The Effect of Imperfect Maintenance on Deterioration Processes

(An investigatory study)

M.Sc. Thesis

Samira Safaei Farahani

Technical University of Delft

Faculty of Electrical Engineering, Mathematics and Computer Science

Department of Applied Mathematics

July 2008

The work was kindly supervised by:

Prof. dr. ir. J. M. van Noortwijk

dr. J. A. M. van der Weide

Delft University of Technology

and

dr. R. P. Nicolai

dr.ir. M. J. Kallen

"HKV Consultants"

Acknowledgments

I would like to express my gratitude to dr. J. A. M. van der Weide, Prof. dr. ir. J. M. van Noortwijk, dr. R. P. Nicolai, and dr.ir. M. J. Kallen as my supervisors for all advises and the devoted time. Also thanks to Prof. dr. M. Dekking for accepting the responsibility of being the chairman of the defense committee. I would like to thank HKV Consultants for giving the opportunity to do my thesis project in their company which is a valuable experience for me. My special thanks are to Roger M. Cooke and Dorota Kurowicka for giving me the chance to spend two wonderful years in the Risk and Environmental Modeling group.

I would also like to thank my parents, my brother, and my dear Behnam for all their love and support. My great appreciation goes out to my best friend Leszek, for all the help, friendly advises and comments. Last but not least, a special word of thanks goes out to all my colleagues and friends for being next to me and letting me share a great, unforgettable time with them.

Abstract

Many metal structures are subject to deterioration such as chemical corrosion, stress corrosion cracking, brittle fracture and fatigue. Therefore, it is required to maintain these structures before they collapse or cause some damages which are not compensable. Many maintenance models have been developed to describe a deterioration process and give optimal solutions to decrease the cost and increase the system reliability of such systems. However, most of these models are based on the ideal assumptions such as perfect inspection and maintenance.

Our interest is in investigating the effect of imperfect maintenance on a deterioration process. Accordingly, in the first part of this project, modeling the deterioration process of coating systems protecting steel structures, which are exposed to outdoor weathering conditions, is studied. A physical corrosion process is simulated in order to enable us to study the deterioration process from the mathematical point of view before and after (the first) imperfect maintenance. This information is useful for finding the optimal maintenance cost or system availability. The second part of the project focuses on the theoretical behavior of the time between maintenance actions considering a large number of maintenance actions. The results of this section can be used as a (decision) criterion for decision makers to determine the average number of required repairs before the system replacement.

Contents

List of Figures

Chapter 1

Introduction

Steel structures such as bridges, tanks and pylons are exposed to outdoor weathering conditions. In order to prevent them from corrosion they are protected by an organic coating system which is subjected to deterioration itself. Corrosion means the breaking down of essential properties in a material due to chemical reactions with its surroundings. Commonly, this means a loss of electrons of metals reacting with water and oxygen. Weakening of iron due to oxidation of the iron atoms is a well-known example of electrochemical corrosion. Consequently, it is always important to model the deterioration (or corrosion as a special case) process to be able to study its characteristics and find a way to control it.

The deterioration modeling for metal structures is used to predict the quality of the coating in time in order to determine when inspections or maintenance should be done and to assess life cycle costs [1]. In general, the deterioration can be modeled in three ways:

- As a black-box statistical time to failure (such as lifetime distribution);
- As a gray-box stress-strength model based on a measurable quantity indicating time-dependent deterioration and failure (such as Brownian motion and the gamma process);
- As a white-box model through simulation of the physics of measurable deterioration and failure.

The first model results in a static (time-invariant) model, the latter two result in a time-dependent stochastic process. Since the black-box approach does not allow for the temporal variability associ-

ated with evolution of degradation, it is less suitable for modeling deterioration of steel structures than the two other ones. Moreover, gray-box models are often preferred over white-box models, because they considerably facilitate the computation of properties such as the lifetime distribution [2]. In the literature on deterioration modeling, both the gamma process and Brownian motion have been used to predict the value of a measurable quantity [3]. Another way of modeling the deterioration is by using a regression model, where the regression function depends on (transformed) time. However, a regression model cannot capture temporal variability associated with evolution of deterioration. As a consequence, the deterioration along a specific sample path is deterministic in the regression model, whereas it varies probabilistically in the Brownian motion and gamma process model [1]. On the other hand, such gray-box models do not give much insight into the physics of the degradation process and since the process is often not analytically tractable, it is useful to have a physical model to have deeper understanding about the process. Nevertheless, in simulating the physical process, the statistical analysis may become very time-consuming which is not desired [2].

At the moment, many maintenance models are available which are based on continuous/partial inspection, perfect/imperfect maintenance, and continuous/discrete time axis. Each of these models attempts to give an optimal solution to increase the system availability and safety, decrease the cost, or both; examples of these models can be found in [1; 4; 5; 6; 2; 7; 8]. However, in all of these papers, either the authors have not considered the maintenance (in particular, imperfect maintenance) or if they did so ([1], and [2]), they have not studied its effect on the deterioration process.

The aim of this thesis is to study the effect of imperfect maintenance on the deterioration process –specifically corrosion process– of steel structures. As a result, we are interested in representing the deterioration process with a mathematical process, such as a gamma or Poisson process, before and after maintenance, in order to be able to explain the corrosion characteristics and properties under the influence of imperfect maintenance. In addition, finding a theoretical interpretation of the behavior of the time between maintenance actions is required to have a good decision criterion for planning the repair and replacement strategies. To fulfill these goals, a white-box model is

designed to describe the deterioration process of a steel surface. Moreover, it is shown theoretically that the time between two maintenance actions has a stationary property (i.e., it converges to a specific value), which can be informative for decision-makers, in order to plan optimal inspection strategies due to the average inter-repair time.

The remainder of the report is organized as follows: In Chapter 2 a physical model which simulates the corrosion process is presented. The underlying assumptions as well as the construction of this model are set out into details besides the simulation scheme. Since we are interested in investigating the effect of (imperfect) maintenance on the deterioration process, a physical model is required to simulate data. In Chapter 3 the model is fitted to these data. This enables us to study the effect of maintenance from the mathematical point of view, and figure out how imperfect maintenance influence the corrosion process; all the corresponding results are presented in this chapter. In Chapter 4, the distribution of inter-repair times is found analytically, based on the asymptotic results of renewal and superposition processes. Moreover, the long term behavior of this distribution is studied for a certain class of distributions. Finally, conclusions are given in Chapter 5.

Chapter 2

Corrosion process simulation

In this chapter, the physics of the corrosion process are explained to give an insight of the real deterioration process of steel structures. Then a physical model is introduced to show how the real process can be simulated by a mathematical model. At the end of this chapter, the proposed model is compared to two other physical models and the differences and similarities are presented.

2.1 Corrosion process

To get a better understanding of the degradation processes, a brief explanation is given on how steel objects are protected against deterioration. Corrosion is a chemical reaction which takes place on the surface of a metal. It involves a reaction with the oxygen in the air producing a metal compound. The speed at which a metal corrodes is dependent on its reactivity; more reactive metals corrode faster than less reactive ones.

To protect the steel from corrosion the galvanization method can be used, i.e. a steel object is dipped into molten zinc giving a thin layer of zinc on the object. The zinc forms a physical barrier against water and oxygen. If this layer is scratched, then it provides sacrificial protection against corrosion (it will corrode instead of the steel). However, in order to prevent a too rapid corrosion of the zinc layer, which would leave the steel unprotected after several years, an organic coating protects the zinc. An intact organic coating shields the zinc from the corrosive environment and prevents it from corroding. Unfortunately, the organic coating is subjected to degradation itself.

Figure 2.1: Corrosion of water pipes

Under the influence of humidity and UV radiation, the protective properties diminish slowly in time, eventually resulting in corrosion of the zinc layer. Too much zinc corrosion is not allowed, because it releases zinc to the environment, causing pollution. Ultimately, the steel corrodes and the structural integrity of the object can be compromised.

Modeling the degradation of a coating system is important in order to optimize inspections and maintenance actions on the basis of the characteristics of data from the corrosion process. Hence, knowing the location of the deterioration is as important as knowing its amount. Note that in contrast to black- and gray-box processes which do not take the location of the corroded spots into account, the white-box processes, which relate more to the physics of degradation, do. On the other hand, black- and gray-box processes may be more convenient to express lifetime statistics of the coating and to allow all kind of optimizations [1].

2.2 Model construction

Recently, Condition-Based Maintenance (CBM) policies have received increasing attention to improve the effectiveness of the preventive maintenance. CBM is a technique according to which the decision of maintaining the system is taken dynamically on the basis of the observed condition of the system. Due to accurate sensors that can continuously provide performance indicators and

process parameters at low cost, such policies are important. In many of the available CBM models, the condition is defined as a preventive threshold; if the deterioration process exceeds this level, then the preventive action (such as repair) will be taken, [9; 10; 2; 11; 12]. In our case, maintenance action is performed according to the corrosion condition of the steel surface.

Visual inspections are used to determine the amount of corrosion deterioration. The amount of deterioration can be expressed as a percentage of the corroded surface or as a percentage of the surface to be repaired. Corrosion usually starts in small spots which grow in time with a more-or-less random pattern on a whole surface. Determining the actual percentage of the corroded surface of large objects by visual inspection is not easy since one has to do a detailed counting. It is much easier to give a rough estimate of the amount of surface to be repainted. According to the experts' information, maintenance action (spot repair or repaint) is performed once 3% of the whole surface under inspection is corroded [1], [2].

There are some white-box models available to simulate the deterioration process. Some examples of these models can be found in [1], [4], and [2]. In these models, corrosion occurs according to a stochastic process such as the Poisson process; however, there are different approaches to model its initiation and expansion procedure, i.e., in some models both of them are stochastic, such as in [2] and [4], while in some others, one is stochastic the other one is deterministic, such as in [1]. Also the propagation process depends on different criteria such as the neighboring conditions or the intensity changes.

The physical model that is proposed in this thesis, is built in such a way that the advantages and strengths of the other models also have been considered. The remainder of this chapter is devoted to the model assumptions and simulation of the corrosion process, follows by some comparison to the other available models.

2.3 Model assumptions

As mentioned, the preventive maintenance (PM) is performed once the percentage of the corroded surface exceeds the preventive threshold 3%. At this time the corroded spots will be repainted; the area to be repaired is about 7 times larger than the corroded area, according to experts information. In our model, each spot is going to be repainted with one level of its neighbors. It is assumed that each spot has eight direct neighbors and up to n levels of neighbors can be considered according

to the model application and the surface qualities or corrosion characteristics. In our model, we consider ten levels of neighbors for each spot.

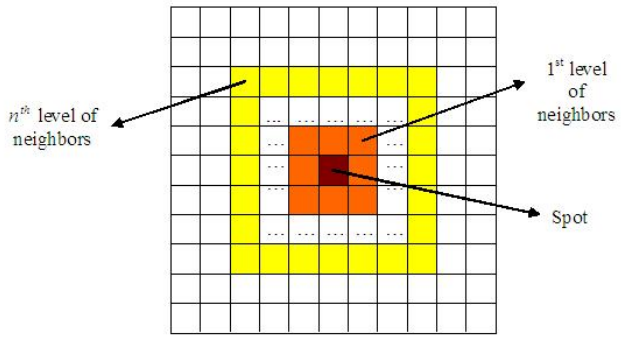

Figure 2.2: n levels of neighbors of each corroded spot

The maintenance process is imperfect in a sense that after each PM action, the surface condition will not be returned to a as good as new condition since we do not paint the whole surface and just do spot repair. However, the repainting process is perfect from the point of view that the repainted spots are considered as good as new cells with intensity rate of the pristine (not corroded) surface, and all the corroded spots will be repainted; i.e. there is no area with corrosion which is not inspected and repaired after the PM action. Note that the repaired spots will be considered as a choice for being corroded again. The other levels of neighbors keep the intensity rate that they possess at the time of PM. In reality, a maintenance action changes the intensity rate of the repaired area to a higher or smaller value depending on the maintenance quality. Nevertheless, due to the complexity of the physical model we neglect this fact and simply reset the intensity rate of the repaired spots to the original value, λ_0. The model is built in such a way that each cell takes the higher intensity rate at each time. So if one cell receives two different rates from two corroded spots at once, its intensity rate will be the larger one; i.e., it will be affected by the corroded spot to which it is closer. This is compatible with what happens in reality.

Another parameter of corrosion that can be taken into account is the depth of the corrosion. In the proposed model this factor has not been considered, and the study is just limited to a two dimensional problem considering the corrosion expansion only on surface and not in depth. This

simplifies our model especially in dealing with different levels of neighbors and intensities. Note that all parameters are estimated by means of expert judgment. If there are some data available, then an alternative candidate for the parameter estimation will be the maximum likelihood estimation or method of moments.

To get a better understanding of how the corrosion starts and expands, there is a brief explanation of the initiation and propagation of the corrosion process in the next parts.

2.4 Spot initiation

The coating layer is represented by a $n \times n$ grid consisting of squares. Due to the non-linear characteristics of the deterioration, it is assumed that spots hit the surface according to a non-homogeneous Poisson process (NHPP). This is in fact because of the dependent structure between neighbors which causes a faster deterioration of the cells which are surrounded by spots that the ones which do no have too many spots around.

Definition 2.4.1. A stochastic process $N(t)$ with the rate parameter λ, is a time-homogeneous, one-dimensional Poisson process if:

1. the numbers of events occurring in two disjoint (non-overlapping) subintervals are independent random variables;

2. the probability of the number of events in some subinterval $[t, t+h]$ is given by

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

A non-homogeneous Poisson process is a Poisson process with rate parameter $m(t)$ such that the rate parameter of the process is a function of time, t. Therefore in this case, λh will be replaced by $\int_t^{t+h} m(s)ds$.

Note that if $X_i \sim \text{Poi}(\lambda_i), i = 1, \ldots, n$, follow a Poisson distribution with parameter λ_i, and X_i's are independent, then $Y = \sum_{i=1}^n X_i \sim \text{Poi}\left(\sum_{i=1}^n \lambda_i\right)$, also follows a Poisson distribution whose parameter is the sum of the component parameters. In our model, the number of arrivals of the spots on the whole surface has a Poisson distribution with parameter $\lambda = n \times n \times \lambda_0$ where $n \times n$ denotes the area of the surface, and λ_0 corresponds to the intensity rate of each square cell for

the initial stage at which the surface has not been corroded yet. It is worthwhile to mention that the intensity rate of the NHPP that is chosen for this model is a time-dependent function $m(t) = q\lambda t^{(q-1)}$ where $q \neq 1$ indicates the non-linearity of the intensity function w.r.t time, and consequently the expected number of corroded cells appearing at each time interval Δt is

$$E(N(\Delta t)) = \int_t^{t+\Delta t} m(s)ds = \lambda\left((t+\Delta t)^q - t^q\right)$$

To start the corrosion process we consider a so called location function with values in the interval $[0, 1]$, which is given in matrix form (the initiation matrix), expressing the probability of the corrosion occurrence on different locations of the surface; for example in pipes, the bottom is more probable to be exposed to corrosion, and the same is true for the area next to welds and connections. In the following figure a sample function is presented which shows that the bottom of the surface is more probable to be corroded. In fact, what is shown in the figure below, is a pipe which is unfolded on a square surface.

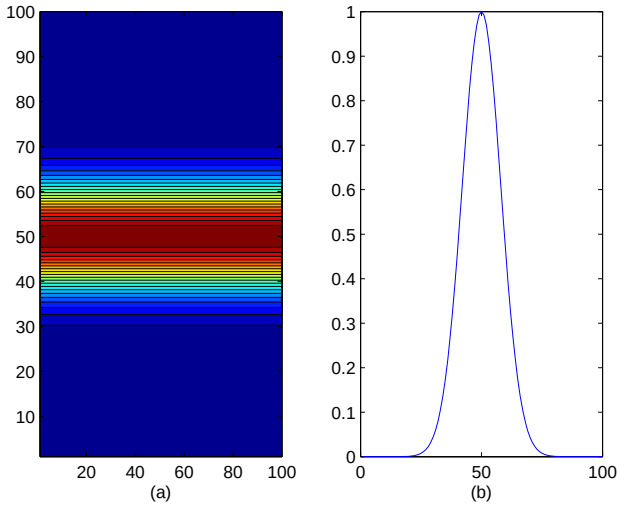

Figure 2.3: A sample function of the initial corrosion locations; (a) The contour plot of the surface showing the most probable corrosion locations, (b) The probability distribution of the surface which is exposed to corrosion

The location function is defined as a cumulative distribution function of the normal distribution

$$f(i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(i-\mu)^2}{2\sigma^2}}$$

where i indicates the y-coordinate of each cell on the surface, μ is the mean which corresponds to the most probable area that can be exposed to corrosion, and σ^2 is the variance which is computed in the following way:

$$\min_{\sigma \geq 0} \left(\int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}} dz - \alpha \right)$$

where x_1 and x_2 corresponds to the error interval that experts consider in expressing the most corrodible location, and α is the confidence level that shows what percent of the corrosion occurs at this location. In figure 2.3, the inputs of the two mentioned functions are as follows:

- The surface is a square of size 10×10 square-meters, i.e., each cell has the size 1×1 dm.
- The most probable location which is exposed to the corrosion is at the bottom of the pipe (or in the middle of the flatted surface, $\mu = 50$) considering $\pm 20m$ error, i.e., $[x_1, x_2] = [30, 70]$.
- The last input is α which indicates the probability that the corrosion will occur at this area; in the figure above 95% of the corrosion is located at the bottom of the pipe considering the range $(30, 70)$.

At each time a random number of squares will be infected by corrosion and removed from the initiation matrix to avoid having corrosion at one cell twice. In other words, the number of corroded spots, $N(\Delta t)$, at each time step Δt for all $t = 1, 2, \ldots, T_{last}$, is defined as

$$N(\Delta t) = N(t + \Delta t) - N(t)$$

and has a Poisson distribution with parameter $\lambda\left((t + \Delta t)^q - t^q\right)$ where in our model $q = 2$, according to experts' information. As soon as the corrosion expansion (explained in the next subsection) starts, those spots will also be removed from the initiation matrix. The initiation process of spots continues until the time at which either 3% of the surface is corroded or there are no more corrodible locations left in the initiation matrix.

2.5 Spot propagation

Once the spot occurs on a surface, the corrosion propagation starts. The corrosion increments follow the Poisson process as well. However, the parameter of the process is different than λ_0. To explain how the intensities change during the corrosion expansion, first we have to introduce the propagation policy. The degradation of a coating layer is modeled as a Markov random field. Before describing the model itself, a few definitions are given to introduce the Markov random field. For decades, Markov Random Fields (MRF) have been used by statistical physicists to explain various phenomena occurring among neighboring particles because of their ability to describe local interactions between them. Basically an MRF model is a spatial-domain extension of a temporal Markov chain where an event at the current time instant depends only on events of a few previous time instants. In MRF the statistical dependence is defined over the neighborhood system, a collection of neighbors, rather than past events as in the Markov chain model [13]. To express the MRF in mathematical language it is required to start with some pre-definitions followed from [13].

Definition 2.5.1. Let X, Y, Z be discrete random variables taking their values in the sets E, F, G respectively. One says that X and Y are conditionally independent given Z if for all x, y, z in E, F, G, respectively, events $\{X = x\}$ and $\{Y = y\}$ are conditionally independent given $\{Z = z\}$.

Consider a stochastic sequence $\{X_n\}_{n\geq 0}$. It is said to have the Markov property if for $n \geq 1$, X_n is statistically independent of $\{X_k, k = 0, 1, \ldots n-2\}$ given X_{n-1}. This suggests a local dependency among adjacent time slots. In more general cases we can extend the notion of dependency to multiple dimensions.

Definition 2.5.2. Let S be a finite set, with elements denoted by s and called sites, and let Λ be a finite set called the **phase space**. A **random field** on S with phases in Λ is a collection $X = \{X(s)\}_{s\in S}$ of random variables $X(s)$ with values in Λ.

A random field can be regarded as a random variable taking its values in the **configuration space** Λ^S. A configuration $x \in \Lambda^S$ is of the form $x = (x(s), s \in S)$, where $x(s) \in \Lambda$ for all $s \in S$. For a given configuration X and subset $H \in S$, define:

$$x(H) = \{x(s), s \in H\},$$

the restriction of x to H. If $S\backslash H$ denotes the complement of H in S, one writes $(x(H), x(S\backslash H))$. In particular, for fixed $s \in S, x = (x(s), x(S\backslash s))$, where $S\backslash s$ is a shorter way of writing $S\backslash\{s\}$. Of special interest are fields characterized by local interactions. Before defining the Markov property for random fields we first need to introduce a **graph** on the sites.

Definition 2.5.3. A neighborhood system on S is a family $N = \{N_s\}_{s\in S}$ of subsets of S such that for all $s \in S$:

1. $s \notin N_s$.

2. $t \in N_s \Rightarrow s \in N_t$.

N_s is called the neighborhood of site s. The pair (S, N) is called a graph.

Definition 2.5.4. The boundary of a subset $H \subset S$ is the set $\partial H = (\cup_{s\in H} N_s)\backslash H$. In the graph interpretation, S denotes the set of vertices and N the set of edges. We say that two sites s and t are neighbors if and only if there exists an edge between them.

Based on the above definitions and setting $\widetilde{N_s} = N_s \cup \{s\}$, we may now define a Markov Random Field:

Definition 2.5.5. The random field X is called a **Markov random field** (MRF) with respect to the neighborhood system N if for all sites $s \in S$ the random variables $X(s)$ and $X(S\backslash\widetilde{N_s})$ are independent given $X(N_s)$:

$$P(X(s) = x(s)|X(S\backslash s)) = P(X(s) = x(s)|X(N_s) = x(N_s))$$

for all $s \in S, x \in \Lambda^S$.

In our model, the set S in definition (2.5.4) can be assigned to the whole surface (the $n \times n$ grid which represents the field); therefore, a site s refers to each of the cells (squares) of the grid. In fact, the random field is the collection of $\{0, 1\}$-valued random variables which presents the condition of each square (being corroded or not). In addition, the neighborhood N_s of each of the squares (site s) is its surrounding area up to the nth level, see figure 2.2. A corroded cell is called a spot and such a spot will infect its neighbors after some time. This infection is not the same for all the neighboring levels and depends on the intensity of each level. The more detailed information about the intensity of the neighbors will be given in the next lines. The MRF indicates that the

probability that one cell in the neighborhood of the corroded spot will be infected is independent of the probability that the other cells out of this neighborhood become infected.

The suggested approach is to consider a distance dependency between the neighbors and the corroded cells; i.e. the further the neighbors are, the less probable it is to be affected by corrosion due to that spot. In fact, according to the MRF property, these n levels of neighbors of each spot will be infected independently of the other cells; however, each of these n levels might be infected with different probabilities corresponding to their distance from the corroded cell. Therefore, there are different parameters λ_i where i indicates the level of the neighbors. By proceeding from level i to level $i+1$, the value of each λ_i decreases such that on the last level it is close to the original parameter λ_0. The function that generates each λ_i is an exponential function

$$\lambda_i = \lambda_0 \left(\exp(-\beta * i) + 1\right) \quad i = 1, \ldots, 10$$

where λ_0 is the intensity rate of each cell at the beginning (when there is no corrosion), i denotes the level of neighbors, and β is a factor that corresponds to the quality of the coating. From the physical point of view, small β means that the coating quality is not good; if a spot occurs, the probability that the neighbors will also be infected is high. Mathematically speaking, small β causes λ_i to converge slowly to λ_0, so a wider area can be affected by the corroded cell (of course with different intensity rates). The visualization of this effect is shown in figure 2.4. The first graph shows the spots with ten levels of their neighbors. The further neighbors have a lighter color. The second graph is the intensity graph which shows different intensities for the different levels of neighbors. As it is noticeable, all the intensity values on the colorbar next to the graph are attained by the different levels of neighbors. This indicates that each spot will exactly affect all ten levels of its neighbors which confirms the bad quality of coating and its permeability.

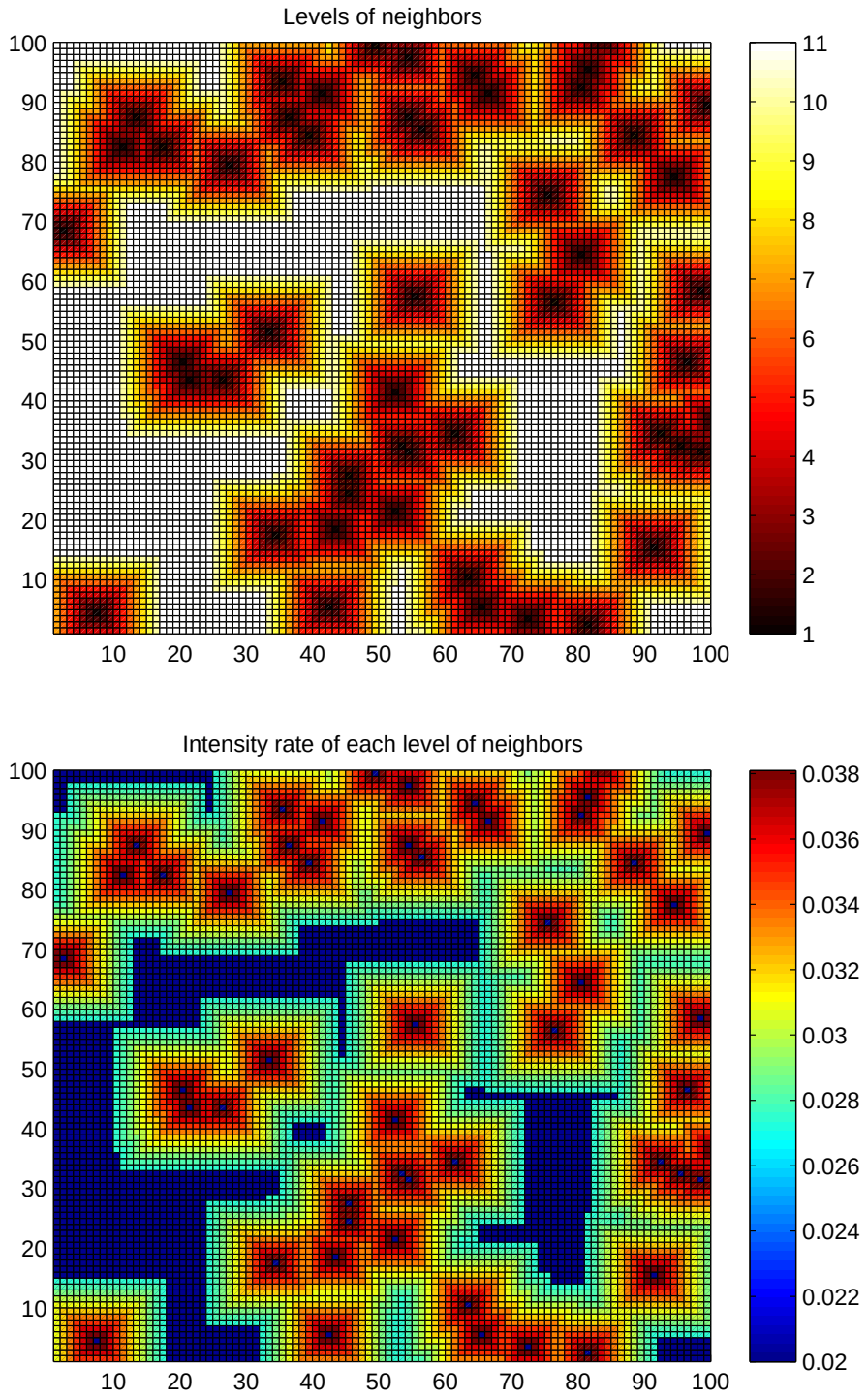

Figure 2.4: Ten levels of neighbors and intensity rate of each level for $\beta = 0.1$

In figure 2.5 it is shown how slowly λ_i converges to λ_0 when β is small.

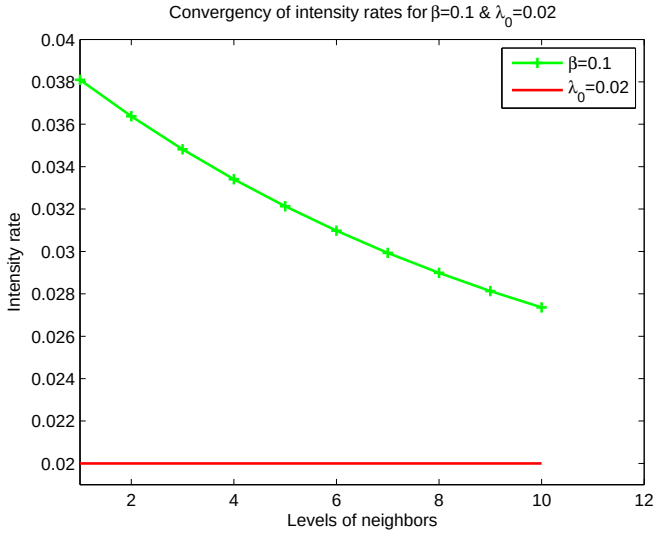

Figure 2.5: Convergence behavior of λ_i for $\beta = 0.1$

As it is visible, for $\beta = 0.1$ the intensity function does not converge to the original intensity after 10 iterations. That is why all levels of neighbors will be affected by the corroded cell.

On the other hand, a large β shows a good quality of coating materials; if a spot occurs, it is less probable that the neighbors will also be infected comparing to the previous case. The λ_i converges faster to the original rate, λ_0, so there are less level of neighbors that might be influenced by corrosion. In figure 2.6 it is shown how a large β can decrease the corrosion influence on neighbors. The first graph is similar to the first one in figure 2.4; however, the intensity graph is totally different from the previous one. Considering the colorbar of the graph, only five levels of neighbors are affected by the corrosion. This indicates that in case of corrosion, its expansion is not as fast as the case with the small β. So either the good quality of coating or the good quality of the steel structure prevents the quick corrosion expansion.

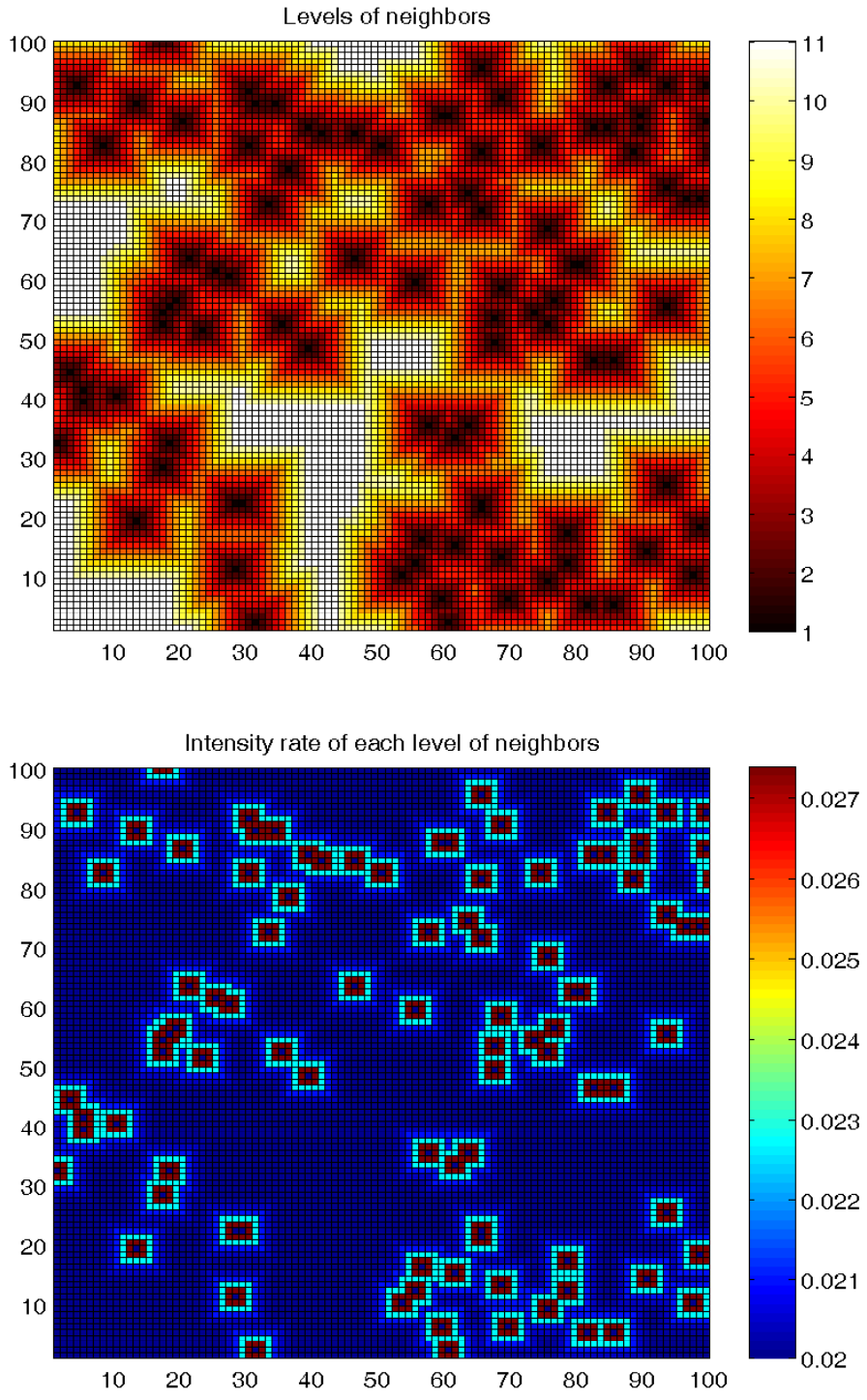

Figure 2.6: Ten levels of neighbors and intensity rate of each level for $\beta = 1$

In figure 2.7, it can be seen that convergence occurs at the fifth level of neighbors.

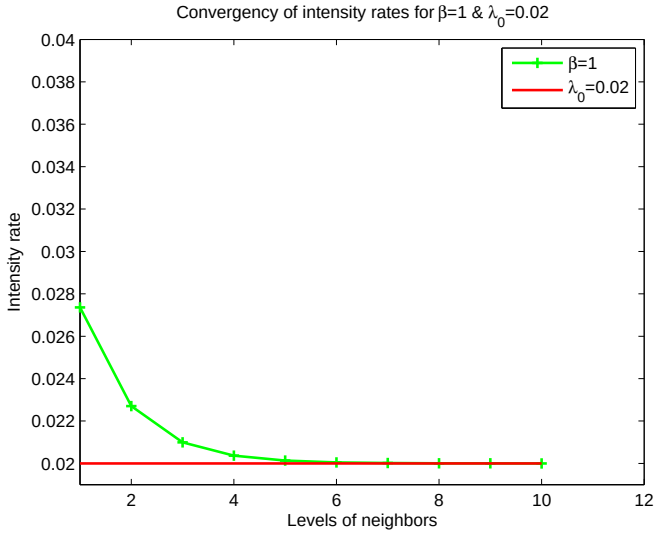

Figure 2.7: Convergence behavior of λ_i for $\beta = 1$

Comparison of figure 2.4 and figure 2.6 proves the importance of the intensity diagrams. In both of these figures the diagrams of levels of neighbors look similar and do not give much information about the possible corrodible area; however, intensity figures show there is a big difference between the cases with small and large β. The parameter β should be chosen according to experts opinion in the field of coating. In figure 2.8 it is shown how fast intensities converge to the original intensity, λ_0, for different values of β.

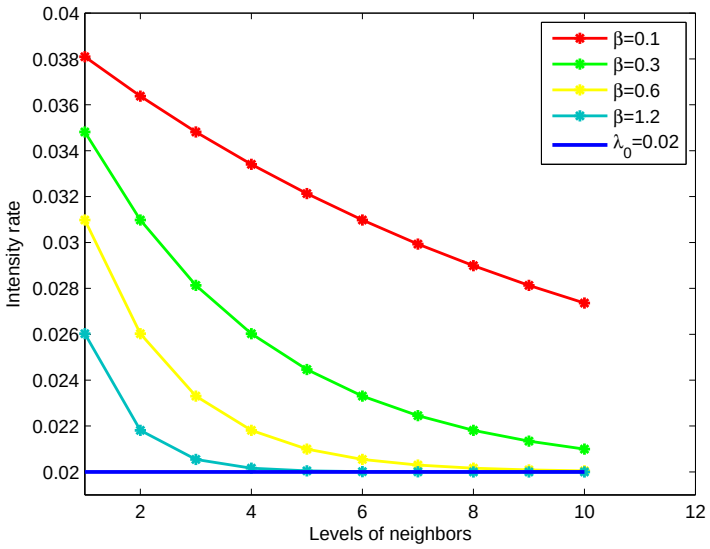

Figure 2.8: Convergence behavior of intensities for different values of β

When the set of the neighbors and their intensities are specified for each corroded spot, some new spots will be added from this set to the set of corroded spots according to a non-homogeneous Poisson process $N_i(t)$ with a rate $\lambda_i\left((t+\Delta t)^q - t^q\right)$ for each level $i = 1, \ldots, n$ of the neighbors (As before, $q = 2$). At each inspection time (i.e., time steps), new corroded cells from the expansion process will be added to the corrosion set and removed from the initial corrosion set (set of possible locations for corrosion).

Similar to the initiation process, the propagation process also continues until either 3% of surface is corroded or all the neighbors are corroded which is rare to occur since the first condition is probably satisfied before this time.

2.6 Model comparison

As stated earlier, there are several physical models that simulate the deterioration process. Here we compare the two mentioned white-box models to our model. As the first model, consider the one which is discussed in [1] and its extension can be found in [2]. This model also considers the spot initiation as a non-homogeneous Poisson process, with a deterministic spot propagation in the first model presented in [1] and a stochastic one in the extended model discussed in [2].

However, in contrast to our model, the stochastic model for corrosion expansion is another point process different than the Poisson process and is more complicated with more computational effort. The expansion process is due to the Markov random field which in fact gave us the inspiration to build the neighborhood structure of the model in this way. In the proposed model in [2], the direct neighbors are four rather than eight in our model and the intensity rate of the neighbors of corroded spots will be changed but just in one level of neighbors. The parameter estimation is done by the method of maximum likelihood using the observed corrosion data of steel gates in the Dutch Haringvliet storm surge barrier. In the models of paper [2] and [1] no maintenance action has been considered.

Another physical model is proposed by A. Ostrowska in her master thesis with the title "Simulating inspections on corroded surfaces", [4]. In this model also two stages are considered for a corrosion process; spot initiation and spot expansion which are both modeled as Poisson processes. The construction of the spot initiation in this model gave us the idea how to design the initiation part of the model. Here also only one level of neighbors, including eight direct neighbors, is considered, comparing to ten levels of neighbors in our model. The spot expansion occurs according to some criteria corresponding to the corrosion depth. The intensity of the direct neighbors will be changed as well if these criteria are satisfied. The same as the previous model, no maintenance action is considered. The parameters of the model were estimated according to experts' data.

It was possible to use any of these models in our case considering the required modifications; however, since none of the models applied maintenance, the modification would be as time consuming as designing a new model. As a result, the proposed model considers the advantages of both former models and gives a new approach to simulating the corrosion process with a white-box model.

2.7 Summary

In this chapter the physical process of corrosion is simulated to generate the required data to be able to study the corrosion behavior. Both of the spot initiation and spot expansion processes are modeled as a non-homogeneous Poisson process; however, the rate of the processes differs between initiation and propagation process. In this model a neighboring network is considered such that for each spot, up to ten levels of neighbors are exposed to corrosion according to the rate of each level and the quality of coating. The proposed model is compared to two other models of which we have tried to consider the advantages. Nevertheless, this model still has place to be improved since we have not taken inot account the depth of the corrosion and other environmental parameters that can effect the corrosion process, such as temperature, humidity, etc.

Chapter 3

Fitting a Gamma process to simulated data

After modeling the corrosion process according to its physical properties, it is possible to work on the simulated data from this model to investigate the deterioration process and the effect of the imperfect maintenance on it. As mentioned before, a white-box model describes the physical process well and is much closer to the real process than a gray- or black-box model. It gives the opportunity to study the deterioration process in 2D or 3D case while with the gray-box models such as gamma process only it is possible to have the amount of deterioration in time. However, a white-box model is much more time-consuming than the other ones, that is the reason it is preferable to express the deterioration process with a gray-box (gamma process in our case) model. In this chapter first the gamma process and some relevant equations are introduced. Afterwards, different methods for parameter estimation of the gamma process are presented, and finally the results of the parameter estimation and fitting, based on the simulated data, are given.

3.1 Gamma process

Since the introduction of the gamma process in the area of reliability in 1975, it has been increasingly used to model stochastic deterioration for optimizing maintenance. Because gamma processes are well suited for modeling the temporal variability of deterioration, they have proved to be useful in determining optimal inspection and maintenance decisions [3]. Gamma process and

Brownian Motion process which are also called gray-box processes are two well-known examples from the class of Lévy processes.

Definition 3.1.1. A Lévy process, is any continuous-time stochastic process $\{X(t), t \geq 0\}$ that

1. admits càdlàg modification,
2. $X(0) = 0$ with probability one, and
3. $X(t)$ has independent, stationary increments.

Definition 3.1.2. A càdlàg function is a function $f : \mathbb{R} \to \mathbb{R}$ that is everywhere right-continuous and has left limits everywhere. Càdlàg functions are important in the study of the stochastic processes that admit (or even require) jumps, unlike Brownian motion which has continuous sample paths.

Definition 3.1.3. A continuous-time stochastic process X assigns a random variable $X(t)$ to each point $t \geq 0$ in time. The increments of such a process are the differences $X(s) - X(t)$ between its values at different times $t < s$. To call the increments of a process **independent** means that increments $X(s) - X(t)$ and $X(u) - X(v)$ are independent random variables whenever the two time intervals do not overlap. To call the increments **stationary** means that the probability distribution of any increment $X(s) - X(t)$ depends only on the length $s - t$ of the time interval.

Recall that a random variable X has a gamma distribution with shape parameter $\nu > 0$ and scale parameter $u > 0$ if its probability density function is given by

$$Ga(x \mid \nu, u) = \frac{u^{\nu}}{\Gamma(\nu)} x^{\nu-1} e^{-ux} I_{(0,\infty)(x)},$$

where $I_A(x) = 1$ if $x \in A$ and 0 otherwise, and $\Gamma(\nu) = \int_{z=0}^{\infty} z^{\nu-1} e^{-z} dz$ is the gamma function for $\nu > 0$. Furthermore, let $\nu(t)$ be a non-decreasing, right continuous, real-valued function for $t \geq 0$ with $\nu(0) = 0$, [3].

Definition 3.1.4. The gamma process with shape function $\nu(t) > 0$ and scale parameter $u > 0$ is a Lévy process $\{X(t), t \geq 0\}$ with the property that $X(\tau) - X(t) \sim Ga(\nu(\tau) - \nu(t), u)$ for all $\tau > t \geq 0$.

Brownian Motion process is not as popular as the gamma process because of the negative increments which does not fit the reality in a sense that the deterioration level can not decrease unless there are errors in the observations. In fact, the reason that the gamma process is used extensively in modeling the deterioration is that the non-negative independent increments of the gamma process describes the deterioration process well. As a matter of fact, we are interested in fitting the simulated data with the gamma distribution before and after imperfect maintenance.

3.2 Parameter estimation for the gamma process

The amount of deterioration at time t can be denoted by $X(t)$ and the probability density function of $X(t)$ can be given by

$$f_{X(t)}(x) = \text{Ga}(x \mid \nu(t), u) \tag{3.1}$$

with

$$\text{E}(X(t)) = \frac{\nu(t)}{u}, \quad \text{Var}(X(t)) = \frac{\nu(t)}{u^2}. \tag{3.2}$$

For the derivation of the expected value and variance, see appendix A. In order to apply the gamma process model to a real case, statistical methods for the parameter estimation of the gamma process are required. A typical data set, for a single deteriorating component, consists of inspection times $0 = t_0 < t_1 < \cdots < t_n$, and corresponding observations of the cumulative amounts of deterioration $0 = x_0 < x_1 < \cdots < x_n$. Consider a gamma process with shape function $\nu(t) = at^b$ and scale parameter u, we are interested in finding the estimators for a, b and u, [3].

Three conventional methods of parameter estimation, namely, maximum likelihood, method of moments, and the Bayesian estimation method, are discussed in the following sections. An interested reader is referred to reference [14] for more details of the derivations in the two former methods and reference [3] and [6] for the last one.

3.2.1 Method of maximum likelihood

The maximum likelihood estimators of a, b and u can be obtained by maximizing the logarithm of the likelihood function of the increments. The likelihood function of the observed deterioration increments $\delta_i = x_i - x_{i-1}$ for $i = 1, \ldots, n$, is a product of independent gamma densities due to the independent increments property of the gamma process.

$$\begin{aligned} \mathcal{L}(\delta_1, \ldots, \delta_n \mid a, b, u) &= \prod_{i=1}^{n} f_{X(t_i)-X(t_{i-1})}(\delta_i) \\ &= \prod_{i=1}^{n} \frac{u^{a[t_i^b - t_{i-1}^b]}}{\Gamma(a[t_i^b - t_{i-1}^b])} \delta_i^{a[t_i^b - t_{i-1}^b]-1} \exp\{-u\delta_i\}. \end{aligned} \tag{3.3}$$

By computing the first partial derivatives of the logarithm of the likelihood function of the increments with respect to a, b and u, the maximum likelihood estimates $\hat{a}, \hat{b}$ and $\hat{u}$ can be solved from

$$\begin{cases} \hat{u} = \frac{\hat{a} t_n^{\hat{b}}}{x_n} \\ \sum_{i=1}^{n} [t_i^{\hat{b}} - t_{i-1}^{\hat{b}}] \{\psi(\hat{a}[t_i^{\hat{b}} - t_{i-1}^{\hat{b}}]) - \log(\delta_i)\} = t_n^{\hat{b}} \log\left(\frac{\hat{a} t_n^{\hat{b}}}{x_n}\right) \\ \sum_{i=1}^{n} \psi(\hat{a}[t_i^{\hat{b}} - t_{i-1}^{\hat{b}}]) + \log \delta_i [\hat{a}(t_i^{\hat{b}} \log t_i - t_{i-1}^{\hat{b}} \log t_{i-1}) - 1] = \hat{a} t_n^{\hat{b}} \log t_n \log \hat{u} \end{cases} \tag{3.4}$$

where the function $\psi(a)$ is the derivative of the logarithm of the gamma function:

$$\psi(a) = \frac{\Gamma'(a)}{\Gamma(a)} = \frac{\partial \log \Gamma(a)}{\partial a}$$

for $a > 0$. The function $\psi(a)$ is called the *digamma* function and can be accurately computed using the algorithm developed by Bernardo (1976). Because cumulative amounts of deterioration are measured, the last inspection contains the most information. This is confirmed by the fact that the expected deterioration, given the maximum likelihood estimator of u in equation above, at the last inspection (at time t_n) equals x_n; i.e.

$$\mathrm{E}(X(t)) = x_n [\frac{t}{t_n}]^b \Rightarrow \mathrm{E}[X(t_n)] = x_n.$$

In our case, the typical data set is defined for a system of multiple components (i.e. corrosion

sample paths) by considering m independent components, for which $n_j,\ j = 1, \ldots, m$, inspections are performed resulting in n_j independent deterioration increments, i.e., (3.3) will be extended to

$$\mathcal{L}(\delta_{11}, \ldots, \delta_{n_j m} \mid a, b, u) \quad = \quad \prod_{j=1}^{m} \prod_{i=1}^{n_j} f_{X(t_{ij})-X(t_{i-1j})}(\delta_{ij}) \tag{3.5}$$

where x_{ij} is the cumulative amount of deterioration at the ith inspection time t_{ij} for the jth component and $\delta_{ij} = x_{ij} - x_{i-1j}$ is the ith deterioration increment for the jth component [1]. Maximizing (3.5) gives an estimation for three mentioned parameters. Therefore, when a large number of data is available, the maximum likelihood method is a feasible tool for parameter estimation.

3.2.2 Method of moments

Back to the single component data set, the parameter estimation in this method is done by using the expectation and variance of the corresponding observations. As before, consider the non-stationary gamma process with the following mean and variance

$$\mathrm{E}(X(t)) = \frac{at^b}{u}, \ \ \mathrm{Var}(X(t)) = \frac{at^b}{u^2}. \tag{3.6}$$

Assuming that the power b is known, the non-stationary gamma process can be easily transformed to a stationary gamma process by performing a monotonic transformation from the calender time t to the transformed or operational time $z(t) = t^b$. Substituting the inverse time transformation $t(z) = z^{1/b}$ gives

$$\mathrm{E}(X(t)) = \frac{az}{u}, \ \ \mathrm{Var}(X(t)) = \frac{az}{u^2}.$$

This results in a stationary gamma process with respect to the transformed time z. Similarly, the transformed inspection times are $z_i = t_i^b, i = 1, \ldots, n$. Let us also define the transformed time between inspections as $w_i = t_i^b - t_{i-1}^b$, and the deterioration increment $D_i = X_i - X_{i-1}$ for $i = 1, \ldots, n$. Note that D_i has a gamma distribution with shape parameter aw_i and scale parameter u for all $i = 1, \ldots, n$, and the increments $D_1, \ldots, D_n$ are independent. In addition, X_i and D_i denote random quantities while x_i and δ_i indicate the corresponding observations, where

$\delta_i = x_i - x_{i-1}$ as defined before in the maximum likelihood method. Following the derivations presented in [14], the method of moments estimates $\hat{a}$ and $\hat{u}$ can be solved from

$$\begin{cases} \frac{\hat{a}}{\hat{u}} = \frac{\sum_{i=1}^{n} \delta_i}{\sum_{i=1}^{n} w_i} = \frac{x_n}{t_n^b} = \bar{\delta}, \\ \frac{x_n}{\hat{u}} \left(1 - \frac{\sum_{i=1}^{n} w_i^2}{[\sum_{i=1}^{n} w_i]^2}\right) = \sum_{i=1}^{n} (\delta_i - \bar{\delta} w_i)^2. \end{cases} \tag{3.7}$$

These are the simple formulae for parameter estimation which can be easily computed. Note that the first equation in (3.7) yields the same result as the first equation in the maximum likelihood estimation in (3.4). The restriction of this method is that only two parameters a and u can be estimated and the power b should be known, for example via expert judgment.

3.2.3 Method of Bayesian statistics

Without the loss of generality, consider the stationary gamma process, i.e., the shape function is linear in time ($b = 1$ and $\nu(t) = at$), with the mean and variance

$$\mathrm{E}(X(t)) = \frac{at}{u} \quad \text{and} \quad \mathrm{Var}(X(t)) = \frac{at}{u^2}.$$

Assuming that the expectation and variance are linear in time, i.e.,

$$\mathrm{E}(X(t)) = \mu t \quad \text{and} \quad \mathrm{Var}(X(t)) = \sigma^2 t,$$

then, the parameters of the process $X(t)$ can be defined as

$$a = \frac{\mu^2}{\sigma^2} \quad \text{and} \quad u = \frac{\mu}{\sigma^2},$$

where μ is the average deterioration rate and σ^2 is the variance of the process. Since both of these parameters are uncertain and assessing both variables is unwieldy, the standard deviation σ is fixed relative to the mean μ through the use of a Coefficient Of Variation (COV) parameter C which is defined as $C = \sigma/\mu$. This variable is preset by expert judgment and fixed in the model. Using COV, the probability density function for $X(t)$ can be written as

$$f_{X(t)}(x) = Ga\left(x \mid \frac{t}{C^2}, \frac{1}{\mu C^2}\right).$$

Now the only uncertain variable is μ which can be specified by assigning the appropriate prior density to it. In [6], the suggested prior density is the inverted gamma density which is given by

$$\mathrm{Ig}(x \mid \nu, u) = \frac{u^{\nu}}{\Gamma(\nu)} \left(\frac{1}{x}\right)^{\nu+1} \exp(-\frac{u}{x}) I_{(0,\infty)}(x).$$

Therefore, it is possible to obtain the posterior density of μ by updating the prior density with an inspection measurement using the continuous version of Bayesian theorem which is given by

$$\pi(\mu \mid x) = \frac{\mathcal{L}(x \mid \mu)\pi(\mu)}{\int_{\mu=0}^{\infty} \mathcal{L}(x \mid \mu)\pi(\mu) d\mu},$$

where $\pi(\mu)$ is the prior density, $\mathcal{L}(x \mid \mu)$ the likelihood function of measurement x given μ, and $\pi(\mu \mid x)$ is the posterior density. For the proof of the theorem an interested reader is referred to [15].

As a result, for perfect inspection the posterior density is again the inverted gamma density according to the property of the gamma and inverted gamma distributions which are a conjugate family for samples from a gamma distribution; i.e.,

$$\pi(\mu \mid x_1, \ldots, x_n) = \mathrm{Ig}\left(\mu \mid \nu + \sum_{i=1}^{n} \frac{t_i - t_{i-1}}{C^2}, u + \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{C^2}\right).$$

which reduces to

$$\pi(\mu \mid x_n) = \mathrm{Ig}\left(\mu \mid \nu + \frac{t_n}{C^2}, u + \frac{x_n}{C^2}\right),$$

assuming that $x_0 = 0$ and $t_0 = 0$. The reason is that since the standard deviation is fixed relative to the mean μ, only the last inspection is needed to calculate the posterior density when using perfect measurements since it contains all the necessary information on the degradation process [6]. If we deal with imperfect inspection, it is also possible to obtain the posterior density using the algorithm given in [6].

Since we use the simulated physical model for generating data, it is possible to have as much data as we want. Also in our data set we consider the amount of corrosion at each step of time. Therefore, we use the maximum likelihood method to estimate the three parameters a, b and u. Moreover, as mentioned, the method of moments gives the estimation of only two parameters and

the power b should be known. In the Bayesian method, if we have a small data set and the prior density from expert judgment, it is possible to estimate parameters a and u assuming that b is known (to estimate b in a Bayesian method see [3]). However, there is no need to deal with the prior and posterior distributions since enough data is available and the maximum likelihood method is feasible, besides the fact that we are interested in estimation of three parameters.

3.3 Data simulation

In this section, we explain how the data simulation is performed and how this data will be used in our experiment. To check the effect of imperfect maintenance on the corrosion process, we consider two sets of data; the one before imperfect maintenance and the one after it. The reason is that firstly, one should know how the behavior of the corrosion is before the maintenance to be able to make a reasonable interpretation of its behavior after the maintenance action. As mentioned in many papers, such as [2; 3; 1; 6], and etc., the deterioration process can be modeled by a gamma process. Therefore, it is reasonable to show that the corrosion process follows a gamma process before any maintenance action. After this checking and finding the parameters of the process, it is time to check the effect of imperfect maintenance on the corrosion data and compare the obtained parameters with the ones before any maintenance.

In our experiment, the model parameters are chosen as follows:

- Size of the surface= 10×10 square-meters (10000 pixels with size 1×1 dm),
- Time interval= $[0, 5]$ years,
- Inspection period= every three weeks, i.e., time step=$1/17$,
- $\lambda_0 = 0.02$, and $\beta = 0.1$.

We simulate data during five years for two reasons: to be sure that 3% corrosion occurs, and to have enough data. During this five-year period no maintenance action is considered and the corrosion exceeds the 3% preventive threshold. In addition the step size is chosen in the way that after this five-year period, we have enough data for the experiment. The parameters λ_0 and β are chosen according to experts' data. As mentioned, the output of the simulation is the number of corroded cells at each inspection, i.e., time step.

3.3.1 Data simulation before imperfect maintenance

After setting the parameters appropriately, the simulation will be run 1000 times to have 1000 sample paths of the corrosion process. In figure 3.1 the percentage of corrosion of the surface is shown; after five years, on average 16% of the surface is corroded which is approximately 5.3 times larger than the 3% preventive threshold.

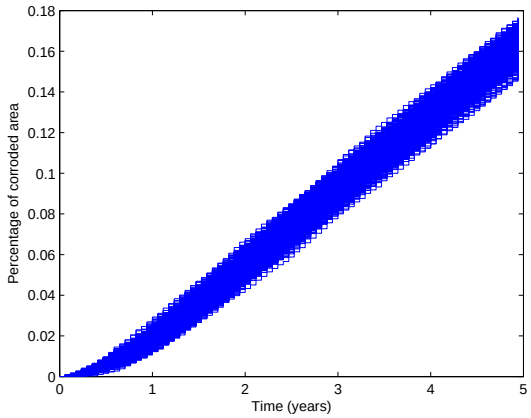

Figure 3.1: 1000 sample paths of the corrosion process without any maintenance action, simulated from the physical model

At this stage, two steps should be considered: Step one is to check whether the paths can be assumed to be drawn from a gamma process. Step two is to find the parameters of the gamma process, of course if in step one it is confirmed that the paths can be assumed to be from a gamma process. In the first step, according to definition 3.1.4, we have to check whether the increments are independent and if they follow the gamma distribution. Checking the independence property is not easy, and to be able to continue the experiment, we have assumed that this property is satisfied. To check the distribution of the increments, we use the Kolmogorov-Smirnov test which is defined below.

Definition 3.3.1. The empirical distribution function F_n for n i.i.d observations X_i is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq x\}}$$

where $I_{X_i \leq x}$ is the indicator function. The Kolmogorov-Smirnov statistic for a given function $F(x)$ is $D_n = \sup_x |F_n(x) - F(x)|$, where $\sup S$ is the supremum of set S.

Definition 3.3.2. The Kolmogorov- Smirnov test (K-S test) can be used to decide whether a sample comes from a population with specified distribution. An attractive feature of this test is that the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the chi-square goodness-of-fit test depends on an adequate sample size for the approximations to be valid). Despite

these advantages, the K-S test has limitations as well: It only applies to continuous distributions; and it tends to be more sensitive near the center of the distribution than at the tails.

The K-S test procedure of testing is as follows [16]:

1. Specify distribution F- associated with theoretical distribution function.
2. Order samples in non-decreasing manner.
3. Calculate $D = \max\left(\max_{1\le i\le n}\left(\frac{i}{n} - F(X_i)\right), \max_{1\le i\le n}\left(F(X_i) - \frac{i-1}{n}\right)\right)$.
4. Calculate p-value from the following formula

$$\text{p} - \text{value} = 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2\lambda^2 k^2} \quad \text{where} \quad \lambda = \max((\sqrt{n} + 0.12 + 0.11/\sqrt{n})D, 0).$$

The Kolmogorov-Smirnov test may also be used to test whether two underlying one-dimensional probability distributions differ. In this case, the Kolmogorov-Smirnov statistic is $D_{n,n'} = \sup_x |F_n(x) - F_{n'}(x)|$, and the null hypothesis is rejected at level α if $\sqrt{\frac{nn'}{n+n'}}D_{n,n'} > K_\alpha$, where n and n' are sample sizes, and K_α is found from $\Pr(K \le K_\alpha) = 1-\alpha$ with K being the Kolmogrov distribution..

Using the K-S test, the null hypothesis that the increments of each corrosion path has a gamma distribution, is not rejected; in other words, all the increments have a gamma distribution before any maintenance action. Due to the large number of sample paths, 1000, the numerical results are omitted. Therefore, we can proceed to the second step to estimate the parameters of the gamma process.

3.3.2 Parameter estimation of the gamma process before imperfect maintenance

Under the independence assumption, and the results of the K-S test, which confirms that all the corrosion increments have a gamma distribution, it is time to estimate the parameters of the gamma process. Consequently, the method of maximum likelihood, considering (3.5) is used to estimate three parameters, a, b and u for all the corrosion paths. The result is as follows:

$$\hat{a} = 139.0792 \quad \hat{b} = 1.2770 \quad \hat{u} = 1.5214 \tag{3.8}$$

There is one point left to be mentioned which is about the covariance matrix. Since from the method of maximum likelihood, we get a set of parameters, it is not possible to calculate the covariance matrix from the samples since there is only one sample. However, there is still a way to obtain it through a term called (Fisher) information matrix which is defined as

$$I(\theta) = E\left(-\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta^T}\right),$$

where θ is the vector that maximizes the likelihood. Therefore, the asymptotic approximation for the covariance matrix is $\Sigma = n(I(\theta))^{-1}$, where n is the sample size [17].

In general, the information matrix is unknown because it depends on the unknown parameter vector θ. Hence it needs to be estimated. In our case we use an estimator $\hat{I}(\cdot)$ that is based on measurements L. We estimate $\frac{\Sigma}{n}$ by

$$\hat{I}(\hat{\theta})^{-1} = \left[\left(-\frac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta^T}\right)^{-1}\right]_{\theta=\hat{\theta}}$$

where $\hat{\theta}$ is the estimated vector of parameters obtained from the method of maximum likelihood [18].

Consequently, the estimated covariance matrix, $S = n[\hat{I}(\hat{\theta})]^{-1}$, of the three estimated parameters from $n = 1000$ samples (sample paths) are:

$$S = \begin{pmatrix} 0.0550 & -4.7627 & -0.0003 \\ -4.7627 & 487.7389 & -0.2345 \\ -0.0003 & -0.2345 & 0.0012 \end{pmatrix}.$$

However, we know that there is a considerable uncertainty in these parameters which is necessary to be specified to validate the estimation.

3.3.3 Uncertainty in estimated parameters (before imperfect maintenance)

Since the estimated parameters from the maximum likelihood method are asymptotically normally distributed [19], it is possible to use the multivariate normal distribution properties to find the appropriate confidence region and intervals for the estimated parameters. The three suggested

methods below are according to [19].

1. Simultaneous confidence intervals

Let θ be a vector of unknown population parameters and Θ be the set of all possible values of θ. A confidence region is a region of likely θ values. This region is determined by the data, and we denote it by $R(X)$, where $X = [X_1, X_2, \ldots, X_n]'$ is the data matrix. The region $R(X)$ is said to be a $100(1-\alpha)\%$ *confidence region* if, before the sample is selected,

$$P[R(X)\text{will cover the true } \theta] = 1 - \alpha$$

Therefore, a $100(1-\alpha)\%$ *confidence region* for the mean μ of a p-dimensional normal distribution is the ellipsoid determined by all μ such that

$$n(\bar{x} - \mu)'S^{-1}(\bar{x} - \mu) \leq \frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha) = c^2$$

where $\bar{x} = \frac{1}{n}\sum_{j=1}^{n}(x_j),\ S = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})(x_j - \bar{x})'$ and $x_1, x_2, \ldots, x_n$ are the sample observations.

While the confidence region $n(\bar{x} - \mu)'S^{-1}(\bar{x} - \mu) \leq c^2$, for c a constant, correctly assesses the *joint* knowledge concerning plausible values of μ, any summary of conclusions ordinarily includes confidence statements about the individual component means. Therefore, we consider the attitude that all of the separate confidence statements should hold simultaneously with a specified high probability.

Let $X_1, X_2, \ldots, X_n$ be a random sample from an $N_p(\mu, \Sigma)$ population with Σ positive definite. Here, μ is the vector of estimated parameters from the method of maximum likelihood and $\Sigma = n[\hat{I}(\hat{\theta})]^{-1}$ is the estimated covariance matrix which is denoted by S. Then, simultaneously for all $\mathbf{a}$, the interval

$$\left(a'\bar{X} - \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)a'Sa},\ \ a'\bar{X} + \sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)a'Sa}\right)$$

will contain $a'\mu$ with probability $1 - \alpha$. The proof of this statement can be found in [19] (p.225). It is convenient to refer to the simultaneous intervals of above statement as T^2-intervals, since the

coverage probability is determined by the distribution of T^2. Simultaneous confidence intervals can be considered as shadows of the confidence ellipsoid which cover a larger area than 95%.

However, when the sample size is large, $n = 1000$ in our case, tests of hypothesis and confidence regions for μ can be constructed without the assumption of a normal population. As a result, $\frac{p(n-1)}{(n-p)}F_{p,n-p}(\alpha)$ is replaced by $\chi^2_p(\alpha)$ since both are approximately equal for n large relative to p. Hence, if $n - p$ is large,

$$a'\bar{X} \pm \sqrt{\chi^2_p(\alpha)}\sqrt{\frac{a'Sa}{n}} \tag{3.9}$$

contains $a'\mu$, for every a, with probability approximately $1 - \alpha$.

So for each parameter, we obtain the simultaneous confidence interval:

$$\begin{aligned}
1.5214 - \sqrt{7.8147}\sqrt{\frac{0.0550}{1000}} \leq \quad & \hat{u} \quad \leq 1.5214 + \sqrt{7.8147}\sqrt{\frac{0.0550}{1000}} \\
\text{or} \qquad 1.5007 \leq \quad & \hat{u} \quad \leq 1.5421 \\
139.0792 - \sqrt{7.8147}\sqrt{\frac{487.7389}{1000}} \leq \quad & \hat{a} \quad \leq 139.0792 + \sqrt{7.8147}\sqrt{\frac{487.7389}{1000}} \\
\text{or} \qquad 137.13 \leq \quad & \hat{a} \quad \leq 141.03 \\
1.2770 - \sqrt{7.8147}\sqrt{\frac{0.0012}{1000}} \leq \quad & \hat{b} \quad \leq 1.2770 + \sqrt{7.8147}\sqrt{\frac{0.0012}{1000}} \\
\text{or} \qquad 1.2739 \leq \quad & \hat{b} \quad \leq 1.2801
\end{aligned}$$

Note that $\chi^2_3(0.05) = 7.8147$.

It is also informative to have a look at the correlation matrix of the three variables. Since an estimate of covariance matrix is available, the correlation matrix can be easily computed as follows:

$$\rho = \begin{pmatrix} 1 & \frac{S_{12}}{\sqrt{S_{11}S_{22}}} & \frac{S_{13}}{\sqrt{S_{11}S_{33}}} \\ \frac{S_{21}}{\sqrt{S_{11}S_{22}}} & 1 & \frac{S_{23}}{\sqrt{S_{22}S_{33}}} \\ \frac{S_{31}}{\sqrt{S_{11}S_{33}}} & \frac{S_{32}}{\sqrt{S_{22}S_{33}}} & 1 \end{pmatrix} = \begin{pmatrix} 1 & -0.92 & -0.037 \\ -0.92 & 1 & -0.31 \\ -0.037 & -0.31 & 1 \end{pmatrix}.$$

where the diagonal entries are variances and the other ones are covariances.

The two parameters $\hat{u}$ and $\hat{a}$ are highly negatively correlated, which means the larger $\hat{a}$ is, the smaller $\hat{u}$ is. This negative correlation is observable from the estimated values, 1.5214 and 139.0792, as well. However, there is not a strong correlation between $\hat{u}$ and $\hat{b}$; this means that any

choice of $\hat{u}$ does not have almost any eefect in the choice of $\hat{b}$. The correlation between $\hat{a}$ and $\hat{b}$ is also considerable and negative which can been seen from the estimated values as well. These correlations are important to be considered when one is choosing the parameters of the gamma process.

2. One-at-a-time (or individual) intervals

An alternative approach to the construction of confidence intervals is to consider the component μ_i one at a time, as

$$a'\bar{x} - t_{n-1}(\frac{\alpha}{2})\frac{\sqrt{a'Sa}}{\sqrt{n}} \leq a'\mu \leq a'\bar{x} + t_{n-1}(\frac{\alpha}{2})\frac{\sqrt{a'Sa}}{\sqrt{n}}$$

with $a' = [0, \ldots, 0, a_i, 0, \ldots, 0]$ where $a_i = 1$. This approach ignores the covariance structure of the p variables and by replacing $t_{n-1}(\frac{\alpha}{2})$ by $z(\frac{\alpha}{2})$ due to the large sample size, it leads to the intervals

$$\bar{x_i} - z(\frac{\alpha}{2})\sqrt{\frac{s_{ii}}{n}} \leq \mu_i \leq \bar{x_i} + z(\frac{\alpha}{2})\sqrt{\frac{s_{ii}}{n}} \tag{3.10}$$

where $\sqrt{s_{ii}}$ is the standard deviation that is taken from the main diagonal of the covariance matrix.

Considering the estimated vector and covariance matrix of our sample, and $z(\frac{\alpha}{2}) = 1.96$, the one-at a time confidence interval for each parameters is as follows:

$$\begin{aligned}
1.5214 \pm 1.96\sqrt{\frac{0.0550}{1000}} \quad &\text{or} \quad 1.5069 \leq \hat{u} \leq 1.5359 \\
139.0792 \pm 1.96\sqrt{\frac{487.7389}{1000}} \quad &\text{or} \quad 137.71 \leq \hat{a} \leq 140.45 \\
1.2770 \pm 1.96\sqrt{\frac{0.0012}{1000}} \quad &\text{or} \quad 1.2749 \leq \hat{b} \leq 1.2791
\end{aligned}$$

In other words, for $1 - \alpha = 0.95, n = 1000$, the multipliers of $\sqrt{s_{ii}/n}$ in (3.9) and (3.10) are

$$\sqrt{\chi^2_3(0.05)} = \sqrt{7.8147} = 2.7955$$

and $z(0.025) = 1.96$, respectively. Consequently, in this case the simultaneous intervals are $100(2.8037 - 1.9625)/1.9625 \approx 43\%$ wider than those derived from the one-at-a-time t method.

The one-at-a-time intervals are too short to maintain an overall confidence level for separate

statements about all p means. Nevertheless, we sometimes look at them as the best possible information concerning a mean, if this is the only inference to be made. The T^2-intervals are too wide if they are applied only to the p component means. To see why, consider the confidence ellipse and the simultaneous intervals; if, for example, μ_1 lies in its T^2-interval and μ_2 lies in its T^2-interval, then (μ_1, μ_2) lies in the rectangle formed by these two intervals. This rectangle contains the confidence ellipse and more. The confidence ellipse is smaller but has the probability 0.95 of covering the mean vector μ with its component means μ_1 and μ_2. Consequently, the probability of covering the two individual means μ_1 and μ_2 will be larger than 0.95 for the rectangle formed by the T^2-intervals. This result leads us to consider a third approach to making multiple comparisons known as the Bonferroni method.

3. The Bonferroni method of multiple comparisons

Often, attention is restricted to small number of individual confidence statements. In these situations it is possible to do better than the simultaneous intervals, T^2-intervals, which gives the simultaneous confidence intervals shorter than the simultaneous T^2-intervals. The alternative method for multiple comparisons is called the *Bonferroni method*, because it is developed from a probability inequality carrying that name. The Bonferroni confidence intervals for each mean component $i, i = 1, \ldots, p$ can be computed as follows:

$$\bar{x_i} \pm z(\frac{\alpha}{2p})\sqrt{\frac{s_{ii}}{n}} \tag{3.11}$$

where $z(\frac{\alpha}{2p})$ is a replacement for $t_{n-1}(\frac{\alpha}{2p})$ due to the large sample size. In our case with $z(0.05/2(3)) = 2.39$, the Bonferroni confidence intervals for the estimated parameters are

$$1.5214 \pm 2.39\sqrt{\frac{0.0550}{1000}} \quad \text{or} \quad 1.5037 \leq \hat{u} \leq 1.5391$$
$$139.0792 \pm 2.39\sqrt{\frac{487.7389}{1000}} \quad \text{or} \quad 137.41 \leq \hat{a} \leq 140.75$$
$$1.2770 \pm 1.9625\sqrt{\frac{0.0012}{1000}} \quad \text{or} \quad 1.2744 \leq \hat{b} \leq 1.2796$$

Since all these confidence intervals from different methods are very narrow, we can conclude that the estimated parameters have small error and so are reliable.

3.3.4 Uncertainty in the gamma process (before imperfect maintenance)

Since the gamma process is a stochastic process, it contains temporal uncertainty. Moreover, due to the uncertainty in the estimated parameters, the gamma process with those parameters has additional uncertainty. To check these uncertainties, we look at the $2.5th$ and $97.5th$ percentiles of the gamma process with the shape parameter $\nu(t) = \hat{a}t^{\hat{b}}$ and scale parameter $\hat{u}$. The parameters of the process are taken from the likelihood estimation, as mentioned in section 3.3.2. The percentiles are shown in figure 3.2. As it is shown, the $97.5th$ and $2.5th$ percentile bands are quite narrow. Hence the temporal uncertainty in the process is not that large. If we compare the quantiles plot with the sample paths in figure 3.1, it is observable that the band width of the quantiles is almost the same as the band width of all 1000 sample paths together, which contains the maximum and minimum number of possible corrosion for different paths.

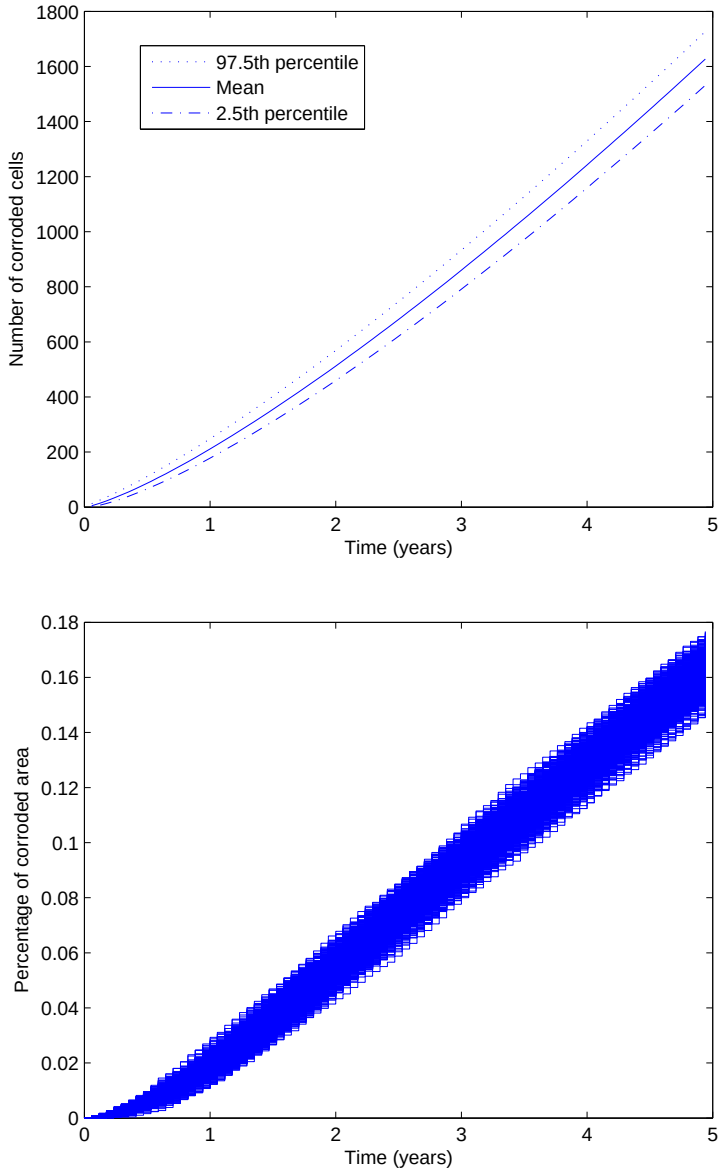

Figure 3.2: From the top: the 97.5th, 50th, and 2.5th percentiles of the gamma process before maintenance for the estimated parameters; the 1000 sample paths of corrosion

To see how the uncertainty in the parameters may influence the gamma process, we choose the upper and lower values of the Bonferroni confidence intervals as the input parameters for a gamma process to check if the percentile bands will be wider or remain as they were before. The parameters from the lower interval are $[1.5037, 137.41, 1.2744]$, and from the upper interval are $[1.5391, 140.75, 1.2796]$. The results are presented in the next figure. As shown, the percentile bands

are quite similar to the ones in figure 3.2, which confirms that the uncertainty in the parameters is not too large to influence the uncertainty in the gamma process severely. In fact, the band width in figure 3.2 is between $[1500, 1700]$, and in figure 3.3 is between $[1500, 1800]$ considering the maximum and minimum of both plots for upper and lower confidence bands values. This difference can be considered as an effect of the parameter uncertainty on the process uncertainty.

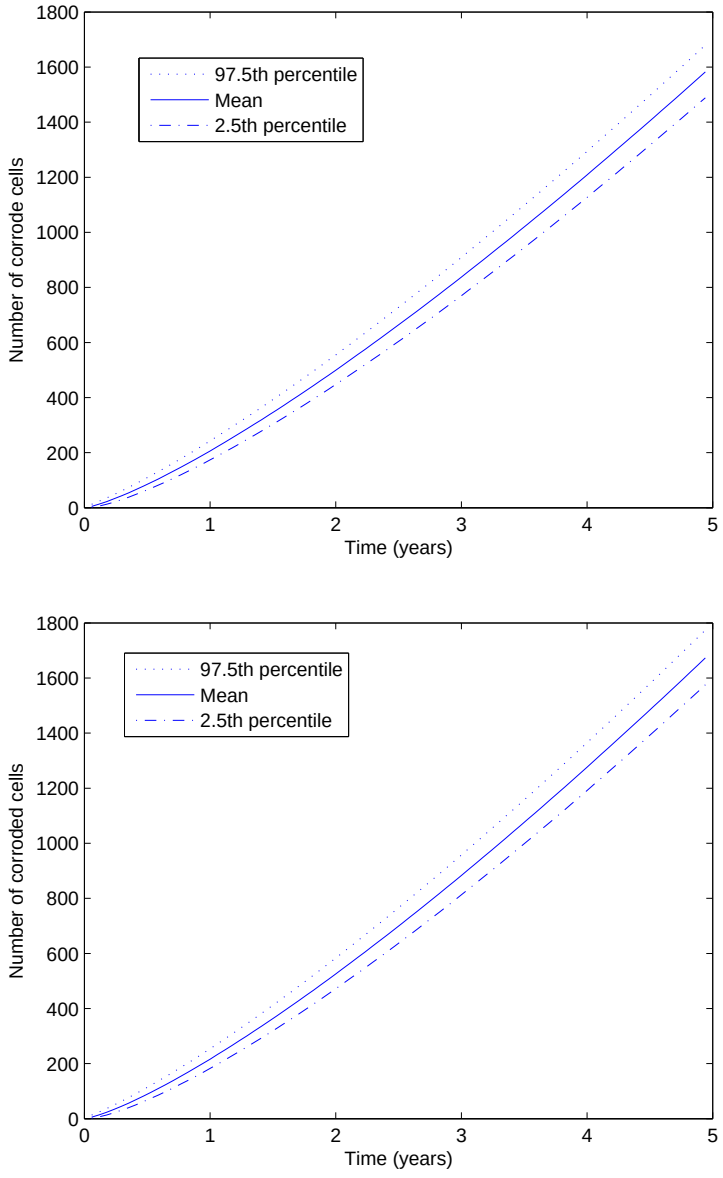

Figure 3.3: The 97.5th, 50th, and 2.5th percentiles of the gamma process before maintenance for the parameters of the lower and upper Bonferroni confidence intervals, respectively

All the results presented here show that we can approximate the corrosion process by a gamma process before an imperfect maintenance. Now we can check the corrosion data after the maintenance action is taken; if the process again is a gamma process then we can compare the parameters with the ones that are estimated before the imperfect maintenance to see how the corrosion behavior may change after the maintenance.

3.3.5 Data simulation after imperfect maintenance

To check the effect of imperfect maintenance on the corrosion process, we simulate 1000 sample paths considering one imperfect maintenance action. All the parameters of the physical model are the same as the ones mentioned before section 3.3.1, except the time interval; in this case, the time interval is larger, $[0, 10]$. The reason is that each path has different maintenance time; in other words, each path reaches the 3% threshold at different time. Note that having different time to reach the threshold is also the case without maintenance since the life time (time at which 3% is reached) is extended. As a result, to be able to have a data set with the same length for all paths, we consider a larger time interval; then for each path, we select the data of five years after the imperfect maintenance action. In this way, all the corrosion data after the maintenance has the same length for each path although there is no time correspondence between the data from different paths.

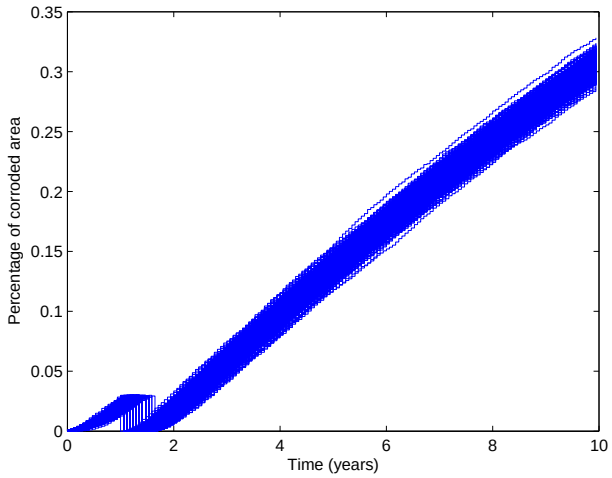

Figure 3.4: 1000 sample paths of the corrosion process considering an imperfect maintenance action, simulated from the physical model

As shown in the figure above, after ten years, on average 30% of the surface is corroded. However, the length of the time interval that we consider in our experiment is five years which includes more or less up to 23% corrosion. This amount of corrosion shows the higher rate in the corrosion process comparing to 16% corrosion in the case without any imperfect maintenance action.

After extracting the appropriate data set from the simulated data, the same procedure is followed to check whether the paths can be compared with the ones from a gamma process. As before, the independence property can not be checked easily and we assume that this condition is satisfied. By applying the Kolmogorov-Smirnov test, it is also confirmed that all increments have a gamma distribution, and hence we can estimate the parameters of the gamma process. This is a promising result since we have a basis now to compare the parameters of the gamma process before and after maintenance.

3.3.6 Parameter estimation and parameter inference of the gamma process after imperfect maintenance

To estimate the parameters of the corrosion process after an imperfect maintenance, the method of maximum likelihood, considering (3.5), is applied as before. The results are as follows:

$$\hat{a} = 190.5646, \quad \hat{b} = 1.1597, \quad \hat{u} = 1.5038. \tag{3.12}$$

Comparing (3.12) to (3.8), parameter $\hat{u}$ has changed slightly, but the two other parameters $\hat{a}$ and $\hat{b}$ are quite different from the ones before imperfect maintenance, and they seem to be shifted.

To obtain the uncertainty in the estimated parameters, we perform the parameter inference applying the three approaches mentioned in the earlier sections. Before finding the confidence intervals, the vector of estimated parameters, the estimated covariance matrix and correlation matrix are presented below:

$$\bar{x} = \begin{pmatrix} \hat{u} \\ \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 1.5038 \\ 190.5646 \\ 1.1597 \end{pmatrix},$$

$$S = \begin{pmatrix} 0.0538 & -6.5474 & -0.0001 \\ -6.5474 & 905.1156 & -0.2471 \\ -0.0001 & -0.2471 & 0.0009 \end{pmatrix}, \rho = \begin{pmatrix} 1 & -0.94 & -0.014 \\ -0.94 & 1 & -0.27 \\ -0.014 & -0.27 & 1 \end{pmatrix}.$$

Similar to the case before the maintenance, there is a strong negative correlation between a and u, and a weak, negative correlation between u and b. Also the negative correlation between a and b is quite considerable. The correlation matrix is quite close to the one before the maintenance. It can be concluded that an imperfect maintenance does not change the correlation between the parameters.

Considering the three mentioned approaches, the confidence intervals are as follow:

1. **Simultaneous confidence intervals**:

   According to the interval (3.9), the T^2-intervals for the estimated parameters, considering the large number of samples, after the maintenance are

$$\begin{aligned} 1.5038 \pm \sqrt{7.8147}\sqrt{\frac{0.0538}{1000}} \text{or } & 1.4833 \leq \quad \hat{u} \quad \leq 1.5243 \\ 190.5646 \pm \sqrt{7.8147}\sqrt{\frac{905.1156}{1000}} \text{or } & 187.91 \leq \quad \hat{a} \quad \leq 193.22 \\ 1.1597 \pm \sqrt{7.8147}\sqrt{\frac{0.0009}{1000}} \text{or } & 1.157 \leq \quad \hat{b} \quad \leq 1.1624 \end{aligned}$$

2. **One-at-a-time or individual intervals**:

   Considering the interval (3.10), the individual intervals for the estimated parameters are

$$\begin{aligned} 1.5038 \pm 1.96\sqrt{\frac{0.0538}{1000}} \quad \text{or} \quad & 1.4894 \leq \hat{u} \leq 1.5182 \\ 190.5646 \pm 1.96\sqrt{\frac{905.1156}{1000}} \quad \text{or} \quad & 188.7 \leq \hat{a} \leq 192.43 \\ 1.1597 \pm 1.96\sqrt{\frac{0.0009}{1000}} \quad \text{or} \quad & 1.1578 \leq \hat{b} \leq 1.1616 \end{aligned}$$

   As before, for $1 - \alpha = 0.95, n = 1000$, the multipliers of $\sqrt{s_{ii}/n}$ in (3.9) and (3.10) are $\sqrt{\chi^2_3(0.05)} = 2.7955$ and $z(0.025) = 1.96$, respectively. Consequently, in this case the simultaneous intervals are $100(2.8037 - 1.9625)/1.9625 \approx 43\%$ wider than those derived from the one-at-a-time t method.

3. **Bonferroni confidence intervals**

   Finally, the last confidence interval is the Bonferroni intervals which is calculated according to (3.11). The confidence intervals are presented below

$$1.5038 \pm 2.39\sqrt{\frac{0.0538}{1000}} \quad \text{or} \quad 1.4863 \leq \hat{u} \leq 1.5213$$
$$190.5646 \pm 2.39\sqrt{\frac{905.1156}{1000}} \quad \text{or} \quad 188.29 \leq \hat{a} \leq 192.84$$
$$1.1597 \pm 2.39\sqrt{\frac{0.0009}{1000}} \quad \text{or} \quad 1.1574 \leq \hat{b} \leq 1.162$$

Comparing the intervals, the one-at-a-time intervals are the smallest ones and the T^2-intervals are the largest ones. Also the length of the intervals shows that the uncertainty in the parameter estimation is not too much, and consequently the results are reliable.

3.3.7 Uncertainty in the gamma process after an imperfect maintenance

The last step is to check the uncertainty in the gamma process. As the case before the maintenance, we consider the 97.5th and 2.5th percentiles as well as the mean of the gamma process with the estimated parameters. As it is shown in figure 3.6, the bands are quite narrow which implies that the uncertainty in the process is not too high. Moreover, since after an imperfect maintenance, the corrosion occurs faster due to the age of the surface, the slope of the plot are quite sharper than the ones in figures 3.3. As before, if we compare the band width of the quantiles to the one from the corrosion sample paths, and due to their compatibility, we can conclude that the uncertainty is not too large and hence the results are reliable.

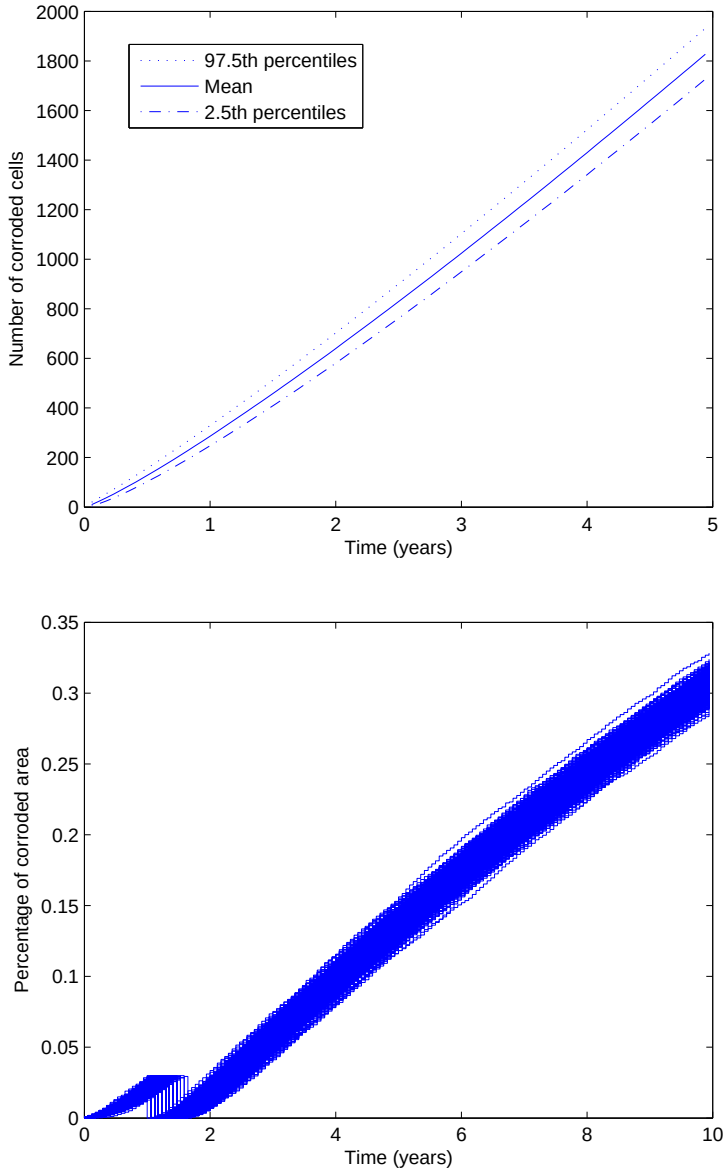

Figure 3.5: From the top: the 97.5th, 50th, and 2.5th percentiles of the gamma process after maintenance for the estimated parameters; the 1000 sample paths of corrosion

Since the uncertainty in the parameters adds additional uncertainty to the gamma process, we present the 97.5th and 2.5th percentiles for the upper and lower Bonferroni confidence interval as well. The confidence band width show that the uncertainty in the parameters does not influence severely the amount of uncertainty in the gamma process. Note that the time interval in these plots is not from zero up to five, but has length five, since we have considered the data of five years after

an imperfect maintenance; the reason that we have started from zero is just for simplicity and more clarity. As before, the band width in figure 3.5 is between $[1700, 1900]$, and in figure 3.6 is between

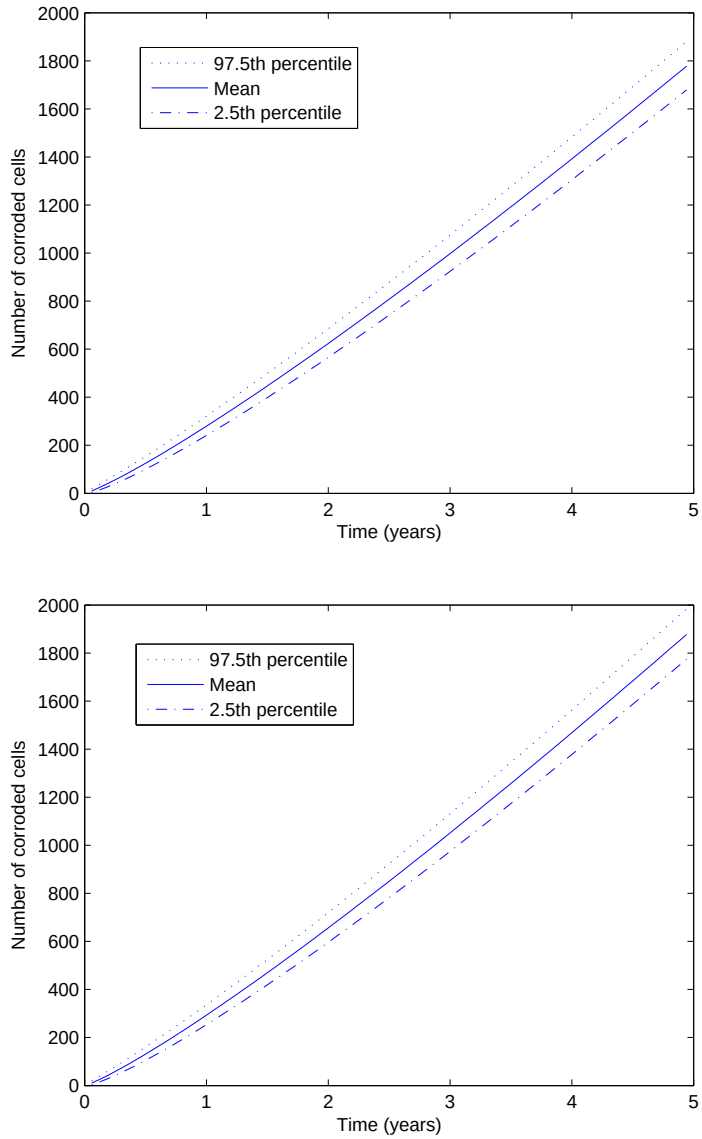

Figure 3.6: The 97.5th, 50th, and 2.5th percentiles of the gamma process after maintenance for the parameters of the lower and upper Bonferroni confidence intervals, respectively

$[1700, 2000]$ considering the maximum and minimum of both plots for upper and lower confidence bands values. This difference can be considered as an effect of the parameter uncertainty on the process uncertainty; it also shows the higher amount of corrosion after maintenance comparing to corresponding results before imperfect maintenance.

3.4 Summary

In this chapter we have checked whether the corrosion process, which is simulated from our physical model, can be represented by a gamma process before and after imperfect maintenance. The results show that, according to the independence assumption of the corrosion increments and the fact that all the increments have a gamma distribution, the corrosion process simulated from the proposed model can be represented by a gamma process before and after imperfect maintenance. However, the parameters of the two processes are different; it seems that the parameters a and b are shifted after the maintenance action. Furthermore, gamma processes before and after imperfect maintenance, are non-stationary due to the estimated parameter b in both cases. Note that although this parameter is very close to one, we can not assume the stationary gamma process due to its narrow confidence interval. In general, all the confidence intervals and confidence bands (percentile) show that the estimated parameters are quite reliable since the intervals and bands are not too wide. The narrow intervals and bands also express that the uncertainty in the estimated parameters and the gamma process are not too large which is another clue to consider the parameters reliable.

Chapter 4

The inter-maintenance time

In this chapter we take a closer look at the inter-maintenance time. As a matter of fact, we consider our case study from another point of view. After explaining the problem, some analytical approaches are suggested as an alternative solution.

4.1 Problem description

As before, consider the steel surface as a square grid and then each cell has its own repair time. Assume that each cell has a lifetime $T_i,\ i = 1, \ldots, n$, which indicates the age of the cell until it becomes corroded; the shorter the life time is, the earlier the cell is corroded. The maintenance action is performed once 3% of the cells are corroded.

At this point the lifetimes of the corroded cells will be reset in a sense that all the corroded cells get the lifetime of $T_{(\alpha n)}$ with $\alpha = 3\%$ plus a new random time; i.e., $T_j = T_{(\alpha n)} + X$ for $j \leq \alpha n$, where X is a random variable that has the same distribution as $T_i,\ i = 1, \ldots, n$ before any corrosion. This sum is actually the new lifetime of the repainted cells; the rest of the cells ($j > \alpha n$) keep their life time. Note that $T_{(\alpha n)}$ is an order statistic.

Definition 4.1.1. Given a sample of n i.i.d random variables $X_1, \ldots, X_n$, reorder them so that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$. Then $X_{(i)}$ is called the ith order statistic; in other words, $X_{(i)}$ is the ith smallest value.

After each maintenance action, the time resetting is performed. However, the problem is that after each time readjustment, the distribution of lifetimes will be changed; i.e., a new distribution

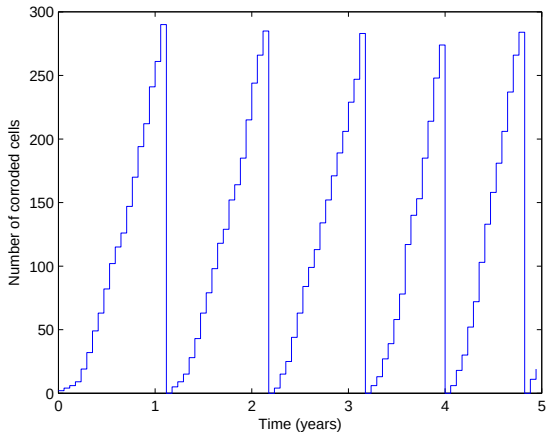

Figure 4.1: Corrosion process considering imperfect maintenance while the process reaches the 3% threshold

will be added to the set of the distributions of the lifetimes after each repair. The reason is that only few distributions are from a conjugate family, such as normal and gamma distributions, and the rest will be changed after summation. In addition, only few of the distributions have known distribution after summation; for example, sum of the exponential random variables has the gamma distribution. Therefore, it will be very complicated to compute the inter-repair time's distribution analytically, after n repairs as $n \to \infty$ which is the aim of this chapter.

4.2 Empirical results

Although it is not easy to get the theoretical results, it is possible to study the inter-maintenance time behavior by means of simulation to have a clue of what exactly happens after many repairs. In figure 4.2, the lifetimes of different distributions are shown. the distributions and other parameters that are used in sampling are as follows: Weibull distribution with shape parameter 2.5 and scale parameter 5, exponential distribution with parameter $\lambda = 10$, and lognormal distribution with parameters $\mu = \log(5)$ and $\sigma^2 = 1.5$, considering 200 repairs and 10000 cells. Note that the lognormal distribution is in fact an alternative for the normal distribution since the inter-repair times can not be negative and therefore, we can not use the normal distribution. The density function of the applied distributions are given in Appendix B.

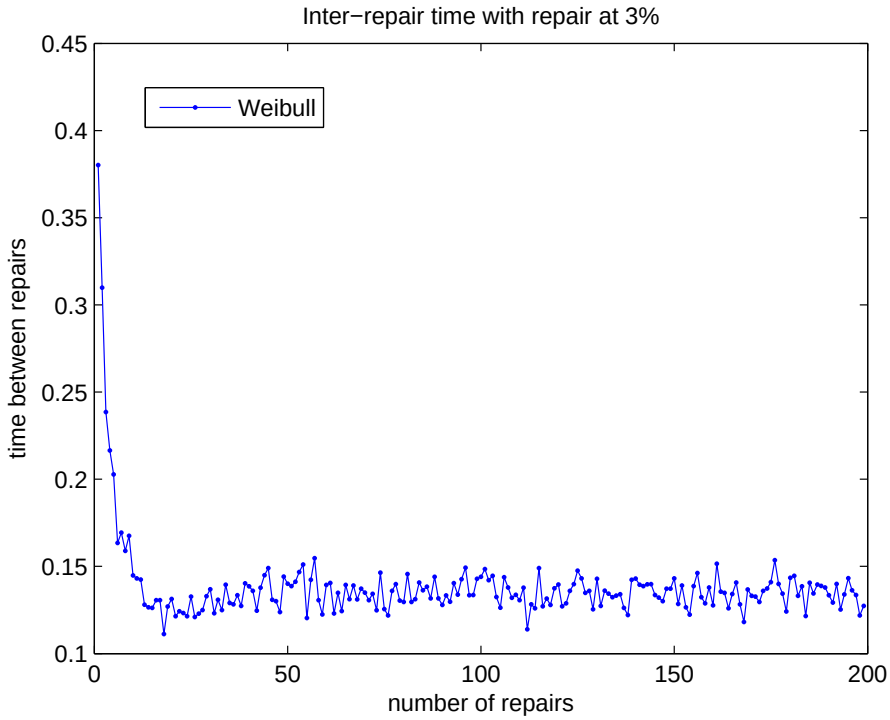

Figure 4.2: Histogram and distribution of the inter-repair time, considering 200 repairs, from the Weibull distribution

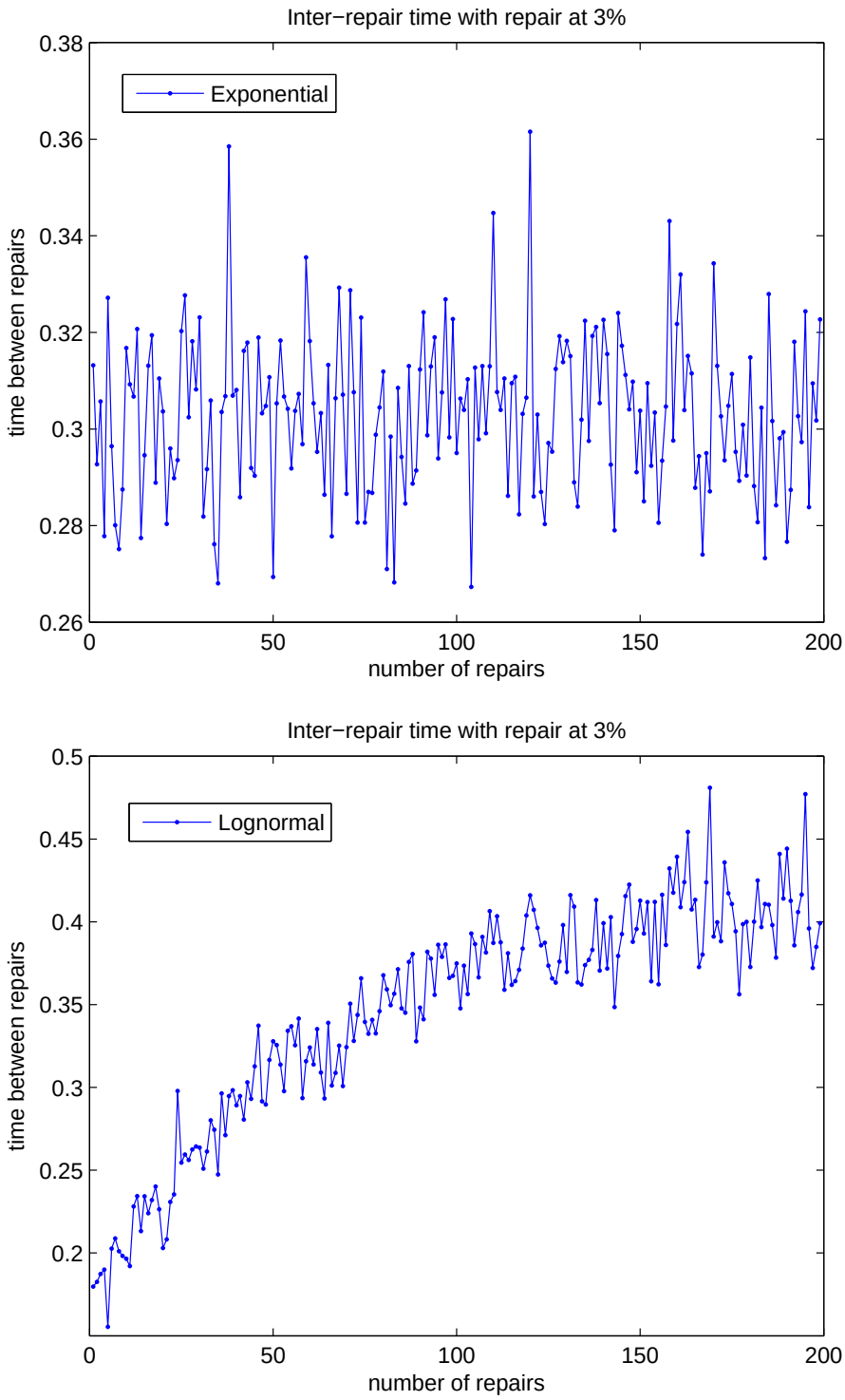

Figure 4.3: Histogram and distribution of the inter-repair time, considering 200 repairs, from the exponential and lognormal distribution

As it is shown in all these figures, if the number of repairs is set large, the inter-repair time has a stationary behavior, i.e., after some number of repairs the time between the repairs converges to a specific value in which we are interested. It is noticeable that the exponential distribution does not converges as the two other do, which is the result of its memory-less property. However, it still oscillates around a specific value.

These experimental results show that we should expect a typical behavior from the inter-repair times after a large number of repairs. This typical behavior can then be useful for decision makers to obtain the average time between the repairs and to define the optimal inspection and maintenance strategies. Consequently, it is of our interest to find the distribution of the inter-maintenance (repair) times to be able to explain its convergent behavior. Nevertheless, as mentioned before, because of the complicated structure of the process and the mixture of distributions, obtaining this distribution is not straight forward and it needs some tactful approaches to be able to deal with its complexity. In the following sections, our approach to this problem and its results are explained.

4.3 Superimposed processes

In general maintenance modeling, it is assumed that the defect occurrence following by the maintenance actions construct a renewal process. This is mainly based on the concept of the perfect maintenance after which the system condition goes back to an as good as a new one. However, in our case, considering the mixture of the distributions after each repair (maintenance action), the process can not be considered as a renewal process anymore –the interarrival times are neither independent nor identically distributed. Our approach to this problem is to consider the process as a superposition process. To introduce this process, we start with the general definition and properties of the renewal process.

4.3.1 Renewal process

As the name of the process indicates, it is used to model *renewals* or replacement of equipment. Let $T_1, T_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) non-negative random variables which with probability 1 are not all zero, $0 < E[T_i] = \mu < \infty$, and $Var(T_i) = \sigma^2 < \infty$. Define for each $n > 0$, $S_n = T_1 + \cdots + T_n$; each S_n is referred to as the "nth" jump time and the intervals $[S_n, S_{n+1}]$ being called renewal intervals. Then the collection of random variables $(N_t)_{t \geq 0}$

given by $N_t = \max\{\, n : S_n \leq t \,\}$ is called a renewal process. Each T_i denotes the ith interarrival (or interoccurrence) time, that is, the time between the $(i-1)$st and ith arrivals.

Note that the Homogeneous Poisson Process (HPP) is a renewal process where the interarrival times are exponentially distributed. A renewal process may thus be considered as a generalization of the HPP.

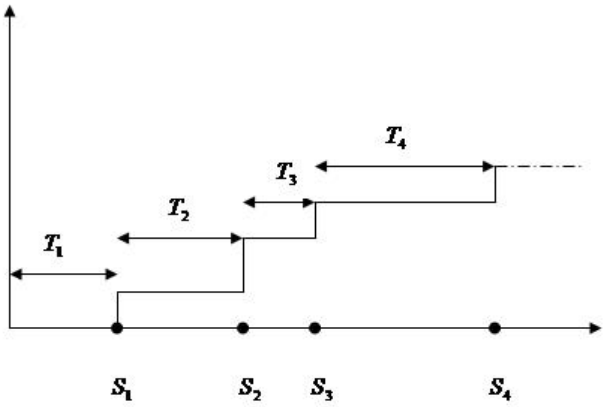

Figure 4.4: A sample renewal process with inter arrivaltime $T_1, T_2, \ldots$

Below we summarize a few of the basic equations which pertain to renewal processes and are presented in [20], [21]. Since the T_i's are i.i.d. with distribution function $F(x)$, $F^{(n)}(t)$ the n-fold convolution of F with itself, is the distribution of S_n. The n-fold convolution of F is defined according to the convolution theorem as follows:

Definition 4.3.1. Let

$$F^{(0)}(t) = \begin{cases} 1, & t \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

and $N(t)$ be the number of events up to time t, and $S_n = X_1 + X_2 + \cdots + X_n$ where X_i's are *i.i.d.* random variables with distribution function F. We have:

$$\begin{aligned} &P(S_1 \leq t) = P(x_1 \leq t) = F(t) = F^{(1)}(t), \\ &P(S_2 \leq t) = P(x_1 + x_2 \leq t) = F^{(2)}(t) = \int_0^t F(t-y)dF(y) \\ &\Rightarrow P(S_n \leq t) = F^{(n)}(t) = \int_0^t F^{(n-1)}(t-y)dF(y) \end{aligned}$$

$F^{(n)}$ is the convolution function. $\square$

In practice, it is often very time-consuming and complicated to find the exact distribution of S_n from the formula above. Often an approximation to the distribution of S_n is sufficient.

The strong law of large numbers holds, that is, with probability 1,

$$\frac{S_n}{n} \to \mu \quad as \quad n \to \infty \tag{4.1}$$

According to the central limit theorem, $S_n = \sum_{i=1}^{n} T_i$ is asymptotically normally distributed:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0,1)$$

or

$$F^{(n)}(t) = P(S_n \leq t) \approx \Phi\left(\frac{t - n\mu}{\sigma\sqrt{n}}\right)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution $\mathcal{N}(0,1)$.

Moreover, from the strong law of large numbers, that is, with probability 1,

$$\frac{N(t)}{t} \to \frac{1}{\mu} \quad as \quad t \to \infty.$$

When t is large, $N(t) \approx \frac{t}{\mu}$. This means that $N(t)$ is approximately a linear function of t when t is large. Therefore, from the definition of $N(t)$ and S_n, it follows that

$$P(N(t) = n) = P(N(t) \geq n) - P(N(t) \geq n+1) = P(S_n \leq t) - P(S_{n+1} \leq t) = F^{(n)}(t) - F^{(n+1)}(t)$$

For large values of n, we can apply the normal approximation for S_n and obtain

$$P(N(t) = n) \approx \Phi\left(\frac{t - n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{t - (n+1)\mu}{\sigma\sqrt{n+1}}\right)$$

Also the *renewal function* $W(t)$ which is the expected number of events in the interval $(0, t]$,

can be expressed as

$$W(t) = E(N(t)) = \sum_{i=1}^{\infty} P(N(t) \geq n) = \sum_{i=1}^{\infty} P(S_n \leq t) = \sum_{n=1}^{\infty} F^{(n)}(t)$$

An integral equation for $W(t)$ may also be obtained as follows

$$\begin{aligned} W(t) &= F(t) + \sum_{r=2}^{\infty} F^{(r)}(t) = F(t) + \sum_{r=1}^{\infty} F^{(r+1)}(t) \\ &= F(t) + \sum_{r=1}^{\infty} \int_0^t F^{(r)}(t-x) dF(x) \\ &= F(t) + \int_0^t \sum_{r=1}^{\infty} F^{(r)}(t-x) dF(x) \\ &= F(t) + \int_0^t W(t-x) dF(x). \end{aligned}$$

Since $W(t)$ is the expected number of renewals in the interval $(0, t]$, the the mean number of failures in this interval is approximately t/μ. We should then expect that

$$\lim_{t \to \infty} \frac{W(t)}{t} = \frac{1}{\mu}.$$

This result is known as the *elementary renewal theorem* and is valid for a general renewal process (the proof may be found in [25] (p.107)). Consequently, the mean number of renewals in the interval $(0, t]$ is

$$E(N(t)) \approx \frac{t}{\mu}$$

when t is large.

Assuming that F is differentiable, the *renewal rate* $w_f(t) = \frac{d}{dt} W(t)$ can also be expressed as $w_f(t) = \sum_{n=1}^{\infty} f^{(n)}(t)$, where $f^{(n)}(t) = \frac{d}{dt} F^{(n)}(t)$.

4.3.2 Superposition processes

In the book of Ascher [20] a superposition process is defined as follows. Assume that n renewal processes $N_1, \ldots, N_n$ with rates $\lambda_1, \ldots, \lambda_n$, which are the rate of occurrence of the events, are operating independently of each other. Then the process formed by the union of all events is known as a *Superposition* (or superimposed) Process (SP), with the rate $\lambda = \sum_{i=1}^{n} \lambda_i$.

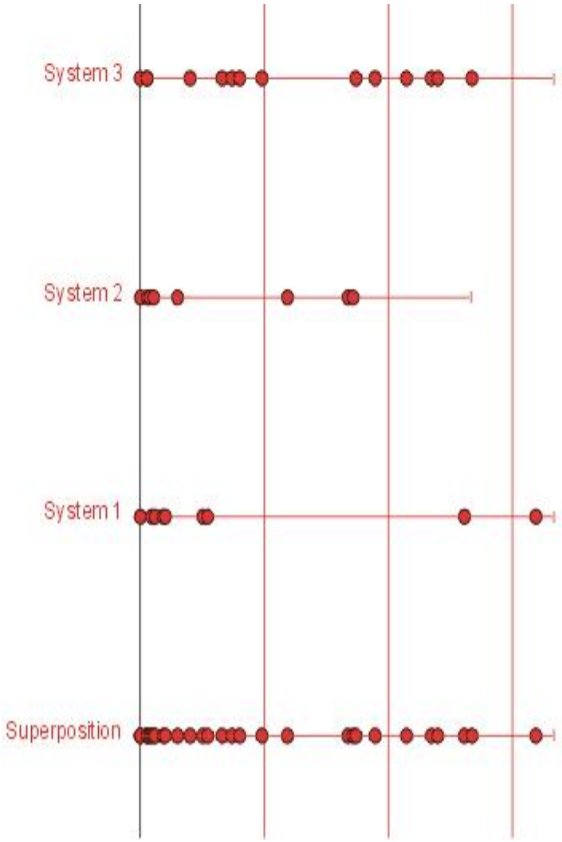

Figure 4.5: A sample superposition process consists of three renewal processes

The following information is based on the text presented in [20]. In general, the SP will not be a renewal process. In fact, if the superposition of two independent renewal processes is a renewal process then all three processes must be Homogeneous Poisson Process (HPP). However, it has been shown by Dernick [22], that the superposition of an infinite number of independent *equilibrium* renewal processes (i.e. asynchronously sampled [1] renewal processes) is an HPP. This result is analogous to the Central Limit Theorem for the sum of the independent random variables.

[1] Asynchronous sampling means that the process starts at $t = -\infty$ but observation begins at $t = 0$. If we begin observation at the occurrence of an event, then the process is being sampled synchronously

The assumption that n is a very large number is often reasonable since many systems are composed of a large number of components. However, it is often unrealistic to assume asynchronous sampling, i.e., to assume that t is extremely large. Therefore, the results of Grigelionis [23] for finite t are of great interest. According to his paper, under some conditions ([20], pages 35,36), the SP will converge to a non-homogeneous Poisson process (NHPP) or an HPP. However, in general, the basic conditions which are proposed by him for convergence even to an NHPP are questionable.

It is often reasonable to assume a renewal process as a model for each of a system components and systems often can be considered to be comprised of parts in series. The SP, therefore, is often a good first order model for system reliability. For large n and $t \to \infty$, the SP is approximately an HPP (Drenick [22], Khintchine [24]); for large n and $t \to 0$, the SP is approximately an NHPP or HPP (Grigelionis [23]). Note that the probabilistic law of the superimposed process, in general, is unknown.

4.4 The alternative approach

It is well known that the superposition of a finite number of independent renewal processes is not renewal anymore because of the dependency between the interarrival times. A renewal approximation can be obtained by simply ignoring the interarrival dependencies and using the interarrival distribution. This is done by Torab and Kamen [26] where he suggested the rate approximation. According to their paper, it is possible to define a rate process that minimizes the mean-squared rate error functional over the set of all renewal processes.

Consider n independent renewal processes with rates $\nu_i(t), i = 1, \ldots, n$, merging together to construct a superposition process. Obviously, the superposition process rate $\nu_s(t)$ equals the sum of the rates for each process

$$\nu_s(t) = \sum_{i=1}^{n} \nu_i(t).$$

Let $\nu_{app}(t)$ be the approximated rate which is proposed to be a replacement for $\nu_s(t)$. Then the mean-squared rate error

$$e\langle \nu(\cdot) \rangle = E\left([\nu_{app}(t) - \nu_s(t)]^2 \right)$$

will be minimized due to this rate approximation.

Before using this idea, we consider a simple case in which all the renewal processes $N_1, \ldots, N_n$ are homogeneous Poisson processes. This guarantees that the superimposed process is also a homogeneous Poisson process [20]. The interarrival times in all the processes are exponentially distributed with parameter λ_i for $i = 1, \ldots, n$. As a result, the interarrival times in the superposition process also has an exponential distribution with parameter $\lambda = \sum_{i=1}^{n} \lambda_i$, as discussed in the previous section. Therefore, the time between the repairs simply equals the sum of the 3% of the interarrival times, $S_{\alpha n} = \sum_{i=1}^{\alpha n} T_i$. Note that $\alpha = 3\%$ and n is chosen large enough so that αn is always an integer.

In this case, thanks to the nice property of the exponential distribution, the distribution of $S_{\alpha n}$ is known and is the gamma (or Erlang) distribution with shape parameter αn and scale parameter $\frac{1}{\lambda}$. This is because of the fact that the exponential distribution with the parameter λ is actually the special case of the gamma distribution with shape parameter 1 and scale parameter $\frac{1}{\lambda}$ and therefore possesses the conjugate property of the gamma distribution, which indicates that the sum of the n independent gamma distributed random variables is again gamma distributed. The proof of this property can be found in Bedford and Cooke [15].

To check if the mentioned theoretical assumptions for the exponential case can be confirmed by the results from the data, a simulation is performed. In this simulation, it is assumed that each lifetime has an exponential distribution and is chosen from a renewal process, and all the renewal processes are independent. Then considering the superposition of all these renewal processes and applying the time readjustment after each repair, which was explained in section 4.1, the time between the repairs is computed and plotted. We consider 10000 independent renewal processes and 200 repairs. On the other hand, as mentioned in the previous part, the superposed process of a homogeneous Poisson processes is itself an HPP process with the exponential interarrival times; hence, the distribution of the inter-repair times is known which is the gamma distribution with the parameters mentioned above.

Considering the mean of the inter-repair times of the renewal processes and αn times the mean of the interarrival times of the superimposed process (0.3044 and 0.3017, respectively) we can conclude that it is convenient to use the properties of the renewal process for computing the inter-maintenance (repair) time distribution when we have the superposition of n independent HPP

processes with exponential interarrival times. Moreover, αn times the mean of the interarrival times of the superimposed HPP process can be considered as the value that the inter-repair times converges to (considering equation (4.1)), after a large number of repairs. However, at this point this has not been proven yet.

In the general case, if the renewal processes are not Poisson processes, the superimposed process in not renewal process any more. However, thanks to the results presented in [26], it is possible to approximate the superimposed process by a renewal process applying a rate approximation. As it is shown in this research, one can approximate the rate of the superposition process in such a way that the new process with this approximated rate represents the renewal process, and the rate-error, $e\langle\nu(\cdot)\rangle = E\left([\nu_{app}(t) - \nu_s(t)]^2\right)$, is sufficiently small enough. For more details, the reader is referred to [26].

For a renewal process with the interarrival density $f(t)$ and survival function $R(t)$, the rate process $\nu(t)$ has the simple form, [27],

$$\nu(t) = \frac{f(t)}{R(t)} = -\frac{d}{dt}\ln R(t).$$

The inverse relations are the familiar reliability equations

$$R(t) = \exp\{-\int_0^t \nu(u)du\} \tag{4.2}$$

$$f(t) = \nu(t)\exp\{-\int_0^t \nu(u)du\} \tag{4.3}$$

which show how to compute the interarrival distribution from the process rate.

Assume that the superposition process is built by n independent renewal processes with the rate process $\nu_i(t)$ for $i = 1, \ldots, n$. According to [26], the suggested approximation rate at each time t can be obtained as follows

$$\nu_{app}(t) = \frac{\sum_{i=1}^n \eta_i(t)\left(\nu_i(t) + \sum_{\substack{j=1 \\ j\neq i}}^n \eta_j(t)\right)}{\sum_{i=1}^n \eta_i(t)} \tag{4.4}$$

where $\eta_i(t)$ is the recurrence rate function of each process. The recurrence rate function is defined

as

$$\eta(t) = \frac{R(t)}{\psi(t)}, \tag{4.5}$$

and $\psi(t)$ is defined as

$$\psi(t) = \int_t^\infty R(u)du$$

with $\psi(0)$ being equal to the mean interarrival time.

It is possible to simplify (4.4) for the case that all the renewal processes have the same distribution function. Assume that all n renewal processes have the same interarrival distribution, $f(t)$, and for each process, the interarrival times $T_i, i = 1, \ldots, n$ are i.i.d and $E(T)$ is their expected value. Therefore, the rate approximation in (4.4) can be simplified as follows:

$$\begin{aligned} \nu_{app}(t) &= \frac{n\eta(t)(\nu(t) + (n-1)\eta(t))}{n\eta(t)} \\ &= \nu(t) + (n-1)\eta(t) \end{aligned} \tag{4.6}$$

Also from (4.5) we can get:

$$\eta(t) = \frac{R(t)}{\psi(t)} = -\frac{d}{dt}\ln\psi(t) \tag{4.7}$$

Considering the density of the interarrival times of the approximated renewal (superimposed) process in (4.3), first we compute the integral of $\nu_{app}(t)$ considering (4.6) and definition of $\nu(t)$.

$$\begin{aligned} \int_0^t \nu_{app}(u)du &= \int_0^t \nu(u)du + (n-1)\int_0^t -\frac{d}{du}\ln(\psi(u))du \\ &= -\ln R(t) - (n-1)(\ln\psi(t) - \ln(E(T))) \\ &= -\left(\ln R(t) + \ln\left(\frac{\psi(t)}{E(T)}\right)^{n-1}\right) \\ &= -\ln\left[R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1}\right] \end{aligned} \tag{4.8}$$

Hence,

$$\begin{aligned} e^{-\int_0^t \nu_{app}(u)du} &= e^{\ln R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1}} \\ &= R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1} \end{aligned} \tag{4.9}$$

Therefore, equation (4.3) simplifies to

$$f(t) = \left(\nu(t) + (n-1)\frac{R(t)}{\psi(t)}\right).\, R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1} \tag{4.10}$$

This simplification speeds up the computational process. To check the accuracy of (4.10), we find the pdf of superimposed process for the exponential case with parameter λ.

$$\begin{aligned} \nu(t) &= \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda, \\ \psi(t) &= \int_t^\infty e^{-\lambda u}du = \frac{1}{\lambda}e^{-\lambda t}, \\ \Rightarrow \quad & \frac{R(t)}{\psi(t)} = \lambda, \\ \Rightarrow \quad & \nu(t) + (n-1)\frac{R(t)}{\psi(t)} = \lambda + (n-1)\lambda = n\lambda, \\ \text{and,} \quad & R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1} = e^{-\lambda t}\left(\frac{\lambda^{-1}e^{-\lambda t}}{\lambda^{-1}}\right)^{n-1} = e^{-n\lambda t}, \\ \Rightarrow \quad & f(t) = \left(\nu(t) + (n-1)\frac{R(t)}{\psi(t)}\right).\, R(t)\left(\frac{\psi(t)}{E(T)}\right)^{n-1} = n\lambda e^{-n\lambda t} \end{aligned}$$

which is exactly what is expected when all the renewal processes have an exponential interarrival time with the same parameters.

As a result, assuming that the interarrival distribution of each of these n renewal processes is known, one can find the rate of the interarrival distribution of the superposition process (which in fact is a renewal processes considering our approximation) by applying (4.4) or (4.6). Note that the probability density function of interarrival times in the approximated renewal process will be computed from (4.3) or (4.10). Accordingly, it is possible to find the distribution of the inter-repair times, $S_{\alpha n}$. However, the interarrival times are not always exponentially distributed and may have some other distributions such as Weibull, lognormal, and etc. In these cases, the distribution of $S_{\alpha n}$ is not known and we need to deal with the n-fold convolution of $F^{(n)}(x)$ which is the distribution

function of interarrival times T_i of each renewal processes. Nevertheless, as it was mentioned in section 4.3.1, dealing with the convolution function is time-consuming and complicated. As a result, we apply two alternative approaches to avoid working with the convolution function.

4.4.1 The Normal approach

The first approach is using the central limit theorem. As it was mentioned in section (4.3.1), S_n is asymptotically normally distributed when n is large enough. This means

$$S_{\alpha n} \approx \mathcal{N}(\alpha n\mu, \alpha n\sigma^2)$$

where $\alpha = 3\%$, μ and σ^2 are the mean and the variance of the interarrival times, $T_i,\ i = 1, \ldots, n$. Hence, After applying the equation (4.4) to find the approximated rate, the pdf of the interarrival times can be computed from equation (4.3) for the approximated renewal process. However, in order to apply the central limit theorem, it is first required to compute the mean and the variance of the interarrival times. This can be done by simply computing the first and the second moments of the interarrival times from their obtained pdf.

For the means of experiment, we choose 5000 cells; each cell represents a single renewal process, i.e. we consider 5000 independent processes. Note that in this part of our research, we neglect the dependency between the cells and assume that each cell corrodes independently. Suppose that the interarrival times of each of these processes have the Weibull distribution with the same parameters. Applying equation (4.10) to obtain the pdf of the interarrival times, enables us to compute their mean, μ, and the variance, σ^2. Therefore, considering the number of processes, $n = 5000$, it follows that $S_{\alpha n} = S_{150} \approx \mathcal{N}(150\mu, 150\sigma^2)$. Like the other experiment with the exponential distribution, we consider 200 repairs and the time interval is $[0, 5]$ with the step size 0.001. To be able to check the accuracy of the results, we also sample 5000 Weibull random variables and apply the time-resetting in the way that was explained in section (4.1).

The mean of the time between the repairs from 5000 renewal processes is 0.1382. In addition, $E(S_{\alpha n}) = \alpha n\mu$, where $\mu = 0.000972$ is the mean of the interarrival times of the process. Consequently in our case, $E(S_{150n}) = 0.1458$ which is close enough to the mean of the inter-repair times of the renewal processes.

We did the same experiments for the case that all renewal processes have the interarrival times

which are lognormally distributed with the same parameters. Using (4.10) to get the pdf and then obtaining the mean and variance of the interarrival times, it is possible to compare the experimental result to the theoretical result. As shown in figure 4.3, the inter-repair times for the lognormal distribution converges approximately to the value 0.45; considering the mean of the interarrival times $\mu = 0.003072$, the mean of the inter-repair times is also $E(S_{\alpha n}) = 150\mu = 0.4608$ which is again very close to the value to which the inter-repair times converge. The results are shown in the table below.

Distribution	Experimental result	Theoretical result
Weibull	0.1382	0.1458
Lognormal	0.4478	0.4608

These results show that the time between the repairs converges to $3\%n$ times the mean of the interarrival times of the superimposed process, using the rate approximation method in [26].

Nevertheless, to deal with the processes with different interarrival times distributions, or even with the same distribution and different parameters, our sampling method and time-resting will not function properly. Therefore, we think of a more general method to be able to tackle these difficulties. The alternative solution is applying the Fast Fourier Transform (FFT) method which is discussed in the next section.

4.4.2 Fast Fourier Transform (FFT) method

A fast Fourier transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) and its inverse which reduces the number of computations needed for N points from $2N^2$ to $2N \log N$. FFTs were first discussed by Cooley and Tukey [28], although Gauss had actually described the critical factorization step as early as 1805 (Bergland [29], Strang [30]).

Suppose we are given a function $f : \mathbb{R} \to \mathbb{R}$ which is in L^1, i.e.,

$$\int_{-\infty}^{+\infty} |f(x)|\, dx < \infty,$$

and if $f(t)$ is continuous, then the Fourier transform of $f(t)$ is defined as:

$$\phi(t) = \mathbb{E}\left(e^{itX}\right) = \int_{-\infty}^{+\infty} e^{itx} f(x)dx = \int_{-\infty}^{+\infty} e^{itx} dF(x),$$

where $t \in \mathbb{R}$. Assuming that $\phi(t)$ is in L^1, the original function can be recovered from its Fourier transform by inversion:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \phi(t)dt.$$

In probability theory a characteristic function of a continouous random variable X, is by definition the Fourier transform of the density of X. Now, suppose that we discretize the domain for x, and t into N grid points, then we consider the vectors $F, f \in \mathbb{C}^N$:

$$F = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_{N-1} \\ F_N \end{pmatrix}, f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N-1} \\ f_N \end{pmatrix}. \tag{4.11}$$

If we let

$$\omega_N = e^{-\frac{2\pi i}{N}},$$

the discretized Fourier Transform- matrix $M \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$ is then defined as:

$$M = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega_N^1 & \omega_N^2 & \dots & \omega_N^{N-1} \\ 1 & \omega_N^2 & \omega_N^4 & \dots & \omega_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega_N^{N-1} & \omega_N^{(N-1)2} & \dots & \omega_N^{(N-1)(N-1)} \end{pmatrix}, \tag{4.12}$$

that is,

$$M_{nk} = \omega_N^{(n-1)(k-1)}.$$

Now, it can be seen that the discrete Fourier transform F of f is then given by the matrix multi-

plication:

$$F = Mf,$$

or equivalently:

$$F_k = \sum_{n=1}^{N} f_n e^{-\frac{2\pi i}{N}(n-1)(k-1)} = \sum_{n=1}^{N} f_n \omega_N^{(n-1)(k-1)}.$$

Lemma 4.4.1 (Gil-Palaez Inversion)**.** *Let $F(x)$ be the cumulative distribution function of some continuous random variable X. Then we have:*

$$F(x) = \frac{1}{2} + \lim_{\delta \to 0} \lim_{T \to \infty} \frac{1}{2\pi} \int_{\delta}^{T} \frac{e^{itx}\phi(-t) - e^{-itx}\phi(t)}{it} dt.$$

Proof. For brevity, the proof is omitted here. The reader is referred to [31] for the complete proof. □

Lemma 4.4.2 (Inversion Lemma)**.** *Let $\phi(t)$ be a characteristic function and $f(x)$ be a probability density function of some continuous random variable X. Then we have:*

$$f(x) = \frac{1}{\pi} \Re \left(\int_0^{\infty} e^{-itx}\phi(t)dt \right)$$

Proof. From the Fourier inverse formula we have:

$$f(x) \quad = \quad \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx}\phi(t)dt = \frac{1}{2\pi} \left(\int_{-\infty}^{0} e^{-itx}\phi(t)dt + \int_0^{+\infty} e^{-itx}\phi(t)dt \right).$$

where the first integral on the RHS can be written as:

$$\int_{-\infty}^{0} e^{-itx}\phi(t)dt = \int_0^{\infty} e^{itx}\phi(-t)dt = \int_0^{\infty} \overline{e^{-itx}\phi(-t)}dt = \int_0^{\infty} \overline{e^{-itx}\phi(t)}dt = \overline{\int_0^{+\infty} e^{-itx}\phi(t)dt}$$

so:

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\left(\int_{-\infty}^{0} e^{-itx}\phi(t)dt + \int_{0}^{+\infty} e^{-itx}\phi(t)dt\right) \\ &= \frac{1}{2\pi}\left(\int_{0}^{+\infty} e^{-itx}\phi(t)dt + \overline{\int_{0}^{+\infty} e^{-itx}\phi(t)dt}\right), \\ &= \frac{1}{\pi}\Re\left(\int_{0}^{\infty} e^{-itx}\phi(t)dt\right). \end{aligned}$$

□

The lemma above shows that we need to find the integral

$$\int_{0}^{\infty} e^{-itx}\phi(t)dt = \int_{0}^{\infty} \gamma(t)dt.$$

Now, we define a trapezoidal integration over domain $[0, \tau]$, for which we have:

$$\begin{aligned} \int_{0}^{\tau} \gamma(t)dt &\approx \frac{\Delta_t}{2}\left[\gamma(t_1) + 2\sum_{n=2}^{N-1}\gamma(t_n) + \gamma(t_N)\right] \\ &= \Delta_t\left[\sum_{n=2}^{N-1}\gamma(t_n) + \frac{1}{2}\left(\gamma(t_1) + \gamma(t_N)\right)\right]. \end{aligned}$$

If we set

$$\tau = N\Delta_t,$$

$$t_n = (n-1)\Delta_t$$

$$x_u = -b + \Delta_x(u-1),$$

where: $u = 1, \ldots, N$ to be the grid in the x-domain, $\Delta_t = t_i - t_{i-1}$, and $\Delta_x = x_i = x_{i-1}$. The constant b is a tuning parameter which can be freely chosen, but here we take:

$$b = \frac{N\Delta_x}{2}.$$

So now, we have:

$$\begin{aligned}\int_0^\tau \gamma(t)dt &\approx \Delta_t\left[\sum_{n=1}^{N} e^{-i[(n-1)\Delta_t][-b+\Delta_x(u-1)]}\phi(t_n) - \frac{1}{2}\left[e^{-ixt_1}\phi(t_1) + e^{ixt_N}\phi(t_N)\right]\right] \\ &= \Delta_t\left[\sum_{n=1}^{N} e^{-i\Delta_x\Delta_t(n-1)(u-1)}e^{i(n-1)b\Delta_t}\phi(t_n) - \frac{1}{2}\left[e^{-ixt_1}\phi(t_1) + e^{ixt_N}\phi(t_N)\right]\right]\end{aligned}$$

Let

$$\Delta_x\Delta_t = \frac{2\pi}{N};$$

therefore,

$$\int_0^\tau \gamma(t)dt \approx \Delta_t\left[\sum_{n=1}^{N} e^{-i\frac{2\pi}{N}(n-1)(u-1)}e^{i(n-1)b\Delta_t}\phi(t_n) - \frac{1}{2}\left[e^{-ixt_1}\phi(t_1) + e^{ixt_N}\phi(t_N)\right]\right].$$

and finally we obtain:

$$\begin{aligned}f(x) &= \frac{1}{\pi}\Re\left(\int_0^\infty e^{-itx}\phi(t)dt\right) \\ &= \frac{1}{\pi}\Re\left\{\Delta_t\left[\sum_{n=1}^{N} e^{-i\frac{2\pi}{N}(n-1)(u-1)}e^{i(n-1)b\Delta_t}\phi(t_n) - \frac{1}{2}\left[e^{-ixt_1}\phi(t_1) + e^{ixt_N}\phi(t_N)\right]\right]\right\} \quad (4.13)\end{aligned}$$

In our case, we are interested in finding the distribution of $S_{3\%n} = \sum_{i=1}^{0.03n} T_i$ applying the FFT algorithm. For brevity, let $3\%n = \alpha n$. So first we have to find the characteristic function of $S_{\alpha n}$ which by i.i.d property of the interarrival times equals:

$$\phi_{S_{\alpha n}}(t) = E(e^{itS_{\alpha n}}) = E(e^{it\sum_{i=1}^{\alpha n} T_i}) = (\phi_{T_1}(t))^{\alpha n} \quad (4.14)$$

Therefore, after using the equations (4.4) and (4.3) (or (4.6) and (4.10)) to obtain the rate and the pdf of the interarrival times, respectively, we apply the FFT algorithm; i.e., firstly we find the characteristic function $\phi(t)$ for the interarrival time T (since all the T_i's are i.i.d we use the general form T); then we apply the inverse function in FFT algorithm, equation (4.13), to attain the probability density function of $S_{\alpha n}$. Note that in equation (4.13) we have to replace $\phi(t)$ by $\phi(t)^{\alpha n}$ as it was shown in equation (4.14).

After finding the density of the inter-repair times, it is of interest to check whether this density

follows any of the known distributions such as the gamma, Weibull, lognormal, normal, etc. To do so, first we have to estimate the parameters of each of these densities. The parameter estimation is done according to the mean and the variance of the obtained density from the FFT algorithm. The normal distribution is the only one which has the property that its parameters are the same as the mean and the variance of the population. For the rest of the distributions, the relation between the parameters and the mean and variance is as follows:

Gamma distribution

Suppose that X is a random variable and has the gamma distribution with parameters a and b; let the mean of X be $E(X) = \mu$ and variance be $Var(X) = \sigma^2$. Then the parameters a and b can be obtained from the mean and the variance as follows:

$$\mu = \frac{a}{b} \quad and \quad \sigma^2 = \frac{a}{b^2} \quad \Rightarrow \quad a = \frac{\mu^2}{\sigma^2} \quad and \quad b = \frac{\mu}{\sigma^2} \tag{4.15}$$

Weibull distribution

If a random variable X with the mean $E(X) = \mu$ and variance $Var(X) = \sigma^2$ has the Weibull distribution with parameters a and b, then the relation between these parameters is as follows:

$$\mu = \frac{1}{a}\Gamma(1 + \frac{1}{b}) \quad and \quad \sigma^2 = \frac{1}{a^2}(\Gamma(1 + \frac{2}{b}) - \Gamma^2(1 + \frac{1}{b})) \tag{4.16}$$

To get the parameters a and b, the system of equations should be solved. We use the MATLAB solve tool to find the answer since the system of equations is not easily solvable by the theoretical methods.

Lognormal distribution

For a random variable X which has the lognormal distribution with the parameters μ and σ^2,

and the mean $E(X)$ and variance $Var(X)$, the parameters can be obtained as follows:

$$E(X) = \exp(\mu + \sigma^2/2) \quad and \quad Var(X) = (\exp(\sigma^2) - 1)\exp(2\mu + \sigma^2)$$

hence

$$\mu = \ln(E(X)) - \frac{1}{2}\ln(1 + \frac{Var(X)}{E^2(X)}) \quad and \quad \sigma^2 = \ln(1 + \frac{Var(X)}{E^2(X)}) \tag{4.17}$$

To perform our experiment, we start with the known distribution, exponential, to check the performance of the FFT method. The number of processes is 100, and it is assumed that all have the exponentially distributed interarrival times with the parameter $\lambda_i = 10$ for $i = 1, \ldots, 100$. The time interval is $[0, 3]$ with the step size 0.001. As it was mentioned at the beginning of the section (4.4), the asymptotic distribution of $S_{\alpha n}$ in this case is a gamma distribution with the parameters αn and $\frac{1}{\lambda}$, where λ is the rate of the superimposed process and equals the sum of the rates of all processes. Remind that this case is a special case where the superimposed process is a renewal process itself. Note also that $\alpha n = 3$, and in equation (4.7) the characteristic function will be to the power 3 to obtain the density of the $S_{\alpha n}$. In figure 4.6 the results are shown; the first plot is the probability density function of the interarrival times in the approximated renewal process, and the second one is the density obtained from the FFT method. As it is visible, the density from FFT and the one from the gamma distribution with the parameters that is mentioned above, fit very well. These results confirm that the FFT method works correctly since the simulation and the theory confirm each other.

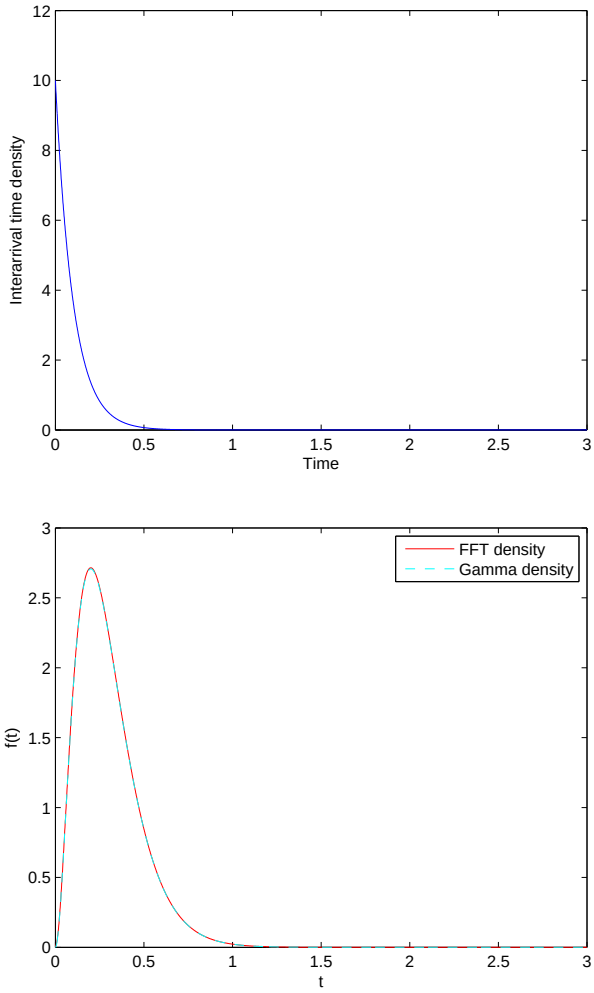

Figure 4.6: From the top: the probability density function of the interarrival times of the estimated renewal process, and the density of the inter-repair times from FFT method (the interarrival times of the processes are exponentially distributed with the same parameters)

For the means of experiment, two other distributions, Weibull and Lognormal, have been chosen. The shape parameter for Weibull is 2.5 and scale parameter is 5. The number of processes is 100 and the time interval and time steps are set as the previous case.

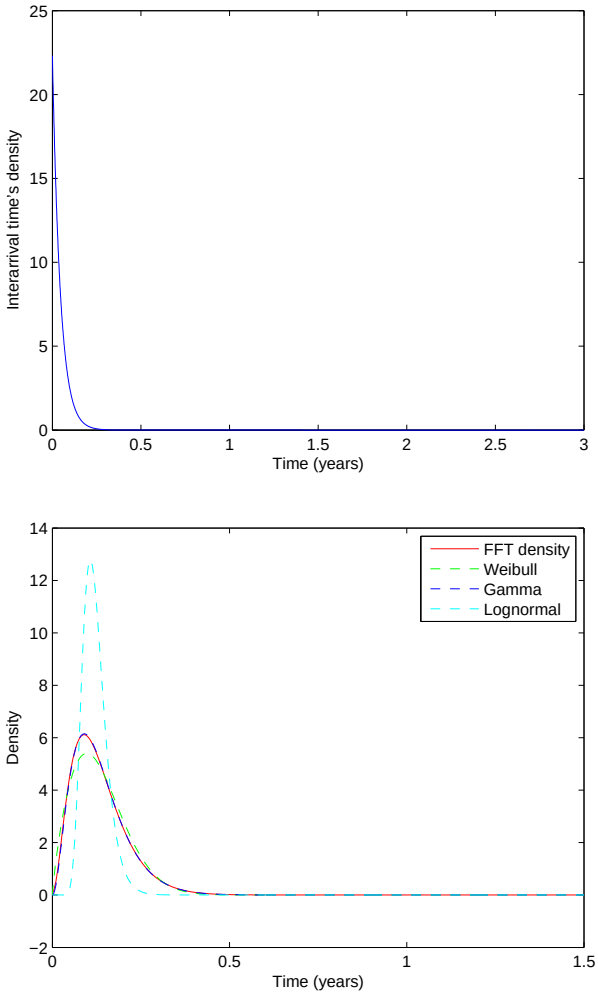

Figure 4.7: From the top: the probability density function of the interarrival times of the estimated renewal process, and the density of the inter-repair times from FFT method (the interarrival times of the processes have Weibull distribution with the same parameters)

As shown it the figure above, the density of the inter-repair time has a gamma distribution since the gamma density fits the density from FFT method perfectly. This means that the interarrival times of the superimposed process have an exponential distribution as it can be seen in the first plot in figure 4.7. In fact the density of the interarrival times in this plot is fitted by an exponential density and the result confirms that this density is exponential.

The same procedure is done for the case that all the renewal processes have a Lognormal distribution with parameters $\mu = \log(5)$ and $\sigma^2 = 1.5$. The number of processes, time interval, and time steps are the same as preveious cases.

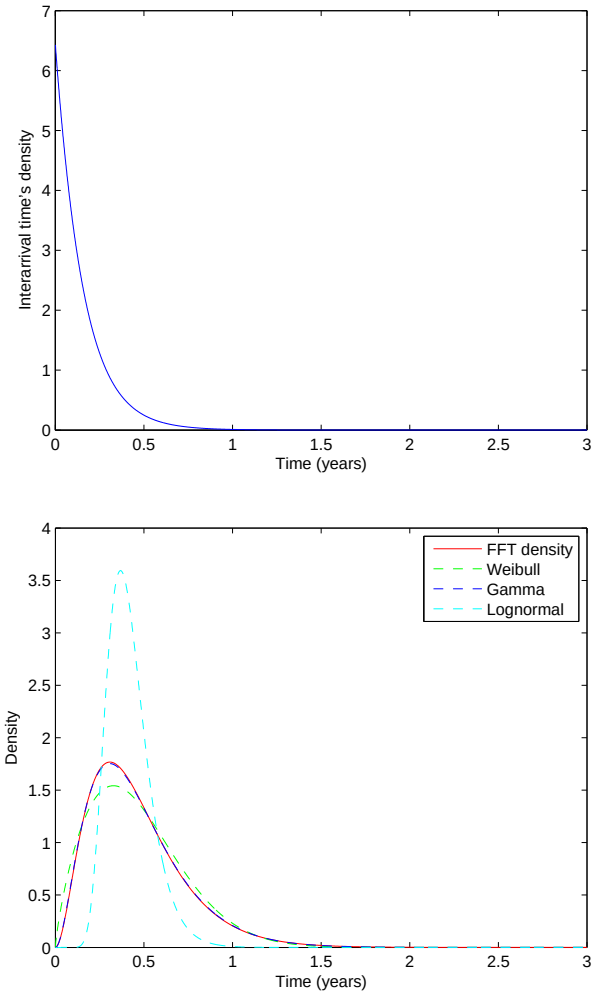

Figure 4.8: From the top: the probability density function of the interarrival times of the estimated renewal process, and the density of the inter-repair times from FFT method (the interarrival times of the processes have lognormal distribution with the same parameters)

In this case also, the best fit is the gamma density as shown in figure 4.8. So the conclusion can be the same as before, i.e., the interarrival times of the superimposed process are exponentially distributed which can be seen in the first plot of figure 4.8 as well. To be more precise, the K-S test is applied to check whether the approximated density of the inter-repair times follows a gamma distribution. In the following table, the results of the K-S test are presented. Considering the level

Distribution	p-values
Weibull	0.1257
Lognormal	0.1164

0.05 for rejecting the null hypothesis, we do not have the basis for rejecting the hypothesis that the inter-repair time density follow a gamma density, for both of the distributions.

Moreover, the obtained results about the distribution of interarrival times of the superimposed process are comparable to the results presented in [32]. According to Cox and Smith [32], if the probability density function of all n renewal processes –assuming that all the renewal processes have the same distribution– has the following property

$$\int_0^t f(x)dx = O(t^\beta), \quad \text{i.e.,} \quad \lim_{t\to 0} \frac{\int_0^t f(x)dx}{t^\beta} = C. \tag{4.18}$$

where C is a constant, then the number of events falling in an interval of length t tends to be distributed in a Poisson distribution as n tends to infinity. In particular, the interarrival times tend to be distributed in an exponential distribution.

Considering the result of this paper, for all the distributions that their pdf satisfy (4.18), the interarrival distribution of the superimposed process are exponentially distributed. This is what we got from the result of the rate approximation of Torab and Kamen [26]. As presented before, all the plots approve that the pdf of the interarrival times of the superimposed process are exponentially distributed according to the density of $S_{\alpha n}$, the inter-repair times. The distributions that we examined were Exponential, Weibull, and lognormal. Obviously, exponential satisfies (4.18). For Weibull with scale parameter α and shape parameter β with the CDF $F(x) = 1 - e^{-(x/\alpha)^\beta}$, by

using the Taylor expansion, we can see:

$$\begin{aligned}
\lim_{t\to 0}\frac{1}{t^\beta}\left(\int_0^t f(x)dx\right) &= \lim_{t\to 0}\frac{1}{t^\beta}\left(1-e^{-(t/\alpha)^\beta}\right)\\
&= \lim_{t\to 0}\frac{1}{t^\beta}\left(1-\left(1-(t/\alpha)^\beta+\frac{(t/\alpha)^{2\beta}}{2!}-\frac{(t/\alpha)^{3\beta}}{3!}+\dots\right)\right)\\
&= \lim_{t\to 0}\frac{1}{\alpha^\beta}-\frac{(t/\alpha)^\beta}{2!}+\frac{(t/\alpha)^{2\beta}}{3!}-\dots\\
&= \frac{1}{\alpha^\beta}=C
\end{aligned}$$

Therefore, Weibull satisfies (4.18) as well. For the case of Lognormal we have:

$$\begin{aligned}
&\int_0^t f(x)dx=\int_0^t \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}dx\\
\Rightarrow\ &\lim_{t\to 0}\frac{1}{t^\beta}\int_0^t f(x)dx=\frac{0}{0}.
\end{aligned}$$

Since $f(x)$ and t^β are both differentiable, by applying the l'Hôpital's rule and assuming that $\mu=0$ and $\sigma^2=1$ for simplicity, we get the following result:

$$\begin{aligned}
\lim_{t\to 0}\frac{1}{t^\beta}\int_0^t f(x)dx &\overset{\text{Hôp}}{=} \lim_{t\to 0}\frac{f(t)}{\beta t^{\beta-1}}\\
&= \lim_{t\to 0}\frac{\frac{1}{t\sqrt{2\pi}}e^{-\frac{1}{2}(\ln t)^2}}{\beta t^{\beta-1}}\\
&\overset{(1)}{=} \frac{1}{\beta\sqrt{2\pi}}\lim_{x\to-\infty}\frac{1}{e^{\beta x+\frac{1}{2}x^2}}=0.
\end{aligned}$$

Note that in equality (1) we apply the parameter transformation as follows:

$$x=\ln t\ \ \Rightarrow t=e^x,\ \text{ Hence, as }\ t\to 0\ \ \Rightarrow x\to-\infty.$$

Therefore, condition (4.18) is satisfied for lognormal distribution as well.

The conclusion is that for the special class of distributions that satisfies (4.18), if the number of renewal processes is large enough, the superimposed process tends to a Poisson process with exponential interarrival time distribution. As observed in our experiment, $n=100$ is already large enough to give the desired result.

In Cox and Smith [32], the probability distribution function of the interarrival times of the

superimposed process is also presented. The mentioned pdf is as follows:

$$f_p(y) = -\frac{d}{dy}\left\{R(y)\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-1}\right\}. \tag{4.19}$$

where $R(x) = 1-F(x)$ is the reliability function. Note that this pdf is defined under the assumption that all the renewal processes have the same interarrival distribution functions, $F(x)$ and mean $E(T) = \mu$ where T is the general notation for an interarrival time. It is also possible to check the accuracy of (4.19) by checking it for the exponential case with parameter λ.

$$\begin{aligned}
& R(y) = e^{-\lambda y} \text{ and } \mu = \lambda^{-1}, \\
\Rightarrow \quad & \int_y^\infty \frac{R(x)}{\mu}dx = e^{-\lambda y}, \\
\Rightarrow \quad & R(y)\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-1} = e^{-\lambda y}(e^{-\lambda y})^{n-1} = e^{-n\lambda y}, \\
\Rightarrow \quad & f(y) = -\frac{d}{dy}\left\{R(y)\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-1}\right\} = n\lambda e^{-n\lambda y}
\end{aligned}$$

which is exactly what is expected for exponential case.

By taking the derivative of (4.19), it is simplified to

$$\begin{aligned}
f_p(y) &= -\frac{d}{dy}\left\{R(y)\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-1}\right\} \\
&= f(y)\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-1} - (n-1)R(y)\frac{-R(y)}{\mu}\left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-2} \\
&= \left[\int_y^\infty \frac{R(x)}{\mu}dx\right]^{n-2}\left[f(y)\int_y^\infty \frac{R(x)}{\mu}dx + (n-1)\frac{(R(y))^2}{\mu}\right]
\end{aligned} \tag{4.20}$$

From (4.20) it is possible to repeat the previous experiment for Weibull and lognormal distributions to see whether it gives the same result as before or not. Let us set everything as it was in the previous case; i.e., the number of processes are 100, time interval is $[0, 3]$ and time step is 0.001. The shape and scale parameters for Weibull distribution are 2.5 and 5 respectively, and for lognormal, $\mu = \log(5)$ and $\sigma^2 = 1.5$. The results shown in figure 4.9 confirm the previous ones in figure 4.7 and 4.8. Note that in this figures the density is not 100% fitted to the gamma distribution which is due to the computational error of the software.

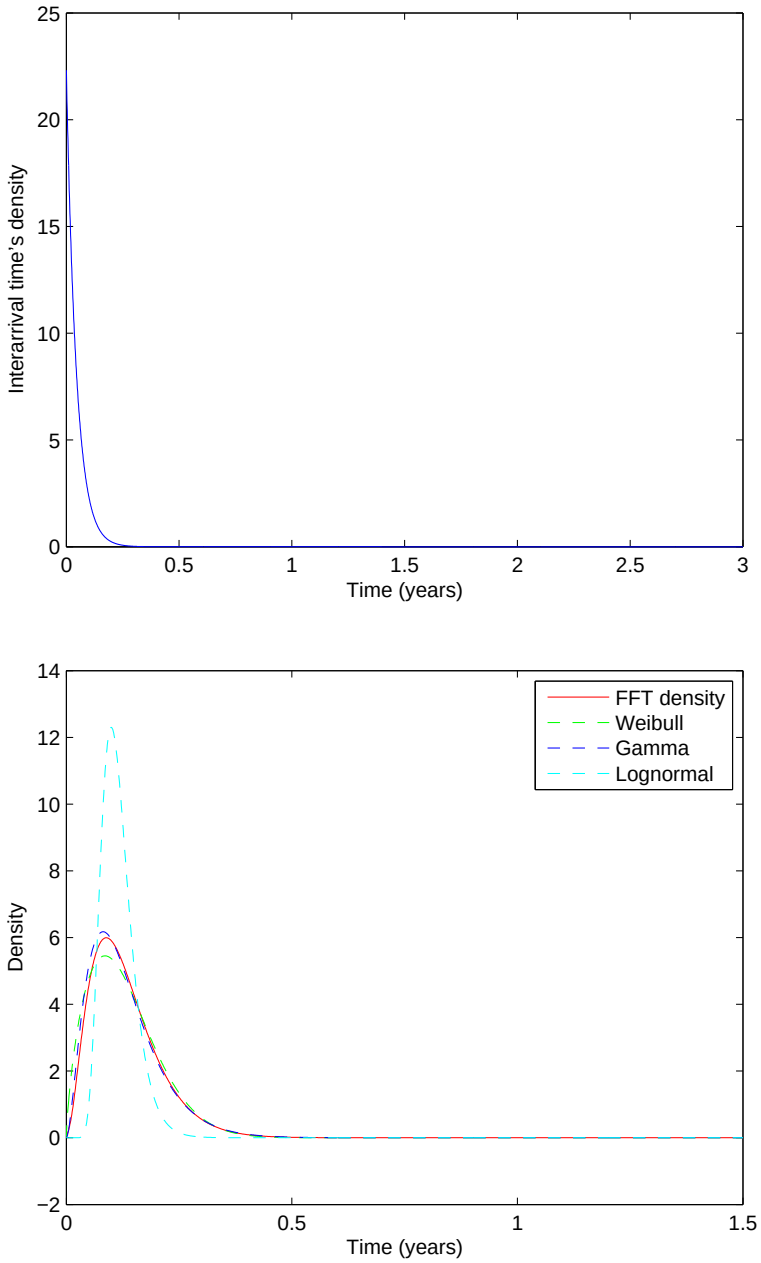

Figure 4.9: From the top: the interarrival density of the superimposed process assuming a Weibull distribution for the interarrival times of all renewal processes; the density of the inter-repair times for the case of Weibull distribution

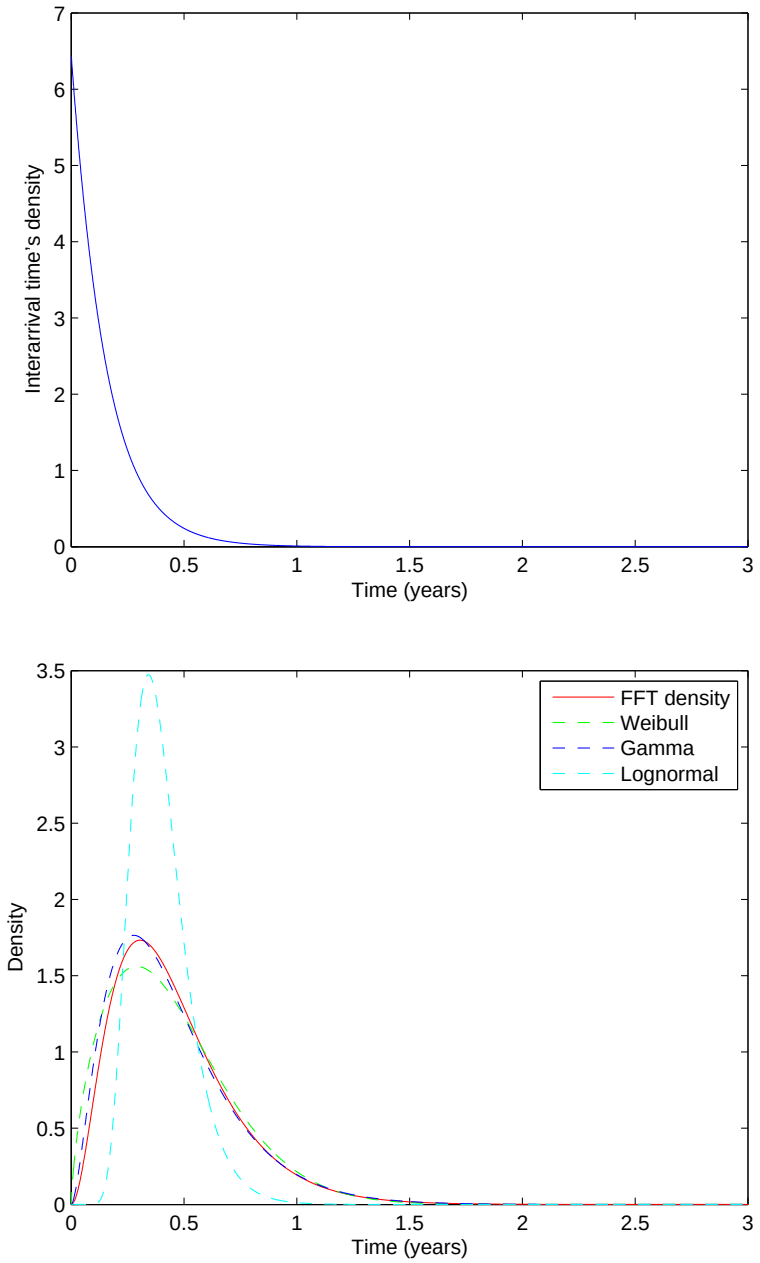

Figure 4.10: From the top: the interarrival density of the superimposed process assuming a lognormal distribution for the interarrival times of all renewal processes; the density of the inter-repair times for the case of lognormal distribution

This comparison can be considered as a reliable tool for validating the results obtained by Torab and Kamen [26], of course for this special class of distributions that satisfies (4.18).

4.5 The average inter-repair time

Approximating the superposition process with a renewal process has given us the chance to estimate the distribution of the inter-repair times even though we could not fit any of the distributions that we considered to the obtained density from the FFT method. Now, we can have a rough estimation of the average of the inter-repair times which can be a criterion for decision makers to adjust the inspection and repair plan with the related costs and expenses considering the required time between the repairs. The mean of the inter-repair times can be simply computed as follows:

$$E[S_{\alpha n}] = E[\sum_{i=1}^{\alpha n} T_i] \overset{(1)}{=} \alpha n \mu$$

where α is the percentage of corrosion, n is the number of cells, μ is the mean of the interarrival time, $E[T_i]$ for $i = 1, \ldots, n$, and the equality (1) comes from the fact that the interarrival times, $T_i,\ i = 1, \ldots, n$, are identically distributed and have the same mean value.

Therefore, the time between the repairs converges to αn times the mean of the interarrival times of the corrosion process, assuming that it can be approximated by a renewal process. This result is comparable to what is shown in figures 4.2 and 4.3; in figure 4.2 the inter-repair times for Weibull distribution converges approximately to 0.14; considering the approximated density from FFT method for the inter-repair time in case of the Weibull distribution, $E[S_{\alpha n}] = 0.03(100)\mu = 3 \times 0.0435 = 0.1305$, which is very close to 0.14. In the following plot these mean values are more visible.

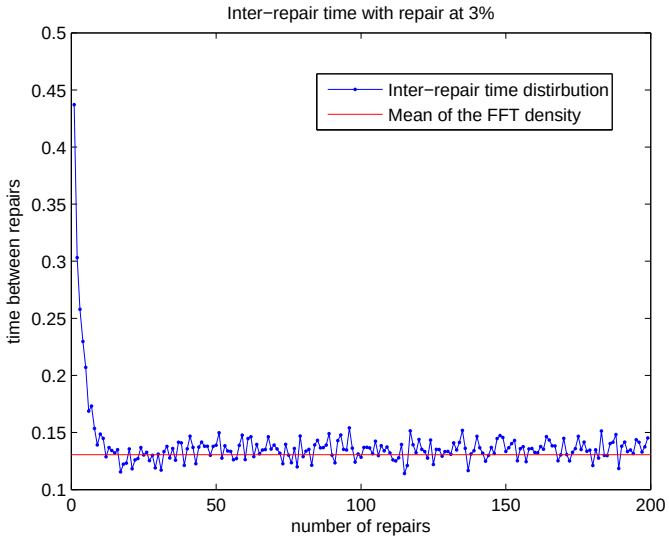

Figure 4.11: The inter-repair time of the Weibull distribution, with the mean value (the solid line) of the interarrival time of the superimposed process

Also in figure 4.3, the inter-repair time for lognormal distribution converges approximately to 0.40; also, the mean of the interarrival times of the superimposed process in the case of the lognormal distribution is $\mu = 0.1540$ and accordingly, $E[S_{\alpha n}] = 3\mu = 0.4620$, which is again quite close to 0.40. In the next figure the result can be seen clearer.

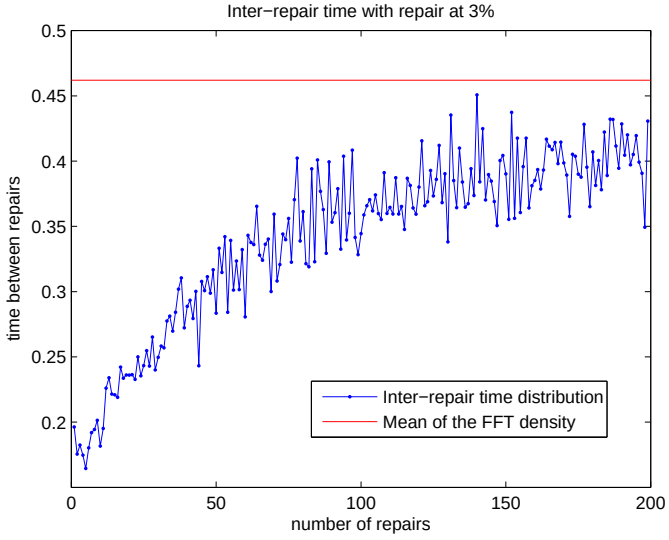

Figure 4.12: The inter-repair time of the Lognormal distribution, with the mean value (the solid line) of the interarrival time of the superimposed process

These figures show a harmony between the theoretical and experimental results since the mean values are close enough to the values that the inter-repair times converge to for a large number of repairs. The small difference, especially in the case of lognormal distribution, is probably due to the approximation methods that we used and computational error, which does not seem too serious to be concerned about.

4.6 Summary

We have tried to interpret the behavior of the inter-repair time in long term, i.e., for large number of repairs. As explained, the problem occurs once we consider the time of all cells together. This makes us to face with a superposition process and dealing with its limitations since the superposition process does not have the nice properties of a renewal process. As an alternative, we approximate this process by a renewal process using a rate-approximation method proposed in [26] and [32]. Considering this results, the superimposed process intends to a Poisson process when the number of processes is large –according to our experiments, 100 processes is large enough to get the results. Therefore, the interarrival times of the superposition process approximately have an exponential distribution, and hence the inter-repair time has a gamma distribution. The value to which the inter-repair times converges to, after a large number of repairs, is $\alpha n \mu$ where α is the corrosion percentage, n is the number of cells, and μ is the mean of the interarrival times.

Chapter 5

Conclusions and future research

In this final chapter, we present a summary of the contributions that have been made in the preceding chapters. We also suggest some possible future directions that are related to the research presented in this thesis.

5.1 Conclusions

The main focus of this project was on two major subjects: the effect of imperfect maintenance on the corrosion process, and the analytical behavior of the time between repair actions. The first topic was mainly based on the simulation of the physical corrosion process for which a model was required to be designed, while the second one was studied from the theoretical point of view.

To check how imperfect maintenance may affect a corrosion process, first we simulated a model that represented a corrosion process on a steel surface. In our model, we considered the surface as a grid consisting of cells, where each cell might be corroded randomly in time and space. The corrosion initiation and propagation were both modeled as a non-homogeneous Poisson process, which means the process rate is not constant in time. In addition, we used the property of Markov random fields that basically means there is a dependent structure between the corroded cells and their neighbors. The closest the neighbors are to the corroded cell, the higher the probability is for infection. Also each level of neighbors had a different rate which also depends on its distance from the corroded cell.

The imperfect maintenance action was performed once 3% of the surface was corroded. This

preventive threshold was defined according to experts opinion. The maintenance action was called imperfect since only the corroded spots were painted and not the whole surface. However, the repairing action is perfect in a sense that each cell would come back to as good as new state. Therefore, after the first repair, some part of the surface became new and the rest had their own age, and that was the occasion that the effect of imperfect maintenance could be studied.

The results of the simulation showed that the corrosion sample path before and after imperfect maintenance followed a gamma process. This was checked according to independence assumption of the corrosion increments and the results of the K-S test which confirmed that all the increments had a gamma distribution. After this approval, we could estimate the parameter of the gamma process before and after imperfect maintenance applying the method of maximum likelihood. The parameters showed that the gamma process in both cases was non-stationary since the power parameter b was not one and its confidence interval was too narrow to be neglected. Also the scale parameter u was very close in both processes while the shape parameters a and b seemed to be shifted after imperfect maintenance. The confidence intervals for all the parameters were narrow enough to show that there is not a big uncertainty in the estimation process. Also the uncertainty in the process itself was not too large according to the 97.5% and 2.5% bands. So we concluded that before and after maintenance, the corrosion process could be represented by a gamma process but with different parameters.

In the second part, we interpreted the behavior of the inter-repair times from the theoretical point of view. We assumed that each cell had its own lifetime; therefore, after the first imperfect maintenance action, some of them kept their old life time and some them got the new ones which caused the mixture of distributions and made us face with a superposition process. To deal with this type of process, we took advantage of the results of two papers to approximate the superimposed process with a renewal process. According to this approximation, the density of the inter-repair time was obtained by applying the FFT method. For two cases of distributions, i.e. Weibull and Lognormal distributions, the obtained density was best fitted to a gamma distribution. Consequently, for a certain class of distributions, the superimposed process had an exponential interarrival times when the number of pooled processes were ralge enough (according to our experiment 100 processes is large enough to get the expected results); subsequently, the inter-repair time, which is defined as the sum of the interarrival times up to 3% corrosion, had a

gamma distribution. As a result, for a large number of repairs, the inter-repair times distribution converged to the percentage of corroded cells multiplied by the mean value of the interarrival times. This could be considered as a criterion for inspection modeling and strategies.

5.2 Future research

Although we tried to have a deep research, there are still some gaps the can be considered as an idea for future research. One of the cases is that in the physical model that we simulated, only one parameter related to the quality of coating was considered. However, in a real case, there are more parameters that can influence the corrosion process such as temperature, pressure, steel quality, and etc. In addition, we just consider the corrosion process in two dimensions and neglect its depth. This can be also considered as a possible option for improvement of the model.

Another expansion can be performed due to the obtained results of the second part of the project. By knowing the value to which the inter-repair times converge, it will be very useful to find an optimal maintenance model with low cost and more optimal inspection strategies, by considering this results. Due to the lack of time, we did not study any cost optimization in this project which can be considered in future research.

Bibliography

[1] Nicolai R.P., Dekker R., van Noortwijk J.M. A comparison of models for measurable deterioration: An application to coatings on steel structures, Reliab Eng Syst Safety, Volume 92, Issue 12, Pages 1635-1650, December 2007.

[2] Nicolai R.P., Koning A.J., A white-box degradation model for steel corrosion, Econometric Institute (EI), report 45, 2006.

[3] van Noortwijk J.M., A survey of the application of gamma processes in maintenance, Reliab Eng Syst Safety, doi:10.1016/j.ress.2007.03.019, under the publication.

[4] Ostrowska A., Simulating inspections on corroded surfaces, M.Sc. thesis, Technical university of Delft, July 2006.

[5] Nicolai R.P., Frenk J.B.G., Dekker R., Modeling and optimizing imperfect maintenance of coatings on steel structures, ERIM Report Series Reference No. ERS-2007-043-LIS, Available at SSRN, March 2007.

[6] Kallen M.J., van Noortwijk J.M., Optimal maintenance decisions under imperfect inspection, Reliab Eng Syst Safety, Volume 90, Issues 2-3, Pages 177-185, November-December 2005.

[7] van Noortwijk J.M., Frangopol D.M., Two probabilistic life-cycle maintenance models for deteriorating civil infrastructures, Reliab Eng Syst Safety, Volume 19, Issue 4, Pages 345-359, October 2004.

[8] Meier-Hirmer C., Sourget F., Roussignol M., Optimizing the strategy of track maintenance, ESREL conference, 2005.

[9] Bérenguer C., Grall A., Dieulle L., Roussignol M., Maintenance policy for a continuously monitored deterioration system, Probability in the Engineering and Informational Sciences, Volume 17, Pages 235-250, Cambridge University Press, 2003.

[10] van Noortwijk J.M., Klatter H.E., Optimal inspection decisions for the block mats of the Eastern-Scheldt barrier, Reliab Eng Syst Safety, Volume 65, Issues 3, Pages 203-211, 1999.

[11] Liao H., Elsayed E.A., Chan L.Y., Maintenance of continuously monitored degrading systems, Reliab Eng Syst Safety, Volume 175, Issue 2, Pages 821-835, December 2006.

[12] Castanier B., Bérenguer C., Grall A., A sequential condition-based repair/replacement policy with non-periodic inspections for a system subject to continuous wear, Appl. Stochastic Models Bus. Ind., 2003.

[13] Varshney P.K., Arora M.K., Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data, Springer, Inc, 2004.

[14] van Noortwijk J.M., Pandey M.D., A stochastic deterioration process for time-dependent reliability analysis, M.A. Maes and L. Huyse, editors, Proceedings of the Eleventh IFIP WG 7.5 Working Conference on Reliability and Optimization of Structural Systems, 2-5 November 2003, Banff, Canada, pages 259-265. London: Taylor & Francis Group, 2004.

[15] Bedford T., Cooke R. M., Mathematical tools for probabilistic risk analysis, Cambridge University Press, 2001.

[16] Grzelak L. A., A statistical approach to determine Microbiologically Influenced Corrosion (MIC) Rates of underground gas pipelines, M.Sc. Thesis, Technical University of Delft, July 2006.

[17] Miller I., Miller M, John E. Freund's mathematical statistics, sixth edition, Prentice-Hall, Inc., 1999.

[18] Greene W. H., Econometric analysis, forth edition, Prentice Hall, Inc, 2000.

[19] Johnson R. A., Wichern D. W., Applied multivariate statistical analysis, fifth edition, Prentice Hall, Inc., 2002.

[20] Ascher H., Feingold H., Repairable systems reliability, Marcel Dekker Inc., 1984.

[21] Rausand M., Hoyland A., System Reliability Theory, John Wiley & Sons, Inc., 2004.

[22] Drenick R. F., The failure law of complex equipment, J. Soc. Indust. Appl. Math., Volume 8, Pages 680-690, 1960.

[23] Grigelionis B. I., Limit theorems for sums of renewal processes, Energy publishing house, Volume 2, Pages 246-266, 1964.

[24] Khintchine A. Y., Mathematical methods in the theory of queueing, Charles Griffin, London, 1960.

[25] Ross S.M., Stochastic Processes, John Wiley & Sons, Inc., 1996.

[26] Torab P., Kamen E. W., On Approximate Renewal Models for the Superposition of Renewal Processes, Communications, 2001. ICC 2001. IEEE International Conference on (Finland), Volume: 9, Pages 2901-2906.

[27] Karr A. F., Point Processes and Their Statistical Inference, Marcel Dekker, Inc., 1986.

[28] Cooley J. W. and Tukey, O. W., An Algorithm for the Machine Calculation of Complex Fourier Series., Math. Comput., Volume 19, Pages 297-301, 1965.

[29] Bergland G. D., A Guided Tour of the Fast Fourier Transform., IEEE Spectrum 6, Pages 41-52, July 1969.

[30] Strang G., Wavelet Transforms Versus Fourier Transforms., Bull. Amer. Math. Soc., Volume 28, Pages 288-305, 1993.

[31] ManWo Hg, Option Pricing via the FFT, M.Sc. Thesis, Technical University of Delft, March 2005.

[32] Cox D. R., Smith W. L., On the superposition of renewal processes, Biometrika, Volume 41, Pages 91-99, 1954.

Appendix A: Mean & variance of the gamma distribution

Suppose that X has a gamma distribution with shape parameter $\nu > 0$ and scale parameter $u > 0$. Considering the definition of probability density function of the gamma distribution given in chapter 3, its mean and the variance can be computed using the moment generating function, as follows:

$$\begin{aligned} M(t) &= \int_0^\infty e^{tx} \frac{u^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-ux} dx \\ &= \int_0^\infty \frac{u^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-(u-t)x} dx \end{aligned}$$

In order to explicitly find the moments of the distribution using the moment-generating function, let

$$y = u(1 - \frac{t}{u})x \qquad dy = u(1 - \frac{t}{u})dx$$

so

$$\begin{aligned} M(t) &= \int_0^\infty \frac{u^\nu}{\Gamma(\nu)} (\frac{y}{u(1-\frac{t}{u})})^{\nu-1} e^{-y} \frac{dy}{u(1-\frac{t}{u})} \\ &= \frac{1}{(1-\frac{t}{u})^\nu \Gamma(\nu)} \int_0^\infty y^{\nu-1} e^{-y} dy \\ &= \frac{1}{(1-\frac{t}{u})^\nu} \end{aligned}$$

Giving the logarithmic moment generating function as

$$\begin{aligned} R(t) &= -\nu ln(1-\frac{t}{u}) \\ R'(t) &= \frac{\nu}{u(1-\frac{t}{u})} \\ R''(t) &= \frac{\nu}{u^2(1-\frac{t}{u})^2} \end{aligned}$$

Therefore, the mean and the variance, respectively, are

$$E(X) = \frac{\nu}{u} \qquad Var(X) = \frac{\nu}{u^2}.$$

Appendix B: Some probability distribution functions

In chapter 4, we used the exponential, Weibull, and lognormal distributions. Due to the fact that thses distributions can be defined in different ways, We present here the probability density functions of these distributions that we used in simulations.

- **Exponential with parameter λ**

$$f(t) = \frac{1}{\lambda} e^{-\frac{t}{\lambda}}$$

- **Weibull with shape parameter b and scale parameter a**

$$f(t) = b a^{-b} t^{b-1} e^{-(\frac{t}{a})^b}$$

- **Lognormal with parameters μ and σ^2**

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}}$$