$ in theorem 3.3.11). Moreover we note that the number of such new components is irrelevant, as in theorem 3.3.11 one may have $\nu_1 \neq \nu_2$. From the theorem, the new Bessel process with drift to be considered is $Bes_0(\nu_1 + \nu_2 + 1, \sqrt{\mu_1^2 + \mu_2^2})$. This result makes Bessel processes with drift attractive for cases where upgrades on the system are to be considered.

3.4 Summary

A system consisting of n-components is considered. The aim was to decide on an appropriate stochastic process that would represent a performance measure of the system, taking into account the deteriorations undergone by each of the components present in the system. The first concern was to decide on an appropriate stochastic process that would be used to model the state of deterioration of a component in the system. The independence property of the increment and the Markov property required for that kind of process led us to the class of Lévy processes. Moreover, among Lévy processes, we preferred continuity of the paths of the process to monotonicity, hence leaving us to deal with Wiener processes: the choice of degradation process for the components was then a Wiener process, where the volatility terms are assumed to be the same but the drift terms may differ. The following step was to consider the corresponding n-dimensional Wiener process, consisting of the n 1-dimensional Wiener processes chosen. This represents the underlaying system's deterioration process but does not really reflects the total amount of deterioration. The performance of the system through time was then chosen to be described by the Euclidean norm of this multi-dimensional process. This corresponds to a Bessel process with drift when the starting point is 0. The reasons why the Euclidean norm was chosen is that it preserves the desired continuity property and its state space is $\mathbb{R}^+$: the point 0 then corresponds to the performance of a new system. Moreover this point is never reached again in the case where $n \geq 2$ suggesting that the performance of the system will never be as good as when it is new, unless maintenance actions are undertaken on the system. Eventually, we note that the performance measure still inherits the non-monotonicity property of the Wiener process. However it is a transient process when $n \geq 3$ and the non-monotonicity of the performance measure is handled by looking at the last hitting times of particular thresholds.

Chapter 4

Methodology

4.1 Introduction

In most earlier work the deterioration state of a system is described by a univariate stochastic process X_t whose performance as represented by 'X_t must meet some specified requirement'. The problem is then usually formulated as repair the system on inspection if its state has not crossed a critical threshold, and to replace if the system has exceeded the critical threshold. The policy is defined as a series of inspection instants with a decision rule that determines the action to take after observing the system. Many authors have restricted the modelling to the family of Lévy processes to retain the Markov property. Since a requirement continuity of sample paths restricts the Lévy process to the non-monotone Wiener process attention has been focused on retaining monotonicity through the use of jump processes and, in particular, the gamma process (this is further discussed in Chapter 3). Others have used the Wiener process but generally force almost monotone behaviour by ensuring that the volatility is much smaller than the drift, [90]. In both the Wiener and gamma process, the first hitting time distributions for the time to cross a critical threshold are readily obtained. The approach can be extended to non-monotone processes by using the maximum process $M_t = \max\limits_{0 \leq s \leq t}\{X_s\}$, [61]. Because the maximum process is monotone, the apparatus of the standard models becomes available. The use of first hitting times also brings a simplification to the modelling because they are stopping times. We further extend and relax the structure of the model to allow for a multi-variate state description, $\mathbf{W}_t \in \mathbb{R}^n$. When the system is inspected a performance measure is calculated. The performance measure is a functional on the underlying process: $R_t = \mathcal{A}(\mathbf{W}_t)$. The performance measure is no longer required to be monotone and to simplify the analysis a different set of criteria for the decision process is introduced. The new approach is to define a critical threshold which determines the response to an inspection. Because we now wish to ensure a minimum level of reliability is maintained we set the critical threshold at an acceptable level and examine the probability that the system will never return to this level after crossing it. The idea is that the system may exceed the critical level, but recover back to or below it. Eventually, the system may cross and never return to the acceptable level. When this occurs, the system is aging in such a way that it needs to be repaired or replaced.

This chapter sets the methodology used to derive the proposed models of Chapters 5 and 6. Since decisions are now based on a last exit time of a process, the first section introduces the notion of non-stopping time in contrast with the usual first hitting time. A thorough description of the considered maintenance actions and inspection strategies then follows.

4.2 Last exit time

The framework proposed in this thesis is based on last exit times. The novelty is that unlike the first hitting time, the last exit time is not a stopping time.

First recall the definition of a stopping time. Let $(\Omega, \mathcal{G}, \mathbb{P}, \{\mathcal{G}_t, t \in \mathbb{R}^+\})$ be a filtered space. A $\{\mathcal{G}_t\}$-stopping time may be defined as

Definition 4.2.1. (Stopping time). A map $T : \Omega \to [0, +\infty]$ is called a $\{\mathcal{G}_t\}$-stopping time if

$$\{T \leq t\} := \{\omega : T(\omega) \leq t\} \in \mathcal{G}_t, \ \forall \ t \leq \infty$$

where $\mathcal{G}_t$ is a sub-σ-algebras of $\mathcal{G}$.

In other words, it should be possible to decide whether or not $T \leq t$ has occurred on the basis of the knowledge of $\mathcal{G}_t$.

From the above definition, we may clearly deduce that the first hitting time is a stopping time. Recall that the last exit time is defined as

$$G_\xi^{x_0} = \inf_{t \in \mathbb{R}^+} \{X_t = \xi | X_0 = x_0\}$$

To see this with the non-monotone stochastic process R_t representing the system's performance, let $\mathcal{G}_t = \sigma\{R_s, s \leq t\}$ be the past history of the process up to time t. If the system is inspected at time $t = \tau$, on the basis of the history of the process up to time τ it is possible to say whether the process has gone over the threshold ξ before τ or not. Hence, using the same notation as above, $G_\xi^{x_0}$ is a $\{\mathcal{G}_\tau\}$-stopping time, $\forall\, \tau > 0$. Thus, stopping times are said to be observable. On the other hand, dealing with the last exit time requires a bit more care. Recall that the first hitting time is defined as

$$H_\xi^{x_0} = \sup_{t \in \mathbb{R}^+} \{X_t \leq \xi | X_0 = x_0\}$$

Knowledge of the past history of the process is now not sufficient to decide whether such a time has occurred or not. Last exit times are therefore non-stopping times. To see this in our particular case, think of the system being inspected at time $t = \tau$. If at inspection time, the process is below the threshold, one may deduce that the last exit time is yet to happen. However, if the process lies above the threshold, the past history only says that the process has escaped and maybe gone back to the interval $[0, \xi)$ a certain amount of time but no information on whether the process will continue doing so is available. Without knowledge of the future history of the process, an up-crossing at time τ cannot be classified as a last exit time : such times are said to be non-observable. However, probabilities such as $\mathbb{P}\left[R_\tau > \xi | H_\xi^0 > \tau\right]$ may be calculated. The particularity and originality of the proposed framework lies in the fact that decisions are taken on the basis of the realization of a non-stopping time. However, we have just mentioned that such times are non-observable, hence the natural question that arises is

'How can a decision be based on the occurrence of such an event if it is not observable?'
To tackle this difficulty, we will base our decision strategy on the probability of occurrence of such events

$$\mathbb{P}\left[H_\xi < \tau\right]$$

This approach is still a natural way to handle the problem and occurs frequently in practice: as decisions cannot be based on non-observable events, these use the computable probabilities of occurrence. Therefore, the models will propose expressions for both the cost of maintenance and the expected cost of maintenance: numerical results for the latter only will be computed.

Remark 4.2.2. This problem does not arise when we consider first hitting times only, since they are stopping times and hence are observable.

4.3 Maintenance actions

4.3.1 Maintenance function: d

The different types of maintenance actions considered are now introduced. As mentioned in section 2.3, these include both replacement and repairs. Replacements of the system simply consist in considering a new system whose performance is assumed to be maximal. After a replacement, the system's performance measure described by the process R_t is set back to the initial value 0, corresponding to the performance of a new system.

The way repairs are modelled requires more attention, as it differs from the usual scheme. The common and natural way consists in considering the state of the process R_τ at inspection time τ and to decrease it to a new level R_{τ^+}. If we assume that repair on the system improves its state of deterioration, the difference $R_\tau - R_{\tau^+}$ is positive and corresponds to the amount of repair undertaken on the system. This amount can be deterministic or random depending on the type of maintenance considered. After such a maintenance action, the state of the system is assumed to be equal to R_{τ^+} which corresponds to a new starting point for the process R_t. When the system is replaced, the process is re-set to the initial value 0, corresponding to a new system. In our particular case, the process chosen is the Euclidean norm of an n-dimensional Wiener process which corresponds to a Bessel process only when the process starts from the initial state 0, as explained in Chapter 3. Hence it is a necessity to always consider the process starting from state 0. This makes the usual repair model, described above, impossible to be considered. This is tackled by considering changes in the value of the critical threshold ξ, rather than a new starting points for the process and hence affects the time taken to traverse the distance to the critical threshold. For this we introduce a repair function which models the amount by which the threshold is lowered after undertaking a repair on the system. The function introduced is denoted by d and if $\{\tau_1, \tau_2, \ldots\}$ refer to the inspection times, d may be defined as

$$\begin{aligned} d : \mathbb{R}^+ &\to \mathbb{R}^+ \\ R_{\tau_i} &\mapsto d\left(R_{\tau_i}\right) \end{aligned} \tag{4.1}$$

It is a function of the performance measure of the system at inspection times. The choice for d is made among the set of bijective functions. The bijective property for d is required when the derived cost functions are numerically evaluated with an appropriate

choice of quadrature points, see Chapter 5.
The idea is that rather than considering R_t starting from a new initial state after the maintenance action with the same threshold value ξ, we reset the value $R_{\mathcal{T}_i}$ to 0 and consider a lower threshold $\xi' = \xi - d\left(R_{\mathcal{T}_i}\right)$. This may also be regarded as a shift of the x-axis of amount $d\left(R_{\mathcal{T}_i}\right)$ upwards. Figure 4.1 is given to illustrate this concept.

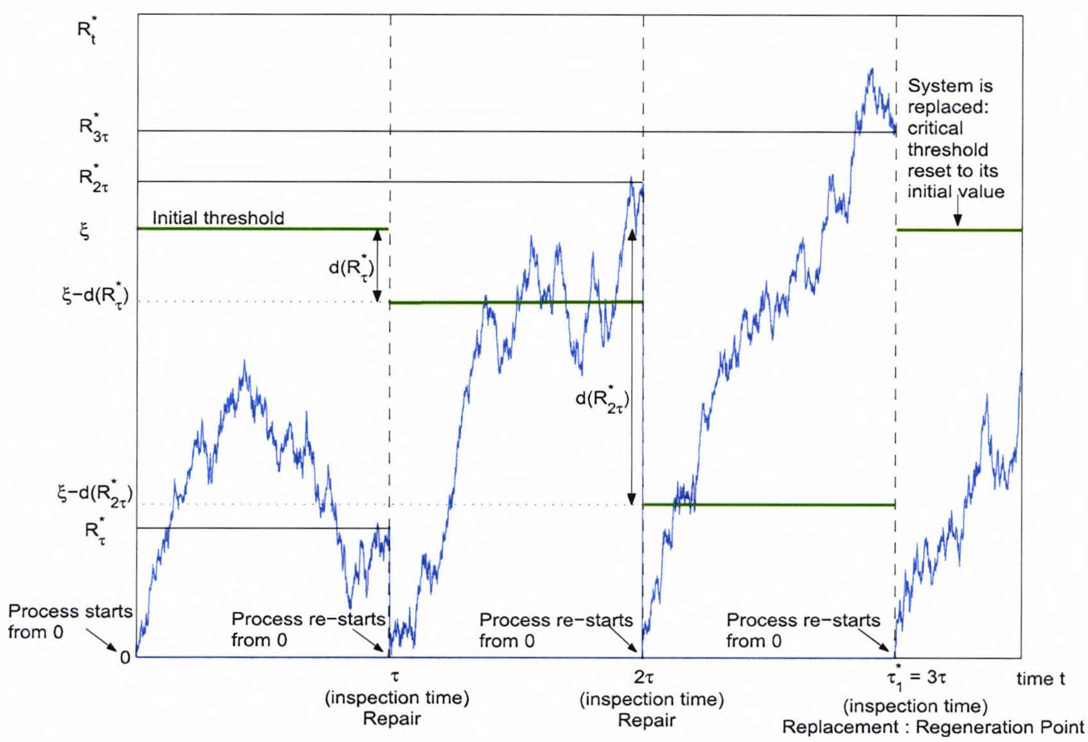


Figure 4.1: Maintenance actions with periodic inspection policy: information available at inspection times only.

Remark 4.3.1.

(i) To model the fact that repair does not worsen the performance of the system, d must satisfy the following property

$$d\left(y\right) \leq y, \forall\ y \in \mathbb{R}^{+} \tag{4.2}$$

(ii) The particular cases of no repair and perfect repair may be considered by choosing

$$\begin{aligned} d\left(y\right) &= y, \forall\ y \in \mathbb{R}^{+} \\ d\left(y\right) &= 0, \forall\ y \in \mathbb{R}^{+} \end{aligned} \tag{4.3}$$

respectively.

The amount by which the threshold is lowered must depend on the repair undertaken on the system and a little more must be said to validate this choice of repair function: this is the matter of interest in the next section. Attention was paid to the fact that the new process $R_t^{'} = R_{t+\tau_i} - R_{\tau_i}$ with initial state 0 does not correspond to a Bessel process starting at 0, and therefore this is not the process chosen to model the performance after the repair. Indeed this process may take negative values as shown in figure 4.2.

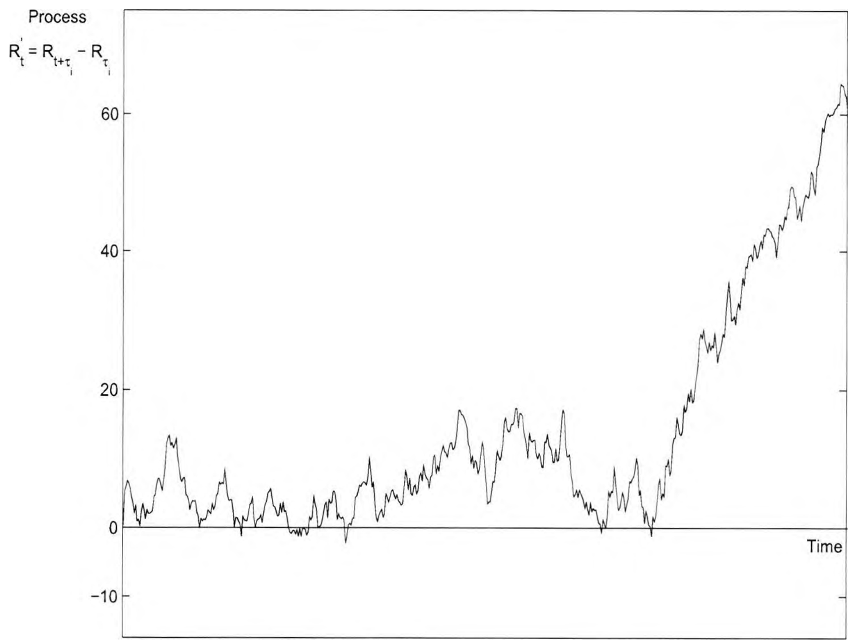


Figure 4.2: The process $R_t^{'} = R_{t+\tau_i} - R_{\tau_i}$ is not a Bessel process.

4.3.2 Component-wise repair

The way repair is physically treated on the system is now explained. The system is inspected and depending on the value of the stochastic process R_t, appropriate repair is undertaken. It is assumed that the repair takes place component-wise. To see this, let's assume that at inspection time τ_i the value of the process, before the repair action, is $R_{\tau_i} = x$. The n-dimensional Wiener process corresponding to the state of degradation of the system at inspection time and before the repair must satisfy:

$$\begin{aligned} \mathbf{W}_{\tau_i} &= \left(W_{\tau_i}^{(1)}, W_{\tau_i}^{(2)}, \ldots, W_{\tau_i}^{(n)}\right) \\ &= (x_1, x_2, \ldots, x_n) \end{aligned} \tag{4.4}$$

with

$$\|\mathbf{W}_{\tau_i}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^{\,2}} \\ = x \tag{4.5}$$

Repair is then undertaken on each of the components, resulting in a new value for $\mathbf{W}_{\tau_i^+}$ (repair is assumed instantaneous):

$$\mathbf{W}_{\tau_i^+} = \left(W_{\tau_i^+}^{(1)}, W_{\tau_i^+}^{(2)}, \ldots, W_{\tau_i^+}^{(n)}\right) \\ = (y_1, y_2, \ldots, y_n) \tag{4.6}$$

with

$$y_i \leq x_i, \ \forall \ i \in \{1, 2, \ldots, n\} \tag{4.7}$$

The value of the repair function is then set to:

$$d\left(R_{\tau_i}\right) = \sqrt{\sum_{i=1}^{n} y_i^{\,2}} \\ = y \tag{4.8}$$

Eventually, the new process considered is

$$R_t^{'} = \|\mathbf{W}_t^{'}\|_2 \\ = \|\mathbf{W}_{t+\tau_i^+} - (y_1, y_2, \ldots, y_n)\|_2 \tag{4.9}$$

The n-dimensional process $\mathbf{W}_t^{'} = \mathbf{W}_{t+\tau_i^+} - (y_1, y_2, \ldots, y_n)$ corresponds to an n-dimensional Wiener process starting at the point 0: $R_t^{'}$ is therefore a Bessel process starting from 0. With this new approach to model repairs, replacements are dealt with by re-setting the process R_t to 0 (as in the repair case) and by re-setting the critical threshold's value to its initial value ξ: the set-up becomes the one considered originally, corresponding to the case of a new system.

4.4 Inspection strategies

All the models presented in the following chapters consider both periodic and non-periodic approaches. The former allows the computation of the expected cost per unit time in an elegant way whereas in the latter the expected total cost is evaluated. The two policies require different techniques and these are explained in this section. The periodic inspection policy on one side exploits the renewal property of the maintained process and utilizes the regeneration points of the process, the non-periodic one on the other side proposes an approach based on an inspection scheduling function that determines the next inspection time.

4.4.1 Periodic inspections

In the case of periodic inspection strategies, times between inspections are assumed to be constant. Over an infinite time horizon, such a policy is denoted by

$$\Pi = \{\tau, 2\tau, 3\tau, \dots\}$$

where τ is the period of inspection. Using the same notation as above, if R_t denotes the process modelling the performance measure of the system between maintenance actions (*i.e.* R_t represents the performance of the un-maintained system), let R_t^* denote the stochastic process modelling the performance of the maintained system. As described above, upon inspection times, repairs on the system are undertaken: these affect both the critical threshold ξ by lowering its value (if required) and the process R_t^* by re-setting it to 0. After a certain amount of time, the performance of the system will fail to meet the prescribed criteria and the system will need to be replaced by a new one. The threshold value is re-set to its initial value ξ and R_t^* to 0: the overall set up is back to the initial one. Let us denote by $(\tau_i^*)_{i\in\mathbb{N}}$ the sequence of times at which the system is replaced.

Remark 4.4.1.

(i) In the particular case when information that the system's performance does not meet the criteria is only available at inspection times, one must have

$$\forall\, i \in \mathbb{N}^*,\ \exists\, j \geq i \in \mathbb{N}^*\ :\ \tau_i^* = j\tau$$

However, this may not be (and will almost never be) the case when this information is available as soon as the performance criteria are not met.

(ii) In the case where repairs are assumed to be perfect, one has

$$\forall\, j \in \mathbb{N}^*,\ \exists\, i \geq j \in \mathbb{N}^*\ :\ \tau_i^* = j\tau$$

When the system is replaced, at the time points τ_i^* ($i \in \mathbb{N}$), the value of the process R_t^* is re-set to 0 and the process restarts itself: it is therefore a regenerative process. Such instants τ_i^* are regeneration points and the sequence $(\tau_i^*)_{i\in\mathbb{N}}$ defines a renewal process. We shall say that a cycle is completed every time a renewal occurs, *i.e.* every time the system is replaced.

The renewal property of the considered process, representing the performance of the maintained system, is exploited to evaluate the expected cost per unit time. This is done evaluating the expected cost of maintenance over a cycle, the expected length of a cycle and using the renewal reward argument, [76]. Let V_ξ, L_ξ, $\mathcal{C}_\xi$ denote the cost of maintenance of a cycle, the length of a cycle and the expected cost per unit time respectively (where the subscript ξ indicates the value for the initial critical threshold), one has

$$\mathcal{C}_\xi = \frac{\mathbb{E}[V_\xi]}{\mathbb{E}[L_\xi]} \tag{4.10}$$

Thus, the problem consists in determining expressions for the expected cost of a cycle and the expected length of a cycle. This is done by considering the different conceivable scenarios at inspection times. Different frameworks are considered depending on the models investigated. The resulting expressions for both the expected cost and length of a cycle are rearranged into Fredholm equations and solved numerically. More details are provided in Chapter 5.

4.4.2 Non-Periodic inspection policy: inspection scheduling function

In the case of non-periodic inspection strategies, the time between two consecutive inspections may vary and hence is assumed not to be constant. Over an infinite time horizon, such a policy is denoted by

$$\Pi = \{\tau_1, \tau_2, \tau_3, \ldots\}$$

with

$$\tau_{i+1} - \tau_i \neq \tau_{j+1} - \tau_j \quad \text{in general.}$$

The main reason for considering non-periodic inspection policies being that they are more general and often more interesting than periodic policies, since they usually result in policies with lower costs. As mentioned in Chapter 2, rather than considering the expected cost per unit time, the expected total cost over a finite time horizon is usually evaluated. A way to deal with such policies is by considering a dynamic programming approach, such as the one proposed by Ross in [76]. This was considered by Dagg and Newby in [26] and [61], with the use of a policy improvement algorithm combined with numerical methods described in Press *et al.* [71].

The approach opted to include non-periodic inspections in our models was proposed by Grall *et al.* in [34], see also [22], [23], [33]. The optimization problem is simplified by defining an inspection scheduling function m which replaces a dynamic programming problem by one in which an optimum is sought with respect to the choice of scheduling function. m is a continuous function of the amount by which the critical threshold is decreased, $d\left(R_{\tau_i}\right)$, after a maintenance action and determines the amount of time until the next inspection time

$$\begin{aligned} m: \quad \mathbb{R}^+ \quad & \rightarrow [m_{min}, m_{max}] \\ d\left(R_{\tau_i}\right) \quad & \mapsto m\left[d\left(R_{\tau_i}\right)\right] \end{aligned} \tag{4.11}$$

with τ_i $(i \in \mathbb{N})$ denoting the times at which the system is inspected and R_{τ_i} its performance. The next inspection time τ_{i+1} is deduced using the relation

$$\tau_{i+1} = \tau_i + m\left[d\left(R_{\tau_i}\right)\right] \tag{4.12}$$

Besides, the choice for m is made among the set of decreasing functions

$$\forall\, i, j \in \mathbb{N}: \, d\left(R_{\tau_i}\right) \leq d\left(R_{\tau_j}\right) \Leftrightarrow m\left[d\left(R_{\tau_i}\right)\right] \geq m\left[d\left(R_{\tau_j}\right)\right] \tag{4.13}$$

This allows to include the idea that the lower the performance of the system is (and hence the lower the value for the new critical threshold after repair is) the more frequently it needs to be inspected. We note that the great advantage with this approach is that it preserves continuity within the model.

Remark 4.4.2. The function m ensures a finite number of inspection in a finite time interval and the transient property of the Bessel process ensures changes in the amount of time between inspections.

The approach here is to optimize the total expected cost with respect to the inspection scheduling function. The inspection functions form a two-parameter family and these two parameters, a and b, are allowed to vary to locate the optimum values. The function can be thus written $m\left[\,.\,|\, a,\, b\right]$ leading to a total expected cost function

$v_\xi(a, b)$ which is optimized with respect to a and b. The two parameters are defined in the following way

$$\begin{aligned} m\left[0 \mid a,\, b\right] &= a, \\ m\left[R_t \mid a,\, b\right] &= \alpha, \text{ if } R_t \geq b, \end{aligned} \tag{4.14}$$

for some fixed chosen value $\alpha \in [0, a]$. From the above, we may deduce that $m_{min} = \alpha$ and $m_{max} = a$. Figure 4.3 is provided to clarify the main features of the inspection scheduling functions investigated. These parameters have physical interpretations:

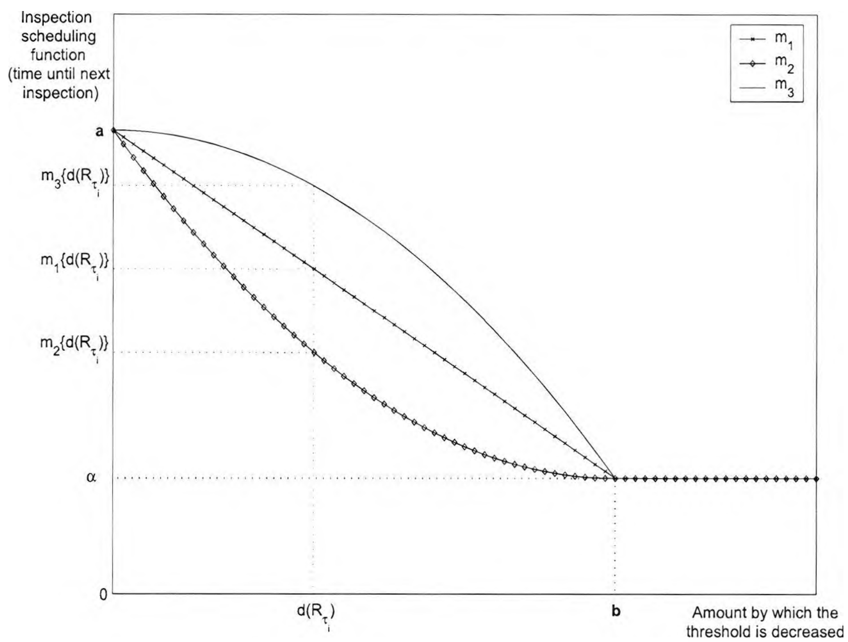


Figure 4.3: Inspection scheduling functions with different convexity properties.

(i) parameter a corresponds to the amount of time elapsed before the first inspection (*i.e.* when the system is new),

(ii) parameter b controls changes in frequency of inspections.

As the choice of inspection scheduling functions is made among the set of decreasing functions, one may deduce

$$\forall\, i \in \mathbb{N},\ \tau_{i+1} - \tau_i \leq a$$

i.e. the amount of time between any two consecutive inspections will not exceed a. Moreover, the parameter b sets a lower bound for the process R_t below which the system's performance is assumed to be insufficient, this therefore justifies a periodic inspection of the system of period α.

Choices for the appropriate inspection scheduling function are made under the main restriction of the decreasing property: this results in more inspections for a system with poor performance. To ensure tractability of the optimization and of the effects of the chosen function on the optimal cost, choices for m are confined within the set of polynomials of order less or equal to 2. We note however that the proposed models are not restricted to this choice of inspection scheduling functions and can be extended to any other type of function. Particular attention will be paid to the convexity or concavity property of m, this allows different rates of inspections as time passes to be considered. Figure 4.3 illustrates the effects of convexity on the inspection times.

4.4.3 Extensions to the inspection scheduling function

Expected observations

As mentioned in [34], the upside of the proposed way to handle non-periodic inspection is that it can be applied in practice: unlike quantities such as the lifetime distribution parameters, the performance measure R_t of the system is a quantity that can be measured on the system. However, we note that the optimal policy only reveals the next inspection time to the planner at the end of an inspection and repair action: the function m and $d(R_t)$ determine the next inspection time. Hence it allows only a 'day to day' response to the revealed system performance. A long term view of the evolution of the plan would be valuable: the whole inspection policy for a life cycle. The expected course of the policy can be indicated by calculating the expected inspection times and repair states. Let $(\tau_i)_{i\in\mathbb{N}}$ be the sequence of inspection times and let s_i be the expected performance before undertaking the maintenance at inspection time τ_i. Define $e_i = \tau_i|s_{i-1}$ as the time to inspection conditional on the performance of the process before maintenance. After the maintenance action, the process re-starts from 0 and the new threshold is set to $\xi - d(s_{i-1})$. Clearly for a process which starts from new $s_0 = 0$. The expected times e_i are calculated recursively by considering the amount by which the thresholds are lowered in the following scheme:

Expected Observation Programme:

Initialization: $s_0 = 0$

$e_1 = \tau_1 | s_0 = m(0) = a$

- - - - - - - - - -

Iteration: $s_n = E[R_{e_n} | s_{n-1}] = \int_0^{+\infty} y f^0_{e_n}(y)\, dy$

$e_{n+1} = \tau_{n+1} | s_n = m[d(s_n)]$

where f^0_τ denotes the transition density function from state 0 in an amount of time τ. The sequence $(e_i)_{i\in\mathbb{N}}$ gives the expected time of occurrence of the inspections and can be used to schedule resources.

Inspections guaranteeing a prescribed level of reliability

We note that the method proposed in 4.4.2 assumes that the system may be inspected and repaired at any moment in its lifetime. This may not be the case, particularly when system inspections and repairs require extensive dismantling or appropriate location. In order to deal with such cases, other types of inspection scheduling functions may be considered. We now propose a way to handle this type of situation by introducing a new inspection scheduling function. We mention however that such an inspection strategy will not be considered in the numerical results presented in Chapters 5 and 6. For this, assume that the amount of time between any two consecutive inspections is restricted to a finite set

$$\begin{aligned} \Pi &= \{\tau_1, \tau_2, \ldots, \tau_n\} \\ \tau_i &< \tau_j,\ \forall i, j \in \{1, 2, \ldots, n\},\ i \neq j \end{aligned} \tag{4.15}$$

Equation (4.15) assumes that any two inspections are distanced by at least τ_1 and at most τ_n units of time. Upon inspection time, the performance of the system is revealed, assume it is equal to x. Maintenance is undertaken and changes the critical threshold value to $\xi - d(x)$. The time τ at which the system will next be inspected is chosen so that the probability of occurrence of the event *'the last exit time of R_t from the interval $[0, \xi - d(x))$ is greater than τ'* is superior to $1 - \epsilon$. Low values for ϵ would usually be chosen and depend on the case studied. For instance, in straight economics ϵ is chosen in order to ensure a minimum level of availability and in the aerospace sector the value of ϵ is chosen with respect to safety regulations: in the case of aircrafts $\epsilon \sim 10^{-6}$ per hour. The proposed inspection function m chooses the greatest time that guarantees a prescribed minimum level of reliability for the system. Hence, one may define m as

$$m(x) = \sup_{\tau \in \Pi \cup \{0\}} \left\{ P\left[H^0_{\xi - d(x)} > \tau \right] > 1 - \epsilon \right\} \tag{4.16}$$

where $H^0_{\xi-d(x)}$ in equation (4.16) denotes the last exit time from the interval $[0, \xi - d(x))$ and ϵ is a fixed constant that imposes the minimum level of reliability on the model. Note that the value 0 is added to the set of possible inspection times: this is to make sure that the function m is well defined. Indeed, cases might require inspection times smaller than τ_1 units of time: in such cases the function returns 0 as an output and the system needs to be replaced.
Dealing with such a function implies the loss of the continuity property since it now corresponds to a step function

$$\begin{aligned} m(x) = {} & \tau_1 \mathbf{1}_{\left\{P\left[H^0_{\xi-d(x)} > \tau_2\right] < 1-\epsilon \leq P\left[H^0_{\xi-d(x)} > \tau_1\right]\right\}} \\ & + \tau_2 \mathbf{1}_{\left\{P\left[H^0_{\xi-d(x)} > \tau_3\right] < 1-\epsilon \leq P\left[H^0_{\xi-d(x)} > \tau_2\right]\right\}} \\ & + \ldots + \tau_n \mathbf{1}_{\left\{P\left[H^0_{\xi-d(x)} > \tau_n\right] \geq 1-\epsilon\right\}} \end{aligned} \tag{4.17}$$

The function is now fully defined and one need not decide on the appropriate shape for it. Moreover, the number of available inspection times τ_i is not restricted, allowing greater flexibility for the model and the inspection planner.

4.5 Summary

The proposed decision framework differs from the ones usually encountered in the way that it is related to the occurrence of the last exit time of the performance process from a given interval. Such a time is not a stopping time and therefore not an observable time and must be handled with care. The proposed approach is therefore based on the probability of occurrence of such an event in order to derive an expected cost of maintenance. Maintenance actions on the system are modelled with the use of a function d. Rather than giving the state of the process after maintenance, the function returns the amount by which the chosen critical threshold must be decreased. The process then restarts from 0 and a new (lower) critical threshold is considered. Both periodic and non-periodic inspections are considered. Whereas the former one is dealt with the usual renewal approach, the latter considers an inspection scheduling function m, which determines the next inspection time on the basis of the value given by the maintenance function d. The considered inspection scheduling functions form a two parameter family, which will be optimized to derive an optimal cost of maintenance.

Chapter 5

Models guaranteeing a prescribed level of reliability

5.1 Introduction

Safety regulations on particular systems require the guarantee of assigned levels of reliability and therefore models preventing a system's failure only are not sufficient: even in a working state a system may not meet the imposed safety criteria. Roads are good examples of such systems: regular maintenance actions are undertaken to keep them in a satisfactory state and even if not considered as failed, reconstructions are usually planned after a certain amount of time for safety issues. The models in this chapter exclusively focus on planning optimal inspections and maintenance actions guaranteeing a prescribed level of reliability for a complex system (the actual failure of the system is not taken into account here but will be considered in Chapter 6). By complex system it is understood a system that consists of several components or subsystems. The approach to model the degradation of the system taking into account the degradation of all of its components is the one described in Chapter 3. The system state is described by a state vector $\mathbf{W}_t$, which is not directly observed, and decisions are based on a performance measure of the system R_t defined as a functional acting on the state vector: its Euclidean norm. The decision maker uses this summary description (the performance measure), known at inspection times only, to plan the level of action which ranges from doing nothing through partial repair to complete replacement. Replacement decisions are based on the last crossings of a critical level, hence implying the use of the probability of occurrence of such an event as mentioned in Chapter 4. The critical level is defined for the performance measure itself and also as the probability of never returning to a satisfactory level of performance. The models thus give a guaranteed level of reliability throughout the life of the system. Replacement defines a regeneration point of the process, which is used, together with the costs and benefits associated with the available actions, to construct a cost function. Both cases of periodic and non-periodic inspection policies are considered. In the periodic case, the optimal strategy is determined by finding the period of inspection minimizing the expected cost per unit time whereas the non-periodic approach consists in determining an optimal inspection scheduling function resulting in a minimal expected total cost.

5.2 Periodic inspections

A periodic inspection policy $\Pi = \{\tau, 2\tau, \ldots\}$ over an infinite time interval is considered. Expressions for the expected cost per cycle and expected length of a cycle are derived in order to determine an optimal period of inspection τ^*, resulting in a minimum expected cost per unit time. The section starts with the features of the model: assumptions and settings for the model are stated. A description of the different scenarios considered in the life cycle of the system is also provided. The expression for the expected cost of maintenance per unit time is then derived and optimized with respect to the period of inspection to determine the optimal policy over an infinite time horizon. Explanations on the method used to evaluate the expression of the expected cost per unit time are given.

5.2.1 Features of the model

Model assumptions

1. The system is assumed to be new at time $t = 0$, *i.e.* $R_0 = 0$, with the value of the critical threshold being equal to ξ.

2. Inspections are periodic, perfect and instantaneous.

3. Maintenance actions are instantaneous.

4. The system's performance is only known at inspection times. However the moment at which the performance does not meet the prescribed criteria is immediately known (self-announcing): in real life situation, this may be interpreted as

an alarm being switched on as soon as the performance criteria are not satisfied. The system is then instantaneously replaced by a new one with cost C_f.

5. Each inspection incurs a fixed cost C_i.

6. Each maintenance action on the system incurs a cost determined by a function of the performance of the system at inspection time. The function is denoted by C_r.

Settings for the model

1. The state space in which the process evolves is partitioned as follows:

$$\mathbb{R}^+ = [0, \xi) \cup [\xi, +\infty), \tag{5.1}$$

where ξ is the critical threshold.

2. The inspection policy is denoted by $\Pi = \{\tau, 2\tau, \ldots, k\tau, \ldots\}$, where τ denotes the period of inspection. τ is the parameter to be optimized in order to minimize the expected cost per unit time.

3. At inspection time $t = \tau$ (prior to any maintenance action), the system's performance is R_τ.

4. Given that the system's initial performance is maximum, *i.e.* $R_0 = 0$, decisions on the level of maintenance (replacement or imperfect maintenance) are made on the basis of the indicator function $\mathbf{1}_{\{H_\xi^0 > \tau\}}$. By this it is meant that decisions on whether to replace the system or not are taken on the basis of the process having definitively escaped from the interval $[0, \xi)$ or not.

5. Deterministic maintenance actions are modelled with the use of the following maintenance function d:

$$d(x) = \begin{cases} x, & x < \frac{\xi}{K} \\ kx, & x \geq \frac{\xi}{K} \end{cases} \tag{5.2}$$

with constants $k \in (0, 1]$ and $K \in (1, +\infty)$. The constant k determines the amount of repair undertaken on the system, K is arbitrarily chosen and sets the region of repairs for the system. Recall that d must belong to the set of bijective functions.

6. The investigated function C_r is defined accordingly to the above maintenance function d:

$$C_r(x) = \begin{cases} 0, & x < \frac{\xi}{K} \\ C_{rep}, & x \geq \frac{\xi}{K} \end{cases} \tag{5.3}$$

with constants $K \in (1, +\infty)$ (same constant as for d), $C_{rep} \in (0, +\infty)$. Note that these are increasing functions of the state of the process upon inspection: the worse the performance of the system is, the more maintenance is required on the system resulting in a higher cost.

The framework

At time $t = 0$ the system's performance is assumed maximum, *i.e.* $R_0 = 0$ and the critical threshold is set to ξ. A periodic inspection policy needs to be chosen in order to minimize the expected cost per unit time. Let τ be the period of inspection that needs to be optimized. Hence, at the first inspection time, the system is in state R_τ. Two cases are considered:

1. $\mathbf{1}_{\{H_\xi^0 > \tau\}} = 1$: the system's performance remains acceptable throughout the time interval $[0, \tau]$. Appropriate maintenance, described in (5.2), is undertaken. We keep in mind that we are considering the radial part of a Brownian motion with drift hence we do not allow the Bessel process to start from any other point than the origin (see subsection 3.3.1). Imperfect maintenance is modelled by considering the same process again starting from the initial state 0 and by changing the value of the critical threshold ξ to the lower value $\xi - d(R_\tau)$, as explained in section 4.3. By doing this, we then assume that the new critical level depends on the amount of repair undertaken on the system component-wise. Therefore, rather than considering a new process starting from state $d(R_\tau)$ and the same threshold value ξ, we consider the same process $Bes_0(\nu, \mu)$ and a different threshold value equal to $\xi - d(R_\tau)$. The process we are dealing with remains a $Bes_0(\nu, \mu)$, and hence a radial Brownian motion with drift.

2. $\mathbf{1}_{\{H_\xi^0 > \tau\}} = 0$: the system's performance fails to meet the prescribed criteria before the planned inspection. The system is replaced: R_t is set back to 0 and the threshold is back to the initial value ξ.

Therefore, after the first inspection the cycle either finishes or continues. In the latter case, another inspection is undertaken at time $t = 2\tau$ but with the new value $\xi - d(R_\tau)$

for the critical threshold : the above two cases need to be considered again. This procedure repeats itself until case 2 eventually happens: this will happen with certainty due to the transience property of the Bessel process. This process describing the performance measure of the maintained system through time is a regenerative process. Times at which the process restarts from 0 (due to replacement) form a renewal process. A cycle is completed every time a replacement of the system occurs. These allow to calculate the expected cost per unit time $\mathcal{C}_\tau$ using the renewal reward argument:

$$\mathcal{C}_\tau(x) = \frac{v_\tau(x)}{l_\tau(x)} \tag{5.4}$$

which is the ratio of the expected cost per cycle over the expected length of a cycle.

5.2.2 Optimal policies

Expected cost per cycle

In this section we propose an expression for the expected cost of maintenance per cycle. The Markov property of the Bessel process allows the total cost to be expressed via a recursive approach: a conditioning argument on the threshold value is considered. The notation $V_\tau(x)$ is used to denote the cost of maintenance per cycle, where x refers to the threshold value $\xi - x$ and τ is the period of inspection. The expression for the cost per cycle is derived by considering the different possible scenarios at inspection time:

$$\begin{aligned} V_\tau(x) &= C_f \times \mathbf{1}_{\{performance\ not\ acceptable\}} \\ &\quad + \{C_i + C_r(R_\tau) + V_\tau(d(R_\tau))\} \times \mathbf{1}_{\{performance\ acceptable\}} \\ &= C_f \times \mathbf{1}_{\{H^0_{\xi-x} \leq \tau\}} + \{C_i + C_r(R_\tau) + V_\tau(d(R_\tau))\} \times \mathbf{1}_{\{H^0_{\xi-x} > \tau\}} \end{aligned} \tag{5.5}$$

We have here expressed the cost of maintenance over an inspection period as the sum of a cost of failure (if the last exit time from the interval $[0, \xi - x)$ is less then the inspection time), an inspection cost, a cost of repair and a future cost made up of the system starting from 0 with the new critical threshold after maintenance (when the last time the process hits the threshold $\xi - x$ is greater than τ). We mention here that the particular case of no maintenance can be taken into account just by setting $d(y) = y$ in the future cost.

We now take the expectation of (5.5), and get the expression for the expected cost of maintenance per cycle:

$$\begin{aligned}
v_\tau(x) &= \mathbb{E}\left[V_\tau(x)\right] \\
&= \mathbb{E}\left[C_f \times \mathbf{1}_{\{H^0_{\xi-x}\leq\tau\}}\right] + \mathbb{E}\left[\{C_i + C_r(R_\tau) + V_\tau(d(R_\tau))\} \times \mathbf{1}_{\{H^0_{\xi-x}>\tau\}}\right] \\
&= C_f \mathbb{P}\left[H^0_{\xi-x} \leq \tau\right] + \int_0^{+\infty} \{C_i + C_r(y) + v_\tau(d(y))\} \mathbb{P}\left[H^0_{\xi-x} > \tau\right] f^0_\tau(y)\, dy \\
&= C_f \int_0^\tau h^0_{\xi-x}(t)dt + \int_0^{+\infty} \{C_i + C_r(y)\} \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) f^0_\tau(y)\, dy \\
&\quad + \int_0^{+\infty} v_\tau(d(y)) \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) f^0_\tau(y)\, dy \\
&= C_f \int_0^\tau h^0_{\xi-x}(t)dt + \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \int_0^{+\infty} \{C_i + C_r(y)\} f^0_\tau(y)\, dy \\
&\quad + \int_0^{+\infty} v_\tau(d(y)) \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) f^0_\tau(y)\, dy
\end{aligned}$$

We note that equality 3 requires the computation of

$$\mathbb{E}\left[V_\tau(d(R_\tau)) \times \mathbf{1}_{\{H^0_{\xi-x}>\tau\}}\right]$$

This is done by noting that

$$\mathbb{P}\left[R_\tau < \alpha | H^0_{\xi-x} > \tau\right] \equiv \mathbb{P}\left[R_\tau < \alpha | R_\tau \text{ `can be in' } (0, \xi - x]\right],$$

hence integrating over the whole range $[0, \infty)$. More details are provided in Appendix C.

Using the fact that the maintenance function d is a bijection (hence d^{-1} exists) and that $v_\tau(x) = 0$ for $x > \xi$ (as this would correspond to a threshold with a negative value $\xi - x$, implying either that the system's performance always or never meets the prescribed criteria), the above equation may be rewritten as:

$$v_\tau(x) = Q(x) + \int_0^{d^{-1}(\xi)} K\{x, y\} v_\tau(d(y))\, dy \tag{5.6}$$

with:

$$\begin{aligned}
Q(x) = C_f \int_0^\tau h^0_{\xi-x}(t)dt + \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) C_i \\
+ \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \int_0^{+\infty} C_r(y) f^0_\tau(y)\, dy
\end{aligned} \tag{5.7}$$

$$K\{x, y\} = \left(1 - \int_0^\tau h^0_{\xi-x}(t)dt\right) f^0_\tau(y)$$

Equation (5.6) is of Fredholm type and is solved numerically, as explained at the end of this subsection.

Expected length of a cycle

The approach we use to formulate the expected length of a cycle is similar to the one adopted for the expected cost per cycle. This is also done by conditioning on the new critical threshold value. Let us first derive the expression for the length of a cycle. The notation $L_\tau(x)$ is used to denote the length of a cycle, given that the threshold value is equal to $\xi - x$ and that the period of inspection is τ.

$$\begin{aligned} L_\tau(x) &= H^0_{\xi-x} \times \mathbf{1}_{\{performance\ not\ acceptable\}} \\ &\quad + \{\tau + L_\tau(d(R_\tau))\} \times \mathbf{1}_{\{performance\ acceptable\}} \\ &= H^0_{\xi-x} \times \mathbf{1}_{\{H^0_{\xi-x} \leq \tau\}} + \{\tau + L_\tau(d(R_\tau))\} \times \mathbf{1}_{\{H^0_{\xi-x} > \tau\}} \end{aligned} \tag{5.8}$$

We now take the expectation of (5.8) to derive the expression for the expected length of a cycle:

$$\begin{aligned} l_\tau(x) &= \mathbb{E}[L_\tau(x)] \\ &= \mathbb{E}\left[H^0_{\xi-x} \times \mathbf{1}_{\{H^0_{\xi-x} \leq \tau\}}\right] + \mathbb{E}\left[\{\tau + L_\tau(d(R_\tau))\} \times \mathbf{1}_{\{H^0_{\xi-x} > \tau\}}\right] \\ &= \mathbb{E}\left[H^0_{\xi-x} \times \mathbf{1}_{\{H^0_{\xi-x} \leq \tau\}}\right] + \tau \times \mathbb{E}\left[\mathbf{1}_{\{H^0_{\xi-x} > \tau\}}\right] + \mathbb{E}\left[L_\tau(d(R_\tau)) \times \mathbf{1}_{\{H^0_{\xi-x} > \tau\}}\right] \\ &= \int_0^\tau t \times h^0_{\xi-x}(t)\, dt + \tau\left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \\ &\quad + \int_0^\infty l_\tau(d(y))\left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) f^0_\tau(y)\, dy \end{aligned}$$

Using the same argument as for the expression of the expected cost per cycle, the above equation can be rewritten as follows:

$$l_\tau(x) = P(x) + \int_0^{d^{-1}(\xi)} K\{x, y\} l_\tau(d(y))\, dy \tag{5.9}$$

with same kernel $K\{x, y\}$ as in (5.7) and:

$$P(x) = \int_0^\tau t \times h^0_{\xi-x}(t)\, dt + \tau\left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \tag{5.10}$$

Equation (5.9) is also a Fredholm type integral equation.

Expected cost per unit time

Both of the expressions for the expected cost per cycle and the expected length of a cycle have been derived, the expected cost per unit time over an infinite horizon $\mathcal{C}_\tau(x)$ can thus be calculated. For this we use the standard renewal reward argument in [76]. We get:

$$\mathcal{C}_\tau(x) = \frac{v_\tau(x)}{l_\tau(x)}, \tag{5.11}$$

with expressions for $v_\tau(x)$, $l_\tau(x)$ given in (5.6) and (5.9) respectively.

Obtaining solutions

We first note that in order to evaluate the above expressions, one needs to know the expression for f_τ^0. This can be derived from the expression of the transition density f_τ^x of a Bessel process with drift (see Appendix A) and is equal to

$$f_\tau^0(y) = \left(\frac{1}{\mu}\right)^\nu \left(\frac{y}{\tau}\right)^{\nu+1} I_\nu(\mu y) e^{-\left(\frac{y^2+\mu^2\tau^2}{2\tau}\right)}, \; \tau > 0 \tag{5.12}$$

Using the expression for h_ξ^0 given in Chapter 3, one may deduce the expressions for K, Q and P required to solve equations (5.6) and (5.9):

$$K\{x,y\} = \left(1 - \int_0^\tau \frac{(\xi-x)^\nu e^{-\left(\frac{(\xi-x)^2+\mu^2t^2}{2t}\right)}}{2t(\mu t)^\nu K_\nu(\mu(\xi-x))} dt\right) \left(\frac{1}{\mu}\right)^\nu \left(\frac{y}{\tau}\right)^{\nu+1} I_\nu(\mu y) e^{-\left(\frac{y^2+\mu^2\tau^2}{2\tau}\right)}$$

$$Q(x) = C_f \int_0^\tau \frac{(\xi-x)^\nu e^{-\left(\frac{(\xi-x)^2+\mu^2t^2}{2t}\right)}}{2t(\mu t)^\nu K_\nu(\mu(\xi-x))} dt + \left(1 - \int_0^\tau \frac{(\xi-x)^\nu e^{-\left(\frac{(\xi-x)^2+\mu^2t^2}{2t}\right)}}{2t(\mu t)^\nu K_\nu(\mu(\xi-x))} dt\right)$$
$$\times \int_0^{+\infty} \{C_i + C_r(y)\} \left(\frac{1}{\mu}\right)^\nu \left(\frac{y}{\tau}\right)^{\nu+1} I_\nu(\mu y) e^{-\left(\frac{y^2+\mu^2\tau^2}{2\tau}\right)} dy$$

$$P(x) = \int_0^\tau t \times \frac{(\xi-x)^\nu e^{-\left(\frac{(\xi-x)^2+\mu^2t^2}{2t}\right)}}{2t(\mu t)^\nu K_\nu(\mu(\xi-x))} dt + \tau \left(1 - \int_0^\tau \frac{(\xi-x)^\nu e^{-\left(\frac{(\xi-x)^2+\mu^2t^2}{2t}\right)}}{2t(\mu t)^\nu K_\nu(\mu(\xi-x))} dt\right) \tag{5.13}$$

Expressions for both the expected cost and expected length of a cycle, (5.6) and (5.9), are of the Fredholm type: discretization of the state space and application of the quadrature rules produce equivalent matrix equations as we now explain. The Nystrom routine with the N point Gauss-Legendre rule at the points $y_j \in [0, d^{-1}(\xi)), \forall j \in \{1, 2, \ldots, N\}$, is applied to (5.6) and (5.9) to get

$$\begin{aligned} v_\tau(x) &= Q(x) + \sum_{j=1}^{N} K\{x, y_j\} v_\tau(d(y_j)) w_j \\ l_\tau(x) &= P(x) + \sum_{j=1}^{N} K\{x, y_j\} l_\tau(d(y_j)) w_j \end{aligned} \tag{5.14}$$

where the set $\{w_j\}$ are the weights of the quadrature rule.

The above equations are then evaluated at the following appropriate points $x_i = d(y_i)$

$$\begin{aligned} v_\tau(x_i) &= Q(x_i) + \sum_{j=1}^{N} K\{x_i, y_j\} v_\tau(d(y_j)) w_j \\ l_\tau(x_i) &= P(x_i) + \sum_{j=1}^{N} K\{x_i, y_j\} l_\tau(d(y_j)) w_j, \end{aligned} \tag{5.15}$$

Since $v_\tau(x_i)$ and $v_\tau(d(y_i))$ (similarly for l_τ) are evaluated at the same points, equations in (5.15) may be rewritten in the following matrix form

$$\begin{aligned} \mathbf{v} &= \mathbf{Q} + \mathbf{K}.\mathbf{v} \\ \boldsymbol{l} &= \mathbf{P} + \mathbf{K}.\boldsymbol{l} \end{aligned} \tag{5.16}$$

where:

$$\begin{aligned} \mathbf{v}_i &= v_\tau(x_i) \\ \boldsymbol{l}_i &= l_\tau(x_i) \\ \mathbf{K}_{i,j} &= K\{x_i, y_j\} w_j \\ \mathbf{Q}_i &= Q(x_i) \\ \mathbf{P}_i &= P(x_i) \end{aligned} \tag{5.17}$$

Rearranging equations (5.16) gives:

$$\begin{aligned} (\mathbf{I}-\mathbf{K}).\mathbf{v} &= \mathbf{Q} \\ (\mathbf{I}-\mathbf{K}).\boldsymbol{l} &= \mathbf{P} \end{aligned} \tag{5.18}$$

where $\mathbf{I}$ denotes the $N \times N$ identity matrix. Equations 5.18 are readily solved numerically (see [71]) by computing the inverse of matrix $\mathbf{I}-\mathbf{K}$.

Remark 5.2.1. $K\{x, y\}$ in (5.13) is the product of a density function by a survival function hence it is bounded by the maximum of the density which, by the Fredholm alternative, ensures that the equations in (5.15) have a solution (*i.e.* $\mathbf{I}-\mathbf{K}$ is invertible).

The optimum policy for the system can then be determined as

$$\tau^* = \underset{\tau \in \mathbb{R}^+}{\text{argmin}} \{\mathcal{C}_\tau(0)\} \tag{5.19}$$

τ^* represents the desired optimum solution: the inspection period resulting in a minimum long run expected cost per unit time.

5.3 Non-periodic inspections

The model proposed in this section constitutes an extension of the previous one to the case of non-periodic inspections. The inspection times are now determined by a deterministic function of the system state: the inspection scheduling function introduced in subsection 4.4.2. As in the previous section, features for the model are provided. A non-periodic policy is developed by evaluating the expected lifetime costs and the optimal policy by an optimal choice of inspection function. Numerical solutions are obtained with a slight reformulation of the expression for the cost: description of the method considered is provided at the end of this section.

5.3.1 Features of the model

Model assumptions

1. Without loss of generality, it is assumed that the system's initial performance is maximum, *i.e.* $R_0 = 0$, with initial critical threshold ξ.
2. Inspections are non-periodic, perfect (in the sense that they reveal the true state of the system) and they are instantaneous.
3. Maintenance actions are instantaneous.
4. The system's performance is only known at inspection times, however the moment at which the performance does not meet the prescribed criteria is immediately known (self-announcing): the system is then instantaneously replaced by a new one with cost C_f.
5. Each inspection incurs a fixed cost C_i.
6. Each maintenance action on the system incurs a cost determined by C_r: it is a function of the performance of the system at inspection time.

Settings for the model

1. The state space in which the process R_t evolves is partitioned by a critical threshold ξ as follows:

$$\mathbb{R}^+ = [0, \xi) \cup [\xi, +\infty) \tag{5.20}$$

   Because the process R_t is non-monotone, the first time at which the process hits the threshold ξ is not considered as the time at which the system fails. Instead, we use the transience and positivity properties of the process, to define the system as unsafe when it has definitely escaped from the interval $[0, \xi)$.

2. The system is inspected at inspection times $\{\tau_1, \tau_2, \ldots\}$. The time between inspections τ_{i-1} and τ_i is T_i, $i \in \mathbb{N}$, and is determined by using the inspection scheduling function m, introduced in Chapter 4. The sequence of inspection times $(\tau_i)_{i \in Z^+}$ is strictly increasing and satisfies:

$$\begin{aligned} \tau_0 &= 0 \\ \tau_i &= \sum_{k=1}^{i} T_k \\ T_i &= \tau_i - \tau_{i-1},\ i \geq 1 \end{aligned} \tag{5.21}$$

   At inspection time τ_i, the corresponding system's state is R_{τ_i} and appropriate maintenance action (repair or do nothing) is undertaken. Let τ_i^* denote the times at which the system is replaced. These times allow us to derive an expression for the expected total cost of inspection and maintenance.

3. At inspection time $t = \tau$ (prior to any maintenance action), the system's performance is R_τ.

4. Given that the system's initial performance is maximum, *i.e.* $R_0 = 0$, decisions on the level of maintenance (replacement or imperfect maintenance) are made on the basis of the indicator function $\mathbf{1}_{\{H_\xi^0 > \tau\}}$. By this it is meant that decisions on whether to replace the system or not are taken on the basis of the process having definitively escaped from the interval $[0, \xi)$ or not.

5. Deterministic maintenance at inspection time is modelled with the use of maintenance function (5.2).

6. As the cost of maintenance strongly depends on the type of maintenance undertaken on the system, the cost function's expression must be related to the maintenance function's expression. The considered function for the cost of maintenance is given in equation (5.3).

Decision rules

The aim of the model proposed in this section is to give an optimal maintenance and inspection policy. The efficiency of the policy entirely depends on the inspection times and the type of maintenance on the system. Therefore, the different decisions considered are on the inspection times and the different maintenance actions, with their corresponding costs.
The proposed model considers a non-periodic inspection policy; the reason for this being that it is a more general approach and often results in policies with lower costs, particularly in cases where high costs of lost production are taken into consideration. Rather than considering a dynamic programming problem, as considered in [61], the optimization problem is simplified by using an inspection scheduling function m as introduced by Grall and his co-authors, [34]. Descriptions of the considered inspection functions are provided in section 4.4.2. The scheduling function is a decreasing function of $d(R_{\tau_i})$, the amount by which the threshold is decreased, and determines the amount of time until the next inspection time. Consequently, it is the state of the performance measure that determines the next inspection time. The inspection times are related in the following way:

$$\tau_{i+1} = \tau_i + m\left[d\left(R_{\tau_i}\right)\right] \tag{5.22}$$

The approach is to optimize the total expected cost with respect to the inspection scheduling function. The inspection functions form a two parameter family and these two parameters are allowed to vary to locate the optimum values. The function can be written $m\left[x \mid a, b\right]$ leading to a total cost function $v_{\xi-x}(a,b)$ which is optimized with respect to a and b. Different forms of inspection functions, based on their convexity property, are considered. The following three expressions for m are investigated

$$m_1\left[x|\,a,\,b\right] = \max\left\{1, a - \frac{a-1}{b}x\right\} \tag{5.23}$$

$$m_2\left[x|\,a,\,b\right] = \begin{cases} \dfrac{(x-b)^2}{b^2}(a-1)+1, & 0 \leqslant x \leqslant b \\ 1, & x > b. \end{cases} \tag{5.24}$$

$$m_3\left[x|\,a,\,b\right] = \begin{cases} -\left(\dfrac{\sqrt{a-1}}{b}x\right)^2 + a, & 0 \leqslant x \leqslant b \\ 1, & x > b \end{cases} \tag{5.25}$$

with $a > 1$ in all cases. Note that if $a = 1$ the policy becomes a periodic inspection policy with period $\tau = a = 1$ and in the case where $a < 1$ the policy inspects less frequently for a more deteriorated system.

Remark 5.3.1. In the rest of the thesis, the notations $m(x)$ and $v_{\xi-x}$ are used rather than $m(x|a,b)$ and $v_{\xi-x}(a,b)$, for clarity.

Plots of the three considered inspection functions are shown in figure 5.1. The function m_1 resembles to the inspection scheduling function considered in the numerical example section of [34] and constitutes a reference for our numerical results. The reasons

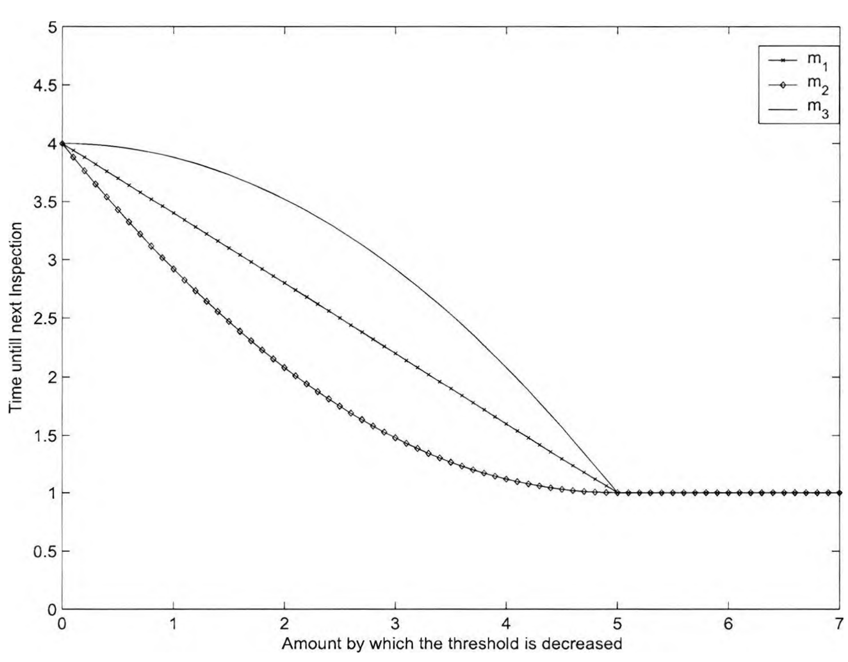


Figure 5.1: Inspection scheduling functions considered (examples with $a = 4, b = 5$).

for this particular choice of functions can be explained as follows. First of all, we note that the three functions differ in the sense that they have different convexity properties. The first one is a straight line up to the performance level $R_t = b$, the second is convex up to the performance level b and the third is concave up to level b. Whereas the time until the next inspection decreases rather quickly when dealing with m_2, m_3 allows greater time between the inspections when the state of the system is still small. The function m_2 might be thought as appropriate for a system experiencing early failures (infant mortality), whereas m_3 is more appropriate for a system that is unlikely to fail in its early age. In the three cases the parameter a corresponds to the time of the first inspection, that is when the system is new:

$$m_i\left[0\right] = a, \ i \in \{1, 2, 3\} \tag{5.26}$$

and parameter b controls change in frequency of inspections

$$m_i\left[R_t\right] = 1, \text{ if } R_t \geq b, \ i \in \{1, 2, 3\} \tag{5.27}$$

As stated in 5.3.1, decision rules for maintenance are made with the help of the maintenance function d. The maintenance actions considered are deterministic and range from doing nothing to partial repair. The case of replacement of the system is considered separately, and is not included in the maintenance function. Decisions about maintenance actions depend entirely on the state of degradation of the system at inspection, as do the related costs.

We note that random maintenance can be included in the proposed model. This can be done as in [61] by introducing an extra random variable representing the amount of repair that has been undertaken on the system. This results in an extra integral term in the expression for the total expected cost of inspection and maintenance.

The framework

The process starts with performance $R_0 = 0$ and at this time the critical threshold considered is at a distance ξ from the current point. After maintenance the system's performance is returned to $R_{\tau_i^+} = x$ (the superscript $+$ in τ_i^+ is here to indicate that it is the time just after the maintenance action, assumed to be instantaneous). The distributions for the radial part of a Brownian motion with drift is only fully specified for paths starting at the origin. Thus we use the Markov semi-group property to develop the unfolding of the distribution in time, the Bessel process is not allowed to start from any point other than the origin, see subsection 3.3.1. Hence, rather than considering the process starting at the new initial state x with critical threshold still being equal to ξ we now treat the problem differently. The following equivalent problem is considered: we assume that the process starts from new and that the value for the critical threshold is now equal to $\xi - x$. As far as the decision problem is concerned, the Markov property of the process is exploited and allows a copy of the original process to be considered:

$$\mathbb{P}[R_t < \xi \mid R_0 = x] = \mathbb{P}[R'_t < \xi - x \mid R'_0 = 0] \tag{5.28}$$

with

$$\begin{aligned} R_t &= \|\mathbf{W}_t\|_2 \\ R'_t &= \|\mathbf{W}_{\tau_i^+ + t} - \mathbf{W}_{\tau_i^+}\|_2 \end{aligned} \tag{5.29}$$

recall that $\mathbf{W}_t$ is the n-dimensional process describing the state of the system and $\|\mathbf{W}_{\tau_i^+}\|_2 = x$, as explained in section 4.3. R'_t is an equivalent process with the same probability structure and starting at the origin. In the more usual notation

$$\mathbb{P}^x[R_t < \xi] = \mathbb{P}^0[R'_t < \xi - x] \tag{5.30}$$

with the superscript indicating the starting point. The time until the next inspection is determined by the inspection scheduling function m and is equal to $m(x)$.

The decisions are made using the exit time from the region of acceptable performance. The time $H^0_{\xi-x}$ can never be known by observation since observing any up-crossing of the threshold reveals a potential exit time but there remains the possibility of a further down-crossing and up-crossing in the future. This is the meaning of the fact that $H^0_{\xi-x}$ is not a stopping time. In a non-probabilistic context, the process $H^0_{\xi-x}$ is described by a non-causal model. The difficulty is readily resolved because the probability that the last exit time occurs before the next inspection is known. In the light of these observations the decision rules are formulated as follows.

- $\mathbf{1}_{\{H^0_{\xi-x} > m(x)\}} = 1$: performance of the system (evaluated with respect to the last time the process hits the critical threshold) meets the prescribed criteria until the next scheduled inspection. Upon inspection, the system's performance is $R_{m(x)}$. The system is inspected, and a cost of inspection C_i is considered. The maintenance brings the system state of degradation back to a lower level $d\left(R_{m(x)}\right)$ with cost $C_r\left(R_{m(x)}\right)$. Future costs enter by looking at the process starting from the origin and with the new critical threshold set up equal to $\xi - d\left(R_{m(x)}\right)$. The system is then next inspected after $m\left[d\left(R_{m(x)}\right)\right]$ units of time.

- $\mathbf{1}_{\{H^0_{\xi-x}>m(x)\}} = 0$: the performance fails to meet the prescribe criteria between two inspections. The system is replaced with cost C_f and the process restarts from the origin. Future costs are then taken into consideration by looking at the process starting from the origin and with the new critical threshold set up equal to ξ.

5.3.2 Optimal policies

The Expected Total Cost

We first give the expression for the total cost and then take the expectation. This is done by considering the above different scenarios.

$$\begin{aligned} V_{\xi-x} &= \left(C_i + V_{\xi-d\left(R_{m(x)}\right)} + C_r\left(R_{m(x)}\right)\right)\mathbf{1}_{\{performance\ acceptable\}} \\ &\quad + \left(C_f + V_\xi\right)\mathbf{1}_{\{performance\ not\ acceptable\}} \\ &= \left(C_i + V_{\xi-d\left(R_{m(x)}\right)} + C_r\left(R_{m(x)}\right)\right)\mathbf{1}_{\{H^0_{\xi-x}>m(x)\}} \\ &\quad + \left(C_f + V_\xi\right)\mathbf{1}_{\{H^0_{\xi-x}\le m(x)\}} \end{aligned} \tag{5.31}$$

Taking the expectation leads to:

$$\begin{aligned} v_{\xi-x} &= \mathbb{E}[V_{\xi-x}] \\ &= \mathbb{E}\left[\left(C_f + V_\xi\right)\mathbf{1}_{\{H^0_{\xi-x}\le m(x)\}}\right] \\ &\quad + \mathbb{E}\left[\left(C_i + C_r\left(R_{m(x)}\right) + V_{\xi-d\left(R_{m(x)}\right)}\right)\mathbf{1}_{\{H^0_{\xi-x}>m(x)\}}\right] \\ &= \left(C_f + v_\xi\right)\mathbb{E}\left[\mathbf{1}_{\{H^0_{\xi-x}\le m(x)\}}\right] \\ &\quad + \mathbb{E}\left[\left(C_i + C_r\left(R_{m(x)}\right) + V_{\xi-d\left(R_{m(x)}\right)}\right)\mathbf{1}_{\{H^0_{\xi-x}>m(x)\}}\right] \\ &= \left(C_f + v_\xi\right)\mathbb{P}\left[H^0_{\xi-x} \le m(x)\right] \\ &\quad + \int_0^{+\infty}\left(C_i + C_r(y) + v_{\xi-d(y)}\right)\mathbb{P}\left[H^0_{\xi-x} > m(x)\right] f^0_{m(x)}(y)\,dy \end{aligned}$$

$$= (C_f + v_\xi) \int_0^{m(x)} h^0_{\xi-x}(t)\, dt$$
$$+ \int_0^{+\infty} \left(C_i + C_r(y) + v_{\xi-d(y)}\right) \left(1 - \int_0^{m(x)} h^0_{\xi-x}(t)\, dt\right) f^0_{m(x)}(y)\, dy$$

$$= C_i \left(1 - \int_0^{m(x)} h^0_{\xi-x}(t)\, dt\right) + (C_f + v_\xi) \int_0^{m(x)} h^0_{\xi-x}(t)\, dt \tag{5.32}$$
$$+ \left(1 - \int_0^{m(x)} h^0_{\xi-x}(t)\, dt\right) \int_0^{+\infty} C_r(y) f^0_{m(x)}(y)\, dy$$
$$+ \int_0^{+\infty} v_{\xi-d(y)} \left(1 - \int_0^{m(x)} h^0_{\xi-x}(t)\, dt\right) f^0_{m(x)}(y)\, dy$$

In (5.32) the expected value $\mathbb{E}\left[V_{\xi-d\left(R_{m(x)}\right)} \mathbf{1}_{\{H^0_{\xi-x} > m(x)\}}\right]$ is required. The expected value is derived by using the conditional independence of $H^0_{\xi-x}$ and R_τ. The independence allows the factorization of the integrals as shown in Appendix C.
Rearranging the above gives:

$$v_{\xi-x} = Q(x) + \lambda(x) v_\xi + \int_0^{d^{-1}(\xi)} v_{\xi-d(y)} K\{x, y\}\, dy \tag{5.33}$$

where:

$$\lambda(x) = \int_0^{m(x)} h^0_{\xi-x}(t)\, dt$$
$$Q(x) = (1 - \lambda(x)) \left(C_i + \int_0^{+\infty} C_r(y) f^0_{m(x)}(y)\, dy\right) + C_f \lambda(x) \tag{5.34}$$
$$K\{x, y\} = \left(1 - \int_0^{m(x)} h^0_{\xi-x}(t)\, dt\right) f^0_{m(x)}(y)$$

Note that now the limit in the integral in (5.33) is finite. The justification for this change of limit is that the expected cost $v_{\xi-x}$ is assumed to be zero when the critical threshold is negative. Indeed, a negative threshold in the model would either mean that the system never reaches a critical state or that it is always in a failed state: hence no maintenance action needs to be considered, setting the expected cost of maintenance to zero.

Determining the Policy

Equation (5.33) is solved numerically in a similar way as the one described in subsection 5.2.2, with a slight reformulation of equation (5.33) required . An approximation to the continuous problem is constructed by discretizing the integrals giving a set of linear matrix equations. The discrete problem is solved using the methods described in [71]. First, note that at $t = 0$ the system is new. Under this condition, we rewrite equation (5.33) as follows:

$$v_{\xi-x} = Q(x) + \lambda(x)\, v_{\xi-x} + \int_0^{d^{-1}(\xi)} v_{\xi-d(y)} K\{x, y\}\, dy \tag{5.35}$$

yielding to the following Fredholm equation:

$$\{1 - \lambda(x)\} v_{\xi-x} = Q(x) + \int_0^{d^{-1}(\xi)} v_{\xi-d(y)} K\{x, y\}\, dy \tag{5.36}$$

Rewriting (5.33) as (5.36) does not affect the solution to the equation and will allow the required solution to be obtained by a homotopy argument based on ξ. Indeed both equation (5.33) and (5.36) are identical when $x = 0$: we therefore solve equation (5.36) and get the solution for $x = 0$. The Nystrom routine with the N-point Gauss-Legendre rule at the points $y_j, j \in \{1, \ldots, N\}$ is applied to (5.36), we get

$$\{1 - \lambda(x)\} v_{\xi-x} = Q(x) + \sum_{j=1}^{N} v_{\xi-d(y_j)} K\{x, y_j\}\, w_j \tag{5.37}$$

We then evaluate the above at the following appropriate points $x_i = d(y_i)$ and obtain:

$$\{1 - \lambda(x_i)\} v_{\xi-x_i} = Q(x_i) + \sum_{j=1}^{N} v_{\xi-d(y_j)} K\{x_i, y_j\}\, w_j \tag{5.38}$$

which, since $v_{\xi-x_i}$ and $v_{\xi-d(y_i)}$ are evaluated at the same points, can be rewritten in the matrix form

$$(\mathbf{D} - \mathbf{K})\, \mathbf{v} = \mathbf{Q}, \tag{5.39}$$

where:

$$\begin{aligned} \mathbf{v}_i &= v_{\xi-x_i} \\ \mathbf{D}_{i,j} &= (1 - \lambda(x_i))\, \mathbf{1}_{\{i=j\}} \\ \mathbf{K}_{i,j} &= K\{x_i, y_j\}\, w_j \\ \mathbf{Q}_i &= Q(x_i) \end{aligned} \tag{5.40}$$

Having obtained the solution at the quadrature points by solving inversion of matrix $\mathbf{D}-\mathbf{K}$ (this can be done using a similar argument as in remark 5.2.1), we get the solution at any other quadrature point x by simply using equation (5.37) as an interpolatory formula. Hence since we are interested in a system which is new at time $t=0$, we just choose the quadrature point $x_i = 0$, which justifies that rewriting (5.33) as (5.36) does not affect the solution to the equation.

5.4 Numerical results, exploring the solutions

This section presents numerical results for both of the models derived in sections 5.2 and 5.3. The numerical results are obtained using the discrete versions of the optimization problem described by equations 5.18 and 5.39, with $N = 20$ points for the Gauss-Legendre rule. The values of the parameters for the process used to model the degradation of the system and the different costs used were chosen arbitrarily to show the important features of the inspection policies. The particular choice of maintenance functions d and inspection scheduling functions m (in the non-periodic case) is made clear throughout the two subsections.

5.4.1 Periodic inspections

The settings for the model are the one described in section 5.2.1. The initial value for the critical threshold is chosen to be $\xi = 10$ and the corresponding Bessel process used is $Bes_0(0.5, 1)$: this corresponds to the radial part of a three dimensional Brownian motion with drift of magnitude $\mu = 1$ and starting from the origin.
The initial values for the cost of inspection and the cost of failure are: $C_i = 50$, $C_f = 250$. The cost of repair is chosen to be dependent on the state of the system found at inspection as follows:

$$C_r(y) = \begin{cases} 0, & y < \frac{\xi}{2} \\ 100, & y \geq \frac{\xi}{2} \end{cases} \tag{5.41}$$

i.e. $C_{rep} = 100$, $K = 2$.
The maintenance actions are modelled with the following maintenance function:

$$d(y) = \begin{cases} y, & y < \frac{\xi}{2} \\ ky, & y \geq \frac{\xi}{2} \end{cases} \tag{5.42}$$

with different considered values for $k \in [0, 1]$.

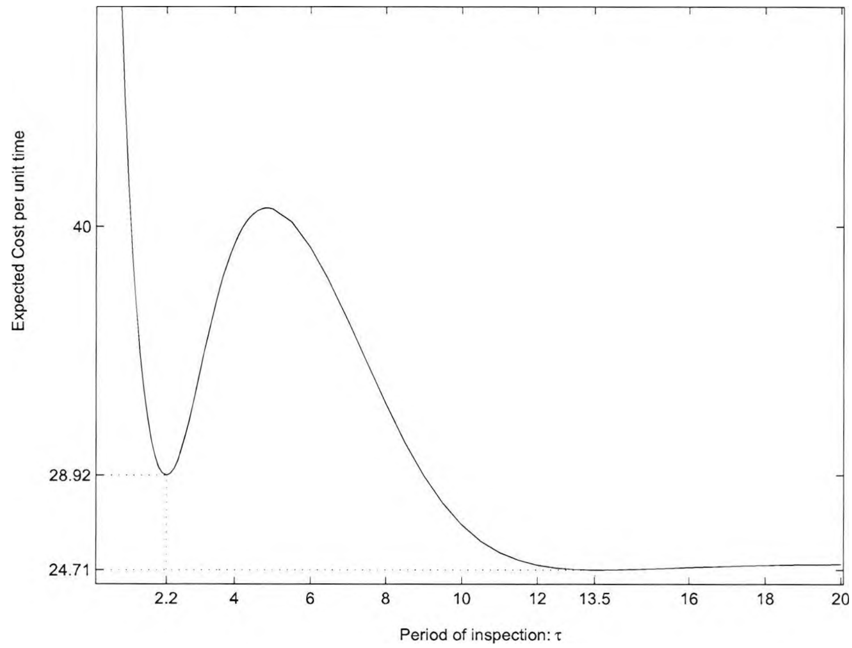


Figure 5.2: Expected cost per unit time, low maintenance action: $k = 0.9$.

Figure 5.2 shows the graph of the expected cost per unit time with respect to the period of inspection. This was obtained in the case where the parameter k in the maintenance function was set equal to 0.9: hence corresponding to a small amount of maintenance. The shape of the graph clearly shows the presence of two local minimums, one of which is a global minimum, allowing the decision maker to chose between two different strategies. The optimal strategy consists in inspecting the system every $\tau^* = 13.5$ units of time, the corresponding cost is then $v_{\tau^*} = 24.71$. However, one may wish to inspect the system more often (when safety regulations are to be considered for instance): the optimal strategy would then consist in inspecting the system every $\tau^* = 2.2$ units of time, yielding to a cost of maintenance equal to $v_{\tau^*} = 28.92$. This strategy is more expensive then the former one but inspects the system more often, hence reducing the risk of an instantaneous failure. When the cost of failure is decreased such that $C_f << C_{rep} + C_i$, the graph for the expected cost of maintenance per unit time, figure 5.3, does not show a minimum and keeps on decreasing. This suggests that for a small cost of failure compared to the sum of the cost of repair with the cost of inspection, the optimal strategy is the 'no maintenance' strategy that chooses not to inspect nor to repair the system but only to replace it upon failure. This makes sense since failure of the system incurs a lower cost C_f than $C_i + C_r$. Note however that this strategy

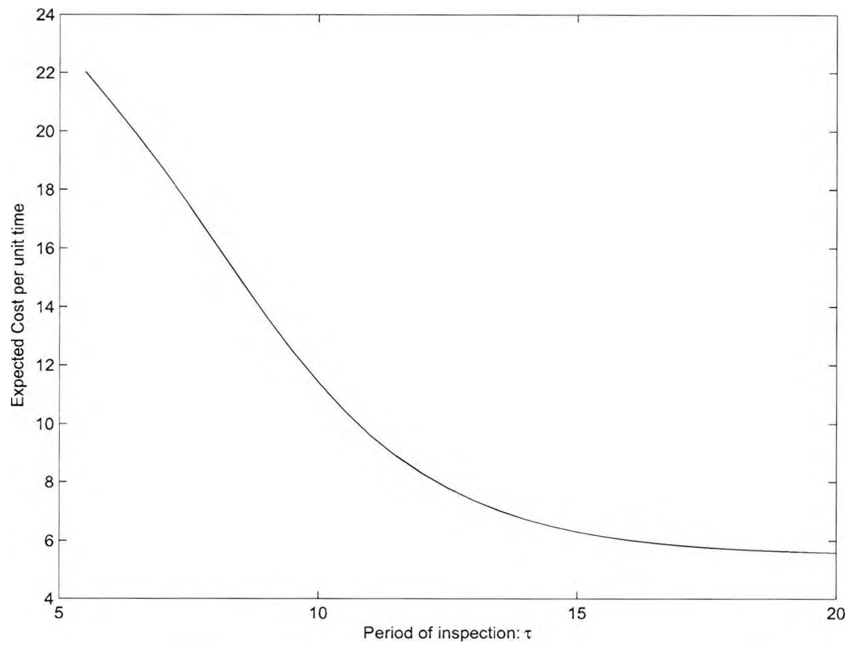
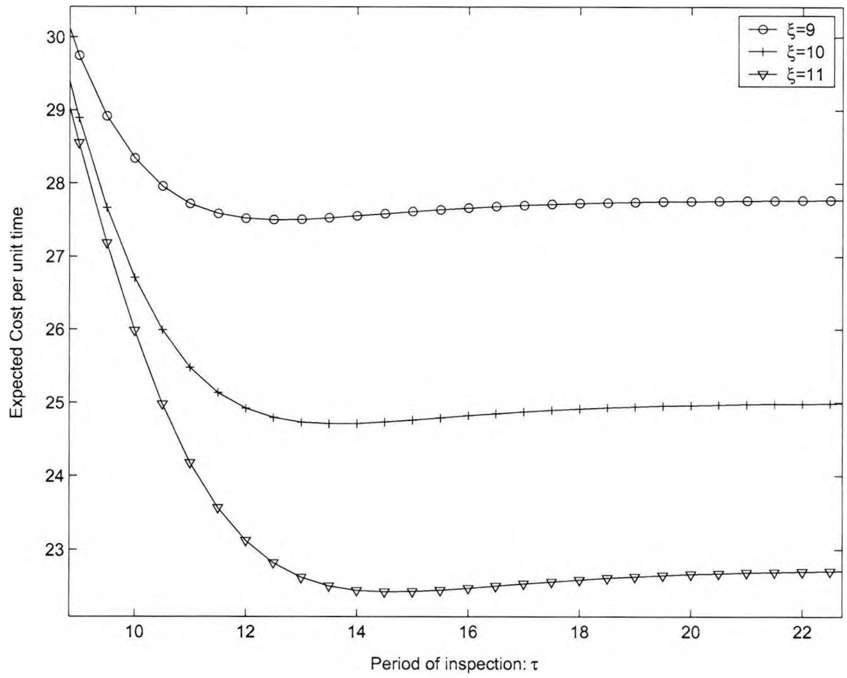
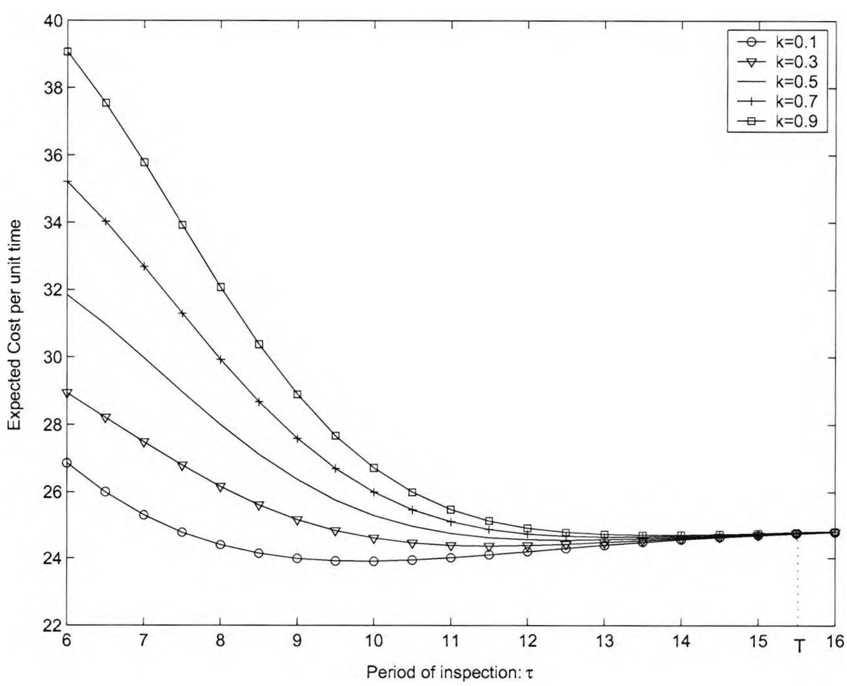


Figure 5.3: Expected cost per unit time with small cost of failure: $(C_f, C_i, C_{rep}) = (55, 50, 100)$.

does not take any safety criterion into consideration. As shown in figure 5.4(a), both the optimal expected cost per unit time and the optimal period of inspection change with the initial critical threshold ξ. Indeed, when the value for ξ increases, the optimal period of inspection increases and the cost of maintenance decreases. This is due to the transience property of the considered Bessel process: the last hitting time of a given threshold increases with the value of the threshold. The optimal expected cost per unit time decreases since for a higher threshold ξ, a greater amount of time to optimize the maintenance strategy is available. The influence of the quality of maintenance on optimal strategy is also treated. This is done by computing the expected cost of maintenance for different values of k. When k is close to 0, the maintenance tends to a 'perfect maintenance' policy (replacement of the system), whereas when k is close to 1, the maintenance approaches a 'minimal repair' policy. It is shown in figure 5.4(b) that the optimal cost decreases when the quality of maintenance increases (*i.e.* when k decreases). However, after a certain amount of time T, no distinction between the different curves can be made: if maintenance of the system can only be undertaken at a time greater than T, one may not need to consider maintaining the system thoroughly.



(a) Expected cost per unit time with different initial critical thresholds



(b) Expected cost per unit time for different maintenance strategies

Figure 5.4: Expected cost per unit time with different initial critical thresholds and for different maintenance strategies.

5.4.2 Non-periodic inspections

The model assumptions are the one described in section 5.3.1. The initial value for the critical threshold is chosen to be $\xi = 5$ and the Bessel process used is $Bes_0(0.5, 1)$. The initial values for the cost of inspection and the cost of failure are $C_i = 50$ and $C_f = 200$ respectively. The maintenance actions are modelled with the following maintenance function:

$$d(y) = \begin{cases} y, & y < \frac{\xi}{2} \\ ky, & y \geq \frac{\xi}{2} \end{cases} \tag{5.43}$$

for $k = 0.9$ (small amount of maintenance) in a first case and $k = 0.1$ (large amount of maintenance on the system) in a second case. The corresponding costs of repair are chosen to be dependent on the state of the system found at inspection as follows:

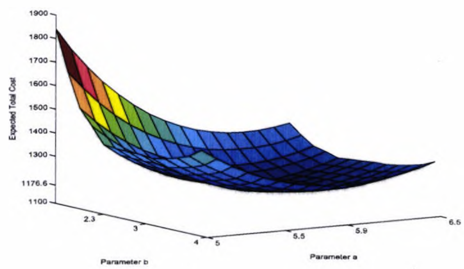
$$C_r(y) = \begin{cases} 0, & y < \frac{\xi}{2} \\ 100, & y \geq \frac{\xi}{2} \end{cases} \tag{5.44}$$

The purpose of the present model is to find an optimal inspection policy for the expected total cost of inspection and maintenance of the system over a finite time interval. Three different types of inspection policies are considered with the use of the three inspection scheduling functions m_1, m_2 and m_3 defined in subsection 5.3.1. The expected total costs are minimized with respect to the two parameters a and b.

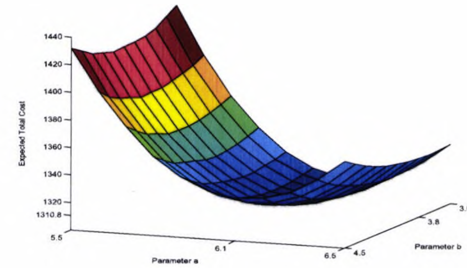
Inspection policies		k=0.9	k=0.1
	a_1^*	5.9	4.5
m_1	b_1^*	2.3	2.3
	v_1^*	1176.6	628.73
	a_2^*	6.1	4.5
m_2	b_2^*	3.8	4.7
	v_2^*	1310.8	631.71
	a_3^*	5.6	4.3
m_3	b_3^*	2.3	1.9
	v_3^*	1089.3	625.67

Table 5.1: Optimal values of the parameters a and b for the three inspection scheduling functions.

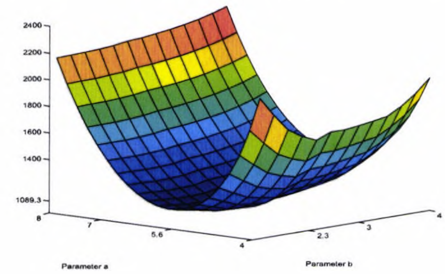
The numerical results for the case of small maintenance on the system ($k = 0.9$) are shown in figures 5.5 and 5.6. In the case of a large amount of maintenance ($k = 0.1$), the numerical results are shown in figure 5.7 and figure 5.8. The optimal values a_i^*, b_i^* and v_i^* ($i = \{1, 2, 3\}$) for a, b and v_ξ respectively, in the different scenarios, are summarized



(a) Inspection function considered: m_1

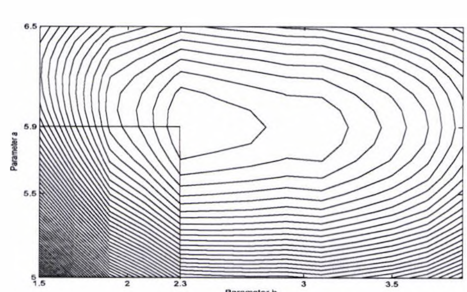


(b) Inspection function considered: m_2

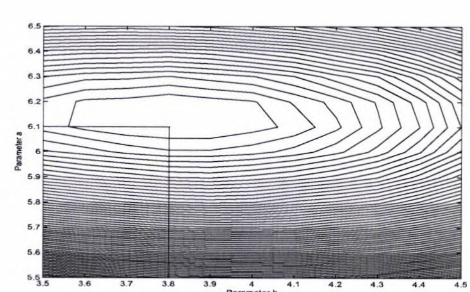


(c) Inspection function considered: m_3

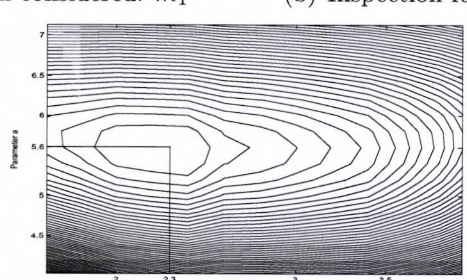
Figure 5.5: Surfaces for the expected total costs with different inspection scheduling functions, $k = 0.9$.



(a) Inspection function considered: m_1

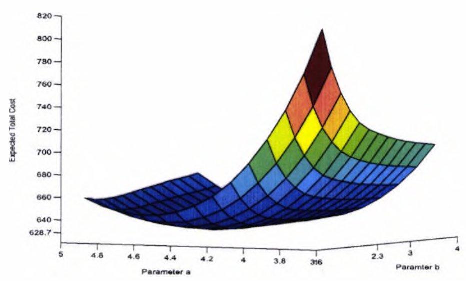


(b) Inspection function considered: m_2

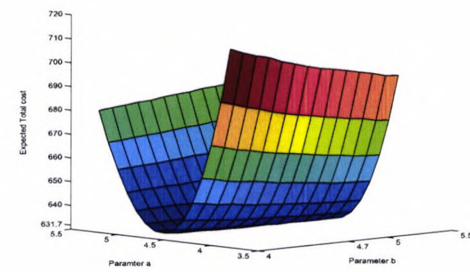


(c) Inspection function considered: m_3

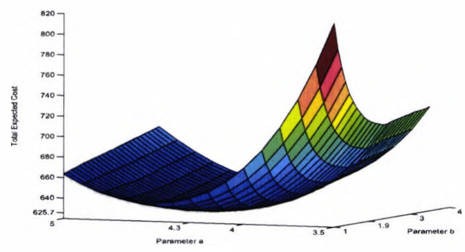
Figure 5.6: Contour representation for the expected total costs with different inspection scheduling functions, $k = 0.9$.



(a) Inspection function considered: m_1

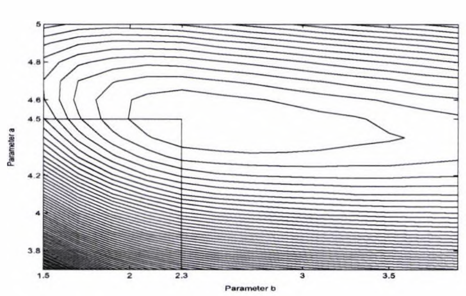


(b) Inspection function considered: m_2

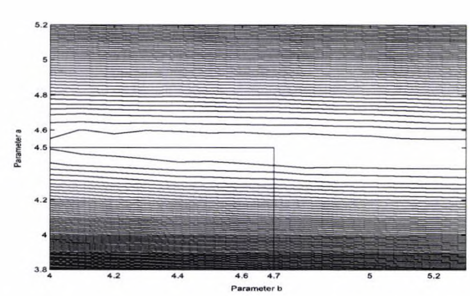


(c) Inspection function considered: m_3

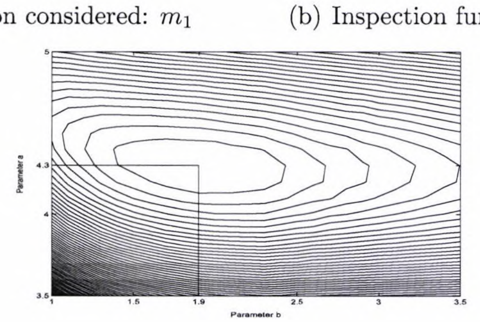
Figure 5.7: Surfaces for the expected total costs with different inspection scheduling functions, $k = 0.1$.



(a) Inspection function considered: m_1



(b) Inspection function considered: m_2



(c) Inspection function considered: m_3

Figure 5.8: Contour representation for the expected total costs with different inspection scheduling functions, $k = 0.1$.

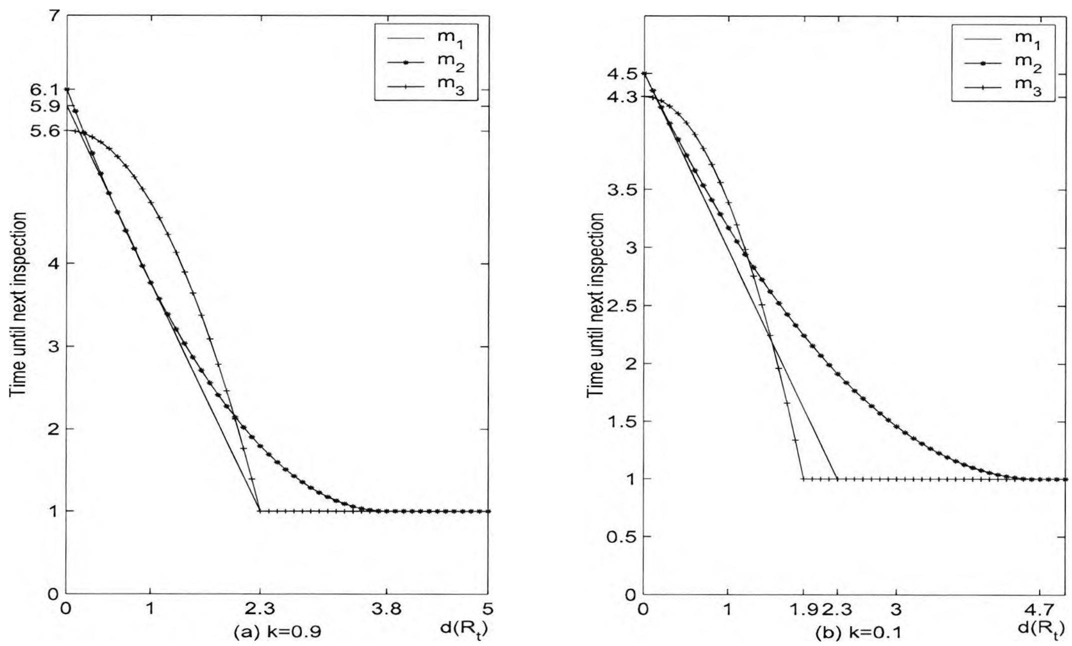


Figure 5.9: Optimal inspection scheduling functions.

in table 5.1.

We first note that the surfaces and the contours plotted in figures 5.5 - 5.8 clearly show the presence of an optimal policy for each inspection function considered. In the case $k = 0.1$ with inspection function m_2, the optimal inspection policy seems to strongly depend on parameter a only, which is the first time of inspection of the system. The choice for b does not seem to be of much importance.

Even if the optimal inspection policy gives a value b^* which is less than ξ, we note that the choice $b > 5\,(= \xi)$ is not meaningless: indeed the value R_{τ_i} of the process at inspection time τ_i may be greater then ξ: it is the last hitting time of ξ by the process that defines the process as unsafe.

From table 5.1, we note that the optimal costs are smaller for $k = 0.1$ than for $k = 0.9$. This makes sense, since in both cases the same values for the costs were considered: the case $k = 0.1$ corresponding to bigger amounts of repair, the system will tend to deteriorate slower and therefore will require less maintenance resulting in a smaller total cost. In both cases $k = 0.9$ and $k = 0.1$, we note that the value for v^* increases with the convexity of the inspection function: $v_3^* < v_1^* < v_2^*$. The plots of the optimal inspection functions in figure 5.9 show that the smallest value for a is a_3, which corresponds to the first inspection time for a new system when inspection function m_3 is considered.

However, when the process reaches some value (rather close to 0), the function m_3 crosses m_1 and m_2 to lie above them. It then crosses m_2 a second time to return below it. We may deduce that for this considered process an optimal policy is first to allow great time between the inspection times, then to change strategy drastically in a small amount of time to an almost periodic inspection policy of period 1. This change of inspection decision within the same policy m_3 happens earlier when $k = 0.1$.

5.5 Summary

The proposed models provide optimal inspection policies for a complex multi- component system whose state is described by a multivariate Wiener process. Decisions are made on the basis of the state of a performance measure defined by the Euclidean norm of the multivariate process and the last exit time from an interval rather than the first hitting time. The models are optimized in the sense that they result in a minimum expected maintenance cost, whose expression uses a conditioning argument on the critical threshold's value.

Cost optimal periodic inspection strategies are derived using a renewal argument. As seen with the numerical experiments, the model proposes two different inspection strategies due to the presence of two local minima, allowing the decision maker greater flexibility in the decision process. Moreover, different types of maintenance can be considered when changing the value of parameter k in the maintenance function: for instance setting $k = 1$ extends the proposed model to a minimal repair strategy. In the non-periodic case, the non-periodicity of the inspection times is modelled with the use of an inspection scheduling function, introduced in [34], which determines the next time to inspect the system based on the value of the performance measure at inspection time. The numerical results obtained show the presence of a cost optimal inspection policy in each of the six cases, where different inspection functions and different amounts of repair are considered. Attention is paid to the influence of the convexity of the inspection function on the optimal expected total cost: the value for v^*, the optimal cost, increases with the convexity of the inspection function.

This chapter outlines an approach and structure which is extended in Chapter 6 by considering a second threshold, whose first hitting time by the process defines failure of the system.

Chapter 6

Incorporating failure: two-threshold models

6.1 Introduction

The models derived in this chapter constitute a natural extension to the ones considered in Chapter 5. These still aim at guaranteeing a prescribed level of reliability but failure of the system is now also considered. Examples of systems thought of are airplanes: regular inspections and repairs are usually undertaken to maintain a certain level of safety. However, unfrequent failures of such systems occur and have to be taken into account due to the catastrophic consequences induced.
Recall that in the previous chapter, one threshold ξ was considered and decisions on whether to replace the system or not were made on the basis of the stochastic process R_t having definitely escaped from the region $[0, \xi)$ or not. Repairs were modelled with the use of a repair function, that considered the state of the process at inspection times. Expressions for the maintenance costs and optimal inspection policies in both the periodic and non-periodic inspection cases were obtained by considering the state of the process after a maintenance action is undertaken. A second threshold $\mathcal{F}$, with $\mathcal{F} \geq \xi$, is now added to the model. This new threshold differs from the initial one: rather than being a critical threshold describing the system's safety, this new threshold corresponds to the value at, and above, which the system is considered as failed. Therefore, the time of failure of the system is described by the first time at which the process, describing the performance measure of the system, hits the new threshold $\mathcal{F}$. Hence, this threshold enables catastrophic failures to be to incorporated in the model. The model now uses

a non-stopping and a stopping time, namely the last exit time from the interval $[0, \xi)$ and the first hitting time of threshold $\mathcal{F}$:

$$\begin{aligned} H_{\xi}^{0} &= \sup_{t \in \mathbb{R}^{+}} \{R_t \leq \xi \mid R_0 = 0\} \\ G_{\mathcal{F}}^{0} &= \inf_{t \in \mathbb{R}^{+}} \{R_t = \mathcal{F} \mid R_0 = 0\} \end{aligned} \tag{6.1}$$

The threshold ξ deals with the system's repair actions. As will be shown, it is incorporated in the maintenance function d with the use of the probability of occurrence of the last exit time, introduced in Chapter 4. Different maintenance actions, depending on this probability, are considered. For this, a parameter ϵ is introduced: this parameter determines the level of reliability for the system below which a repair is required. The model therefore ensures that the probability of escaping from a performance region is bounded above and thus the reliability bounded below. The threshold $\mathcal{F}$, on the other hand, deals with failure of the system and hence with its replacement. The sequence of times $G_{\mathcal{F}-x}^{0}$, $x \in [0, \mathcal{F})$, constitutes a renewal process, since at these particular times the state of the process is sent back to zero and the thresholds' values are set back to the initial values (due to the replacement of the whole system). These times are then used to derive analytical expressions for both the expected cost per unit time over an infinite time horizon (for the periodic inspection policy) and the expected total cost (for the non-periodic inspection policy). The expected total cost is optimized with respect to both the inspection policy and the repair threshold ξ: the model thus provides an optimal inspection and maintenance strategy for the considered system.

6.2 Periodic inspections

6.2.1 Features of the model

Model assumptions

1. The system is assumed to be new at time $t = 0$, *i.e.* $R_0 = 0$, with values for the thresholds being equal to ξ and $\mathcal{F}$, with $\xi \leq \mathcal{F}$.

2. Inspections are periodic, perfect and instantaneous.

3. Maintenance actions are instantaneous.

4. The system's performance is only known at inspection times. However, times at which the system fails, defined by the first hitting time of the threshold $\mathcal{F}$ by

the process, are instantaneously revealed (self-announcing failures): this assumption allows to incorporate catastrophic failure in the model. The system is then instantaneously replaced by a new one with cost C_f.

5. Each inspection incurs a fixed cost C_i.

6. Each maintenance action on the system incurs a cost determined by a cost function C_r.

Settings for the model

1. The state space in which the process evolves is partitioned as follows:

$$\mathbb{R}^+ = [0, \xi) \cup [\xi, \mathcal{F}) \cup [\mathcal{F}, +\infty), \tag{6.2}$$

where ξ is the threshold related to repairs of the system and $\mathcal{F}$ is the critical threshold defining failure of the system.

2. The inspection policy is denoted by $\Pi = \{\tau, 2\tau, \ldots, k\tau, \ldots\}$, where τ denotes the period of inspection.

3. τ and ξ are the parameters to be optimized in order to minimize the expected total cost. The optimal inspection policy is derived by choosing the period of inspection τ^* which minimizes the cost and an optimal maintenance policy (repairs) is now also provided by choosing the appropriate value ξ^* that induces a minimum expected cost.

4. At inspection time $t = \tau$ (prior to any maintenance action), the system's performance is R_τ.

5. Given that the system' s initial performance is maximum, *i.e.* $R_0 = 0$, decisions on the level of maintenance (replacement or imperfect maintenance) are made on the basis of the indicator function $\mathbf{1}_{\{G^0_{\mathcal{F}} > \tau\}}$. By this it is meant that decisions on whether to replace the system or not are made on the basis of the occurrence of the first hitting time of threshold $\mathcal{F}$ by the process R_t.

6. Deterministic maintenance actions are modelled with the use of a maintenance functions d, which determines the amount by which both of the thresholds' values

are decreased. The choice for the maintenance function differs from the previous ones investigated since it now incorporates the time H_ξ^0 as follows:

$$d(x) = \begin{cases} x, & \mathbb{P}[H^0_{\xi-x} \leq \tau] \leq 1-\epsilon \\ kx, & \mathbb{P}[H^0_{\xi-x} \leq \tau] > 1-\epsilon \end{cases} \tag{6.3}$$

with constants $\epsilon, k \in [0,1]$. Recall that d must belong to the set of bijective functions.

Remark 6.2.1. Note that setting $\epsilon = 0$ in the maintenance function d allows the case of no maintenance on the system to be considered, whereas setting $k = 1$ corresponds to a minimal repair scenario (as bad as old) and $k = 0$ to a perfect repair (good as new).

7. To each maintenance action, a corresponding cost is associated. As maintenance is modelled with the use of a function, the cost of maintenance is expressed with a corresponding cost function depending on the amount by which the thresholds' values are decreased:

$$C_r(d(x)) = \begin{cases} 0, & \mathbb{P}[H^0_{\xi-d(x)} \leq \tau] \leq 1-\epsilon \\ C_{rep}, & \mathbb{P}[H^0_{\xi-d(x)} \leq \tau] > 1-\epsilon \end{cases} \tag{6.4}$$

Note that C_r is no longer a function of the state of the process upon inspection but a function of the amount by which both of the thresholds are decreased after a repair on the system. The transience property of the Bessel process implies that C_r is well defined, *i.e.*:

$$\forall \epsilon \in (0,1), \exists \tau^* \in \mathbb{R}^+ : \begin{cases} \forall \tau \leq \tau^*, \mathbb{P}[H^0_{\xi-x} \leq \tau] \leq 1-\epsilon \\ \forall \tau > \tau^*, \mathbb{P}[H^0_{\xi-x} \leq \tau] > 1-\epsilon \end{cases} \tag{6.5}$$

The framework

At time $t = 0$ the system's performance is assumed maximum, *i.e.* $R_0 = 0$. The threshold responsible for handling repairs on the system is set to ξ and the critical threshold value representing failure of the system is set to $\mathcal{F}$. A periodic inspection policy needs to be chosen in order to minimize the expected cost per unit time. As in the previous model with periodic inspections of Chapter 5, let τ be the period of inspection that needs to be optimized. Assume that at inspection time t_1 say, and prior to any maintenance action, the performance measure's value is equal to x. Maintenance on the system therefore lowers the initial thresholds value ξ and $\mathcal{F}$ to $\xi - d(x)$ and $\mathcal{F} - d(x)$. The two scenarios considered now correspond to whether the process will hit the failure threshold $\mathcal{F} - d(x)$ prior to the next inspection or not:

1. $\mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}} = 1$: this scenario assumes that failure of the system does not happen before the next planned inspection in τ units of time. Hence, it is assumed that the system will be inspected at time $t_1 + \tau$ inducing a cost C_i. The cost of repair to be considered at this next inspection time is determined now by the function C_r and is equal to $C_r(d(x))$. The performance measure's value at time $t_1 + \tau$ is R^0_τ and will determine the amount by which the thresholds' value will be lowered.

2. $\mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}} = 0$: the process defining the performance measure of the system hits the threshold $\mathcal{F} - d(x)$ before the next inspection planned at time $t_1 + \tau$. As soon as this happens, the system is considered as failed and is instantaneously replaced with a new one, incurring a cost of failure C_f. Such times, at which the system fails (and is replaced) define a renewal process which is used to derive expression for the expected costs.

Hence, a cycle corresponds to the occurrence of case 1 a certain amount of time and ends with case 2: the system fails and is replaced corresponding to the end of a cycle.

6.2.2 Optimal periodic inspection policy

Expected cost per cycle

An expression for the expected cost per cycle is now derived. The technique used to derive the following expression differs from the one used in the previous section. As explained above, rather than looking at the process after the maintenance action, the state of the process at inspection time prior to any maintenance action is considered. If the system's performance measure at inspection time prior to any maintenance actions is $R_\tau = x$, its value after maintenance is set to zero and the thresholds' values $\mathcal{F}$ and ξ are set to $\mathcal{F} - d(x)$ and $\xi - d(x)$ at time τ^+ (as in the previous chapter τ^+ denotes the time just after a maintenance action, where maintenance is assumed to be instantaneous). Hence, looking at the value of the process R_τ, prior to a maintenance action, and considering the future costs (first the ones included until the next inspection and then the remaining ones with a recursive argument), an analytical expression for the cost of inspection and maintenance per cycle may be derived. The notation V^x_τ

denotes the cost per cycle with periodic inspection of period τ, given that at inspection time (prior to any maintenance action) $R_\tau = x$.

$$\begin{aligned} V_\tau^x &= C_f \times \mathbf{1}_{\{system\ fails\ given\ d(x)\}} \\ &\quad + \left[C_i + C_r\left(d\left(x\right)\right) + V_\tau^{R_\tau^0}\right] \times \mathbf{1}_{\{system\ dos\ not\ fail\ given\ d(x)\}} \\ &= C_f \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} \leq \tau\}} \\ &\quad + \left[C_i + C_r\left(d\left(x\right)\right) + V_\tau^{R_\tau^0}\right] \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}} \end{aligned} \tag{6.6}$$

The term $V_\tau^{R_\tau^0}$ indicates the future cost, given that at the next inspection time the process will be in state R_τ^0, since it started from state 0. Recall that maintenance is modelled by decreasing the thresholds' values by $d\left(x\right)$.

Taking the expectation of the above leads to the following expression for the expected cost per cycle:

$$\begin{aligned} v_\tau^x &= \mathbb{E}\left[V_\tau^x\right] \\ &= \mathbb{E}\left[C_f \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} \leq \tau\}}\right] + \mathbb{E}\left[\left(C_i + C_r\left(d\left(x\right)\right) + V_\tau^{R_\tau^0}\right) \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}}\right] \\ &= C_f \mathbb{P}\left[G^0_{\mathcal{F}-d(x)} \leq \tau\right] \\ &\quad + \left(C_i + C_r\left(d\left(x\right)\right) + \int_0^{\mathcal{F}-d(x)} v_\tau^y f_\tau^0\left(y\right) dy\right) \times \mathbb{P}\left[G^0_{\mathcal{F}-d(x)} > \tau\right] \\ &= C_f \int_0^\tau g^0_{\mathcal{F}-d(x)}\left(y\right) dy \\ &\quad + \left(C_i + C_r\left(d\left(x\right)\right) + \int_0^{\mathcal{F}-d(x)} v_\tau^y f_\tau^0\left(y\right) dy\right) \times \left(1 - \int_0^\tau g^0_{\mathcal{F}-d(x)}\left(y\right) dy\right) \\ &= C_f \int_0^\tau g^0_{\mathcal{F}-d(x)}\left(y\right) dy + \left\{C_i + C_r\left(d(x)\right)\right\} \left(1 - \int_0^\tau g^0_{\mathcal{F}-d(x)}\left(y\right) dy\right) \\ &\quad + \left(1 - \int_0^\tau g^0_{\mathcal{F}-d(x)}\left(y\right) dy\right) \int_0^{\mathcal{F}-d(x)} v_\tau^y f_\tau^0\left(y\right) dy \end{aligned} \tag{6.7}$$

The third equality requires the computation of $\mathbb{E}\left[V_\tau^{R_\tau^0} \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}}\right]$: this is done using a conditional independence argument, as explained in Appendix C.

Rearranging equation (6.7) gives:

$$v_\tau^x = Q(x) + \lambda(x) \int_0^{\mathcal{F}-d(x)} v_\tau^y f_\tau^0(y)\, dy \tag{6.8}$$

with:

$$\begin{aligned} \lambda(x) &= 1 - \int_0^\tau g^0_{\mathcal{F}-d(x)}(y)\, dy \\ Q(x) &= (1-\lambda(x))\, C_f + \lambda(x) \{C_i + C_r(d(x))\} \end{aligned} \tag{6.9}$$

Expected length of a cycle

Similarly, an expression for the expected length of a cycle may be derived. Let's first write the expression for the length of a cycle. The notation L_τ^x is used to denote the length of a cycle given that the value of the process prior to a maintenance action is x and that the period of inspection is equal to τ.

$$\begin{aligned} L_\tau^x &= G^0_{\mathcal{F}-d(x)} \times \mathbf{1}_{\{system\ fails\ given\ d(x)\}} \\ &\quad + \left[\tau + L_\tau^{R_\tau^0}\right] \times \mathbf{1}_{\{system\ dos\ not\ fail\ given\ d(x)\}} \\ &= G^0_{\mathcal{F}-d(x)} \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} \leq \tau\}} \\ &\quad + \left[\tau + L_\tau^{R_\tau^0}\right] \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > \tau\}} \end{aligned} \tag{6.10}$$

The term $L_\tau^{R_\tau^0}$ indicates the future length, given that at the next inspection time the process will be in state R_τ^0, since it started from state 0. Taking the expectation leads to the following expression for the expected length of a cycle:

$$l_\tau^x = P(x) + \lambda(x) \int_0^{\mathcal{F}-d(x)} l_\tau^y f_\tau^0(y)\, dy \tag{6.11}$$

with λ defined in (6.9) and

$$P(x) = (1-\lambda(x)) \int_0^\tau y g^0_{\mathcal{F}-d(x)}(y)\, dy + \lambda(x)\, \tau \tag{6.12}$$

Expected cost per unit time

The use of the standard renewal reward argument allows to obtain the expression for the expected cost per unit time over an infinite horizon:

$$C_\tau^x = \frac{v_\tau^x}{l_\tau^x} \tag{6.13}$$

with expressions for v_τ^x and l_τ^x given in (6.7) and (6.11) respectively.

Obtaining solutions

Obtaining solutions to equations (6.8) and (6.11) requires the expression for $g^0_{\mathcal{F}}$, the probability density function for the first hitting time of threshold $\mathcal{F}$ by a Bessel process with drift. As mentioned in Chapter 3, such an expression is not known explicitly, and several attempts to derive an expression for such a density function did not produce any convincing results. However, the expression for the Laplace transform of $G^0_{\mathcal{F}}$ is available in the literature. This expression, given in Chapter 3, is recalled here:

$$\mathbb{E}[e^{-\frac{1}{2}\beta^2 G^0_{\mathcal{F}}}] = \left(\frac{\sqrt{\beta^2+\mu^2}}{\mu}\right)^{\nu} \frac{I_\nu\left(\mu\mathcal{F}\right)}{I_\nu\left(\mathcal{F}\sqrt{\beta^2+\mu^2}\right)}, \ \nu \geq 0 \tag{6.14}$$

see [70] and [91] for proofs. This motivated the idea to invert this expression back to the time domain. Much research in the literature (such as [32, 36, 49, 54, 64, 72]) was undertaken in order to derive an analytical expression for the inverse Laplace transform, but was unfruitful. The difficulty with finding the inverse Laplace transform of such an expression arises from the presence of a modified Bessel function of the first kind at the denominator of the fraction:

$$\frac{I_\nu\left(\mu\mathcal{F}\right)}{I_\nu\left(\mathcal{F}\sqrt{\beta^2+\mu^2}\right)}$$

Hence, solutions to (6.8) were obtained by performing numerical inversions of this Laplace transform. The method chosen to perform such an inversion is the EULER method from Abate and Whitt, described in [2]. The numerical results obtained were compared to results from simulations procedure and proved to be more than satisfactory. More details on the method and the simulations undertaken may be found in Appendix B.

Note that equations (6.8) and (6.11) are Volterra type integral equations, due to the presence of the variable x in the integral limit. These equations are reformulated as Fredholm type integral equations by incorporating an indicator function inside the integral term:

$$\begin{aligned} v^x_\tau &= Q\left(x\right) + \lambda\left(x\right)\int_0^{\mathcal{F}} K\left\{x,y\right\} v^y_\tau \, dy \\ l^x_\tau &= P\left(x\right) + \lambda\left(x\right)\int_0^{\mathcal{F}} K\left\{x,y\right\} l^y_\tau \, dy \end{aligned} \tag{6.15}$$

with Q, λ defined in (6.9), P defined in (6.12) and

$$K\left\{x,y\right\} = \mathbf{1}_{\{y\leq\mathcal{F}-d(x)\}} f^0_\tau\left(y\right) \tag{6.16}$$

The above equations are then solved numerically: the Nystrom routine with the N point Gauss-Legendre rule at the points $y_j \in [0, \mathcal{F})$, $\forall j \in \{1, 2, \ldots, N\}$, is applied. The discretized versions of (6.15) are then evaluated at the points $x_i = y_i, \forall i \in \{1, \ldots, N\}$, in a similar way as in subsection 5.2.2.

Remark 6.2.2. Note that for $x_i \in (\xi, \mathcal{F}]$, $d(x_i)$ is not defined since $\xi - x_i < 0$. For such values, it is assumed that $d(x_i) = kx_i$, *i.e.* repair is undertaken on the system. If $\xi - d(x_i) < 0$, a cost of repair is automatically included at the next inspection time, *i.e.* $C_r(d(x_i)) = C_{rep}$. This seems to be a physically reasonable assumption since R_t is positive and will therefore always stay above such a threshold with negative value, meaning that the last exit time has already happened and hence that repair must be considered.

A little more must be said to explain how the expected length of a cycle may be calculated. Indeed, this requires the evaluation of the following integral:

$$\int_0^\tau y g^0_{\mathcal{F}-d(x)}(y)\, dy$$

Using integration by parts, one gets

$$\int_0^\tau y g^0_{\mathcal{F}-d(x)}(y)\, dy = \tau\, G^0_{\mathcal{F}-d(x)}(\tau) - \int_0^\tau G^0_{\mathcal{F}-d(x)}(y)\, dy,$$

where $G^0_{\mathcal{F}}(t) = \int_0^t g^0_{\mathcal{F}}(s)ds$ (the probability distribution function). The first term is obtained with the EULER method described in Appendix B. As far as the second term is concerned, one may note that:

$$\int_0^\tau G^0_{\mathcal{F}-d(x)}(y)\, dy = \mathcal{L}^{-1}\left[\frac{\mathcal{L}\left[g^0_{\mathcal{F}-d(x)}\right]}{s^2}\right]$$

Hence, rather than numerically integrating $G^0_{\mathcal{F}}(t)$ which is rather computationally demanding since it requires to run the EULER algorithm n times (where n is the number of points used for the numerical integration), the integral term is obtained by numerically inverting $\frac{1}{M}\frac{\mathcal{L}\left[g^0_{\mathcal{F}-d(x)}\right]}{s^2}$ with the use of the EULER method. M is any constant satisfying $M > \tau$ ensuring convergence of the EULER method, since:

$$\left|\frac{1}{M}\int_0^\tau G^0_{\mathcal{F}-d(x)}(y)\, dy\right| \leq \frac{\tau}{M} < 1$$

Hence:

$$\int_0^\tau G^0_{\mathcal{F}-d(x)}(y)\, dy = M\mathcal{L}^{-1}\left[\frac{1}{M}\frac{\mathcal{L}\left[g^0_{\mathcal{F}-d(x)}\right]}{s^2}\right]$$

The minimum value required for M may be determined by the values of threshold $\mathcal{F}$ and the Bessel process' drift term μ. Indeed these provide sufficient information on the time required to hit the critical threshold to be able to give an upper bound for the optimal period of inspection.

The optimal period of inspection and repair threshold can then be determined as:

$$(\tau^*,\, \xi^*) = \underset{(\tau,\,\xi)\in\mathbb{R}^+\times[0,\mathcal{F}]}{\text{argmin}} \{\mathcal{C}^0_\tau\}$$

6.3 Non-periodic inspections

6.3.1 Features of the model

An extension of the above model to the case of non-periodic inspection policies is now presented. In order to avoid any redundancy, most of the features to the model are not stated: these are just the non-periodic version of the ones stated in subsection 6.2.1, just as the features for the non-periodic inspection policy with one threshold were to the corresponding periodic inspection policy. The evolution of the performance measure, the different values for the two thresholds and inspection times over one life cycle are illustrated in figure 6.1. We note that the process R_t is considered here, and not the process R^*_t representing the maintained system. Attention must be paid to the considered maintenance function d and cost function C_r. The time at which the system is inspected now depends on the performance measure's value: a slight modification to the repair function is applied. Let x be the state of the process R_t at inspection time before any maintenance is undertaken on the system, the amount of repair considered is calculated as follows:

$$d\,(x) = \begin{cases} x, & \mathbb{P}[H^0_{\xi-x} \leq m\,(x)] \leq 1-\epsilon \\ kx, & \mathbb{P}[H^0_{\xi-x} \leq m\,(x)] > 1-\epsilon \end{cases} \tag{6.17}$$

with constants $\epsilon, k \in [0,1]$. The probabilities involved in 6.17 should instead be $\mathbb{P}[H^0_{\xi-x} \leq m\,(d\,(x))]$, making $d\,(x)$ impossible to be calculated. The amount of maintenance undertaken being determined, the next inspection time can be computed and will happen in $m\,(d\,(x))$ units of time. The cost of repair at the next inspection is therefore equal to $C_r\,(d\,(x))$, where:

$$C_r\,(d\,(x)) = \begin{cases} 0, & \mathbb{P}[H^0_{\xi-d(x)} \leq m\,(d\,(x))] \leq 1-\epsilon \\ C_{rep}, & \mathbb{P}[H^0_{\xi-d(x)} \leq m\,(d\,(x))] > 1-\epsilon \end{cases} \tag{6.18}$$

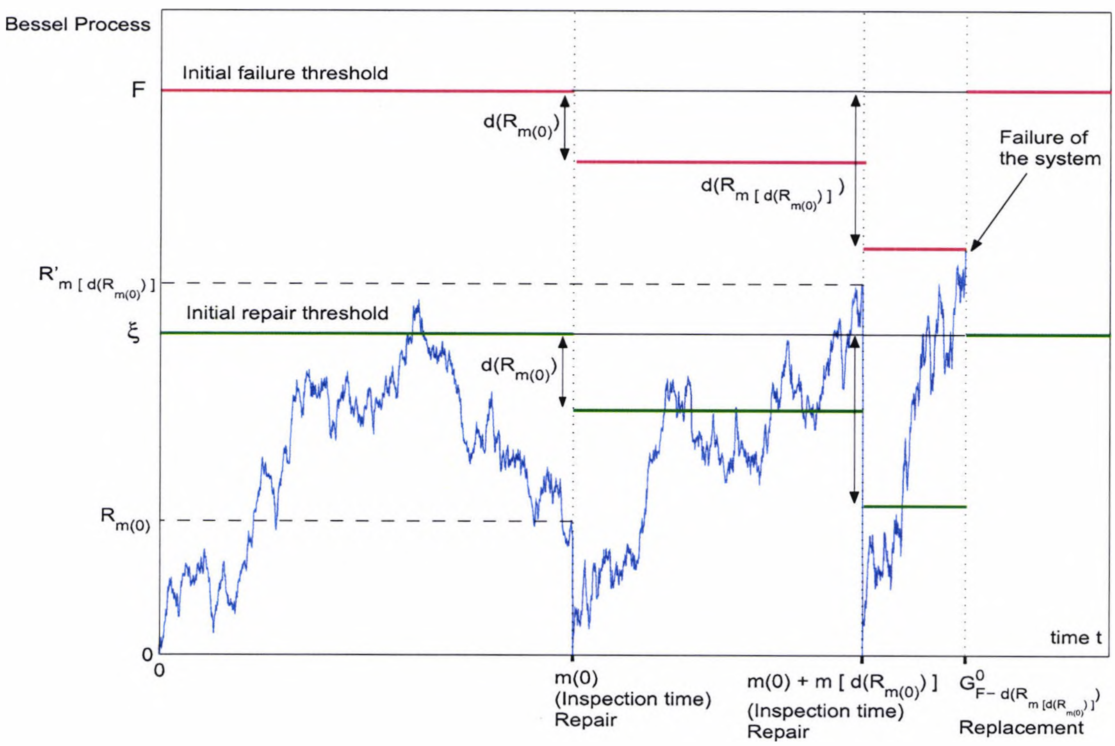


Figure 6.1: Evolution of the system's performance measure over one life cycle, with changes in the thresholds' value.

6.3.2 Expected total cost

Similarly to the way the expression for the expected cost per unit time with a periodic inspection policy was derived, an expression for the expected total cost over a finite time horizon in the non-periodic case may be deduced with the use of the inspection scheduling function. Recall that for this, the value of the process R_t at inspection time and prior to any maintenance action is considered. Hence, if at inspection time τ_1 say, $R_{\tau_1} = x$, the system is next inspected after $m\left(d\left(x\right)\right)$ units of time. Thus, the expression for the total cost of inspection and maintenance is:

$$\begin{aligned} V^x &= \left[C_f + V^0\right] \times \mathbf{1}_{\{system\ fails\ given\ d(x)\}} \\ &\quad + \left[C_i + C_r\left(d\left(x\right)\right) + V^{R^0_{m(d(x))}}\right] \times \mathbf{1}_{\{system\ dos\ not\ fail\ given\ d(x)\}} \\ &= \left[C_f + V^0\right] \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} \leq m(d(x))\}} \\ &\quad + \left[C_i + C_r\left(d\left(x\right)\right) + V^{R^0_{m(d(x))}}\right] \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} > m(d(x))\}} \end{aligned} \tag{6.19}$$

Taking the expectation of the above expression gives the expected total cost of inspection and maintenance of the system. The intermediate steps are omitted to avoid any redundancy: these are similar to the ones with periodic inspections. We get:

$$\begin{aligned}
v^x &= \mathbb{E}\left[V^x\right] \\
&= \left(C_f + v^0\right) \int_0^{m(d(x))} g^0_{\mathcal{F}-d(x)}\left(y\right) dy + \left\{C_i + C_r\left(d(x)\right)\right\} \left(1 - \int_0^{m(d(x))} g^0_{\mathcal{F}-d(x)}\left(y\right) dy\right) \\
&\quad + \left(1 - \int_0^{m(d(x))} g^0_{\mathcal{F}-d(x)}\left(y\right) dy\right) \int_0^{\mathcal{F}} \mathbf{1}_{\{y \leq \mathcal{F}-d(x)\}} v^y f^0_{m(d(x))}\left(y\right) dy
\end{aligned} \tag{6.20}$$

which may be re-arranged as

$$v^x = Q\left(x\right) + \left(1 - \lambda\left(x\right)\right) v^0 + \lambda\left(x\right) \int_0^{\mathcal{F}} K\left\{x, y\right\} v^y dy \tag{6.21}$$

with:

$$\begin{aligned}
\lambda\left(x\right) &= 1 - \int_0^{m(d(x))} g^0_{\mathcal{F}-d(x)}\left(y\right) dy \\
Q\left(x\right) &= \left(1 - \lambda\left(x\right)\right) C_f + \lambda\left(x\right) \left\{C_i + C_r\left(d\left(x\right)\right)\right\} \\
K\left\{x, y\right\} &= \mathbf{1}_{\{y \leq \mathcal{F}-d(x)\}} f^0_{m(d(x))}\left(y\right)
\end{aligned} \tag{6.22}$$

6.3.3 Obtaining solutions

Using the same homotopy argument as the one proposed in subsection 5.3.2, equation (6.21) may be rewritten as:

$$v^x = Q\left(x\right) + \left(1 - \lambda\left(x\right)\right) v^x + \lambda\left(x\right) \int_0^{\mathcal{F}} K\left\{x, y\right\} v^y dy \tag{6.23}$$

Equation (6.23) is then solved numerically as described in the previous model, allowing the solution at $x = 0$ to be obtained.

6.4 Discounted cost criterion using Laplace transforms

In this section, expressions for the costs of maintenance under particular inspection policies are given. These results are well-known and can be found in [61] for instance. The reason why they are included in the thesis is that rather than being based on the probability density function of the first hitting time, the cost expressions are derived with the use of the Laplace transform of the desired density, denoted by $\tilde{g}^0_{\mathcal{F}}$. Hence

solutions to the given cost expressions, with our particular choice of process, may be obtained without having to perform any numerical inversion (formally required to invert the Laplace transform). These models use the Laplace transform of the desired density but yield expressions for the costs in the time domain: the idea is to consider discounted costs, hence introducing an exponential term in the integral. The discounted cost criterion allows future costs to be discounted, which is a real significant factor in life. This discounting reflects the time value of money. Benefits and costs are worth more if they are experienced sooner. The higher the discount rate, the lower is the present value of future cash flows.
The considered models are the 'no inspection' and the 'continuous condition monitoring' models. Note however that these special cases only deal with the failure threshold $\mathcal{F}$, since replacements of the system only are considered.

6.4.1 No inspection policy with discounted cost

The inspection and maintenance policy considered in this particular case are the following: the system starts in a new state and is never inspected. It runs until it fails, at that particular time only it is replaced by a new one (this is the kind of maintenance generally used for light bulbs). The system fails when the process R_t hits the threshold $\mathcal{F}$ for he first time, hence at time $G^0_{\mathcal{F}}$.
The total discounted cost of maintenance is denoted by $V_{\infty,\delta}$, where the subscripts ∞ and δ refer to the 'no inspection strategy' and the discount rate respectively. The expression for the total cost of maintenance is given by:

$$\begin{aligned} V_{\infty,\delta} &= (C_f + V_{\infty,\delta}) \times e^{-\delta G^0_{\mathcal{F}}} \mathbf{1}_{\{system\ fails\}} \\ &= (C_f + V_{\infty,\delta}) \times e^{-\delta G^0_{\mathcal{F}}} \mathbf{1}_{\{G^0_{\mathcal{F}} < \infty\}} \end{aligned} \tag{6.24}$$

Taking the expectation gives the total expected cost:

$$\begin{aligned} v_{\infty,\delta} &= \mathbb{E}\left[V_{\infty,\delta}\right] \\ &= \mathbb{E}\left[(C_f + V_{\infty,\delta}) \times e^{-\delta G_{\mathcal{F}}} \mathbf{1}_{\{G^0_{\mathcal{F}} < \infty\}}\right] \\ &= \int_0^\infty (C_f + v_{\infty,\delta}) \times e^{-\delta t} g^0_{\mathcal{F}}(t) dt \\ &= C_f \int_0^\infty e^{-\delta t} g^0_{\mathcal{F}}(t) dt + v_{\infty,\delta} \int_0^\infty e^{-\delta t} g^0_{\mathcal{F}}(t) dt \\ &= C_f \widetilde{g}^0_{\mathcal{F}}(\delta) + v_{\infty,\delta} \widetilde{g}^0_{\mathcal{F}}(\delta) \end{aligned} \tag{6.25}$$

Hence, it can be deduced:

$$v_{\infty,\delta} = \frac{C_f \widetilde{g}^0_{\mathcal{F}}(\delta)}{1 - \widetilde{g}^0_{\mathcal{F}}(\delta)} \tag{6.26}$$

6.4.2 Condition monitoring policy with discounted cost

This other special case of condition monitoring policy assumes that the value of the process R_t is continuously observed. Thus, the model includes a fixed cost per unit time C_{ci} representing the cost of continuous inspection, and a cost of replacement of the system C_{repl}. As the system is continuously monitored it is never let to fail but is replaced at an appropriate time, chosen by the decision maker. It is assumed that the decision maker decides to change the system when the process R_t reaches the level $\mathcal{F}^-$, $\mathcal{F}^- < \mathcal{F}$, for the first time.

Using a similar notation to the one above, the expression for the total expected cost is:

$$\begin{aligned} V_\delta &= \int_0^{G^0_{\mathcal{F}^-}} C_{ci} \times e^{-\delta t} dt + (C_{repl} + V_\delta) \times e^{-\delta G^0_{\mathcal{F}^-}} \\ &= C_{ci} \left[-\frac{e^{-\delta t}}{\delta} \right]_0^{G^0_{\mathcal{F}^-}} + (C_{repl} + V_\delta) \times e^{-\delta G^0_{\mathcal{F}^-}} \\ &= \frac{C_{ci}}{\delta} + \left(C_{repl} - \frac{C_{ci}}{\delta} + V_\delta \right) \times e^{-\delta G^0_{\mathcal{F}^-}} \end{aligned} \tag{6.27}$$

Thus, the expression for the expected total cost is

$$\begin{aligned} v_\delta &= E\left[V_\delta\right] \\ &= \int_0^{+\infty} \left(C_{repl} - \frac{C_{ci}}{\delta} + v_\delta \right) \times e^{-\delta t} g^0_{\mathcal{F}^-}(t) dt + \frac{C_{ci}}{\delta} \\ &= \left(C_{repl} - \frac{C_{ci}}{\delta} + v_\delta \right) \times \int_0^{+\infty} e^{-\delta t} g^0_{\mathcal{F}^-}(t) dt + \frac{C_{ci}}{\delta} \\ &= \frac{C_{ci}}{\delta} + \left(C_{repl} - \frac{C_{ci}}{\delta} + v_\delta \right) \widetilde{g}^0_{\mathcal{F}^-}(\delta) \end{aligned} \tag{6.28}$$

Hence

$$v_\delta = \frac{\frac{C_{ci}}{\delta} + \left(C_{repl} - \frac{C_{ci}}{\delta} \right) \widetilde{g}^0_{\mathcal{F}^-}(\delta)}{1 - \widetilde{g}^0_{\mathcal{F}^-}(\delta)} \tag{6.29}$$

6.4.3 Obtaining solutions

As mentioned above, these two particular inspection policies allow the expected total costs (6.26) and (6.29) to be expressed in terms of the Laplace transform of the first hitting time of a fixed threshold by a Bessel process with drift. Hence, solutions to these equations may be obtained without having to numerically invert the Laplace transform, as was required in the previous models. Numerical results for such inspection policies are not provided but may readily be obtained by simply inserting the required values for the considered costs and discount rate.

6.5 Numerical results

As in the previous section, the results for both the periodic and non-periodic inspection policy were obtained with an $N = 20$ points Gauss-Legendre rule.

6.5.1 Periodic inspection policy

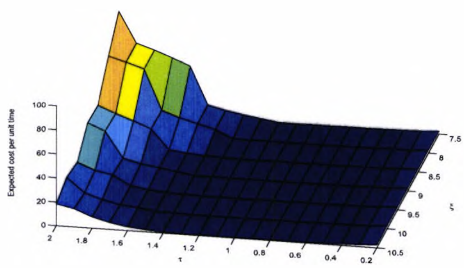
The numerical results in this section were obtained with $Bes_0\,(0.5, 2)$ as the choice for R_t and a failure threshold $\mathcal{F} = 12$. Unless stated otherwise, the choice for the different costs and for the maintenance function's parameters are

$$(C_i, C_{rep}, C_f) = (50, 200, 500)$$
$$(k, \epsilon) = (0.9, 0.5)$$

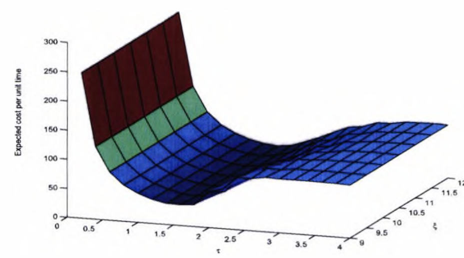
These, however, are allowed to change in order to investigate the behaviour of the proposed model. This choice of Bessel process, the values for the considered thresholds, the costs and the maintenance function parameters were chosen arbitrarily to show some important features of both the inspection and maintenance policies.

The influence of the costs

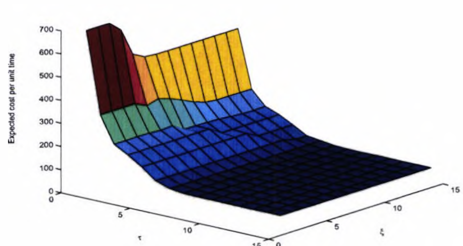
In this subsection, the behaviour of the model as values for the considered costs change is examined. The considered values are: $C_i \in \{0.5, 50, 500\}$, $C_{rep} \in \{2, 200, 2000\}$ and $C_f \in \{5, 500, 5000\}$. The optimal period of inspection τ^*, repair threshold ξ^* and expected cost per unit time $C_\tau^0{}^*$ (with superscript 0 to indicate that the system starts from new) were obtained in each case and are summarized in table 6.1. The expected cost per cycle and expected length of a cycle resulting in the optimal expected cost per unit time are denoted by $v_\tau^0{}^*$ and $l_\tau^0{}^*$ respectively. Figure 6.2 shows the surfaces representing the expected cost per unit time as a function of ξ and τ, in the seven different cost scenarios investigated.



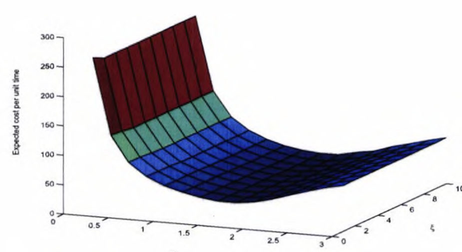
(a) $(C_i, C_{rep}, C_f) = (0.5, 200, 500)$



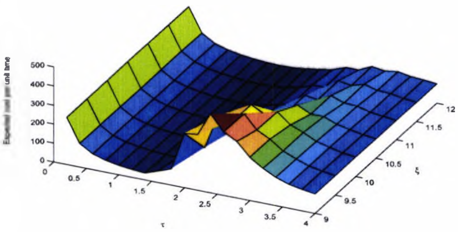
(b) $(C_i, C_{rep}, C_f) = (50, 200, 500)$



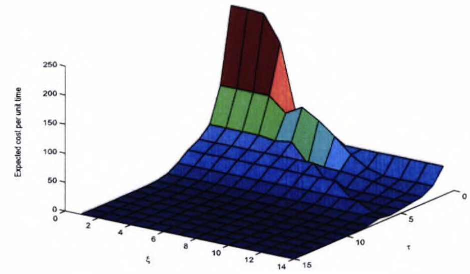
(c) $(C_i, C_{rep}, C_f) = (500, 200, 500)$



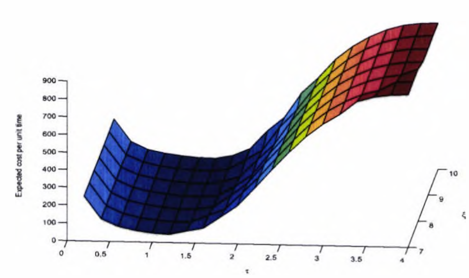
(d) $(C_i, C_{rep}, C_f) = (50, 2, 500)$



(e) $(C_i, C_{rep}, C_f) = (50, 2000, 500)$



(f) $(C_i, C_{rep}, C_f) = (50, 200, 5)$



(g) $(C_i, C_{rep}, C_f) = (50, 200, 5000)$

Figure 6.2: Expected cost per unit time for different cost values ($k = 0.9$, $\epsilon = 0.5$).

(C_i, C_{rep}, C_f)	(τ^*, ξ^*)	$v_{\tau}^{0\ *}$	$l_{\tau}^{0\ *}$	$C_{\tau}^{0\ *}$
$(0.5, 200, 500)$	$(1.2, 9)$	9255.1	20781	0.45
$(50, 200, 500)$	$(1.6, 9.5)$	6972	208.81	33.39
$(500, 200, 500)$	(max, any)	500	5.83	85.89
$(50, 2, 500)$	$(1.8, 7.5)$	2640.2	81.77	32.29
$(50, 200, 500)$	$(1.6, 9.5)$	6972	208.81	33.39
$(50, 2000, 500)$	$(1.6, 10.5)$	5542.3	163.69	33.86
$(50, 200, 5)$	(max, any)	5	5.83	0.86
$(50, 200, 500)$	$(1.6, 9.5)$	6972	208.81	33.39
$(50, 200, 5000)$	$(1.4, 8.5)$	90295	2350.5	38.42

Table 6.1: Optimal period of inspection, repair threshold, expected cost and length per cycle and expected cost per unit time with different values for the maintenance costs ($k = 0.9$, $\epsilon = 0.5$).

- Changing the value for C_i : as C_i increases, both the optimal expected cost per unit time and the optimal period of inspection increase. Increasing C_i consists in considering more expensive inspections: the optimal strategy therefore decides to inspect the system less frequently. This also results in a decrease for $l_{\tau}^{0\ *}$: as the system is less frequently inspected, knowledge of the performance measure's evolution is restrained and the occurrence of a failure is therefore more likely implying the end of a cycle. For a high cost of inspection (*i.e.* $C_i = C_f$) the optimal strategy decides not to inspect the system and to let it run until it fails: the resulting cost C_f being smaller than the sum $C_i + C_{rep}$. The expected cost per cycle is therefore equal to C_f and the expected length of a cycle corresponds to $\mathbb{E}[G_{\mathcal{F}}^0]$ $(= 5.83)$, where $G_{\mathcal{F}}^0$ denotes the first hitting time of threshold $\mathcal{F}$. The value for the repair threshold does not matter here, since the system is never repaired.

- Changing the value for C_{rep} : an increase in C_{rep} does not really affect the optimal period of inspection and results in higher values for the optimal repair threshold. By increasing ξ^*, the optimal strategy decides to reduce the frequency of occurrence of repairs hence preventing from high spending in repair costs. As for τ^*, the optimal expected cost per unit time is not really affected by a change in C_{rep}: the optimal strategy really seems to be driven by the repair threshold, controlling the frequency of maintenance actions on the system and hence preserving an optimal expected total cost.

- Changing the value for C_f : as C_f decreases, τ^* and ξ^* increase. For a small cost of failure (*i.e.* $C_f << C_i + C_{rep}$), as in the case of a high inspection cost, the optimal strategy decides to let the system fail and then replace it resulting in a lower cost than a repair or a simple inspection ($v_\tau^{0\,*} = C_f$ and $l_\tau^{0\,*} = \mathbb{E}\left[G_{\mathcal{F}}^0\right] = 5.83$). The value for ξ does not affect the optimal strategy since the system is never repaired. As C_f increases from 500 to 5000, τ^* and ξ^* decrease to prevent an expensive failure cost. However this results in a higher value for $v_\tau^{0\,*}$ and $l_\tau^{0\,*}$ due to the occurrence of more frequent inspections and repairs on the system.

Investigating the maintenance actions

The maintenance function considered so far is:

$$d(x) = \begin{cases} x, & \mathbb{P}[H_{\xi-x}^0 \leq \tau] \leq \frac{1}{2} \\ 0.9x, & \mathbb{P}[H_{\xi-x}^0 \leq \tau] > \frac{1}{2} \end{cases} \tag{6.30}$$

i.e. with the choice of parameters $(k, \epsilon) = (0.9, 0.5)$. The effects of parameters ξ, k and ϵ on the model in the case $(C_i, C_{rep}, C_f) = (50, 200, 500)$ are now the matter of interest. Recall that in this cost configuration, the optimal expected cost per unit time was obtained with the following period of inspection and repair threshold $(\tau^*, \xi^*) = (9.5, 1.6)$.

(i) **The repair threshold:**

The dependence of the optimal solution on the maintenance function is first investigated by allowing ξ to vary and keeping τ^*, k and ϵ constant. The results obtained are shown in figure 6.3. The figure clearly shows the impact of the repair threshold's chosen value on the optimal solution. An inappropriate choice for ξ may result in an enormous increase for C_τ^0, particularly if $\xi \in [0, 9.5)$. Choosing a value greater than 9.5 slightly increases the expected cost per unit time. However, it must be pointed out that even a slight improvement in C_τ^0 may result in enormous savings at the end of a whole cycle (cost may be expressed in thousands of pounds with a large expected cycle length).

(ii) **Level of repair:**

The level of repair in the considered model (associated to the amount by which both of the thresholds are decreased after the maintenance) is represented by parameter $k \in [0, 1]$ in the cost function. The effect on the level of maintenance is

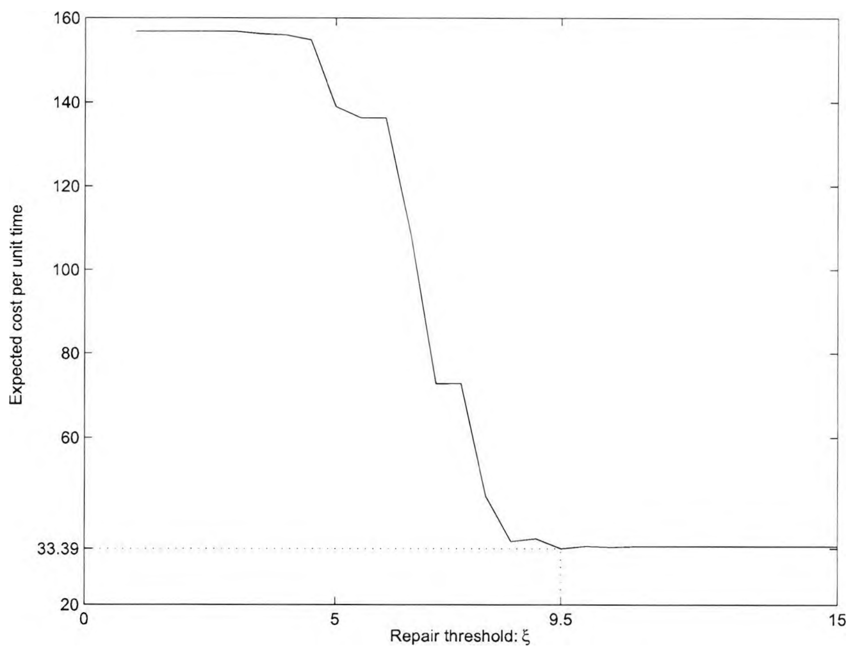


Figure 6.3: Effect of parameter ξ on the optimal solution $C_\tau^{0\,*}$ with parameters $(k, \epsilon, \tau^*) = (0.9, 0.5, 1.6)$.

now analyzed by letting k vary. The different values considered for this parameter are $\{0, 0.1, \ldots, 1\}$: the value zero representing a perfect repair, and the value 1 the minimal repair strategy. Optimal values for the period of inspection and the expected costs per unit time are summarized in table 6.2.

The first conclusion that can be drawn concerns the effect of parameter k on τ^*: as k increases the optimal period of inspection decreases. An increase in k corresponding to less maintenance undertaken on the system, the occurrence of a failure in a shorter amount of time becomes more likely. The optimal strategy therefore chooses to inspect more frequently when the amount of maintenance is low, preventing a failure of the system which is expected to occur sooner and induces a higher cost (recall that $C_f > C_r + C_i$). Secondly, as parameter k increases, $C_\tau^{0\,*}$ increases. This may be explained by noting that as the level of maintenance on the system gets smaller, more frequent repairs need to be undertaken in order to prevent an expensive failure: as repair are more frequent, the total cost induced increases. A more realistic approach would be to let the cost of repair depend on the level of maintenance: indeed a higher level of maintenance may require more time or more staff, usually associated with a higher cost.

(iii) **Decisions on repair:**

k	τ^*	$C_\tau^{0\ *}$
0	3.4	15.72
0.1	3.2	17.04
0.2	3	18.80
0.3	2.8	20.37
0.4	2.4	23.08
0.5	2.4	25.18
0.6	2	28.59
0.7	2	30.02
0.8	1.6	32.78
0.9	1.6	33.39
1	1.4	36.21

Table 6.2: Optimal period of inspection and expected cost per unit time for different values of parameter k ($\epsilon = 0.5$).

The occurrence of repair at inspection times entirely depends on the probability

$$\mathbb{P}[H^0_{\xi-x} \leq \tau]$$

Hence, considering different values for the parameter ϵ in the repair function allows various cases in the decision making process to be considered. The different values studied for ϵ are 0.1, 0.5 and 0.9: these may be associated to a decision maker's attitude towards repair. A value close to 1 corresponds to the occurrence of a repair on the system almost surely and as the value decreases to 0, the decision maker decides to repair less often corresponding to a riskier option. The numerical results for the optimal inspection periods, repair thresholds and total costs per unit time were obtained with $k = 0.9$ and are given in table 6.3. Figure 6.4 shows the evolution of C^0_τ as a function of the period of inspection in the three cases considered.

The first striking conclusion is that, as parameter ϵ varies, the optimization seems to redistribute the actions and effort to meet the reliability criterion without changing the overall cost. The model adapts itself to the different decision maker's attitudes to repair (the change in ϵ) by choosing different values for the optimal repair thresholds. As ϵ increases and therefore as repairs will be considered more often, ξ^* increases: by doing so, the optimal strategy seems to be heading to keep the frequency of repairs constant corresponding to the optimal maintenance strategy. Whereas the optimal expected cost per unit time remains constant in the three cases studied , figure 6.4 clearly shows that this is not the case for periods of inspection $\tau \in (\tau^*, \tau_{\epsilon=0.1}]$, where $\tau_{\epsilon=0.1}$ is defined as

$$\forall\, t > \tau_{\epsilon=0.1} : \mathbb{P}\left[H_\xi^0 < t\right] > 1 - 0.1$$

For most values in this interval, the expected cost per unit time increases with ϵ: the model penalizes a costly strategy that favors too many repairs. For a period of inspection greater than $\tau_{\epsilon=0.1}$, the expected costs per unit time are identical since in all three cases the approach towards repair is similar: the system will be repaired with certainty ($\mathbb{P}\left[H_\xi^0 < t\right] > 0.9 \Rightarrow \mathbb{P}\left[H_\xi^0 < t\right] > 0.5 \Rightarrow \mathbb{P}\left[H_\xi^0 < t\right] > 0.1$).

Decision towards repair	ξ^*	τ^*	$C_\tau^{0\,*}$
$\epsilon = 0.1$	8	1.6	33.39
$\epsilon = 0.5$	9.5	1.6	33.39
$\epsilon = 0.9$	11	1.6	33.39

Table 6.3: Optimal repair threshold, period of inspection and expected cost per unit time for different values of ϵ ($k = 0.9$).

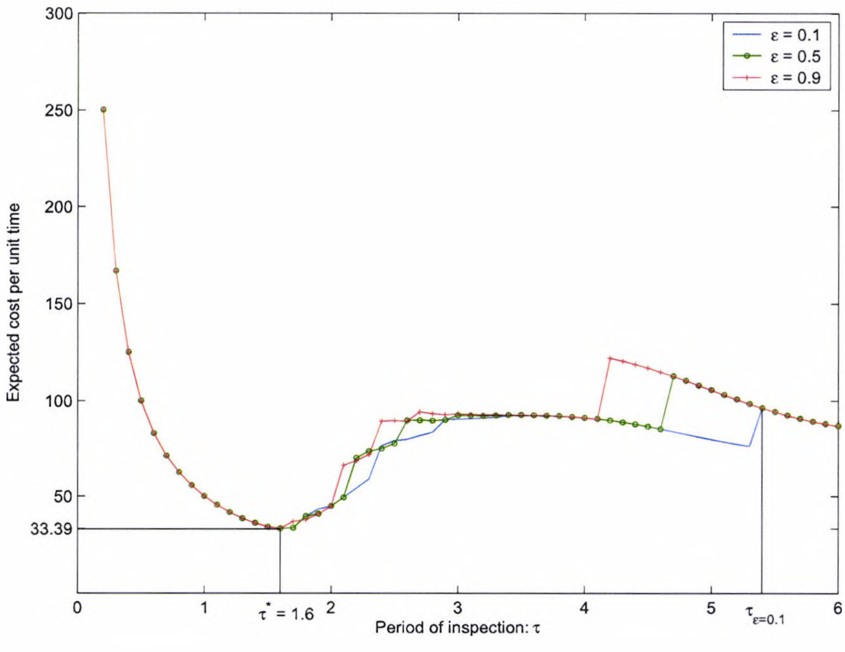


Figure 6.4: Effect of parameter ϵ on the optimal solution $C_\tau^{0\,*}$ with parameters $(k, \xi^*, \tau^*) = (0.9, 9.5, 1.6)$.

Simulation of a cycle

To illustrate the strategy followed by the proposed model, a simulation of the behaviour of the maintained performance measure, the repair thresholds and failure thresholds was undertaken and is shown in figure 6.5. The case considered is the one with the cost configuration $(C_i, C_{rep}, C_f) = (50, 2, 500)$: the reason for this being that this scenario produces a reasonable expected length for the cycle enabling to obtain a clear representation. The value for the parameters are $(k, \epsilon, \tau^*, \xi^*) = (0.9, 0.5, 1.8, 7.5)$. The cycle length obtained is equal to 89.89 units of time and therefore consists of $\lfloor \frac{89.89}{1.8} \rfloor = 49$ inspections. The two thresholds values evolve through time and are set with the use of the maintenance function d. Eventually the process R_t hits the higher threshold resulting in a system failure and ends the cycle.

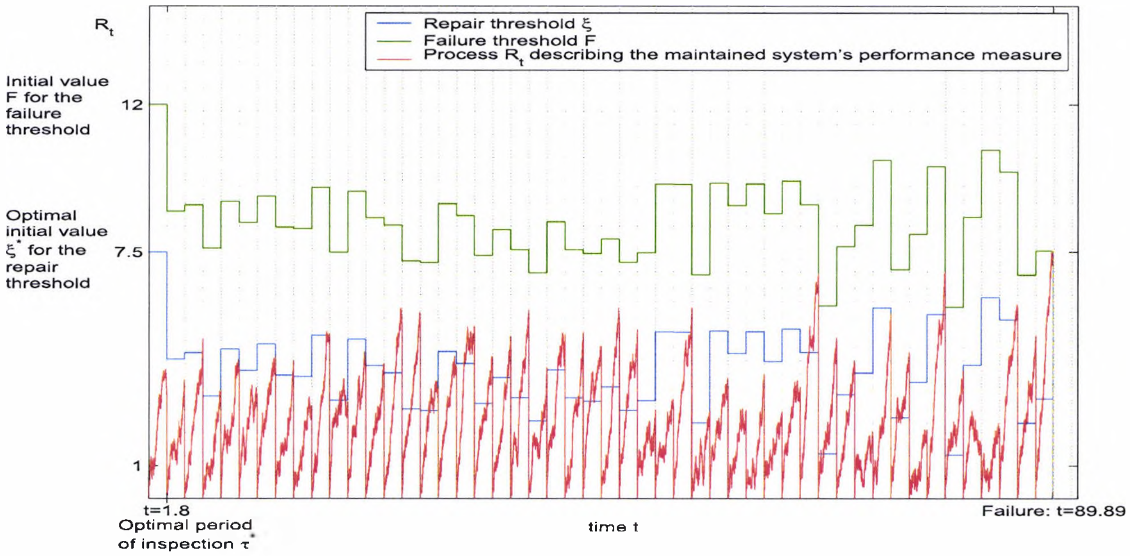


Figure 6.5: Simulation of the proposed model over a cycle, with $(C_i, C_{rep}, C_f) = (50, 2, 500)$ and $(k, \epsilon, \tau^*, \xi^*) = (0.9, 0.5, 1.8, 7.5)$.

6.5.2 Non-periodic inspection policy

The numerical results in this section were obtained with $Bes_0\,(0.5, 2)$ as the choice for R_t and a failure threshold $\mathcal{F} = 5$. Unless stated otherwise, the choice for the different costs and for the maintenance function's parameters are

$$(C_i, C_{rep}, C_f) = (50, 100, 200)$$
$$(k, \epsilon) = (0.1, 0.5)$$

respectively. These, however, are allowed to change in order to investigate the behaviour of the proposed model. As previously, the choice of Bessel process, the values for the considered thresholds, the costs and the maintenance function parameters were chosen arbitrarily to show some important features of both the inspection and maintenance policies.

Three different types of inspection policies are considered with the use of the three inspection scheduling functions m_1, m_2 and m_3 defined in subsection 5.3.1. The expected total costs over a finite time horizon are minimized with respect to the two parameters a and b. The notation a_i^* is used to denote the optimal value of parameter a with inspection scheduling function m_i, $i \in \{1, 2, 3\}$, similarly for b and v.

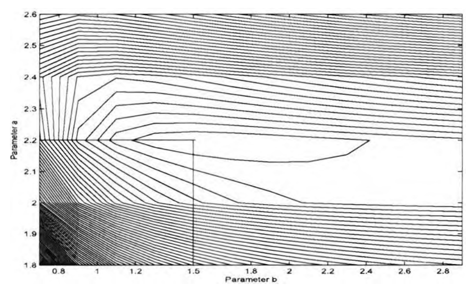
The optimal maintenance policy

The parameters leading to the optimal inspection policy with the corresponding cost (for each of the three inspection scheduling functions) are listed in table 6.4. The effects of the repair threshold on the optimal solution are also investigated by considering different values for $\xi \in [1, \mathcal{F}]$. We first note the presence of an optimal inspection strategy in all the fifteen cases considered. This strategy strongly depends on the choice of parameters a and b as can be seen in figures 6.6 - 6.10. Moreover, it can be noticed that the optimal expected total cost value is more sensitive to a change in ξ than a change in the inspection strategy. Indeed as ξ increases from 1 to 5, v^* increases from the order of 10^3 to 10^6, but for a given value for ξ, a change in the inspection scheduling function slightly affects the optimal expected cost.

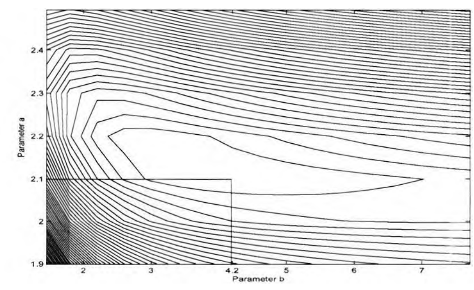
For ξ fixed, a^* (which corresponds to the optimal first inspection time) does not seem to be highly affected by a change in the inspection scheduling function: *e.g.* when $\xi = 3, a_1^* = a_3^* = 2.4$ and $a_2^* = 2.5$. However, the value b^* does get affected and in all the encountered cases $b_3^* \leq b_1^* \leq b_2^*$. This gives greater flexibility to the inspection planner, since different types of inspection strategies may be considered in order to obtain an almost optimal expected total cost. For instance, we note that the overall optimal cost $v^* = 1169.5$ corresponds to the choice of repair threshold $\xi = 1$ and inspection scheduling function m_2, with a large value for b^* implying that the inspection policy switches to a periodic inspection policy only for relatively large values of $d(R_t)$, and therefore after a considerable amount of time. An alternative to this is to adopt the inspection strategy given by m_1, which chooses a fast switch to a periodic inspection policy (since in this case $b_1^* = 0.9$) only resulting in a slightly higher cost $v^* = 1170.4$. A value for $\xi \geq 4$ induces a considerable increase in the optimal expected total cost. For such values of the repair threshold (close to the failure threshold), failure of the system is more likely to happen before a repair is considered. The optimal inspection therefore chooses a high value for a^* in order to prevent unnecessary costs of inspection, with the risk of letting the system fail. In the extreme case where $\xi = \mathcal{F}$, a large value for a^* is accompanied with a small value for b^*: the optimal strategy chooses to let a new system run for a sufficiently long time with the possibility of a failure, reducing the value for a^* incurs unnecessary inspection costs without the certainty of preventing a failure with a repair since $\xi = \mathcal{F}$. Another important fact that must be mentioned is that in all the fifteen different cases, the optimal cost is more sensitive to parameter a than to b: it is the first time at which the system is being inspected that is most responsible for the optimality of the proposed strategy. This is illustrated in figures 6.6 - 6.10 were one may notice that the optimal expected total cost can be reached for a smaller range of values for a than for b.

Repair threshold	Inspection policy	a^*	b^*	v^*
	m_1	2.2	1.5	1171.7
$\xi = 1$	m_2	2.1	4.2	1169.5
	m_3	2.1	0.9	1170.4
	m_1	2.2	1.7	1189.1
$\xi = 2$	m_2	2.2	2.9	1194
	m_3	2.1	1	1189.9
	m_1	2.4	2.5	1546.3
$\xi = 3$	m_2	2.5	2.8	1572.1
	m_3	2.4	1	1547.8
	m_1	5.2	3.7	2.3283×10^5
$\xi = 4$	m_2	5.2	3.8	2.34×10^5
	m_3	5.2	1.9	2.3264×10^5
	m_1	6.5	0.5	3.8437×10^6
$\xi = 5$	m_2	6.5	0.7	3.8437×10^6
	m_3	6.5	0.5	3.8437×10^6

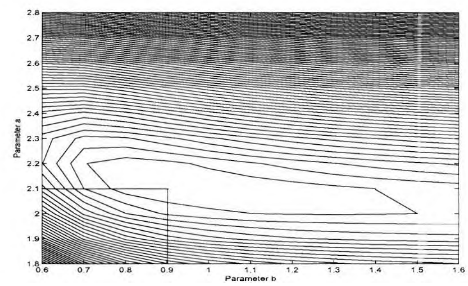
Table 6.4: Optimal expected total cost with corresponding optimal parameters a and b, for different repair thresholds ξ.



(a) Inspection scheduling function: m_1

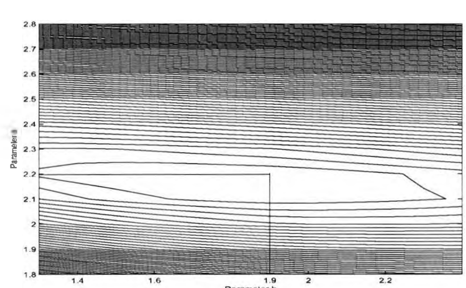


(b) Inspection scheduling function: m_2

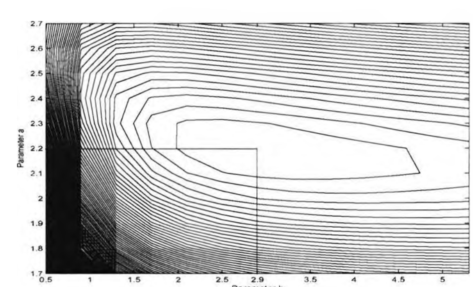


(c) Inspection scheduling function: m_3

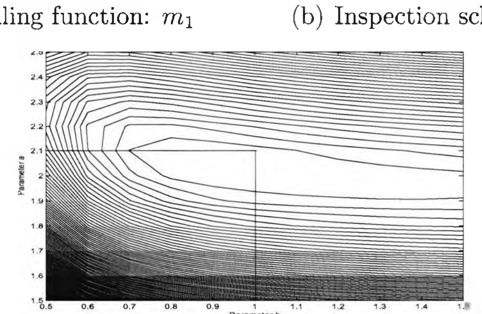
Figure 6.6: Contour representation for the expected total cost with different inspection scheduling function, $\xi = 1$.



(a) Inspection scheduling function: m_1

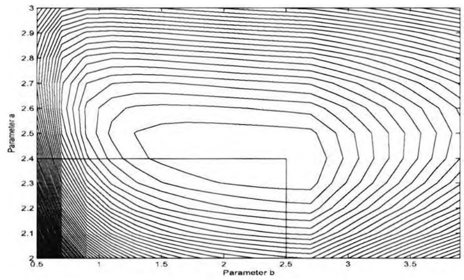


(b) Inspection scheduling function: m_2

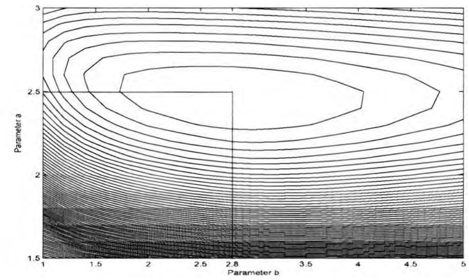


(c) Inspection scheduling function: m_3

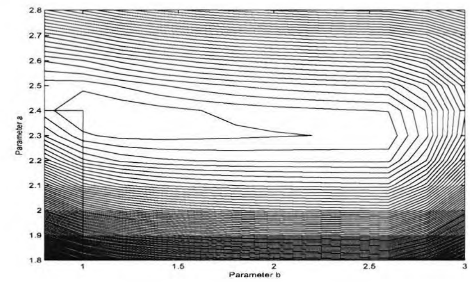
Figure 6.7: Contour representation for the expected total cost with different inspection scheduling function, $\xi = 2$.



(a) Inspection scheduling function: m_1

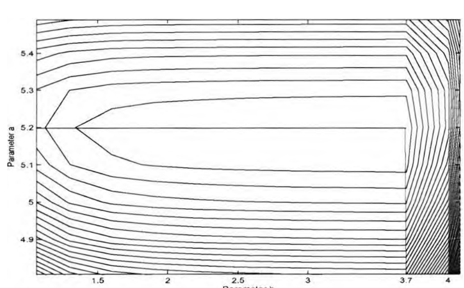


(b) Inspection scheduling function: m_2

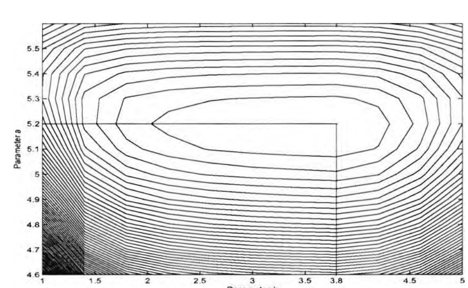


(c) Inspection scheduling function: m_3

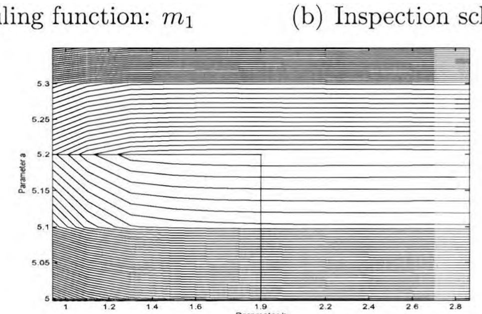
Figure 6.8: Contour representation for the expected total cost with different inspection scheduling function, $\xi = 3$.



(a) Inspection scheduling function: m_1

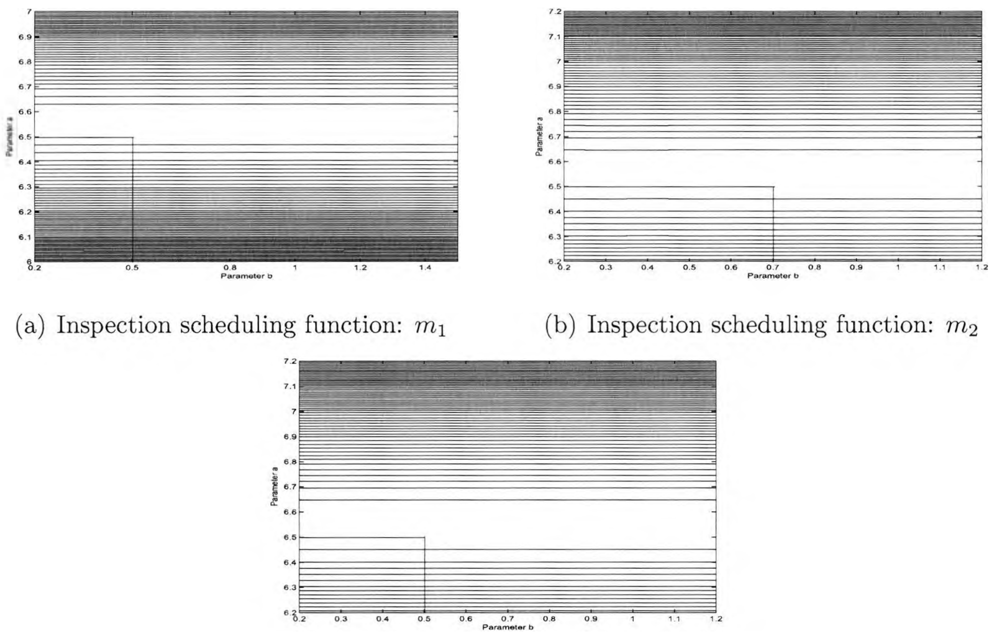


(b) Inspection scheduling function: m_2



(c) Inspection scheduling function: m_3

Figure 6.9: Contour representation for the expected total cost with different inspection scheduling function, $\xi = 4$.



(a) Inspection scheduling function: m_1

(b) Inspection scheduling function: m_2

(c) Inspection scheduling function: m_3

Figure 6.10: Contour representation for the expected total cost with different inspection scheduling function, $\xi = 5$.

The influence of the costs

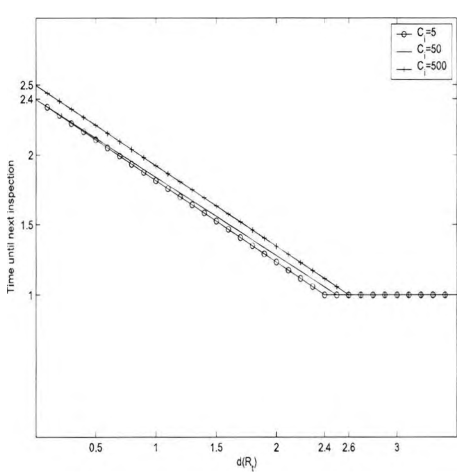
As in the previous example, different values for the considered costs are considered. The case studied is the one with inspection scheduling function m_1 and repair threshold $\xi = 3$. The optimal parameters a^* and b^* obtained are summarized in table 6.5 and so are the resulting expected total costs. Figure 6.11 illustrates the corresponding optimal inspection scheduling functions. It must first be mentioned that as expected, an increase in any of the cost results in an increase for v^*. Furthermore, as shown in figure 6.11, the inspection strategy adapts itself to a change in any of the considered cost. As C_i increases, both of the optimal values for a and b increase: the policy chooses to inspect less frequently when the cost of inspection increases. The highest values for parameters a^* and b^* are obtained when $C_i = 50$, $C_r = 1000$ or $C_f = 2$: the optimal strategy seems to favor long inspection intervals leading to a faster failure of the system in order to avoid numerous expensive costs of inspection or repair.

Investigating the maintenance actions

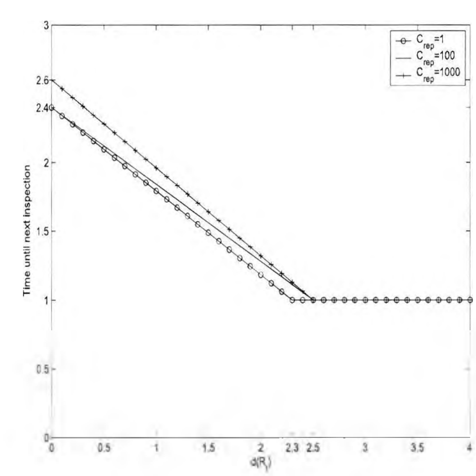
(i) **Level of repair:** Figure 6.12 illustrates the behaviour of the optimal expected total cost as a function of the level of maintenance (parameter $k \in [0, 1]$). The

(C_i, C_{rep}, C_f)	a^*	b^*	v^*
(5, 100, 200)	2.4	2.4	1467.6
(50, 100, 200)	2.4	2.5	1546.3
(500, 100, 200)	2.5	2.6	2296.7
(50, 1, 200)	2.4	2.3	1377.3
(50, 100, 200)	2.4	2.5	1546.3
(50, 1000, 200)	2.6	2.5	2971.9
(50, 100, 2)	3.2	2.9	203.67
(50, 100, 200)	2.4	2.5	1546.3
(50, 100, 2000)	2.4	2.2	13134

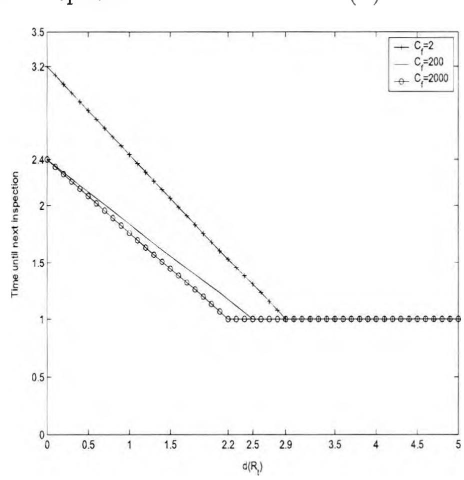
Table 6.5: Optimal expected total cost with corresponding optimal parameters a and b for different values of the maintenance costs, $\xi = 3$.



(a) Different costs of inspection



(b) Different costs of repair



(c) Different costs of failure

Figure 6.11: Optimal inspection scheduling functions m_1 for different inspection and maintenance costs, $\xi = 3$.

results are obtained for the three optimal inspection scheduling functions and with repair threshold's value $\xi = 3$. As one might have expected, in all three cases the expected total cost is increasing with k: increasing the value for k implies a reduction in the amount of maintenance undertaken on the system at repair times. The system will therefore require more frequent inspections and repairs or will fail sooner implying an increase in the total expected cost value.

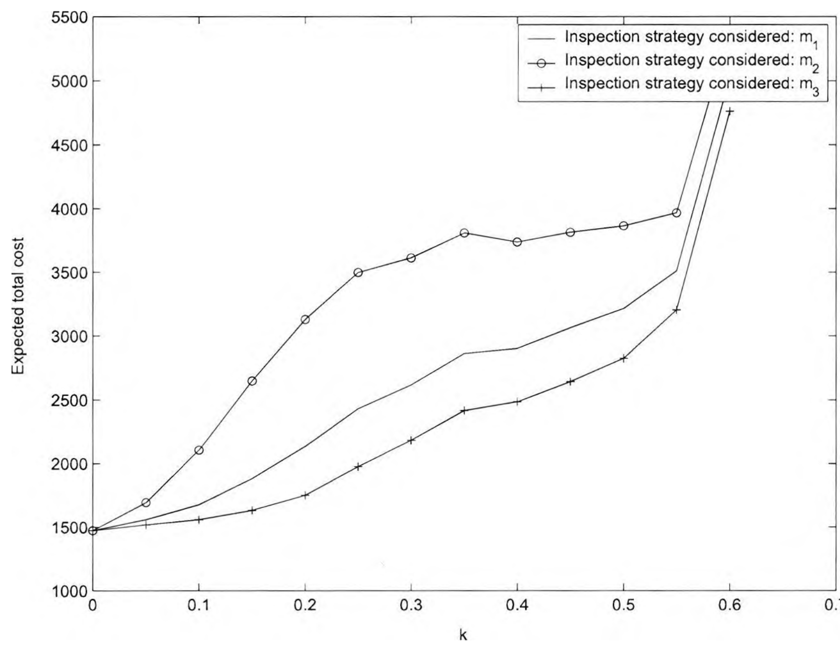


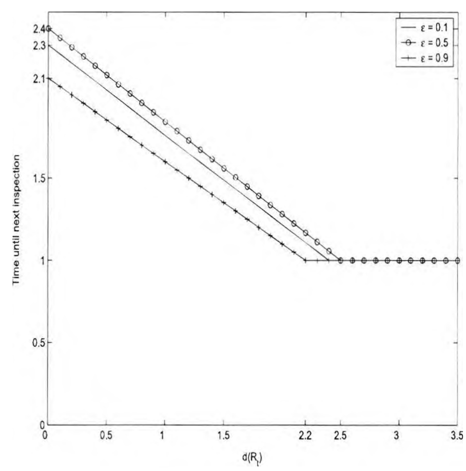
Figure 6.12: Optimal expected total cost as a function of k for the three inspection strategies.

(ii) **Decisions on repair:** As in the periodic case, the attitude of the decision maker towards repair is investigated through parameter $\epsilon \in [0, 1]$. Three values for this parameter are considered here, $\epsilon = 0.1, 0.5, 0.9$, and the optimal expected costs obtained with corresponding optimal parameters (for the three inspection strategies) are summarized in table 6.6. Plots of the resulting optimal inspection functions are shown in figure 6.13. Recall that letting ϵ increase to one means that the decision maker tends to a safer maintenance approach as the constraints on undertaking a repair action are relaxed. The repair threshold's value considered is $\xi = 3$.

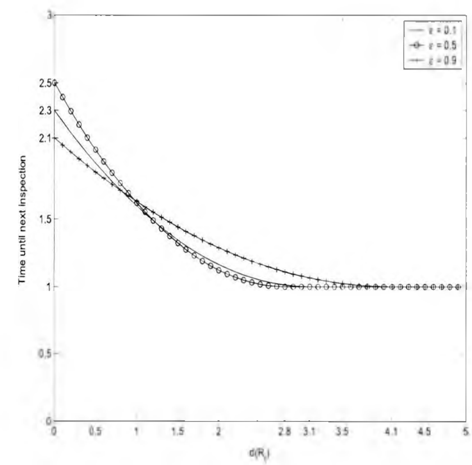
As can be seen, changes in the value for ϵ induce changes in both the optimal inspection policy (figure 6.13) and the resulting optimal expected total cost (table 6.6). In all three cases for ϵ, m_1 produces the minimum value for v^*. Further more, the optimal strategy seems to favor frequent repairs since the minimum values for v^* are obtained for $\epsilon = 0.9$.

Decision towards repair	Inspection policy	a^*	b^*	v^*
	m_1	2.3	2.4	1535
$\epsilon = 0.1$	m_2	2.3	3.1	1622.7
	m_3	2.3	2.1	1560.3
	m_1	2.4	2.5	1546.3
$\epsilon = 0.5$	m_2	2.5	2.8	1572.1
	m_3	2.4	1	1547.8
	m_1	2.1	2.2	1169.1
$\epsilon = 0.9$	m_2	2.1	4.1	1169.1
	m_3	2.1	0.9	1169.9

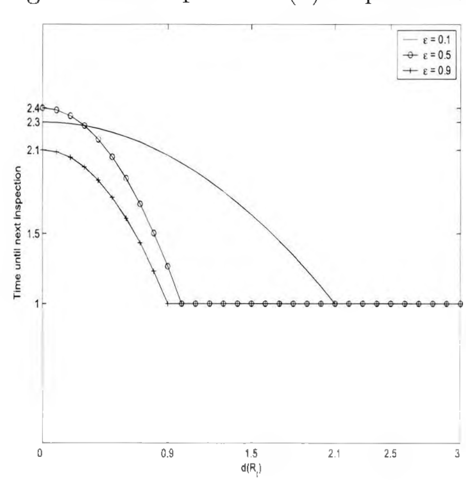
Table 6.6: Optimal expected total cost with corresponding optimal parameters a and b for different values of parameter ϵ, $\xi = 3$.



(a) Inspection scheduling function: m_1



(b) Inspection scheduling function: m_2



(c) Inspection scheduling function: m_3

Figure 6.13: Optimal inspection scheduling functions for different values of parameter ϵ, $\xi = 3$.

6.6 Summary

The aim of the models derived and investigated in the present chapter was to guarantee a prescribed level of reliability while considering the occurrence of catastrophic failures of the system. This was done by introducing a second threshold $\mathcal{F}$: failure of the system is considered when the stochastic process describing the performance measure of the system hits $\mathcal{F}$ for the first time. The threshold ξ, initially investigated in Chapter 5, is still taken into account and is now incorporated in the repair function d: repair on the system at inspection time is now related to the last exit time of the considered process from the interval $[0, \xi)$. Decision on whether to repair the system or not entirely depends on the probability of occurrence of this last exit time before the next inspection. The models proposed hence include both a stopping time (the first hitting time) and a non-stopping time (the last exit time).

Expressions for the expected cost per unit time (periodic inspection policy) and the expected total cost (non-periodic inspection policy) were derived in a slightly different way than the one proposed in the previous chapter. Rather than being based on the amount by which the critical thresholds are decreased, the recursive argument considers the state of the process at inspection time, before a repair is undertaken. The probability density function of the first hitting time for a Bessel process with drift being not known explicitly, it had to be obtained numerically: a numerical inversion of the Laplace transform of the first hitting time's density function was required. However, models for special types of inspection policies are derived using the available Laplace transform of the first hitting time's density and a discounted cost criterion.

Numerical results for both periodic and non-periodic inspection policies were obtained. In the periodic case, the expected cost per unit time was optimized with respect to both the period of inspection and the repair threshold value, hence leading to an optimal maintenance policy (τ^*, ξ^*). Changes in the inspection and maintenance costs resulted in changes in the period of inspection, the repair threshold and the optimal expected cost per unit time. The effects of the repair threshold, the level of maintenance (described by parameter k in the repair function) and the decision taken towards repair (described by parameter ϵ) were investigated. The numerical results revealed a strong influence of the threshold's value ξ and parameter k on both the optimal period of inspection and the optimal expected cost per unit time. Letting parameter ϵ vary produced changes in the optimal repair threshold only, suggesting that the optimal strategy aims at keeping a relatively constant frequency of repairs. A simulation of a cycle showing the evolution of the process describing the maintained system and the changes for the two thresholds' values is provided. In the non-periodic case, the three inspection strategies m_1, m_2 and m_3 were considered. The total expected cost was optimized with respect to the inspection scheduling function's parameters a and b. Different repair thresholds, levels of repair and decisions on repairs were also studied. Optimal inspection policies in all the different cases encountered were derived.

Chapter 7

Conclusions

The research reported in this thesis focused on inspection and maintenance policies guaranteeing an optimal performance for multi-component systems. The models derived permit the investigation of various inspection and maintenance scenarios by choosing appropriate values for the parameters. The methodological development opted to achieve this work necessitated the consideration of new concepts to the field of reliability and maintenance. First of all, the multi-dimensionality aspect of the problem is assessed by introducing several stochastic processes describing the deteriorations undergone by the system's components. Whereas these deteriorations are assumed not to be directly observable, information on the state of the system is available through values taken by a performance metric. This performance metric is represented by a functional acting on the multi-variate process describing the system's state. Secondly, whilst most models proposed in the literature tend to favor monotone processes, all the processes considered here are non-monotone but have continuous paths. The choice of Wiener processes with drift to describe the deterioration of the components is justified by assuming that these include minor repairs which improve the system and do not need to be implicitly modelled. The non-monotonicity of the Bessel process was tackled with the consideration of last exit times: attention was therefore paid to non-stopping times whose occurrence in time cannot be known with certainty. The Markov property of the Bessel process allowed a recursive formulation for the expected costs of inspection and maintenance. The derived Fredholm equations representing the expected costs were solved numerically with, when required, the use of an homotopy argument. Some cases also required a numerical inversion of the Laplace transform of the first hitting time's density function. Moreover, the consideration of particular properties of the Laplace transform allowed improvement of the numerical technics, in particular for the computation of integrals in the time domain.

The models were derived keeping in mind their applicability to concrete problems in the field of reliability and maintenance: roads are examples of systems thought of for the first type of models (regular maintenance actions and occasional reconstructions for safety issues) and airplanes for the second type (regular inspections and repairs with consideration of occurrence of unfrequent catastrophic failures). However, the proposed frameworks may also apply to different areas such as supply chain management where ensuring a minimum availability of products in the different nodes of the chain might be of interest, optimal arbitrage trading where traders are concerned with optimal position towards an asset's price changing through time, [18].

We note that the stochastic process considered in the decision procedure, the Bessel process with drift , does not have independent increments and is therefore not a Lévy process but a diffusion process. This brings up the question of how far the approach might be extended to other diffusions? Is an approach to the problem via stochastic differential equations worth considering?

The output aimed was to obtain an optimal inspection policy and maintenance strategy for a system, whose components' degradations are all taken into account. Solutions to the task set were obtained but required simplifications to the models. Indeed, the different deteriorations were assumed to be independent, when dependency seems more plausible. All maintenance actions considered were assumed deterministic whereas in practice the state of a system after a repair is not truly known. Consideration of random maintenance may be included with the use of a random variable modelling the effect of repair on the system's state. However, such an approach implies an extra integral term in the derived expressions for the costs and getting solutions becomes extremely computationally demanding. Moreover, the stochastic processes considered to describe the components' deteriorations through time were all of the same kind: this assumes similar types of deterioration for all the present parts of the system. Furthermore, as mentioned also in sections 5.4 and 6.5, letting the cost of repair depend on both the value of the performance measure at inspection time and the amount of repair undertaken (*i.e.* $C_{rep}(R_\tau, k)$) would improve the model significantly. These assumptions were made in order to be able to obtain solutions to the problem. Ideas of extensions to the models in order to remedy these shortcomings have been thought of and are now stated:

(i) **Improving the maintenance function:** A natural and rather straight forward extension concerning repairs on the system would be to propose different levels of repair depending on the state of the performance measure. This may be done by partitioning the state space $[0, \mathcal{F})$ into n intervals ,$[0, \mathcal{F}) = [0, \xi_1) \cup [\xi_1, \xi_2) \cup \ldots \cup [\xi_{n-1}, \mathcal{F})$, and considering the following repair function

$$d(x) = \begin{cases} x, & \mathbb{P}[H^0_{\xi_1 - x} \leq \tau] \leq 1 - \epsilon \\ a_1 x, & \mathbb{P}[H^0_{\xi_1 - x} \leq \tau < H^0_{\xi_2 - x}] > 1 - \epsilon \\ \ldots & \\ a_{n-1} x, & \mathbb{P}[H^0_{\xi_{n-1} - x} \leq \tau] > 1 - \epsilon \end{cases} \tag{7.1}$$

$a_i \in [0, 1]$, $\forall i \in \{1, \ldots, n-1\}$, with different costs associated to the different repair types.

(ii) **More generality in the deterioration processes:** Recall that the stochastic process proposed to describe the deterioration of a component is a Wiener process with a drift term and a volatility term. A simplifying assumption was made by considering n Wiener processes with different drift terms $\mu_i, i \in \{1, \ldots, n\}$, but similar volatility terms σ. An extension to investigate would be to consider different volatility terms, hence considering the processes

$$W_t^{(i)} = \mu_i t + \sigma_i B_t^{(i)}, \forall i \in \{1, \ldots, n\} \tag{7.2}$$

In such a way, each component's degradation could uniquely be described, enabling a more realistic approach. A possible way to handle this new scenario may be to consider the square of the Euclidean norm of $\left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(n)}\right)$ rather than the Euclidean norm itself. The idea being that this new process would correspond to a weighted sum of non-centrally χ^2 distributed random variables, [56]. An appealing approach is to look at the trace of non-central Wishart matrices, [40], [50], [51], [52]. Moreover, this would also allow to deal with situation where more importance to certain components in the system is wished to be paid: attaching weights to the processes describing the different deteriorations is a way of incorporating a hierarchical partitioning of the components with respect to their relevance in the functioning of the system.

(iii) **Dependence in the deteriorations:** A strong assumption made was the independence in the different components' degradations. This was modelled by considering independent Brownian motions and may be relaxed. It seems more realistic to include cases were the degradation of a component may affect the degradations of others. A way to incorporate this scenario is by considering dependent Brownian motions through a Wishart distribution with a non-diagonal covariance matrix.

(iv) **Dealing with covariates:** A multi-dimensional approach was considered here as the degradations of all the components were wished to be considered. A similar argument may be used if covariates acting on the way the system/component deteriorates are wished to be studied. The degradation of the system in its normal state is described by a Wiener process with drift and volatility terms, $W_t = \mu t + \sigma B_t$. Covariates (temperature, pressure, humidity, etc...) or the history of the component affect the degradation process as follows. Let $z_1, z_2, \ldots, z_{n-1}$ be the $n-1$ covariates, and $I_1, I_2, \ldots, I_{n-1}$ the corresponding critical sets: if $z_i \in I_i$, the degradation of the system switches to a new Wiener process

$$W_t^{(i)} = \mu_i t + \sigma B_t^{(i)}$$

To obtain a full description which allows overlapping time intervals during which $z_i \in I_i$, the following n-dimensional Wiener process may be considered

$$\mathbf{W}_t = \left(\left\{ \prod_{i=1}^{n-1} \mathbf{1}_{\{z_i \notin I_i\}} W_t \right\}, \mathbf{1}_{\{z_1 \in I_1\}} W_t^{(1)}, \mathbf{1}_{\{z_2 \in I_2\}} W_t^{(2)}, \ldots, \mathbf{1}_{\{z_{n-1} \in I_{n-1}\}} W_t^{(n-1)} \right)$$

with corresponding Bessel process

$$R_t = \sqrt{ \left\{ \prod_{i=1}^{n-1} \mathbf{1}_{\{z_i \notin I_i\}} \right\} [W_t]^2 + \sum_{i=1}^{n-1} \mathbf{1}_{\{z_i \in I_i\}} \left[W_t^{(i)} \right]^2 }$$

Dealing with the above process allows us to treat any possible combination for the presence of covariates. The indicator functions

$$\prod_{i=1}^{n-1} \mathbf{1}_{\{z_i \notin I_i\}}, \text{ and } \mathbf{1}_{\{z_i \in I_i\}}, \; i = 1 \ldots n-1$$

act as switches as the environment evolves. The decisions are then based on the Bessel process with drift associated with the underlying Wiener processes.

Appendix A

Transition density for the Bessel process with drift

We show how the transition density from the state $x = 0$ to the state $y > 0$ may be derived from the case $x > 0$. For this, let $f_\tau^x(y)$ denote the transition density from state x to state y. Using results in [70], for $x, y > 0$, we have:

$$\begin{aligned} f_\tau^x(y) &= \frac{y}{t}\frac{I_\nu(\mu y)}{I_\nu(\mu x)}\, I_\nu\left(\frac{xy}{t}\right) e^{-\frac{x^2+y^2+\mu^2 t^2}{2t}} \\ &= \frac{y}{t}\, I_\nu(\mu y)\, e^{-\frac{x^2+y^2+\mu^2 t^2}{2t}}\, \frac{I_\nu(\frac{xy}{t})}{I_\nu(\mu x)} \end{aligned} \tag{A.1}$$

Using the series expansion for the modified Bessel function of the first kind, [54],

$$I_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{+\infty} \frac{\left(\frac{x}{2}\right)^{2k}}{k! \times \Gamma(k+\nu+1)}$$

allows us to conclude that

$$\begin{aligned} \frac{I_\nu(\frac{xy}{t})}{I_\nu(\mu x)} &= \frac{\left(\frac{xy}{2t}\right)^\nu \sum_{k=0}^{+\infty} \frac{\left(\frac{xy}{2t}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}}{\left(\frac{\mu x}{2}\right)^\nu \sum_{k=0}^{+\infty} \frac{\left(\frac{\mu x}{2}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}} \\ &= \left(\frac{y}{\mu t}\right)^\nu \frac{\sum_{k=0}^{+\infty} \frac{\left(\frac{xy}{2t}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}}{\sum_{k=0}^{+\infty} \frac{\left(\frac{\mu x}{2}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}} \end{aligned} \tag{A.2}$$

Hence using equations (A.1) and (A.2):

$$f_\tau^x(y) = \frac{y}{t}\, I_\nu(\mu y)\, e^{-\frac{x^2+y^2+\mu^2 t^2}{2t}} \left(\frac{y}{\mu t}\right)^\nu \frac{\sum_{k=0}^{+\infty} \frac{\left(\frac{xy}{2t}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}}{\sum_{k=0}^{+\infty} \frac{\left(\frac{\mu x}{2}\right)^{2k}}{k!\times\Gamma(k+\nu+1)}} \tag{A.3}$$

Letting x tend to zero in (A.3), we get:

$$f_{\tau}^{0}(y) = \left(\frac{1}{\mu}\right)^{\nu} \left(\frac{y}{t}\right)^{\nu+1} I_{\nu}(\mu y) e^{-\frac{y^2+\mu^2 t^2}{2t}} \tag{A.4}$$

Appendix B

First hitting time for a Bessel process with drift , numerical approximation

This appendix first aims at giving a description of the algorithm considered in this thesis to numerically invert the Laplace transform of a probability distribution function. This numerical inversion was needed to obtain the probability distribution function for the first hitting time of a Bessel process with drift. As far as we know, no analytical expression has been found. However, we mention that such an expression is available in the simpler case of a Bessel process (with no drift), see for example [19]. The description given follows the lines of [2] and is given here for consistency purposes. We then apply the method to the desired function and compare the results with the one obtained via simulations of the first hitting time.

B.1 Describing the method

The chosen algorithm is one proposed by Abate and Whitt, [2], called the EULER method. We first note that two methods are provided in the referenced paper, namely the EULER method and the POST-WIDDER method, both of which have been tested for our particular case. We will only consider the EULER method here, since it is the method that provided the most accurate results when compared to the one obtained via simulations.

Let $f(t)$ be a real-valued function of a positive real variable t and $\widehat{f}(s)$ its Laplace transform. The method uses the following argument that can be found in [28]

$$f(t) = \frac{1}{2\pi i}\int_{a+i\infty}^{a-i\infty} e^{st}\widehat{f}(s)\,ds$$
$$= \frac{2e^{at}}{\pi}\int_{0}^{+\infty} Re\left(\widehat{f}(a+iu)\right)\cos(ut)\,du \tag{B.1}$$

where $i^2 = -1$ and a is chosen such that $\widehat{f}(s)$ has no singularities on or to the right of it.

The above integral is then numerically evaluated by means of the trapezoidale rule. If a step size h is used, this gives:

$$f(t) \approx f_h(t) \equiv \frac{he^{at}}{\pi}Re\left(\widehat{f}(a)\right) + \frac{2he^{at}}{\pi}\sum_{k=1}^{+\infty} Re\left(\widehat{f}(a+ikh)\right)\cos(kht) \tag{B.2}$$

Letting $h = \frac{\pi}{2t}$ and $a = \frac{A}{2t}$, we obtain

$$f_h(t) = \frac{e^{A/2}}{2t}Re\left(\widehat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{\frac{A}{2}}}{t}\sum_{k=1}^{+\infty}(-1)^k Re\left(\widehat{f}\left(\frac{A+2k\pi i}{2t}\right)\right) \tag{B.3}$$

The discretization error associated with equation (B.3) is then identified with the use of the Poisson summation formula. The idea is to replace the damped function $g(t) = e^{-bt}f(t)$, $b > 0$, by the periodic function $g_p(t) = \sum_{k=-\infty}^{+\infty} g\left(t + \frac{2\pi k}{h}\right)$ of period $\frac{2\pi}{h}$. This can be done when $|f(t)| \leq 1, \forall\ t$.The periodic function $g_p(t)$ can also be represented by its complex Fourier series as:

$$g_p(t) = \frac{h}{2\pi}\sum_{k=-\infty}^{+\infty} \widehat{f}(b+ikh)\,e^{ikht} \tag{B.4}$$

Combining these two representations for $g_p(t)$ gives:

$$g_p(t) = \sum_{k=-\infty}^{+\infty} g\left(t+\frac{2\pi k}{h}\right) = \sum_{k=-\infty}^{+\infty} f\left(t+\frac{2\pi k}{h}\right)e^{-b\left(t+\frac{2\pi k}{h}\right)}$$
$$\text{and} \tag{B.5}$$
$$g_p(t) = \frac{h}{2\pi}\sum_{k=-\infty}^{+\infty} \widehat{f}(b+ikh)\,e^{ikht}$$

Letting $h = \frac{\pi}{t}$ and $b = \frac{A}{2t}$, we may deduce from the above:

$$f(t) = \frac{e^{\frac{A}{2}}}{2t} \sum_{k=-\infty}^{+\infty} (-1)^k Re\left(\widehat{f}\left(\frac{A+2k\pi i}{2t}\right)\right) - \sum_{k=1}^{+\infty} e^{-kA} f((2k+1)t) \tag{B.6}$$

Equation (B.6) and equation (B.3) give the error associated with the trapezoidal rule as

$$e_d = \sum_{k=1}^{+\infty} e^{-kA} f((2k+1)t) \tag{B.7}$$

If $|f(t)| \leq 1$, the error is bounded by:

$$|e_d| \leq \frac{e^{-A}}{1-e^{-A}}$$

Remark B.1.1. We note that this is the reason why the probability distribution function is inverted rather than the density function. Indeed, working with a probability distribution function F ensures $|F(t)| \leq 1$ and hence enables control on the error term.

As proposed in the referenced paper, the chosen value for A was 18.4 to achieve a discretization error of order 10^{-8}.

The next step is to numerically calculate (B.3) which includes an infinite sum. The EULER summation is the proposed option. This can be described as the weighted average of the last m partial sums by a binomial probability distribution with parameters m and $p = 0.5$. Let s_n be the approximation of $f_h(t)$ to n terms

$$s_n(t) = \frac{e^{\frac{A}{2}}}{2t} Re\left(\widehat{f}\left(\frac{A}{2t}\right)\right) + \frac{e^{\frac{A}{2}}}{t} \sum_{k=1}^{n} (-1)^k a_k(t) \tag{B.8}$$

where

$$a_k(t) = Re\left(\widehat{f}\left(\frac{A+2k\pi i}{2t}\right)\right) \tag{B.9}$$

EULER summation is applied to m terms after an initial n, so that the EULER sum for $s_n(t)$ is

$$E(m,n,t) = \sum_{k=0}^{m} \begin{pmatrix} m \\ k \end{pmatrix} 2^{-m} s_{n+k}(t) \tag{B.10}$$

In order to estimate the error associated with the EULER summation, it is suggested to use the difference of successive terms

$$E(m,n+1,t) - E(m,n,t)$$

We started with the values $m = 11$ and $n = 15$ as proposed in the referenced paper, and increased when necessary.

B.2 The algorithm in action

The function $\widehat{f}(s)$ that is wished to be numerically inverted is the Laplace transform of the density function for $G^0_{\mathcal{F}}$: the first hitting time of a Bessel process with drift starting from 0. Its expression was derived by Pitman and Yor, [70], and also by Yin in [91]

$$\mathbb{E}_0[e^{-\frac{1}{2}\beta^2 G^0_{\mathcal{F}}}] = \left(\frac{\sqrt{\beta^2+\mu^2}}{\mu}\right)^{\nu} \frac{I_\nu(\mu\mathcal{F})}{I_\nu\left(\mathcal{F}\sqrt{\beta^2+\mu^2}\right)}, \quad \nu \geq 0 \tag{B.11}$$

where the subscript 0 indicates that the process starts from the point 0.
Using the change of variable $s = \frac{1}{2}\beta^2$, equation (B.11) may be rewritten as:

$$\widehat{f}(s) \doteq \mathbb{E}_0[e^{-sG^0_{\mathcal{F}}}] = \left(\frac{\sqrt{2s+c^2}}{c}\right)^{h} \frac{I_h(c\mathcal{F})}{I_h\left(\mathcal{F}\sqrt{2s+c^2}\right)} \tag{B.12}$$

Note that the EULER method was proposed for a probability function and assumptions based on that fact ensured the good functioning of the algorithm. Hence rather than inverting $\widehat{f}(s)$ numerically, $\dfrac{\widehat{f}(s)}{s}$ is considered: this expression corresponds to the Laplace transform of the probability distribution function for the first hitting time. This is shown by noting that if a function $f(t)$ has its Laplace transform equal to $\widehat{f}(s)$, then

$$\mathcal{L}\left[\int_0^t f(t)\right] = \frac{\widehat{f}(s)}{s}$$

where $\mathcal{L}[.]$ denotes the Laplace operator.
On the other hand, simulations for the first hitting time of a Bessel process with drift were performed using the mathematical software Matlab. The results of the simulations shown in figure B.1 are based on the process $Bes_0\left(0.5, \frac{4}{\sqrt{3}}\right)$ with critical threshold $\xi = 4$ and use a sample of size 2000.
The cumulative distribution function corresponding to the simulation results was then plotted alongside the one obtained via the numerical approximation method, see figure B.2. Moreover, 50 simulations each of sample size 2000 where undertaken. After each simulation, the value $A_i = \sup|F_i - F^*|$ $(i = 1, \ldots, 50)$ was evaluated, where F_i and F^* denote the probability distribution function obtained with the simulation procedure and with the numerical approximation respectively. Over the 50 simulations, the average value obtained was: $\overline{A} = \dfrac{\sum_{i=1}^{50} A_i}{50} = 0.0689$. The accuracy of the numerical method convinced us to use this approximation for the cost models requiring the probability distribution function of the first hitting time for a Bessel process with drift .

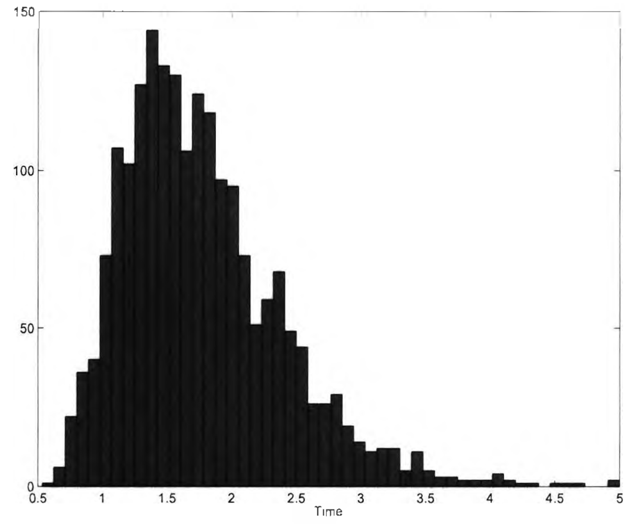


Figure B.1: Simulations of the first hitting time G_4^0 for $Bes_0\left(0.5, \frac{4}{\sqrt{3}}\right)$.

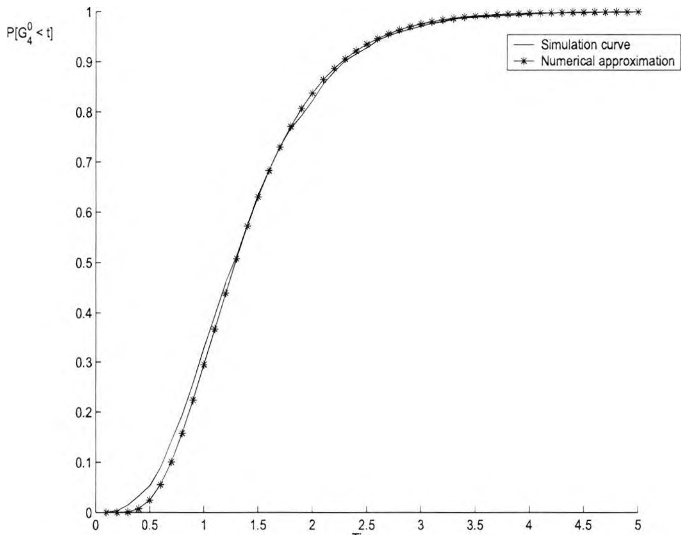


Figure B.2: Probability distribution functions obtained with the simulation and the EULER method.

Appendix C

Conditional independence argument

C.1 One threshold models

C.1.1 Periodic case:

Let $f_{R_\tau, H^0_{\xi-x}}$ be the joint probability density function of the process at time τ and the last exit time from the interval $[0, \xi - x)$. We may deduce:

$$\begin{aligned}
\mathbb{E}\left[V_\tau\left(d(R_\tau)\right) \times \mathbf{1}_{\{H^0_{\xi-x}>\tau\}}\right] &= \int_0^{+\infty}\int_0^{+\infty} v_\tau\left(d(y)\right) \times \mathbf{1}_{\{t>\tau\}} f_{R_\tau, H^0_{\xi-x}}(y,t)\, dydt \\
&= \int_0^{+\infty}\int_0^{+\infty} v_\tau\left(d(y)\right) \times \mathbf{1}_{\{t>\tau\}} f_{R_\tau|H^0_{\xi-x}=t}(y)\, h^0_{\xi-x}(t)\, dydt \\
&= \int_\tau^{+\infty}\int_0^{+\infty} v_\tau\left(d(y)\right) f_{R_\tau|H^0_{\xi-x}>\tau}(y)\, h^0_{\xi-x}(t)\, dydt \\
&= \int_\tau^{+\infty} h^0_{\xi-x}(t) \int_0^{+\infty} v_\tau\left(d(y)\right) f_{R_\tau|H^0_{\xi-x}>\tau}(y)\, dydt \\
&= \int_\tau^{+\infty} h^0_{\xi-x}(t)\, dt \int_0^{+\infty} v_\tau\left(d(y)\right) f_{R_\tau}(y)\, dy \\
&= \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \int_0^{+\infty} v_\tau\left(d(y)\right) f_{R_\tau}(y)\, dy \\
&= \left(1 - \int_0^\tau h^0_{\xi-x}(t)\, dt\right) \int_0^{+\infty} v_\tau\left(d(y)\right) f^0_\tau(y)\, dy
\end{aligned}$$

Where the change from $f_{R_\tau|H^0_{\xi-x}>\tau}$ to f_{R_τ} in the fifth equality comes from a conditional independence argument: as $H^0_{\xi-x} > \tau$, this actually means that the process may still be in the region $[0, \xi - x)$, hence the region of integration to be considered remains $[0, +\infty)$ and not $[\xi - x, +\infty)$.

C.1.2 Non-Periodic case:

Let $f_{R_{m(x)},H^0_{\xi-x}}$ be the joint probability density function of the process at time $m(x)$ and the last exit time from the interval $[0,\xi-x)$. We may deduce:

$$
\begin{aligned}
E\left[V_{\xi-d\left(R_{m(x)}\right)} \times \mathbf{1}_{\{H^0_{\xi-x}>m(x)\}}\right] &= \int_0^{+\infty}\int_0^{+\infty} v_{\xi-d(y)} \times \mathbf{1}_{\{t>m(x)\}} f_{R_{m(x)},H^0_{\xi-x}}(y,t)\,dydt \\
&= \int_0^{+\infty}\int_0^{+\infty} v_{\xi-d(y)} \times \mathbf{1}_{\{t>m(x)\}} f_{R_{m(x)}|H^0_{\xi-x}=t}(y) \\
&\quad \times h^0_{\xi-x}(t)\,dydt \\
&= \int_{m(x)}^{+\infty}\int_0^{+\infty} v_{\xi-d(y)} f_{R_{m(x)}|H^0_{\xi-x}>m(x)}(y)\, h^0_{\xi-x}(t)\,dydt \\
&= \int_{m(x)}^{+\infty} h^0_{\xi-x}(t) \int_0^{+\infty} v_{\xi-d(y)} f_{R_{m(x)}|H^0_{\xi-x}>m(x)}(y)\,dydt \\
&= \int_{m(x)}^{+\infty} h^0_{\xi-x}(t)\,dt \int_0^{+\infty} v_{\xi-d(y)} f_{R_{m(x)}}(y)\,dy \\
&= \left(1-\int_0^{m(x)} h^0_{\xi-x}(t)\,dt\right)\int_0^{+\infty} v_{\xi-d(y)} f_{R_{m(x)}}(y)\,dy \\
&= \left(1-\int_0^{m(x)} h^0_{\xi-x}(t)\,dt\right)\int_0^{+\infty} v_{\xi-d(y)} f^0_{m(x)}(y)\,dy
\end{aligned}
$$

The conditional independence allows the replacement of $f_{R_{m(x)}|H^0_{\xi-x}>m(x)}$ by $f_{R_{m(x)}}$: as $H^0_{\xi-x}>m(x)$, the process may still be in the region $[0,\xi-x)$ and hence the region of integration remains $[0,+\infty)$.

C.2 Two thresholds models

C.2.1 Periodic case:

Let $f_{R_\tau^0, G_\mathcal{F}^0}$ be the joint probability density function of the process at time τ and the first hitting time of threshold $\mathcal{F}$. We may deduce:

$$\begin{aligned}
E[V_\tau^{R_\tau^0} \times \mathbf{1}_{\{G^0_{\mathcal{F}-d(x)} \geq \tau\}}] &= \int_0^{+\infty} \int_0^{+\infty} \{v_\tau^y \times \mathbf{1}_{\{t>\tau\}}\} f_{R_\tau^0, G^0_{\mathcal{F}-d(x)}} (y,t)\, dydt \\
&= \int_\tau^{+\infty} \int_0^{+\infty} v_\tau^y f_{R_\tau^0 | G^0_{\mathcal{F}-d(x)} > \tau} (y)\, g^0_{\mathcal{F}-d(x)} (t)\, dydt \\
&= \int_\tau^{+\infty} g^0_{\mathcal{F}-d(x)} (t) \int_0^{+\infty} v_\tau^y f_{R_\tau^0 | G^0_{\mathcal{F}-d(x)} > \tau} (y)\, dydt \qquad \text{(C.1)} \\
&= \int_\tau^{+\infty} g^0_{\mathcal{F}-d(x)} (t)\, dt \int_0^{\mathcal{F}-d(x)} v_\tau^y f_{R_\tau^0} (y)\, dy \\
&= \left(1 - \int_0^\tau g^0_{\mathcal{F}-d(x)} (t)\, dt\right) \int_0^{\mathcal{F}-d(x)} v_\tau^y f_\tau^0 (y)\, dy
\end{aligned}$$

The fourth equality is obtained using the following conditional argument: given that the process R_t has not crossed the failure threshold $\mathcal{F}$ during the time interval $[0, \tau)$, we know that the process has remained in the interval $[0, \mathcal{F} - d(x))$: the region of integration for the transition density f_τ^0 is then $[0, \mathcal{F} - d(x))$.

Bibliography

[1] British standards institution. *BS3811 Glossary of maintenance terms in Terotechnology*, 1984.

[2] J. Abate and W. Whitt. Numerical inversion of Laplace transforms of probability distribution. *ORSA Journal on Computing*, 7(1):36–43, 1995.

[3] M. Abdel-Hameed. A gamma wear process. *IEEE Transaction on Reliability*, 24(2):152–153, 1975.

[4] T. Aven and R. Dekker. A useful framework for optimal replacement models. *Reliability Engineering System Safety*, 58(1):61–67, 1997.

[5] T. Aven and U. Jensen. *Stochastic Models in Reliability.* Springer, 1999.

[6] C. T. Barker and M. J. Newby. Inspection and maintenance planning for complex multi-component system. *Proceedings of the IEEE SMC UK-RI Chapter Conference 2005 on Applied Cybernetics, London, UK*, pages 176–181, 2005.

[7] C. T. Barker and M. J. Newby. Inspection and maintenance planning for complex multi-component systems. *Proceedings of ALT'2006, ISTIA, Angers, France*, pages 299–304, 2006.

[8] C. T. Barker and M. J. Newby. Non-periodic inspections to guarantee a prescribed level of reliability. *Proceedings of the International Conference 'Statistical Methods for Biomedical and Technical Systems', Limassol, Cyprus*, pages 43–48, 2006.

[9] C. T. Barker and M. J. Newby. Optimal inspection and maintenance to meet prescribed performance standards for a multi-variate degradation model. *To appear in Proceedings of the ESREL 2006 Conference, Estoril, Portugal*, 2006.

[10] C. T. Barker and M. J. Newby. Optimal non-periodic inspections for a multi-variate degradation model. *To appear in Reliability Engineering and System Safety (RESS) International Journal*, pages 475–481, 2006.

[11] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability.* New York: John Wiley and Sons, 1965.

[12] R. E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing Probability Models.* Holt, Rinehart and Winston, Inc., 1975.

[13] A. Barros. *Maintenance des systèmes multicomposants sous surveillance imparfaite: modélisation stochastique et optimisation.* Ph.D thesis, Université de Technologie de Troyes, Troyes, 2003.

[14] A. Barros, C. Bérenguer, and A. Grall. On the hazard rate process for imperfectly monitored multi-unit systems. *Reliability Engineering and System Safety*, 90:169–176, 2004.

[15] R. Betensky. A boundary crossing probability for Bessel process. *Adv. in Appl. Prob.*, 30:807–830, 1998.

[16] N. H. Bingham. Probability models based on processes with independent increments. *Postgraduate Course, Royal Statistical Society, Wolfson Court, Cambridge*, September 2004.

[17] C. H. Blickenstorfer. *http://www.pencomputing.com/frames/obd2.html.*

[18] E. Boguslavskaya. *Optimization Problems to Financial Mathematics: Explcit Solutions for Diffusion Models.* Ph.D thesis, Universiteit van Amsterdam, Amsterdam, 2006.

[19] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion - Facts and Formulae.* Probability and Its Applications. Birkhäuser, second edition, 2002.

[20] L. Breiman. *Probability.* Addison-Wesley, Reading, 1968.

[21] J. M. Calvin. Efficient simulation for discrete path-dependent option pricing. *Proceedings of the 2001 Winter Simulation Conference*, pages 325–328, 2001.

[22] B. Castanier. *Modélisation stochastique et optimisation de la maintenance consitionelle des systèmes à degradation graduelle.* Ph.D thesis, Université de tecnologie de Troyes, Troyes, 2001.

[23] B. Castanier, C. Bérenguer, and A. Grall. A sequential condition-based repair/replacement policy with non-periodic inspections for a system subject to continuous wear. *Applied Stochastic Models in Business and Industry*, 19:327–347, 2003.

[24] D. R. Cox. Regression models and life tables. *Journal of the Royal Statistics Society*, B(34):187–202, 1972.

[25] D. R. Cox and H. D. Miller. *The Theory of Stochastic Processes.* Methuen and Co Ltd, 1965.

[26] R. Dagg. *Optimal Inspection and Maintenance for Stochastically Deteriorating Systems.* Ph.D thesis, City University, London, 2000.

[27] R. Dekker. Applications of maintenance optimization models: a review and analysis. *Reliability Engineering and System Safety*, 51:229–240, 1996.

[28] G. Doetsch. *Introduction to the Theory and Application of the Laplace Transformation.* Springer-Verlag, NewYork, 1974.

[29] D. Dufresne. Bessel processes and a functional of brownian motion. *Lecture given by the author at the Workshop on Mathematical Methods in Finance, Melbourne,* 2004.

[30] G. Dullerud and F. Paganini. *A Course in Robust Control Theory.* Springer-Verlag, 2000.

[31] A. Going-Jaeschke and M. Yor. A survey and some generalizations of Bessel processes. *Bernoulli*, 9(2):313–349, 2003.

[32] I. S. Gradshteyn and I. M. Ryzhik. *Tables of Integrals, Series, and Products.* Academic Press Inc., 1965.

[33] A. Grall, C. Berenguer, and L. Dieulle. A condition-based maintenance policy for stochastically deteriorating systems. *Reliability Engineering and System Safety*, 76:167–180, 2002.

[34] A. Grall, L. Dieulle, C. Berenguer, and M. Roussignol. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE*, 51(2):141–150, 2002.

[35] M. Green and D. Limebeer. *Linear Robust Control.* Prentice-Hall, 1995.

[36] M. Healy. *Tables of Laplace, Heaviside, Fourier, and Z transforms.* W. and R. Chambers Ltd., 1967.

[37] I. Itô and H. P. McKean. *Diffusion Processes and their Sample Paths.* Springer-Verlag, second edition, 1974.

[38] N. P. Jewell. Marker processes in survival analysis. *Lifetime Data Analysis*, 2:15–29, 1996.

[39] J. Kent. Some probabilistic properties of Bessel functions. *The Annals of Probability*, 6(5):760–770, 1978.

[40] S. Kourouklis and P.G. Moschopoulos. On the distribution of the trace of a non-central Wishart matrix. *Metron*, XLIII(1-2):85–92, 1985.

[41] D. Kumar and B. Klefsjö. Proportional hazards model: a review. *Reliability Engineering and System Safety*, 44:177–188, 1994.

[42] D. Kumar, B. Klefsjö, and U. Kumar. Reliability analysis of power transmission cables of electric mine loaders using the proportional hazards model. *Reliability Engineering and System Safety*, 37:217–222, 1991.

[43] D. Kumar and U. Westberg. Proportional hazards modelling of time-dependent covariates using linear regression: A case study. *IEEE Transactions on Reliability*, 45(3):386–391, 1996.

[44] J. Lawless and M. Crowder. Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10(3):213–227, 2004.

[45] A. Lehmann. Degradation-failure models with repair. *Proceedings of the first International French-Russian Workshop 'Longevity, Aging and Degradation Models in Reliability, Public Health, Medicine and Biology', St. Petersburg, Russia*, pages 195–204, 2004.

[46] A. Lehmann. Degradation threshold-schock models. *Inserm Workshop 154, Recent Developments in Censored Data Analysis, Toulon, France*, 2004.

[47] A. Lehmann. Joint modelling of degradation and failure time data. *Proceedings of ALT'2006, ISTIA, Angers, France*, pages 26–32, 2006.

[48] V. Linetsky. The spectral representation of Bessel processes with constant drift: Application in queueing and finance. *J. Appl. Prob*, 41(2):327–344, 2004.

[49] Y. L. Luke. *Integrals of Bessel functions.* McGraw-Hill Book Company, Inc., 1962.

[50] A. M. Mathai. The exact noncentral distribution of the generalized variance. *Annals of the institute of Statistical Mathematics*, 24:53–65, 1972.

[51] A. M. Mathai. Moments of the trace of a noncentral Wishart matrix. *Communications in Statistics - Theory and Methods*, A9(8):795–801, 1980.

[52] A. M. Mathai and K. C. S. Pillai. Further results on the trace of a noncentral Wishart matrix. *Communications in Statistics - Theory and Methods*, 11(10):1077–1086, 1982.

[53] H. P. Jr. McKean. A winding problem for a resonator driven by white noise. *J. Math. Kyoto Univ.*, 2:227–235, 1963.

[54] N. W. McLachlan. *Bessel functions for Engineers.* Oxford University Press, 1934.

[55] C. Meier-Hirmer, F. Sourget, and M. Roussignol. Optimising the strategy of track maintenance. *Advances in Safety and Reliability: Proceedings of the ESREL 2005 Conference, Tri City, Poland*, 2:1385–1391, 2005.

[56] R. J. Muirhead. *Aspects of Multivariate Statistical Theory.* John Wiley and Sons, Wiley Series in Probability and Mathematical Statistics, 1982.

[57] M. J. Newby. Estimating of Paris-Erdogan law parameters and the influence of environmental factors on crack growth. *Int. J. Fatigue*, 13:187–198, 1991.

[58] M. J. Newby. Analysing crack growth. *Proc. ImechE.*, Part G:157–166, 1998.

[59] M. J. Newby. *Probability, Statistics and Modeling in Public Health*, chapter Bivariate Decision Processes, pages 364–377. Springer-Verlag, 2006.

[60] M. J. Newby and C. T. Barker. A bivariate process model for maintenance and inspection planing. *International Journal of Pressure Vessels and Piping*, 83:270–275, 2006.

[61] M. J. Newby and R. Dagg. Optimal inspection and perfect repair. *IMA Journal of Management Mathematics*, 15(2):175–192, 2004.

[62] R. P. Nicolai, G. Budai, R. Dekker, and M. Vreijling. Modelling the deterioration of the coating on steel structure. *Advances in Safety and Reliability: Proceedings of the ESREL 2005 Conference, Tri City, Poland*, 2:1443–1448, 2005.

[63] R. P. Nicolai, G. Budai, R. Dekker, and M. Vreijling. Modelling the deterioration of the coating on steel structure: a comparison of methods. *Proceedings of MIMAR 2004 - Fifth IMA International Conference, University of Salford, UK*, 2:197–202, 2005.

[64] F. Oberhettinger and L. Badii. *Tables of Laplace Transforms.* Springer-Verlag, 1973.

[65] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer, fifth edition, 2000.

[66] P. Patie. *On some first passage time problems motivated by financial applications.* Ph.D thesis, Swiss Federal Institute of Technology, Zurich, 2004.

[67] D. F. Percy and B. M. Alkali. Grim scheduling applications. *Proceedings of MIMAR 2004*, pages 213–217, 2004.

[68] J. W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. Appl. Prob.*, 7:511–526, 1975.

[69] J. W. Pitman and M. Yor. Processus de Bessel, et mouvement brownien, avec drift. *C. R. Acad. Sc. Paris*, Série A:151–153, 1980.

[70] J. W. Pitman and M. Yor. Bessel process and infinitely divisible laws, stochastic integrals. *Lecture Notes in Mathematics 851, Springer Berlin*, pages 285–370, 1981.

[71] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C, 2nd Edition*, volume 5. Cambridge Univerity Press, 1982.

[72] A. P. Prudnikov. *Integrals and Series: Inverse Laplace Transforms.* Gordon and Breach, 1992.

[73] D. Revuz and M. Yor. *Continuous Martingale and Brownian Motion.* Springer Verlag, 1991.

[74] L. C. Rogers and D. Williams. *Diffusions, Markov processes and Martingales*, volume Volume 1: Foundations (2nd Edition). John Wiley and Sons, 1994.

[75] L. C. G. Rogers and J. W. Pitman. Markov functions. *Annals of Probability*, 9:573–582, 1980.

[76] S. Ross. *Applied Probability Models with Optimization Applications.* Holden-Day, 1970.

[77] P. A. Scarf and M. Deara. On the development and application of maintenance policies for a two-component system with failure dependence. *IMA Journal of Mathematics Applied in Business and Industry*, 9:91–107, 1998.

[78] P. A. Scarf and M. Deara. Block replacement policies for a two-component system with failure dependence. *Wiley InterScience*, 50:70–87, 2002.

[79] X. Shao and C. Yin. On the joint distribution of the future infinimum and its location for a transient Bessel process. *Journal of Mathematical Research and Exposition*, 21(3):344–348, 2001.

[80] T. Shiga and S. Watanabe. Bessel diffusions as a one-parameter family of diffusion process. *Z. Wahrscheinlicjkeittheory verw.*, 27:37–46, 1973.

[81] N. D. Singpurwalla. Survival in dynamic environments. *Statistical Science*, 10(1):86–103, 1995.

[82] K. Sobczyk and B. F. Spencer Jr. *Random Fatigue, From Data to Theory.* Academic Press, Inc., 1992.

[83] J. Soszynska. Reliability of large parallel-series system in variable operation conditions. *Advances in Safety and Reliability: Proceedings of the ESREL 2005 Conference, Tri City, Poland*, 2:1877–1884, 2005.

[84] L. Speijker, J. M. van Noortwijk, M. Kok, and R. Cooke. Optimal maintenance decisions for dikes. *Probability in the Engineering and Informational Sciences*, 14:101–121, 2000.

[85] J. M. van Noortwijk, R. M. Cook, and M. Kok. A bayesian failure model based on isotropic deterioration. *European Journal of Operational Research*, 82:270–282, 1995.

[86] J M. van Noortwijk, M. Kallen, and M. D. Pandey. Gamma processes for time-dependent reliability of structures. *Advances in Safety and Reliability: Proceedings of the ESREL 2005 Conference, Tri City, Poland*, 2:1457–1464, 2005.

[87] J. M. van Noortwijk and H. E. Klatter. Optimal inspection decisions for the block mats of the eastwern-scheldt barrier. *Reliability Engineering and System Safety*, 65:203–211, 1999.

[88] J. M. van Noortwijk, M. Kok, and R. Cooke. Optimal maintenance decisions for the sea-bed protection of the eastern-scheldt barrier. *Engineering Probabilistic Design and Maintenance for Flood Protection*, pages 25–56, 1997.

[89] J. M. van Noortwijk and P. van Gelder. Optimal maintenance decisions for berm breakwaters. *Structural Safety*, 18(4):293–309, 1996.

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

[91] C. Yin. The joint distribution of the hitting time and place to a sphere or spherical shell for Brownian motion with drift. *Statistics and Probability Letters*, 42:367–373, 1999.

[92] M. Yor. On square-root boundaries for Bessel processes, and pole-seeking brownian motion. *Stochastic Analysis and Applications (Lecture Notes Math. 1095)*, pages 100–107, 1984.

[93] R. Zequeira and C. Bérenguer. A block replacement policy for a periodically inspected two-unit parallel standby safety system. *Advances in Safety and Reliability: Proceedings of the ESREL 2005 Conference, Tri City, Poland*, 2:2091–2098, 2005.