Extension of the gamma process and its application in prescriptive and condition-based maintenance under imperfect degradation information

Nicola Esposito

COLLEGE SCIENCES
DOCTORAL DE L'INGENIERIE
PAYS DE LA LOIRE ET DES SYSTEMES

université angers

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ D'ANGERS
délivré conjointement avec
L'UNIVERSITÀ DEGLI STUDI
DI NAPOLI FEDERICO II

ÉCOLE DOCTORALE N° 602
Sciences pour l'Ingénieur
Spécialité : *Electronique, signal, génie industriel*

Par
Nicola ESPOSITO

Extension of the gamma process and its application to prescriptive and condition-based maintenance under imperfect information

Thèse présentée et soutenue à Angers, le 04/12/2023
Unité de recherche : LARIS

Rapporteurs avant soutenance :

Mitra FOULADIRAD — Professeur des Universités, Centrale Marseille
Antoine GRALL — Professeur des Universités, Université de Technologie de Troyes

Composition du Jury :

Président :	Christophe BÉRENGUER	Professeur des Universités, Université Grenoble Alpes
Examinateurs :	Mitra FOULADIRAD	Professeur des Universités, Centrale Marseille
	Antoine GRALL	Professeur des Universités, Université de Technologie de Troyes
	Zhiguo ZENG	Maître de conférences HDR, CentraleSupélec
Dir. de thèse :	Bruno CASTANIER	Professeur des Universités, Université d'Angers
Co-dir. de thèse :	Massimiliano GIORGIO	Professeur des Universités, Università degli Studi di Napoli Federico II

Acknowledgement

I owe my deepest thanks to my supervisors, Massimiliano Giorgio and Bruno Castanier.
Often PhD supervisors are thanked for their "guidance", but in my experience their support went much beyond mere guidance. Very often, especially in the first years when I needed it the most, but also in later years, they were coming at me with suggestions, ideas, they were reading papers on my behalf, digging through the literature together with me, developing mathematical methods for me to implement. Then, when writing papers, they did not stop at the mere "reviewing" or "editing" or "giving some advices", but very often got their hand dirty and heavily modified, or outright wrote entire chapters of some of our papers. I felt like they were carrying me on their back.
And throughout this entire process, I never felt like I was being tutored by some gurus. They always treated me like I was one of they peer, they always tried to make me feel like I was an equal to them. And I want to stress the "tried" part, because despite their best efforts, their knowledge about the scientific literature, their competence, and especially their hunger to know more, to read more, to study more even after decades in the field, made it so that 3 years of PhD is not nearly enough time to catch up to them.

Then, I want to thank my family. A PhD is a long process, 3 years is a long time, especially when you are relatively young and far from home. When I started the PhD I was 24 and clueless, now I am 27 and slightly less clueless, and it was only during these years that I genuinely realized how much my family really loved me and how much they were willing to sacrifice for me and my well-being.
It was heartwarming to see dad's face light up whenever I got off from the train that brought me back home after some months in France.
It was lovely to pull up my phone and see mom's message writing me that she had been to the grocery store and had bought the particular brand of chocolate that she knows I like, in preparation for my return home a week later.
It was touching to receive a message from my sister sending me a photo of our wonderful dog, telling me that they both missed me.
You all have been there for me all along, but it took living away from home for me to

realize it. I love you all so much.

To my colleagues here in Angers: Sabrine, Bassel, Salim, Fatima, Algassimou, Patty, Guilherme, Rima, Jaber, and Guillaume. When I arrived in Angers three years ago, I had nothing outside of work. Over the years, you have become my life here. I loved getting to hang out with people coming from all over the world, getting to know you and your culture, tasting your cuisine, and making you taste some italian dishes.
And I am especially grateful that, with some of you, I have been able to develop a personal relationship that goes beyond the office, and I can really call you a friend.

To all of my friends back home in Italy. Thank you for making me feel the support from afar. We could go for months without talking to each other, and then randomly meet somewhere in Europe and feel like not a day had passed from the last time we met. I am very grateful for this, and for the fact that, despite the distance, our relationship never changed.

Table of Contents

General introduction

Maintenance is a vital component that holds significant value in all areas of life, ranging from personal health to material possessions such as cars, homes, or industrial equipment. It encompasses the regular upkeep and care of something to prevent damage, breakdowns, or even more significant problems. Overlooking maintenance can lead to costly repairs, decrease in performance, or complete failure of the equipment. By regularly maintaining something, one can ensure its longevity, safety, and optimal functioning. Maintenance also helps to identify potential problems before they become critical, saving time, money, and stress in the long run. Moreover, maintenance activities are also often required to meet regulatory standards.

The importance of maintenance has led to significant research interest, with scholars investigating various aspects of the practice. Research in this area has explored different models, frameworks, and tools aimed at improving the efficiency and effectiveness of maintenance activities. The increasing complexity of industrial machines and equipment has also led to a more sophisticated approach to maintenance, with a focus on strategies that, rather than correcting and minimizing the consequences of a failure (that, depending on the application, can be particularly severe), allow to predict and prevent failures from happening in the first place. To achieve this, it can be useful to subject the system under study to condition monitoring and regular maintenance interventions.

Considering also that in many practical applications monitoring the state of the system and/or performing a repair/replacement may require specially trained personnel or special equipment, it might be needed to schedule maintenance interventions well in advance. Ideally, there would be a mathematical function that, given a certain system in input, returns the exact time at which it will fail, so that maintenance activities will be scheduled accordingly. In reality, this is of course not possible, but the closest thing to this hypothetical function is a model that provides a prediction of the failure time of a system, given all the information available on its state.

As a rule of thumb, if maintenance interventions are carried out regularly and frequently, the risk of failure is vastly mitigated. However, each maintenance action may entail lengthy interruptions of operation and frequent replacements, which may result in

a lower utilization of the equipment over its lifetime. This generates an underlying contrast in maintenance optimization, where the objective is, on the one hand, to intervene on the equipment as little as possible and maximize its utilization, and on the other hand keeping the risk of failure as low as possible.
The classical approach to formulate such a model resorts to lifetime distributions. A lifetime distribution is essentially a random variable that, to each given input (that is usually time, or some other measure of operating lifetime, such as rotating cycles for a drilling rig) associates the probability that the system under study will be in working order at that input. This kind of models are usually calibrated with ad hoc estimation procedures that generally take into account failure time data only. The specific distribution is then chosen from a limited list of pre-packaged black-box models, with model selection being performed via goodness of fit tests and possibly some physics-based considerations about the aging characteristics of the system. However, relying on failure data only presents several problems. Firstly, failure data are often scarce, as running equipment until failure can be expensive and/or time consuming, particularly in the case of highly reliable systems. Historical failure data can be useful in this case, but the records can often be inaccurate, incomplete, and are likely to be significantly right-censored if preventive maintenance actions are performed. Moreover, an approach based on lifetime models is not able to capitalize on all pieces of information that might be available in a practical application, such as degradation measurements.

To overcome these limitations, numerous alternative approaches have been proposed in the literature. One avenue being explored is the integration of artificial intelligence (AI) techniques, such as machine learning and deep learning. These approaches leverage the analysis of large datasets to provide accurate diagnoses of system degradation and are particularly adept at identifying subtle indications of potential failures ("weak signals"). However, one drawback of these methods is their limited predictive capability, which diminishes their effectiveness in maintenance optimization. Moreover, training AI-based approaches requires large amounts of data, which are not always easy to obtain.

Another avenue focuses on the use of stochastic processes. Indeed, a stochastic process, differently than a lifetime distribution, for any given input not only returns the probability of the system being in working order at that input, but also a random variable that describes the distribution of its predicted degradation level. Its use allows, at the same time:

— to formulate models that can incorporate the technological information available

on the degradation/failure causing mechanism (which can be converted in specific features of the stochastic models),

— to use historical degradation data to calibrate the models and evaluate their fitting abilities even without directly observing any failures, and

— to use degradation data collected in real time (via condition monitoring) to update and implement condition-based maintenance strategies.

Objectives of the research activity

The main goal of this research activity is to (i) formulate stochastic degradation processes that accurately model the degradation of the system under study and (ii) leverage their use in ad-hoc developed maintenance strategies.
These models should perform accurate prediction of the failure time of the system, taking into account as much uncertainty as possible and capitalizing on all available pieces of information, while having a simple mathematical structure and effective estimation procedures. In fact, uncertainty on the estimation process and on the prediction abilities of a model can be caused by several sources. Some can be endogenous (such as randomness of the environment, which induces differences in the evolution of the degradation of nominally identical systems), some others can be exogenous (such as the presence of measurement error). Accounting for these kind of uncertainties may lead to the loss of some particularly convenient mathematical properties that can exacerbate not only modeling and inferential problems, but also the use of well-established maintenance strategies. For example, the presence of measurement error requires appropriate modeling solution and very computational- and time-intensive estimation algorithms. Moreover, it prevents from observing the true degradation state of the system, which complicates the use of several maintenance strategies that rely on direct degradation measurements.

Therefore, attention will be focused on developing stochastic degradation processes that can properly take into account all these sources of uncertainty. Secondly, advanced maintenance strategies that exploit the use of these models should also be devised. Nevertheless, the contribution of this manuscript with respect to the maintenance strategies will not be limited at this latter aspect. Indeed, as already mentioned, one of the core issues in maintenance optimization is the contrast between maximum utilization of the system and managing the risk of unexpected breakdowns. In the very recent literature some new maintenance paradigms, such as prescriptive maintenance, have been proposed

as a new way of dealing with this issue. Prescriptive maintenance aims to develop an integrated decision-making framework where all functionalities of a system (i.e., not only the strictly maintenance-related aspects but also economic and operational considerations) are taken into account simultaneously to achieve a more affordable global trade-off. Part of the contribution of this research activity is consecrated to exploring this new research avenue.

The rest of this manuscript is organized as follows:

— Chapter 1 is devoted to a general introduction of the study of maintenance. The general concepts and definitions are laid out. The different types of maintenance actions and strategy are introduced, along with some examples of rigorous performance criteria that can be adopted to optimize said strategies;

— Chapter 2 focuses on the stochastic degradation processes adopted in this manuscript. After a brief introduction of the mathematical background, the specific stochastic processes proposed and adopted in the remainder of the manuscript are illustrated in detail. Moreover, several mathematical tools and numerical algorithms, necessary to perform model calibration and predictions, are discussed;

— Chapter 3 presents two examples of application of a newly developed stochastic process that is able to simultaneously account for several forms of uncertainty, showing (in some experimental scenarios) its superiority to more classical approaches. This chapter is based on the results published in Esposito, Mele, et al. (2022);

— Chapter 4 reports the results of a large misspecification study where the issue of model selection between two stochastic degradation processes with similar characteristics is analyzed, investigating also how measurement error impacts this issue. This chapter is based on the results published in Castanier, Esposito, Giorgio, et al. (2020), Esposito, Mele, et al. (2021b), and Esposito, Mele, et al. (2023b);

— Chapter 5 is devoted to presenting three new maintenance policies that take advantage of the new stochastic processes proposed in Chapter 2. Some example of application of the new policies are developed. Moreover, the misspecification study of Chapter 4 is expanded to investigate its effect on maintenance optimization. This chapter is based on the results published in Esposito, Mele, et al. (2021a), Esposito, Castanier, and Giorgio (2022b), Esposito, Mele, et al. (2023a), and Esposito, Castanier, Giorgio, et al. (2022);

— In Chapter 6 the concept of prescriptive maintenance is tackled. After a brief dis-

cussion about its definition, the proposition of two newly developed examples of prescriptive maintenance strategies follows. This chapter is based on the results published in Esposito, Castanier, and Giorgio (2022a) and Esposito, Castanier, and Giorgio (2023);

— Chapter 7 deals with a preliminary application of prescriptive maintenance to the case of a production system. This chapter is based on the results published in Esposito, Castanier, and Giorgio (2022d) and Esposito, Castanier, and Giorgio (2022c).

Chapter 1

Introduction to maintenance

1.1 Introduction

According to the "Association Française de Normalisation" AFNOR (the french Standardization Association), maintenance is the "combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function".

Within the industrial landscape, in the search for performance optimization and enhancing organizational competitiveness, the role of maintenance has transformed, expanding to encompass all facets of a system's operation, spanning equipment selection, design, and performance monitoring. To execute maintenance effectively, a combination of technical prowess and organizational capabilities, including resource management and task scheduling, is indispensable. The selection and implementation of a maintenance strategy necessitates careful consideration of numerous factors, with a clear understanding of the long-term implications on the overall performance of the system being of utmost importance.

Moreover, maintenance is also a balancing act between two competing goals. On the one hand, frequent interventions and an overall less aggressive utilization of the system can reduce the occurrences of unexpected breakdowns, but can also entail frequent interruption of operations and lower exploitation of the system. On the other hand, infrequent interventions and aggressive usage can extract more economic value from the system at the expense of a higher probability of unexpected breakdowns.

In the remainder of this chapter, some general concepts regarding maintenance will be laid out.

1.2 General concepts

1.2.1 The state of the system and its failure

A key concept in maintenance is the definition of the state of the system under analysis and what it means for it to fail. Both concepts are linked to the aptitude of a system to accomplish a certain function (the *intended mission*) under certain operating conditions, with satisfactory performances.
A system is in a *working state* when it is in a state that allows it to carry out the intended mission. The *new* (also *initial*) state is a working state. If the system is in a *degraded* state, but it is still able to complete the intended mission, than also that degraded state is a working state.
The system is in a *failed state* when it can no longer complete the intended mission.
The *failure* of a system is the event of transitioning from a working state to a failed state. The failure can be *hard* if the system transitions abruptly from the new state to a failed state, or it can be *soft* if there exist some intermediate states between the new state and the failed state.
Failure may be *self-announcing*, meaning that the transition from a working state to the failed state is marked by some signal. For example, the failure of a lightbulb which suddenly blows out and stops working can be envisaged as self-announcing. Alternatively, failure can be *not self-announcing* (e.g., see Bautista, Castro, and Landesa (2022) and Bismut, Pandey, and Straub (2022)) if no signals of its occurrence are given beforehand. In this case, the only way to detect a failure is by means of an ad-hoc inspection revealing the state of the system. This is often the case when failure is defined by the system being deemed too degraded to continue functioning in a safe and effective manner. For instance, a car tyre may be declared as failed and replaced when the treads have eroded more than a preassigned value (that can be defined by the manufacturer or by some regulatory body). This kind of failure is not self-announcing, because it is impossible to detect it unless the thickness of the treads is specifically measured via an appropriate inspection method.

1.2.2 Maintenance actions

A maintenance action is an intervention that can be carried out on the system with the intent of maintaining it. A clear distinction can be made between *monitoring* actions and *curative* actions.

Monitoring actions

A monitoring action, also referred to as *inspection*, is a maintenance action that has the aim of assessing the state of the system.
Inspections can differ on several characteristics, such as their quality (*perfect* or *imperfect*) and frequency (*continuous* or *periodic*).

An inspection is *perfect* when the information that it allows to obtain can be thought of as representing exactly and without error the current state of the system. Conversely, it is *imperfect* when the obtained information is affected by measurement error, or when the state of the system cannot be directly observed and only partial information is available. To properly account for the information gathered by means of an imperfect inspections, appropriate tools must be put in place, such as a filtering procedure, or a model of the system that integrates measurement error.

Inspections are *continuous* when information about the state of the system is available at all times. While continuously inspecting a system may be useful to promptly detect failures, the cost of such procedure can also be quite hefty. Moreover, dealing with a continuous stream of data is not a trivial task. If, on the other hand, monitoring information is only available at some specified times, then inspections are *periodic*.

In addition, inspections can be *destructive* if performing one alters the state of the system in some way, or *non-destructive* if they do not. Finally, if an inspection can be carried out in a time length that is negligible with respect to the lifetime of the system, it is said to be *instantaneous*.

Curative actions

A *curative action* is an action that has the goal of improving the state of the system. Curative actions can be separated between *replacement* and *repair*. A *repair* is an intervention on the system that improves its state without changing any other characteristic (i.e., the behavior of the system is the same before and after the repair). Repairs can come in different types:

— A *total repair* (also *perfect repair*) restores the system to the new state. This is also called "As Good As New (AGAN)" state;

— A *partial repair* (also *imperfect repair*) restores the system to an intermediate state between the failed one and the new one;

— A *minimal repair* is a partial repair where the system is restored to the working state

it was in immediately before failing. This is also called "As Bad As Old (ABAO)" state. Replacing only a small part of a complex system (for example, replacing a faulty pipe on a car engine) could be assimilated to a minimal repair, if the impact of this replacement is negligible onto the whole system's reliability metrics, such as the failure rate;

— A *worse repair* is a repair where the system post-repair is in a worse state than pre-repair. This situation can arise when a poorly-executed repair is carried out and the state of the system is not directly observable after the intervention (either because it is not actually possible to obtain a measurement or because it is affected by error).

A *replacement* is a curative action where the system is substituted by an equivalent system that has the exact same performances and characteristics. This replacing system can be in the new state or in some other degraded state. Depending on the particular application, it is also possible that the replacing system is already in a failed state (replacing a flat car tyre, only to discover that the spare one is flat as well, can be an example of this). Replacements can be *total* (the entire system is replaced) or *partial* (only some part of it is replaced).

1.2.3 Performance criterion

In order to effectively find the optimal trade-off between maintenance's competing goals, it is necessary to define a rigorous performance criterion. The optimal maintenance strategy will then be the one that optimizes this criterion. In the literature, we can distinguish between three widely used criteria: economic, availability, and safety/reliability.

Economic criterion

The economic criterion is the most widely used. It is based on the notion that every maintenance action performed on the system entails some direct and some indirect costs, which the criterion should take into account. The evaluation of the cost of each maintenance action is complex and dependent on various factors such as the type of maintenance action (inspection, repair, replacement), its quality (perfect or imperfect), its duration, and the state of the system at the time when the maintenance action is carried out (for example, replacing a failed system is usually more expensive than replacing a system that is still in a working state). Then, depending on the application, there might also be other

costs involved, such as a logistic cost (that can include the cost of setting up maintenance equipment, transporting maintenance crew on site, among others).

Finally, performing a maintenance action can also entail some indirect costs, such as a lost production (in case operations need to be halted while maintenance is carried out), or a missed opportunity cost (for example, replacing a system that is still in a working state means wasting, by definition, at least some fraction of its operational life).

Availability criterion

If assessing the economic impact of a given maintenance action is particularly difficult, it might be useful to resort to a different performance criterion, optimizing the availability. According to the AFNOR, availability is "the ability of an item to be in a state to perform a required function under given conditions at a given instant of time or during a given time interval, assuming that the required external resources are provided."

Optimizing availability means minimizing the time that a system spends in a non-operational condition, be it planned downtime (such as when the system is under maintenance) or unplanned downtime (such as when the system is in a failed state).

Safety/Reliability criterion

When tight safety regulations are put in place by some regulatory body (such as in the aeronautics industry, or in the space launch industry), or when the consequences of failure are catastrophic and avoidance of failure is of the utmost importance (such as in the nuclear industry), it can be more fruitful to develop a maintenance strategy that prioritizes safety. Primary focus would be, in this case, to optimize reliability, even if this might come at a higher economic cost.

1.3 Types of maintenance policies

1.3.1 Corrective maintenance

Corrective Maintenance (CM) refers to the collection of maintenance actions carried out when the system is already in a failed state (also *reactive maintenance* or *run-to-failure maintenance*). CM is very easy to implement in practice, but, by its own definition, cannot avoid failures. This means that the useful lifetime of the system is systematically

exploited to the fullest, but also that failures will systematically happen. Since maintenance activities are not pre-planned, there is no risk of over-maintaining (i.e., intervening on the system too frequently). However, depending on the complexity on the curative action to be implemented there might be long periods of time where the system is offline awaiting maintenance actions to be carried out (for example waiting for spare parts to be acquired), which results in high maintenance costs. Systematically observing failures might also result in a higher risk of other nearby systems failing (for example, the failure of a pump in a pumping station may put additional load on the other pumps and increase the risk of them failing too). CM is best suited in practical applications where the consequences of failure are very mild and/or replacing/repairing items is a relatively quick and inexpensive process.

For example, household lightbulbs are commonly maintained adopting a corrective approach and are usually replaced only when they go out. In this specific application, a corrective approach can be convenient as the consequences of failure are very mild and replacing a household lightbulb is a quick process that requires limited technical abilities.

1.3.2 Preventive maintenance

Preventive maintenance (PM) refers to the collection of maintenance actions taken on a system that has not yet failed. The main advantage of PM is that the occurrences of failure can be reduced. On the other hand, this also means that some portion of the useful lifetime of the system will not be exploited.
When the consequences of failure are particularly severe, avoiding them might be an interesting proposition. The safety of personnel, other equipment, and the environment can be improved by reducing the number of observed failures. PM is built on the underlying assumption that the failure of the system can somehow be predicted via some model of its state, such as a lifetime distribution or a stochastic process.

The main difference between CM and PM is that the former is a reactive approach while the latter is proactive. However, no proactive approach can prevent all failures. Therefore, even in a nominally preventive strategy, some corrective maintenance actions are to be expected. Maintenance can then become a matter of compromising between frequent interventions (which allow to prevent failures, but can also be very expensive) and maximizing equipment utilization (which can be more economically profitable, but also incurs a higher risk of failure).

Preventive maintenance can be carried out in several ways. In increasing order of

complexity, *systematic*, *condition-based*, and *predictive*.

Systematic maintenance

Systematic (also *time-based* or *age-based*) maintenance is an approach where maintenance actions are performed based on a predetermined intervention calendar. This calendar can be redacted in terms of time, age of the unit, or some other more appropriate measure of the lifetime of the system, such as rotating cycles or items produced. Optimizing this kind of approach boils down to finding the optimal interval between successive maintenance interventions.

Condition-based maintenance

An approach where maintenance interventions are determined based on some information about the state of the system is called *condition-based maintenance* (CBM). This information can be, for example, a measurement of degradation such as the wall thickness of a corroding pipeline or the percentage loss luminescence of an LED light.
Obviously, a condition-monitoring scheme, with appropriate inspection procedures, must be put in place. In addition, also a model of the state of the system that is able to accommodate real-time state updates, such as a stochastic process-based model, should be adopted. Condition-based approaches, under the right circumstances, can be very advantageous from an economic point of view, albeit the implementation can be more complex with respect to systematic maintenance.

Predictive maintenance

Building on condition-based maintenance, if the interventions can be dynamically scheduled taking into account some prediction of the lifetime of the system, the approach is called predictive maintenance.

1.4 Conclusions

The constant evolution of industrial maintenance, transitioning from a simple repair of malfunctioning equipment to a proactive approach focused on anticipation and prediction of failures, reflects a profound shift in how we approach the management of complex systems. This evolution is driven by the desire to better control the operation of industrial

systems, optimize resources, minimize unexpected downtime, and, most importantly, gain competitiveness in an increasingly demanding global market.
One of the main challenges in this endeavor is uncertainty. Uncertainty in maintenance manifest in various forms, such as data-related uncertainty, environmental-related uncertainty, and model-related uncertainty. These uncertainties are often difficult to quantify and effectively manage in practice.
It is worth to emphasize that, even in the presence of uncertainty, decisions must be made to maintain the availability and reliability of industrial systems and that the potential consequences of poor maintenance decisions can be significant, ranging from financial costs to safety concerns.

Thus, to analyze and enhance the performance of maintenance strategies, it can be useful to rely on stochastic models of the relevant system, capable of capturing its random behavior and inherent uncertainty. However, developing and selecting appropriate stochastic models is only the first step. Equally crucial is the need to develop relevant decision rules (i.e., maintenance policies) that can be rigorously evaluated to ensure their effectiveness and their ability to optimize maintenance performance.
In the following chapters, we will delve into detail on some stochastic processes that can be used to model some sources of uncertainty (tackling the data-related and environmental-related uncertainty), investigate the risk and consequences of potential errors in the model selection process (tackling the model-related uncertainty), and then explore new policies that can integrate the aforementioned processes within the framework of maintenance optimization.

Chapter 2

Mathematical framework for perturbed stochastic processes

2.1 Introduction

As remarked in Chapter 1, a cornerstone of the success of the application of advanced maintenance strategies is the idea that the failure of the system under analysis can be predicted or estimated. Historically, this has been achieved by adopting lifetime distributions that directly provide, for each instant of the operational life of the system, the probability that failure will occur before or after this instant.

A more powerful and detailed approach can be obtained by adopting stochastic processes. Under this approach, the probability of failure of the system is estimated by modeling its degradation process. In this context, failure is usually defined by the first passage time of the aforementioned degradation process to a preassigned threshold. It follows that the successful implementation of this approach is based on the ability of the stochastic process to accurately describe the degradation phenomenon under analysis.

A common drawback of the use of stochastic processes is the mathematical complexity that can quickly arise. In fact, formulating and calibrating these models often relies on computationally-heavy and time-intensive procedures that can easily run into convergence problems.

A substantial amount of literature has been devoted to the study of the application of stochastic processes in the field of degradation modeling and maintenance (see, for example, the extensive review in Van Noortwijk (2009)). In the remainder of this chapter, we will focus our attention on the particular stochastic processes that will be adopted in the rest of the thesis, as well as the procedures developed and adopted to calibrate the considered degradation models and perform failure-time predictions on the basis of real-time data. In particular, Chapter 2.2 gives some general background on stochastic processes, Chapter 2.3 introduces a gamma and a new inverse Gaussian processes that can

both account for measurement error, and Chapter 2.4 presents a new gamma process that, in addition to measurement error, also takes into account unit-to-unit variability. Then, the different numerical algorithms adopted to calibrate and perform predictions with the aforementioned processes are illustrated in Chapters 2.5, 2.6, 2.7, and 2.8: respectively, the formulation of the likelihood function, the remaining useful life, the EM algorithm, and the particle filter algorithm.

2.2 Background

Stochastic processes are a mathematical object extensively used to model a broad variety of phenomena characterized by randomness and uncertainty. They find application in numerous fields ranging from finance, physics, biology, and engineering. The term "random function" is also used to refer to this objects.
A stochastic process $\{W(t); t \geq 0\}$ is defined as a collection of random variables indexed by a parameter $t \in T$. The index t is often (but not necessarily, although it will be for the rest of the manuscript) interpreted as time. The set T is called the *index set* of the process. If the index set is a countable set, then the stochastic process is said to be a *discrete-time process*, whereas if T is some interval of the real number line, then $\{W(t); t \geq 0\}$ is said to be a *continuous-time* process. For any given value of the index $t \in T, W(t)$ is a random variable and is referred to as the *state* of the process at that time. The set of all possible values that the state $W(t)$ can assume is called the *state space*. If the state space is a discrete (continuous) set, then $\{W(t); t \geq 0\}$ is called a *discrete-state* (*continuous-state*) stochastic process.

Therefore, a stochastic process can be envisaged as a function that, to any time $t \in T$, associates a random variable that describes the state of the system under analysis at that time.
An *increment* of a stochastic process is the amount by which it changes between two time inputs. A *realization* is the function $\{w(y), y \leq t\}$ that describes the outcome (i.e., the *observed trajectory*, or *observed path*) of the process until time t, where $y, t \in T$. A classical example of a discrete-time, discrete-state process would be the process describing the earnings of a gambler that wins (loses) one when a fair coin is flipped heads (tails). In this case, the index set would be the set of non-negative integers, with t being the number of coin flips, whereas the state space would be the set of all rational numbers $\mathbb{Z}$. The state $W(t)$ would describe the current (i.e., after t coin flips) gambler's earnings. In the

remainder of this manuscript, we will focus out attention on continuous-time, continuous-state stochastic processes.

The function $F_{W(t)}(w) = P[W(t) \leq w]$ is called the *cumulative distribution function* of $W(t)$, and the set $\{F_{W(t)}(w), t \in T\}$ is the collection of one-dimensional cumulative distribution functions. The derivative of $F_{W(t)}(w)$, $f_{W(t)}(w)$, (if it exists) is called the one-dimensional *probability density function* of $W(t)$. Assigning the set $\{F_{W(t)}(w), t \in T\}$ does not suffice to completely define a stochastic process probabilistically. Indeed, the joint (n-dimensional) distribution function of all the n-tuple of random variables $W(t_1), W(t_2), \ldots, W(t_n)$, with $n = 1, 2, \ldots$ and $t_1 < t_2 < \ldots < t_n$ is required. Denoting by $\boldsymbol{W}(\boldsymbol{t}_n)$ the set $\boldsymbol{W}(\boldsymbol{t}_n) = \{W(t_1), W(t_2), \ldots, W(t_n)\}$ and by $\boldsymbol{w_n}$ the set $\boldsymbol{w_n} = \{w_1, w_2, \ldots, w_n\}$, the joint ($n$-dimensional) distribution function can be expressed as:

$$F_{\boldsymbol{W}(\boldsymbol{t}_n)}(\boldsymbol{w_n}) = P\left[\bigcap_{i=1}^{n} W(t_i) \leq w_i\right] \tag{2.1}$$

The n-order derivative of Eq. (2.1):

$$f_{\boldsymbol{W}(\boldsymbol{t}_n)}(\boldsymbol{w_n}) = \frac{\partial}{\partial w_1 \partial w_2 \ldots \partial w_n} F_{W(t_1), W(t_2), \ldots, W(t_n)}(w_1, w_2, \ldots, w_n) \tag{2.2}$$

if it exists, is called the *n-order probability density function* of $\{W(t); t \geq 0\}$.
The mean function of a stochastic process can be expressed as:

$$\mu_{W(t)} = E\{W(t)\} = \int_S w \cdot f_{W(t)}(w) \cdot dw \tag{2.3}$$

where S is the state space, while its variance is:

$$\sigma_{W(t)} = E\left\{\left(W(t) - \mu_{W(t)}\right)^2\right\} = \int_S \left(w - \mu_{W(t)}\right)^2 \cdot f_{W(t)}(w) \cdot dw \tag{2.4}$$

Eq. (2.2) can also be expressed in terms of conditional distributions as:

$$f_{\boldsymbol{W}(\boldsymbol{t}_n)}(\boldsymbol{w_n}) = f_{W(t_1)}(w_1) \cdot f_{W(t_2)|\boldsymbol{W}(\boldsymbol{t}_1)}(w_2|\boldsymbol{w_1}) \cdot \ldots \cdot f_{W(t_n)|\boldsymbol{W}(\boldsymbol{t}_{n-1})}(w_n|\boldsymbol{w_{n-1}}) \tag{2.5}$$

Direct assignment of the distributions in Eq. (2.1) is often impossible in practice. Indeed, it requires a very detailed knowledge of the underlying physical mechanisms that cause the degradation process. Nevertheless, in many applications some simplifying assumption make the task much easier, such as the *Markov property*, *homogeneity*, or *independence*

of increments.

Markov property

The Markov property consists in assuming that the future evolution of the state of the process only depends on its state at the current time. Formally, denoting the *history* up to t $H_t = \{W(y), 0 \leq y \leq t\}$ the Markov property consists in assuming that the state $W(t + \Delta t)$ of the process at time $t + \Delta t$ depends on the past history only through the value of its state at t.

In other words, a stochastic process is a *Markov process* if one can make prediction on its future state based solely on its current state $W(t)$ and discarding the *past history* $H_{t^-} = \{W(y), 0 \leq y < t\}$.

The Markov property can also be expressed by this equality of probabilities:

$$\begin{aligned} P\left[W(t+\Delta t) \in S | H_t = h_t\right] &= P\left[W(t+\Delta t) \in S | H_{t^-} = h_{t^-}, W(t) = w\right] \\ &= P\left[W(t+\Delta t) \in S | W(t) = w\right] \end{aligned} \tag{2.6}$$

where $h_t = \{w(y), 0 \leq y < t\}$ is the realization of H_t and $h_{t^-} = \{w(y), 0 \leq y < t\}$ is the realization of H_{t^-}. Noting that Eq. (2.6) can also be expressed as:

$$f_{W(t_n)|\boldsymbol{W}(\boldsymbol{t}_{n-1})}(w_n|\boldsymbol{w_{n-1}}) = f_{W(t_n)|W(t_{n-1})}(w_n|w_{n-1}),$$

under a Markov process, the joint pdf in Eq. (2.2) can be simplified to:

$$\begin{aligned} f_{\boldsymbol{W}(\boldsymbol{t}_n)}(\boldsymbol{w_n}) = f_{W(t_1)} \cdot f_{W(t_2)|W(t_1)}(w_2|w_1) \cdot \ldots \cdot f_{W(t_{n-1})|W(t_{n-2})}(w_{n-1}|w_{n-2}) \times \\ \times f_{W(t_n)|W(t_{n-1})}(w_n|w_{n-1}) \end{aligned} \tag{2.7}$$

Homogeneity

Let $\Delta W(t, t + \Delta t) = W(t + \Delta t) - W(t)$ denote the increment of a stochastic process. Then, a Markov process is called *homogeneous* if the conditional pdf $f_{\Delta W(t,t+\Delta t)|W(t)}(\delta|w_t)$ of the increment $\Delta W(t, t + \Delta t)$ given $W(t) = w_t$ depends solely on the value of Δt and not on the specific value of t. This is equivalent to assuming that $f_{\Delta W(t,t+\Delta t)|W(t)}(\delta|w_t) = f_{W(0,\Delta t)|W(t)}(\delta|w_t)$. A process that does not satisfy this property is called *non-homogeneous*.

Independence of increments

A stochastic process has *independent increments* if the conditional distribution $f_{\Delta W(t,t+\Delta t)|W(t)}(\delta|w_t)$ of the increment $\Delta W(t, t + \Delta t)$ coincides with the unconditional

one $f_{\Delta W(t,t+\Delta t)}(\delta)$. If this property is satisfied, then the pdf in Eq. (2.2) simplifies to:

$$f_{\boldsymbol{W}(\boldsymbol{t_n})}(\boldsymbol{w_n}) = f_{W(t_1)}(\delta_1) \cdot f_{\Delta W(t_1,t_2)}(\delta_2) \cdot \ldots \cdot f_{\Delta W(t_{n-2},t_{n-1})}(\delta_{n-1}) \cdot f_{\Delta W(t_{n-1},t_n)}(\delta_n) \quad (2.8)$$

Conceptually, this means that the distribution of the increment between time t and $t+\Delta t$ is not influenced by the value of the process at time t.

2.2.1 Definition of a perturbed stochastic process

In the experimental scenarios where measurements are collected via in-service, non-destructive, or even indirect inspection methods, it is common to obtain measurements which are contaminated by random errors. Consequently, adopting a degradation process that is able to also take these errors into account may lead to better modeling performance. In the literature, this situation is commonly tackled with a *perturbed stochastic process*, denoted by $\{Z(t); t \geq 0\}$ and formulated as:

$$\{Z(t) = W(t) + \varepsilon(t);\ t \geq 0\}, \quad (2.9)$$

where $\{W(t); t \geq 0\}$ is the *hidden* (also *actual*, or *true*) degradation process and $\varepsilon(t)$ is a *perturbing* (also *error*) term, here intended as measurement error. To completely define the stochastic process $\{Z(t); t \geq 0\}$ it is necessary to specify the hidden process $\{W(t); t \geq 0\}$, the perturbing term $\varepsilon(t)$ and their mutual stochastic relationship.

In the overwhelming majority of the literature, the perturbing term is assumed to be Gaussian-distributed and independent of the hidden degradation process (e.g., see Lu, Pandey, and W.-C. Xie (2013), Le Son, Fouladirad, and Barros (2016), Bordes, Paroissin, and Salami (2016), S. Hao, J. Yang, and Bérenguer (2019b), X. Chen et al. (2019), Kallen and Van Noortwijk (2005), Whitmore (1995)), which usually leads to more easily mathematically tractable formulations. Some other authors (Zhai and Ye (2017) have also experimented with non-Gaussian based (but still independent of the hidden process) formulations. However, in some applications it might be interesting to assume that measurement error is stochastically dependent on the actual degradation. To this aim, Pulcini (2016) modeled measurement error via a state-dependent Gaussian random variables. Similar solutions are also used by Sun et al. (2021) and Oumouni and Schoefs (2021). Finally, some non-Gaussian state-dependent formulations, such as in Giorgio, Mele, and Pulcini (2019) have also been investigated.

While considering non-Gaussian and/or state-dependent solutions can lead to more accurate modeling, it also leads to increased mathematical complexity and the need for ad-hoc model calibration algorithms.

In the remainder of this manuscript, the following (common) assumption will be made regarding the relationship between $\varepsilon(t)$ and $W(t)$:

— for any $n > 1$, the measurement error $\varepsilon(t_j)$, given $W(t_j)$, is conditionally independent both of $\varepsilon(t_k)$ and $W(t_k)$ $\forall k \neq j$ $(j, k = 1, \ldots, n)$. Thus, equivalently, the perturbed observation $Z(t_j)$ given $W(t_j)$ is conditionally independent both of $Z(t_k)$ and $W(t_k)$ $\forall k \neq j$, $(j, k = 1, \ldots, n)$.

2.3 The perturbed gamma and the perturbed inverse Gaussian processes

In this manuscript, the perturbed gamma and perturbed inverse Gaussian processes are formulated as (see also Esposito, Mele, et al. (2023b)):

$$\{Z(t) = W(t) + \varepsilon(t);\ t \geq 0\}. \tag{2.10}$$

where:

(i) the hidden process might be either a gamma process or an inverse Gaussian one. Both processes have independent increments and hence can be fully defined by an initial condition (here $W(0) = 0$) and the probability density function (pdf) of its generic increment $\Delta W(t, t + \Delta t) = W(t + \Delta t) - W(t)$. In the case of the gamma process, the pdf of $\Delta W(t, t + \Delta t)$ can be expressed as:

$$f_{\Delta W(t,t+\Delta t)}(\delta) = \frac{\delta^{\Delta\eta(t,t+\Delta t)-1}}{\Gamma[\Delta\eta(t, t + \Delta t)] \cdot \theta^{\Delta\eta(t,t+\Delta t)}} \cdot e^{-\frac{\delta}{\theta}}, \qquad \delta \geq 0, \tag{2.11}$$

where $\Gamma(y) = \int_0^\infty u^{a-1} \cdot e^{-u} du$ is the complete gamma function, θ ($\theta > 0$) is the scale parameter, $\eta(t)$ is a non-negative monotonic increasing function, here referred to as the age function, and $\Delta\eta(t, t + \Delta t) = \eta(t + \Delta t) - \eta(t)$.
Whereas in the case of the inverse Gaussian process, the pdf of $\Delta W(t, t + \Delta t)$ can

be expressed by using the following special parameterization:

$$f_{\Delta W(t,t+\Delta t)}(\delta) = \frac{\Delta\eta(t,t+\Delta t)}{\sqrt{2\cdot\pi\cdot\theta^{-1}\cdot\delta^3}}\cdot e^{\frac{(\delta-\theta\cdot\Delta\eta(t,t+\Delta t))^2}{2\cdot\pi\cdot\delta}}, \qquad \delta\geq 0, \tag{2.12}$$

which can be easily obtained from the classical functional form of the pdf of the increment of the inverse Gaussian process C.-Y. Peng (2015):

$$f_{\Delta W(t,t+\Delta t)}(\delta) = \sqrt{\frac{\omega\cdot[\Delta\Omega(t,t+\Delta t)]^2}{2\pi\delta^3}}\cdot e^{-\frac{\omega\cdot[\delta-\mu\cdot\Delta\Omega(t,t+\Delta t)]^2}{2\mu^2\delta}} \tag{2.13}$$

where $\Omega(t)$ is a non-negative monotonic increasing function, $\Delta\Omega(t,t+\Delta t)=\Omega(t+\Delta t)-\Omega(t)$, and μ and ω are positive valued parameters, by setting $\mu^2/\omega=\theta$ and $\mu\cdot\Delta\Omega(t,t+\Delta t)=\Delta\eta(t,t+\Delta t)$. The main advantage of adopting the special form in Eq. (2.12) is that it allows the hidden gamma and the hidden inverse Gaussian process to share the same parameters and the same functional form of the mean $V\{W(t)\}$ and variance $V\{Z(t)\}$ functions. These latter functions can be formulated as:

$$E\{W(t)\} = \eta(t)\cdot\theta \tag{2.14}$$

$$V\{W(t)\} = \eta(t)\cdot\theta^2 \tag{2.15}$$

whereas under the classical parameterization of Eq. (2.13) the mean and variance functions of the hidden inverse Gaussian process would have been equal to $V\{W(t)\}=\mu\cdot\Delta\Omega(t,t+\Delta t)$ and $V\{W(t)\}=[\mu\cdot\Delta\Omega(t,t+\Delta t)]^3/\omega$, respectively.

From Eq. (2.11) the cumulative distribution function (cdf) of the increment $\Delta W(t,t+\Delta t)$ of the hidden gamma process is:

$$F_{\Delta W(t,t+\Delta t)}(\delta) = \frac{\gamma\left[\Delta\eta(t,t+\Delta t),\frac{\delta}{\theta}\right]}{\Gamma[\Delta\eta(t,t+\Delta t)]}, \qquad \delta\geq 0. \tag{2.16}$$

where $\gamma(\cdot)$ is the lower incomplete gamma function. Similarly, from (2.12), under

the hidden inverse Gaussian process the cdf can be expressed as:

$$F_{\Delta W(t,t+\Delta t)}(\delta) = \Phi\left[\frac{\delta - \theta\cdot\Delta\eta(t,t+\Delta t)}{\sqrt{\theta\cdot\delta}}\right] + e^{\theta\cdot\Delta\eta(t,t+\Delta t)}\cdot\Phi\left[-\frac{\delta+\theta\cdot\Delta\eta(t,t+\Delta t)}{\sqrt{\theta\cdot\delta}}\right] \quad (2.17)$$

where $\Phi(\cdot)$ is the standard normal cdf.
To complete the definition of the hidden models it is necessary to provide the functional form of the age function. In the remainder of this manuscript, the very flexible and largely adopted power-law expression $\eta(t) = (t/a)^b$ is used.

(ii) To model the perturbing term $\varepsilon(t)$, in this manuscript two options are considered:

1) following Giorgio, Mele, and Pulcini (2019), $\varepsilon(t)$ is assumed to depend, in stochastic sense, on the hidden degradation level $W(t)$ and that, given $W(t)$, is conditionally distributed as an inverse gamma random variable whose pdf can be expressed as:

$$f_{\varepsilon(t)|W(t)}(\varepsilon|w) = \frac{(\alpha(w))^{\beta(w)}\cdot(\varepsilon+w)^{-\beta(w)-1}}{\Gamma(\beta(w))}\cdot e^{-\frac{\alpha(w)}{\varepsilon+w}}, \quad \varepsilon \geq -w \quad (2.18)$$

where $\beta(w) = \varphi\cdot w^{2-\nu} + 2,\ \varphi > 0, -\infty < \nu < \infty$, and $\alpha(w) = (\beta(w)-1)\cdot w$. This modeling option ensures that the resulting perturbed measurement $Z(t) = W(t) + \varepsilon(t)$ is non-negative (indeed, this feature that the hidden gamma and inverse Gaussian processes possess is preserved also in the perturbed models) and enables to model scenarios where the measurement error depends (in stochastic sense) on the measured degradation level.
Under this setting, the conditional mean and variance of $\varepsilon(t)$, given $W(t) = w$, can be formulated as:

$$E\{\varepsilon(t)|W(t)=w\} = \frac{\alpha(w)}{\beta(w)-1} - w = 0, \quad (2.19)$$

$$V\{\varepsilon(t)|W(t)=w\} = \frac{(\alpha(w))^2}{(\beta(w)-1)^2\cdot(\beta(w)-2)} = \frac{w^\nu}{\varphi}. \quad (2.20)$$

Therefore, from Eq. (2.10) the observed perturbed measurement $Z(t)$, given

$W(t) = w$, has conditional pdf:

$$f_{Z(t)|W(t)}(\varepsilon|w) = \frac{(\alpha(w))^{\beta(w)} \cdot z^{-\beta(w)-1}}{\Gamma(\beta(w))} \cdot e^{-\frac{\alpha(w)}{\varepsilon+w}}, \quad z \geq 0, \tag{2.21}$$

conditional mean:

$$E\{Z(t)|W(t) = w\} = E\{\varepsilon(t)|W(t) = w\} + w = w, \tag{2.22}$$

and conditional variance:

$$V\{Z(t)|W(t) = w\} = V\{\varepsilon(t)|W(t) = w\} = \frac{w^{\nu}}{\varphi}. \tag{2.23}$$

2) inspired by Pulcini (2016), $\varepsilon(t)$ is assumed to be Gaussian-distributed, with zero mean and variance that depends on $W(t)$. In this case, the conditional pdf of $\varepsilon(t)$ can be expressed as:

$$f_{\varepsilon(t)|W(t)}(\varepsilon|w) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2(w)}} \cdot e^{-\frac{1}{2} \cdot \frac{\varepsilon^2}{\sigma^2(w)}}, \qquad -\infty < \varepsilon < +\infty, \tag{2.24}$$

where $\sigma^2(w) = w^{\nu}/\varphi$, $\varphi > 0$, and $-\infty < \nu < +\infty$.
Consequently, under option 2) $\varepsilon(t)$, given $W(t) = w$, has the same conditional mean and variance as under option 1):

$$E\{\varepsilon(t)|W(t) = w\} = 0, \tag{2.25}$$

$$V\{\varepsilon(t)|W(t) = w\} = \sigma^2(w) = \frac{w^{\nu}}{\varphi}. \tag{2.26}$$

Hence, from Eq. (2.10) the observed perturbed measurement $Z(t)$, given $W(t) = w$, has conditional pdf:

$$f_{Z(t)|W(t)}(z|w) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2(w)}} \cdot e^{-\frac{1}{2} \cdot \frac{(z-w)^2}{\sigma^2(w)}}, \qquad -\infty < z < +\infty, \tag{2.27}$$

conditional mean:

$$E\{Z(t)|W(t) = w\} = E\{\varepsilon(t)|W(t) = w\} + w = w, \tag{2.28}$$

and conditional variance:

$$V\{Z(t)|W(t)=w\} = V\{\varepsilon(t)|W(t)=w\} = \frac{w^{\nu}}{\varphi}. \tag{2.29}$$

This specific formulation has been chosen so that $\varepsilon(t)$ and $Z(t)$, given $W(t) = w$, have the same conditional mean and variance that they have under option 1). A perturbed gamma process that adopts this modeling solution for the perturbing term has been proposed in Pulcini (2016). Sun et al. (2021) adopted it to formulate a perturbed inverse Gaussian process.

Under these assumptions, from Eqs. (2.14), (2.19), (2.22), (2.25), and (2.28), via the law of total mean, the marginal means of $\varepsilon(t)$ and $Z(t)$, under all considered processes (i.e., under both hidden processes and under both error modeling options 1) and 2)), can be expressed as:

$$E\{\varepsilon(t)\} = E\{E\{\varepsilon(t)|W(t)\}\} = E\{0\} = 0, \tag{2.30}$$

and:

$$E\{Z(t)\} = E\{E\{Z(t)|W(t)\}\} = E\{W(t)\} = \theta \cdot \eta(t). \tag{2.31}$$

Similarly, by using the law of total variance, from Eqs. (2.15), (2.20), (2.22), (2.23), (2.26), (2.28), and (2.29), under both modeling option 1) and 2), the marginal variances of $\varepsilon(t)$ and $Z(t)$ are equal to:

$$V\{\varepsilon(t)\} = V\{E\{\varepsilon(t)|W(t)\}\} + E\{V\{\varepsilon(t)|W(t)\}\} = V\{0\} + \frac{E\{(W(t))^{\nu}\}}{\varphi} \tag{2.32}$$

and:

$$V\{Z(t)\} = V\{E\{Z(t)|W(t)\}\} + E\{V\{Z(t)|W(t)\}\} = V\{W(t)\} + \frac{E\{(W(t))^{\nu}\}}{\varphi}, \tag{2.33}$$

where the fractal moment $E\{(W(t))^{\nu}\}$ (in general) depends on the hidden process. In the case of the gamma process it can be expressed as;

$$E\{(W(t))^{\nu}\} = \theta^{\nu} \cdot \frac{\Gamma(\eta(t)+\nu)}{\Gamma(\eta(t))}, \tag{2.34}$$

whereas in the case of the inverse Gaussian process it is given by this integral:

$$E\{(W(t))^{\nu}\} = \frac{\eta(t)}{\sqrt{2 \cdot \pi \cdot \theta^{-1}}} \cdot \int_0^{\infty} w^{\nu - \frac{3}{2}} \cdot e^{-\frac{[w - \theta \cdot \eta(t)]^2}{2 \cdot \theta \cdot w}} \cdot dw \tag{2.35}$$

that (in general) is not available in closed form. Nevertheless, in case $\nu = 0$, 1, or 2, $V\{\varepsilon(t)\}$ and $V\{Z(t)\}$ have simple expressions and do not depend on the hidden model. Specifically, if $\nu = 0$ it is, under both processes:

$$V\{\varepsilon(t)\} = \frac{1}{\varphi}$$

and:

$$V\{Z(t)\} = \theta^2 \cdot \eta(t) \cdot \frac{1}{\varphi}.$$

Likewise, if $\nu = 1$ it is:

$$V\{\varepsilon(t)\} = \frac{\theta \cdot \eta(t)}{\varphi}$$

and:

$$V\{Z(t)\} = \theta^2 \cdot \eta(t) + \frac{\theta \cdot \eta(t)}{\varphi}.$$

In addition, in this case the ratio between $V\{Z(t)\}$ and $V\{W(t)\}$ does not depend on t:

$$\frac{V\{Z(t)\}}{V\{W(t)\}} = \frac{\theta^2 \cdot \eta(t) + \frac{\theta \cdot \eta(t)}{\varphi}}{\theta^2 \cdot \varphi} = 1 + \frac{1}{\theta \cdot \varphi}. \tag{2.36}$$

Again, by using similar arguments, if $\nu = 2$, noting that under the considered special parameterization the two hidden processes share the same mean and variance functions (and hence second moment) it is:

$$E\left\{(W(t))^2\right\} = \theta^2 \cdot \eta(t) \cdot (\eta(t) + 1)$$

and therefore:

$$V\{\varepsilon(t)\} = \theta^2 \cdot \frac{\eta(t) \cdot (\eta(t) + 1)}{\varphi}$$

and:

$$V\{Z(t)\} = \theta^2 \cdot \left[\eta(t) + \frac{\eta(t) \cdot (\eta(t) + 1)}{\varphi}\right].$$

Under the current setup, it is also possible to express the distribution of $F_{W(t_1)}(w_1)$ conditional to $W(t_2) = w_2$, where t_1 and t_2 be two generic reference times, with $t_1 < t_2$. In fact, being:

$$\begin{aligned} F_{W(t_1)|W(t_2)}(w_1|w_2) &= \int_0^{w_1} f_{W(t_1)|W(t_2)}(w|w_2) \cdot dw \\ &= \int_0^{w_1} \frac{f_{W(t_2)|W(t_1)}(w_2|w) \cdot f_{W(t_1)}(w)}{f_{W(t_2)}(w_2)} \cdot dw \\ &= \int_0^{w_1} \frac{f_{\Delta W(t_1,t_2)|W(t_1)}(w_2 - w|w) \cdot f_{W(t_1)}(w)}{f_{W(t_2)}(w_2)} \cdot dw \end{aligned}$$

from Eqs. (2.11) and (2.12), under the gamma process it is:

$$F_{W(t_1)|W(t_2)}(w_1|w_2) = \mathbb{B}\left(\frac{w_1}{w_2}; \eta(t_1), \Delta\eta(t_1, t_2)\right) \tag{2.37}$$

whereas under the inverse Gaussian it can be computed by numerically solving the integral:

$$\begin{aligned} &F_{W(t_1)|W(t_2)}(w_1|w_2) \\ &= \int_0^{w_1} \frac{\Delta\eta(t_1, t_2) \cdot \eta(t_1)}{\eta(t_1)} \cdot \sqrt{\frac{\theta}{2 \cdot \pi} \cdot \left[\frac{w_2}{w_1 \cdot (w_2 - w_1)}\right]^3} \cdot e^{-\frac{\theta \cdot [w_2 \cdot \eta(t_1) - w_1 \cdot \eta(t_2)]^2}{2 \cdot w_1 \cdot w_2 \cdot (w_2 - w_1)}} \cdot dw. \end{aligned} \tag{2.38}$$

Finally, it is worth to remark that, under the considered settings:

— the perturbed process $\{Z(t); t \geq 0\}$ is non-Markovian.

— under error model 1), when $\nu = 0$, despite the conditional mean and variance of $\varepsilon(t)$ not depending on $W(t)$, $\varepsilon(t)$ still stochastically depends on $W(t)$ because its support depends on it (indeed, from Eq. (2.21), it must be $\varepsilon \geq -w$).

— under error model 2), when $\nu = 0$, $\varepsilon(t)$ is stochastically independent of $W(t)$. Hence, in this case, error model 2) reduces to the classical model adopted in the bulk of existing literature on perturbed degradation modeling (see Lu, Pandey, and W.-C. Xie (2013), Le Son, Fouladirad, and Barros (2016), Bordes, Paroissin, and Salami (2016), S. Hao, J. Yang, and Bérenguer (2019a), S. Hao, J. Yang, and Bérenguer (2019b), and X. Chen et al. (2019)).

2.4 The perturbed gamma process with random effect

The considered perturbed gamma process with random effect is formulated as (see also Esposito, Mele, et al. (2022)):

$$\{Z(t) = W(t) + \varepsilon(t); t \geq 0\} \tag{2.39}$$

where:

(i) $\{W(t); t \geq 0\}$ is the non-homogeneous gamma process with random effect originally suggested in Lawless and Crowder (2004). As in Giorgio, Mele, and Pulcini (2019), given a unit randomly chosen from a heterogeneous population, its hidden degradation process is described by a gamma process. Yet, Lawless and Crowder (2004) assume that the rate parameter of the unit-specific gamma process varies randomly from unit to unit.
Given the (unknown) value λ of the rate parameter Λ, the pdf of the degradation increment $\Delta W(t, t+\Delta t) = W(t+\Delta t) - W(t)$ of a generic unit is expressed as:

$$f_{\Delta W(t,t+\Delta t)|\Lambda}(\delta|\lambda) = \frac{\lambda^{\Delta\eta(t,t+\Delta t)} \cdot \delta^{\Delta\eta(t,t+\Delta t)-1}}{\Gamma\left(\Delta\eta(t,t+\Delta t)\right)} \cdot e^{-\lambda\cdot\delta}, \qquad \delta \geq 0. \tag{2.40}$$

The rate parameter Λ is assumed to have the gamma pdf:

$$g_{\Lambda}(\lambda) = \frac{c^d \cdot \lambda^{d-1}}{\Gamma(d)} \cdot e^{-c\cdot\lambda}, \quad \lambda, c, d > 0 \tag{2.41}$$

(ii) The perturbing term is modeled by adopting the inverse gamma random variable presented in Eq. (2.18).
In addition, it is assumed that $\varepsilon(t)$, given $W(t)$ is conditionally independent on the parameter Λ. Under this setup, the conditional mean of $\varepsilon(t)$ is:

$$E\{\varepsilon(t)|W(t) = w\} = \frac{\alpha(w)}{\beta(w) - 1} - w = 0$$

and conditional variance equal to:

$$V\{\varepsilon(t)|W(t) = w\} = \frac{(\alpha(w))^2}{(\beta(w) - 1)^2 \cdot (\beta(w) - 2)} = \frac{w^{\nu}}{\varphi}.$$

One of the main advantages of this model is that the hidden marginal process $\{W(t); t \geq 0\}$ is still Markovian and that the conditional pdf of its increment can be expressed in closed form. In fact, the conditional pdfs of $\Delta W(t_l, t_l + \Delta t)$ and Λ, given $W(t_1) = w_1, \ldots, W(t_l) = w_l$, can be expressed as in Eqs. (2.42)-(2.43), respectively:

$$\begin{aligned}
&f_{\Delta W(t_l,t_l+\Delta t)|W(t_1),\ldots,W(t_l)}(\delta|w_1,\ldots,w_l) \\
&= \frac{\int_0^\infty f_{\Delta W(t_l,t_l+\Delta t)|\Lambda}(\delta|\lambda)\cdot\prod_{j=1}^{l} F_{\Delta W(t_{j-1},t_j)|\Lambda}(w_j - w_{j-1}|\lambda)\cdot g_\Lambda(\lambda)\cdot d\lambda}{\int_0^\infty \prod_{j=1}^{l} f_{\Delta W(t_{j-1},t_j)|\Lambda}(w_j - w_{j-1}|\lambda)\cdot g_\Lambda(\lambda)\cdot d\lambda} \\
&= \frac{\int_0^\infty f_{\Delta W(t_l,t_l+\Delta t)|\Lambda}(\delta|\lambda)\cdot f_{W(t_1),\ldots,W(t_l)|\Lambda}(w_1,\ldots,w_l|\lambda)\cdot g_\Lambda(\lambda)\cdot d\lambda}{f_{W(t_1),\ldots,W(t_l)}(w_1,\ldots,w_l)} \\
&= \frac{\int_0^\infty f_{\Delta W(t_l,t_l+\Delta t)|\Lambda}(\delta|\lambda)\cdot g_{\Lambda|W(t_1),\ldots,W(t_l)}(\lambda|w_1,\ldots,w_l)\cdot f_{W(t_1),\ldots,W(t_l)}(w_1,\ldots,w_l)\cdot d\lambda}{f_{W(t_1),\ldots,W(t_l)}(w_1,\ldots,w_l)} \\
&= \int_0^\infty f_{\Delta W(t_l,t_l+\Delta t)|\Lambda}(\delta|\lambda)\cdot g_{\Lambda|W(t_1),\ldots,W(t_l)}(\lambda|w_1,\ldots,w_l)\cdot d\lambda,
\end{aligned} \tag{2.42}$$

$$\begin{aligned}
g_{\Lambda|W(t_1),\ldots,W(t_l)}(\lambda|w_1,\ldots,w_l) &= \frac{\prod_{j=1}^{l} f_{\Delta W(t_{j-1},t_j)|\Lambda}(w_j - w_{j-1}|\lambda)\cdot g_\Lambda(\lambda)}{\int_0^\infty \prod_{j=1}^{l} f_{\Delta W(t_{j-1},t_j)|\Lambda}(w_j - w_{j-1}|\lambda)\cdot g_\Lambda(\lambda)\cdot d\lambda} \\
&= \frac{\prod_{j=1}^{l} \frac{\lambda^{\eta(t_{j-1},t_j)}\cdot(w_j-w_{j-1})^{\eta(t_{j-1},t_j)-1}}{\Gamma[\eta(t_{j-1},t_j)]}\cdot e^{-\lambda\cdot(w_j-w_{j-1})}\cdot\frac{c^d\cdot\lambda^{d-1}}{\Gamma(d)}\cdot e^{-c\cdot\lambda}}{\int_0^\infty \prod_{j=1}^{l} \frac{\lambda^{\eta(t_{j-1},t_j)}\cdot(w_j-w_{j-1})^{\eta(t_{j-1},t_j)-1}}{\Gamma[\eta(t_{j-1},t_j)]}\cdot e^{-\lambda\cdot(w_j-w_{j-1})}\cdot\frac{c^d\cdot\lambda^{d-1}}{\Gamma(d)}\cdot e^{-c\cdot\lambda}\cdot d\lambda} \\
&= \frac{\lambda^{\eta(t_l)+d-1}\cdot e^{-(w_l+c)\cdot\lambda}}{\int_0^\infty \lambda^{\eta(t_l)+d-1}\cdot e^{-(w_l+c)\cdot\lambda}\cdot d\lambda} \\
&= \frac{(c+w_l)^{\eta(t_l)+d}\cdot\lambda^{\eta(t_l)+d-1}\cdot e^{-(w_l+c)\cdot\lambda}}{\Gamma[\eta(t_l)+d]},
\end{aligned} \tag{2.43}$$

where $t_0 = w_0 = 0$, $t_1 < \ldots < t_l$, and $w_1 < \ldots < w_l$. Eq. (2.42) also shows that $\Delta W(t_l, t_l + \Delta t)$ depends on the history $H_t = \{W(t_1), \ldots, W(t_l)\}$ only through $g_{\Lambda|W(t_1),\ldots,W(t_l)}(\lambda|w_1,\ldots,w_l)$, which, in turn (as shown by Eq. (2.43)), depends on H_t only through the value of $W(t_l)$. Therefore, it is clear that, given the value $W(t_l)$ of the degradation at time t_l, the distribution of the increment $\Delta W(t_l, t_l + \Delta t)$ is conditionally independent of the past history $H_{t^-} = \{W(t_1), \ldots, W(t_{l-1})\}$. It is worth noticing that Eq. (2.43) also shows that, given $W(t_l) = w_l$, Λ is still conditionally gamma distributed, with rate parameter $w_l + c$ and shape parameter $\eta(t_l) + d$.

Consequently, from Eqs. (2.40), (2.42), and (2.43), the conditional pdf of the increment $\Delta W(t, t+\Delta t)$, given $W(t) = w$, can be expressed as:

$$\begin{aligned} f_{\Delta W(t,t+\Delta t)|W(t)}(\delta|w) &= \int_0^\infty f_{\Delta W(t,t+\Delta t)|\Lambda}(\delta|\lambda) \cdot g_{\Lambda|W(t)}(\lambda|w) \cdot d\lambda \\ &= \int_0^\infty \frac{\lambda^{\Delta\eta(t,t+\Delta t)} \cdot \delta^{\Delta\eta(t,t+\Delta t)-1} \cdot e^{-\lambda\delta}}{\Gamma\left[\Delta\eta(t,t+\Delta t)\right]} \cdot \frac{(w+c)^{\eta(t)+d} \cdot \lambda^{\eta(t)+d-1} \cdot e^{-(w+c)\cdot\lambda}}{\Gamma\left[\eta(t)+d\right]} \cdot d\lambda \\ &= \frac{(w+c)^{\eta(t)+d} \cdot \delta^{\Delta\eta(t,t+\Delta t)-1}}{\Gamma[\eta(t)] \cdot \Gamma[d]} \cdot \int_0^\infty \lambda^{\eta(t+\Delta t)+d-1} \cdot e^{-(w+c)\cdot\lambda} \cdot d\lambda \\ &= \frac{1}{B\left[\Delta\eta(t,t+\Delta t), \eta(t)+d\right]} \cdot \frac{(w+c)^{\eta(t)+d} \cdot \delta^{\Delta\eta(t,t+\Delta t)-1}}{(w+\delta+c)^{\eta(t+\Delta t)+d}} \end{aligned} \tag{2.44}$$

Then, from Eq. (2.44), the conditional cdf of $\Delta W(t, t+\Delta t)$, given $W(t) = w$, can be expressed as:

$$F_{\Delta W(t,t+\Delta t)|W(t)}(\delta|w) = \mathbb{B}\left[\frac{\delta}{\delta+w+c}; \Delta\eta(t,t+\Delta t), \eta(t)+d\right] \tag{2.45}$$

where:

$$\mathbb{B}(y;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} \cdot \int_0^y u^{\alpha-1} \cdot (1-u)^{\beta-1} \cdot du \tag{2.46}$$

is the regularized beta function and:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

is the beta function.

Eq. (2.44), together with the initial condition (here $W(0) = 0$) and the functional form of the age function (the power-law function $\eta(t) = (t/a)^b$) fully specifies the hidden model $\{W(t); t \geq 0\}$. From Eqs. (2.44) and (2.45) it also follows that the conditional pdf and cdf of $\Delta W(t, t+\Delta t)$, given $W(t) = w$, can be expressed as:

$$\begin{aligned} f_{W(t+\Delta t)|W(t)}(\delta+w|w) = &\frac{1}{B\left[\Delta\eta(t,t+\Delta t), \eta(t)+d\right]} \\ &\times \frac{(w+c)^{\eta(t)+d} \cdot [(w+\delta)-w]^{\Delta\eta(t,t+\Delta t)-1}}{(w+\delta+c)^{\eta(t+\Delta t)+d}} \end{aligned} \tag{2.47}$$

and:

$$F_{W(t+\Delta t)|W(t)}(\delta + w|w) = \mathbb{B}\left[\frac{(\delta + w) - w}{(\delta + w) + c}; \Delta\eta(t, t + \Delta t), \eta(t) + d\right]. \tag{2.48}$$

Finally, from Eq. (2.43), it results that $\{W(t); t \geq 0\}$ has marginal pdf:

$$f_{W(t)}(w) = \frac{c^d \cdot w^{\eta(t)-1}}{B(\eta(t), d) \cdot (w + c)^{\eta(t)+d}}, \quad w \geq 0, \tag{2.49}$$

marginal cdf:

$$F_{W(t)}(w) = \mathbb{B}\left(\frac{w}{w + c}; \eta(t), d\right), \quad w \geq 0, \tag{2.50}$$

mean (that exists for $d > 1$):

$$E\{W(t)\} = \frac{c \cdot \eta(t)}{d - 1}, \tag{2.51}$$

and variance (that exists for $d > 2$):

$$V\{W(t)\} = \frac{c^2 \cdot \eta(t) \cdot [\eta(t) + d - 1]}{(d - 1)^2 \cdot (d - 2)}. \tag{2.52}$$

Let t_1 and t_2 be two generic reference times, with $t_1 < t_2$. Under the current setup, it is also possible to express in closed form the distribution of $F_{W(t_1)}(w_1)$ conditional to $W(t_2) = w_2$. In fact, being

$$\begin{aligned} F_{W(t_1)|W(t_2)}(w_1|w_2) &= \int_0^{w_1} f_{W(t_1)|W(t_2)}(w|w_2) \cdot dw \\ &= \int_0^{w_1} \frac{f_{W(t_2)|W(t_1)}(w_2|w) \cdot f_{W(t_1)}(w)}{f_{W(t_2)}(w_2)} \cdot dw \\ &= \int_0^{w_1} \frac{f_{\Delta W(t_1,t_2)|W(t_1)}(w_2 - w|w) \cdot f_{W(t_1)}(w)}{f_{W(t_2)}(w_2)} \cdot dw \end{aligned}$$

from Eqs. (2.46), (2.47), and (2.49):

$$
\begin{aligned}
F_{W(t_1)|W(t_2)}(w_1|w_2) &= \\
&= \int_0^{w_1} \frac{\frac{(w+c)^{\eta(t_1)+d}\cdot(w_2-w)^{\Delta\eta(t_1,t_2)-1}}{B(\Delta\eta(t_1,t_2),\eta(t_1)+d)\cdot(w_2+c)^{\eta(t_2)+d}} \cdot \frac{c^d\cdot w^{\eta(t_1)-1}}{B(\eta(t_1),d)\cdot(w+c)^{\eta(t_1)+d}}}{\frac{c^d\cdot w_2^{\eta(t_2)-1}}{B(\eta(t_2),d)\cdot(w_2+c)^{\eta(t_2)+d}}} \cdot dw \\
&= \frac{B(\eta(t_2),d)}{B(\Delta\eta(t_1,t_2),\eta(t_1)+d)\cdot B(\eta(t_1),d)} \cdot \int_0^{w_1} \frac{(w_2-w)^{\Delta\eta(t_1,t_2)-1}\cdot w^{\eta(t_1)-1}}{w_2^{\eta(t_2)-1}} \cdot dw \\
&= \frac{1}{B(\eta(t_1),\Delta\eta(t_1,t_2)} \cdot \int_0^{w_1} \left(\frac{w}{w_2}\right)^{\eta(x)-1} \cdot \left(1-\frac{w}{w_2}\right)^{\Delta\eta(t_1,t_2)-1} \cdot \frac{1}{w_2} \cdot dw \\
&= \frac{1}{B(\eta(t_1),\Delta\eta(t_1,t_2)} \cdot \int_0^{w_1/w_2} y^{\eta(t_1)-1}\cdot(1-y)^{\Delta\eta(t_1,t_2)-1}\cdot dy \\
&= \mathbb{B}\left(\frac{w_1}{w_2};\eta(t_1),\Delta\eta(t_1,t_2)\right), \qquad (2.53)
\end{aligned}
$$

Concerning the perturbed process, from Eqs. (2.40) and (2.21), by variable transformation, $Z(t)$, given $W(t) = w$, is distributed as an inverse gamma random variable with conditional pdf:

$$
f_{Z(t)|W(t)}(z|w) = \frac{(\alpha(w))^{\beta(w)}\cdot z^{\beta(w)-1}}{\Gamma(\beta(w))}\cdot e^{-\frac{\alpha(w)}{z}}, \qquad z \geq 0, \qquad (2.54)
$$

conditional cdf:

$$
F_{Z(t)|W(t)}(z|w) = \frac{\Gamma(\beta(w),\alpha(w)/z)}{\Gamma(\beta(w))}, \qquad z \geq 0, \qquad (2.55)
$$

conditional mean:

$$
E\{Z(t)|W(t)=w\} = E\{\varepsilon(t)|W(t)=w\} + w = w, \qquad (2.56)
$$

and conditional variance:

$$
V\{Z(t)|W(t)=w\} = V\{\varepsilon(t)|W(t)=w\} = \frac{w^\nu}{\varphi}, \qquad (2.57)
$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function. The marginal distribution of $Z(t)$ is not available in closed form. However, from Eqs. (2.49), (2.54), and (2.55), the marginal

pdf and cdf of $Z(t)$ can be numerically computed via the following integrals:

$$\begin{aligned} f_{Z(t)}(z) &= \int_0^\infty f_{Z(t)|W(t)}(z|w) \cdot f_{W(t)}(w) \cdot dw \\ &= \int_0^\infty \frac{(\alpha(w))^{\beta(w)} \cdot z^{\beta(w)-1} \cdot e^{-\frac{\alpha(w)}{z}}}{\Gamma(\beta(w))} \cdot \frac{c^d \cdot w^{\eta(t)-1}}{B(\eta(t),d) \cdot (w+c)^{\eta(t)+d}} \cdot dw, \end{aligned} \tag{2.58}$$

$$\begin{aligned} F_{Z(t)}(z) &= \int_0^\infty F_{Z(t)|W(t)}(z|w) \cdot f_{W(t)}(w) \cdot dw \\ &= \int_0^\infty \frac{\Gamma(\beta(w), \alpha(w)/z)}{\Gamma(\beta(w))} \cdot \frac{c^d \cdot w^{\eta(t)-1}}{B(\eta(t),d) \cdot (w+c)^{\eta(t)+d}} \cdot dw. \end{aligned} \tag{2.59}$$

Moreover, from Eqs. (2.51), (2.52), (2.56), and (2.57), by using the laws of total expectation and total variance, the marginal mean of $Z(t)$ (that exists for $d > 1$ can be expressed as:

$$E\{Z(t)\} = E\{E\{Z(t)|W(t)\}\} = E\{W(t)\} = \frac{c \cdot \eta(t)}{d-1} \tag{2.60}$$

and its conditional variance:

$$\begin{aligned} V\{Z(t)\} &= V\{E\{Z(t)|W(t)\}\} + E\{V\{Z(t)|W(t)\}\} \\ &= V\{W(t)\} + E\left\{\frac{(W(t))^\nu}{\varphi}\right\} = \frac{c^2 \cdot \eta(t) \cdot [\eta(t)+d-1]}{(d-1)^2 \cdot (d-2)} + \frac{E\{(W(t))^\nu\}}{\varphi} \end{aligned} \tag{2.61}$$

where the term $E\{(W(t))^\nu\}$, that exists for $d > \nu$, is equal to:

$$E\{(W(t))^\nu\} = \frac{c^\nu \cdot \Gamma(d-\nu)}{\Gamma(d)} \cdot \frac{\Gamma(\eta(t)+\nu)}{\Gamma(\eta(t))}. \tag{2.62}$$

Finally, it is worth to remark that, under the considered setup:

— the perturbed process $\{Z(t); t \geq 0\}$ is not a Markov process.

— the marginal hidden process $\{W(t); t \geq 0\}$ is age and state dependent. This feature is also preserved in the particular case where the temporal variability of the unit-specific paths is described by homogeneous gamma processes (see Giorgio and Pulcini (2018)). Indeed, the dependence on the state arises from the existence within the population of units whose degradation rate evolve over time at diverse paces (i.e., units with elevated degradation levels are anticipated to experience heightened future deterioration, reflecting their inherent weakness). This distinction between

weaker and stronger units hinges on the recognition and comprehension of this interdependence.

2.5 Formulation of the likelihood function

Let us consider m identical units operating under homogeneous conditions and suppose that the degradation level of the unit i $(i = 1, ..., m)$ is measured at times $t_{i,1}, ..., t_{i,n_i}$, and that measurements are contaminated by random errors. Finally, let us denote by $Z_{i,j} = Z(t_{i,j})$ the perturbed degradation level of the unit i at the measurement time $t_{i,j}$, by $z_{i,j}$ its realization, and by $\boldsymbol{\xi}$ the vector of model parameters. Then, under these setting, the likelihood function of the considered processes can be formulated as:

$$L(\boldsymbol{\xi}; \boldsymbol{z}) = \prod_{i=1}^{m} \prod_{j=1}^{n_i} f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}\left(z_{i,j}|\boldsymbol{z}_{i,j-1}\right) \tag{2.63}$$

where $\boldsymbol{z} = \{\boldsymbol{z}_{1,n_1}, \ldots, \boldsymbol{z}_{m,n_m}\}$ is the realization of the whole set of available noisy measurements $\boldsymbol{Z} = \{\boldsymbol{Z}_{1,n_1}, \ldots, \boldsymbol{Z}_{m,n_m}\}$, $\boldsymbol{Z}_{i,j} = \{Z_{i,1}, \ldots, Z_{i,j}\}$ is the set of perturbed measurements of the degradation state of the unit i collected up to time $t_{i,j}$, $\boldsymbol{z}_{i,j} = \{z_{i,1}, \ldots, z_{i,j}\}$ is the realization of $\boldsymbol{Z}_{i,j}$, $t_{i,0} = 0$, $\boldsymbol{Z}_{i,0}$ and $\boldsymbol{z}_{i,0}$ are the empty set, and $f_{Z_{i,1}|\boldsymbol{Z}_{i,0}}(z_{i,1}|\boldsymbol{z}_{i,0}) = f_{Z_{i,1}}(z_{i,1})$.

Given that, under all considered processes, $\{Z(t); t \geq 0\}$ is not a Markov process, Eq. (2.63) cannot be expressed in closed form and its computation is very demanding. Nevertheless, the needed pdfs can be efficiently computed, for any $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, by using the recursive equations:

$$\begin{aligned} f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1}) = \int_0^{w_{i,j}} & f_{\Delta W_{i,j}|W_{i,j-1}}(\Delta w_{i,j}|w_{i,j-1}) \times \\ & \times f_{W_{i,j-1}|\boldsymbol{Z}_{i,j-1}}(w_{i,j-1}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j-1} \end{aligned} \tag{2.64}$$

$$f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1}) = \int_0^{\infty} f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j} \tag{2.65}$$

$$f_{W_{i,j}|\boldsymbol{Z}_{i,j}}(w_{i,j}|\boldsymbol{z}_{i,j}) = \frac{f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1})}{f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1})} \tag{2.66}$$

where $W_{i,j} = W(t_{i,j})$ denotes the hidden (true) degradation level of the unit i at time $t_{i,j}$, $\Delta W_{i,j} = \Delta W(t_{i,j-1}, t_{i,j}) = W_{i,j} - W_{i,j-1}$ is the hidden (true) degradation increment of the unit i in the interval $(t_{i,j-1}, t_{i,j})$, $w_{i,j}$ is the realization of $W_{i,j}$, and $\Delta w_{i,j} = w_{i,j} - w_{i,j-1}$ is the realization of $\Delta W_{i,j}$.

In fact, Eq. (2.64) can be derived as:

$$\begin{aligned} f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1}) &= \int_0^{w_{i,j}} f_{W_{i,j},W_{i,j-1}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}, w_{i,j-1}|\boldsymbol{z}_{i,j}) \cdot dw_{i,j-1} \\ &= \int_0^{w_{i,j}} f_{W_{i,j}|W_{i,j-1},\boldsymbol{Z}_{i,j-1}}(w_{i,j}|w_{i,j-1}, \boldsymbol{z}_{i,j-1}) \cdot f_{W_{i,j-1}|\boldsymbol{Z}_{i,j-1}}(w_{i,j-1}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j-1} \\ &= \int_0^{w_{i,j}} f_{W_{i,j}|W_{i,j-1}}(w_{i,j}|w_{i,j-1}) \cdot f_{W_{i,j-1}|\boldsymbol{Z}_{i,j-1}}(w_{i,j-1}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j-1}, \end{aligned}$$

where the last equality is possible because $\{W(t); t \geq 0\}$ is Markovian and hence the state $W_{i,j}$, given $W_{i,j-1}$, is conditionally independent on the perturbed measurements $\boldsymbol{Z}_{i,j-1}$ collected up to and including time $t_{i,j-1}$.
Eq. (2.65) can be derived as follows:

$$\begin{aligned} f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1}) &= \int_0^{\infty} f_{Z_{i,j},W_{i,j}|\boldsymbol{Z}_{i,j}}(z_{i,j}, w_{i,j}|\boldsymbol{z}_{i,j}) \cdot dw_{i,j} \\ &= \int_0^{\infty} f_{Z_{i,j}|W_{i,j},\boldsymbol{Z}_{i,j}}(z_{i,j}|w_{i,j}, \boldsymbol{z}_{i,j}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j} \\ &= \int_0^{\infty} f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1}) \cdot dw_{i,j}, \end{aligned}$$

where the last equality holds due to the assumption laid out in Chapter 2.2.1. Similarly, Eq. (2.66) can be obtained through the following chain of equalities:

$$\begin{aligned} f_{W_{i,j}|\boldsymbol{Z}_{i,j}}(w_{i,j}|\boldsymbol{z}_{i,j}) &= \frac{f_{\boldsymbol{Z}_{i,j},W_{i,j}}(\boldsymbol{z}_{i,j}, w_{i,j})}{f_{\boldsymbol{Z}_{i,j}}(\boldsymbol{z}_{i,j})} = \frac{f_{Z_{i,j},W_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}, w_{i,j}|\boldsymbol{z}_{i,j-1}) \cdot f_{\boldsymbol{Z}_{i,j-1}}(\boldsymbol{z}_{i,j-1})}{f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1}) \cdot f_{\boldsymbol{Z}_{i,j-1}}(\boldsymbol{z}_{i,j-1})} \\ &= \frac{f_{Z_{i,j},W_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}, w_{i,j}|\boldsymbol{z}_{i,j-1})}{f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1})} \\ &= \frac{f_{Z_{i,j}|W_{i,j},\boldsymbol{Z}_{i,j-1}}(z_{i,j}|w_{i,j}, \boldsymbol{z}_{i,j-1}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1})}{f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1})} \\ &= \frac{f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) \cdot f_{W_{i,j}|\boldsymbol{Z}_{i,j-1}}(w_{i,j}|\boldsymbol{z}_{i,j-1})}{f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1})} \end{aligned}$$

Under the perturbed processes adopted in this manuscript, the likelihood function cannot be computed in closed form. Indeed, it is computed numerically via the particle filter described in Chapter 2.8, which takes advantage of Eqs. (2.64)-(2.66). The value of $\boldsymbol{\xi}$ that maximizes (over the parameter space) the likelihood function 2.63 is by definition the Maximum Likelihood (ML) estimate $\widehat{\boldsymbol{\xi}}$ of $\boldsymbol{\xi}$. However, the direct numerical maximization of this function carries a particularly heavy computational burden and also poses serious convergence issues. Therefore, in this manuscript the ML estimates of model parame-

ters are retrieved by using the new expectation-maximization particle filter algorithm described in Chapter 2.7 that massively alleviates these issues.

2.6 Cdf of RUL

A unit is assumed to fail when its true degradation level passes a preassigned failure threshold w_M. Hence, the lifetime X of the unit is defined as the first passage time of the hidden degradation process to the threshold w_M:

$$X = inf\{x : W(x) > w_M\}.$$

Accordingly, its remaining useful life $RUL(t)$ at time t is defined as:

$$RUL(t) = max\{0, X - t\}.$$

In this manuscript, it is assumed that:

— Failures are not self-announcing (Bautista, Castro, and Landesa (2022), Bismut, Pandey, and Straub (2022)) and that a failed unit may continue to operate, albeit with (possibly) reduced performance.

— Available measurements only consist of perturbed measurements.

Consequently, even conditional to the available data, it is not possible to assess with certainty whether a unit is failed or not. Under these assumptions, and given also that under all considered perturbed processes, the hidden processes are all monotonic increasing, the conditional cdf $F_{RUL(t)|\boldsymbol{Z}(t)}\left(\tau|\boldsymbol{z}(t)\right)$ of the $RUL(t)$ given $\boldsymbol{Z}(t) = \boldsymbol{z}(t)$, is formulated as:

$$\begin{aligned} &F_{RUL(t)|\boldsymbol{Z}(t)}\left(\tau|\boldsymbol{z}(t)\right) = P\left[RUL(t) \leq \tau|\boldsymbol{Z}(t) = \boldsymbol{z}(t)\right] = P\left[W(t+\tau) > w_M|\boldsymbol{Z}(t) = \boldsymbol{z}(t)\right] \\ &= 1 - F_{W(t+\tau)|\boldsymbol{Z}(t)}(w_M|\boldsymbol{z}(t)) = 1 - \int_0^{w_M} F_{\Delta W(t,t+\tau)|W(t)}(w_M - w|w) \cdot f_{W(t)|\boldsymbol{Z}(t)}(w|\boldsymbol{z}(t)) \cdot dw \end{aligned} \tag{2.67}$$

where $\boldsymbol{Z}(t) = \{Z(t_j); j \geq 1, t_j \leq t\}$ denotes the set of measurements gathered up to and included the time t and $\boldsymbol{z}(t) = \{z(t_j); j \geq 1, t_j \leq t\}$ is its realization.
Coherently with the previous assumptions, because measurement error is present and failures are not self-announcing, it results $F_{RUL(t)}(0|\boldsymbol{z}(t)) > 0$. It is worth noticing that,

if the set $\boldsymbol{Z}(t)$ contains l perturbed measurements, Eq. (2.67) is fully equivalent to

$$\begin{aligned} F_{RUL(t)|\boldsymbol{Z}(t)}\left(\tau|\boldsymbol{z}(t)\right) &= P\left[RUL(t_l) \leq \tau + t - t_l|\boldsymbol{Z}(t_l) = \boldsymbol{z}(t_l)\right] \\ &= P\left[W(t+\tau) > w_M|\boldsymbol{Z}(t_l) = \boldsymbol{z}(t_l)\right] = 1 - F_{W(t+\tau)|\boldsymbol{Z}(t_l)}(w_M|\boldsymbol{z}(t_l)) \\ &= 1 - \int_0^{w_M} F_{\Delta W(t,t+\tau)|W(t_l)}(w_M - w_l|w_l) \cdot f_{W(t_l)|\boldsymbol{Z}(t_l)}(w|\boldsymbol{z}(t_l)) \cdot dw_l \end{aligned} \quad (2.68)$$

where $\boldsymbol{Z}(t_l) = \{Z(t_1), \ldots, Z(t_l)\}$ is the set of perturbed measurements made on the considered unit up to the time t_l, $\boldsymbol{z}(t_l) = \{z_1, \ldots, z_l\}$ is its realization, t_l is the epoch of the lth measurements, and $F_{W(t+\tau)|W(t_l)}(\cdot|w_l)$ is the conditional cdf of the actual degradation level at $t + \tau$ given $W(t_l) = w_l$. Hence, before the first measurement is taken (i.e., for $t < t_1$) $\boldsymbol{Z}(t)$ should be intended as the empty set so that, in this case, $F_{W(t+\tau)|\boldsymbol{Z}(t)}(w_M|\boldsymbol{z}(t))$ coincides with $F_{W(t+\tau)}(w_M)$. From Eq. (2.67), given $\boldsymbol{Z}(t) = \boldsymbol{z}(t)$, the conditional mean of the $RUL(t)$, here intended as the conditional mean $E\{RUL(t)|\boldsymbol{Z}(t) = \boldsymbol{z}(t)\}$ and denoted as $MRUL\left(t|\boldsymbol{Z}(t) = \boldsymbol{z}(t)\right)$, can be computed as:

$$\begin{aligned} MRUL(t|\boldsymbol{Z}(t) = \boldsymbol{z}(t)) &= \int_0^{\infty} \left(1 - F_{RUL(t)|\boldsymbol{Z}(t)}\left(\tau|\boldsymbol{z}(t)\right)\right) \cdot d\tau \\ &= \int_0^{\infty} F_{W(t+\tau)|\boldsymbol{Z}(t)}\left(w_M|\boldsymbol{z}(t)\right) \cdot d\tau. \end{aligned} \quad (2.69)$$

Unfortunately, under all the considered models, the cdf $F_{W(t+\tau)|\boldsymbol{Z}(t)}(w_M|\boldsymbol{z}(t))$ is not available in closed form. Hence, both the cdf of $RUL(t)$ in Eq. (2.67) and the related conditional mean $MRUL(t|\boldsymbol{Z}(t) = \boldsymbol{z}(t))$ in Eq. (2.69) are computed by using the particle filter algorithm illustrated in Chapter 2.8.

Further results which involve the lifetime X are derived here (these will be important for the computations of maintenance costs in Chapter 5). Let $f_{X|W(\tau)}(x|w_\tau)$ denote the conditional pdf of the lifetime X given $W(\tau) = w_\tau$, where τ is a generic reference time, and w_τ is the true value of the degradation level at time τ. Then, it is possible to readily obtain the conditional cdf of X given $W(\tau) = w_\tau$, in the cases where $w_\tau \leq w_M$, as:

$$\begin{aligned} F_{X|W(\tau)}(x|w_\tau) &= P\left[X \leq x|W(\tau) = w_\tau\right] = P\left[W(x) > w_M|W(\tau) = w_\tau\right] \\ &= \begin{cases} 0, & w_\tau \leq w_M,\ x < \tau \\ 1 - F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau), & w_\tau \leq w_M,\ x \geq \tau \end{cases} \end{aligned} \quad (2.70)$$

where the first equality can be explained by observing that, since (under all stochastic processes considered in this manuscript) the process $\{W(t); t \geq 0\}$ is monotone increasing,

the event $\{X \leq x\}$ is equivalent to the event $\{W(x) > w_M\}$.
Differently, in the cases where $w_\tau > w_M$ the conditional cdf of X given $W(\tau) = w_\tau$ can be expressed as:

$$\begin{aligned} F_{X|W(\tau)}(x|w_\tau) &= P\left[X \leq x | W(\tau) = w_\tau\right] = P\left[W(x) > w_M | W(\tau) = w_\tau\right] \\ &= \begin{cases} 1 - F_{W(x)|W(\tau)}(w_M|w_\tau), & w_\tau > w_M,\ x \leq \tau \\ 1, & w_\tau > w_M,\ x > \tau. \end{cases} \end{aligned} \tag{2.71}$$

Then, the conditional mean of X, given $W(\tau) = w_\tau$, can be expressed as:

$$\begin{aligned} E\{X|W(\tau) = w_\tau\} &= \int_0^\tau x \cdot f_{X|W(\tau)}(x|w_\tau) \cdot dx \\ &= \left[-x \cdot \left(1 - F_{X|W(\tau)}(x|w_\tau)\right)\right]_0^\tau + \int_0^\tau \left(1 - F_{X|W(\tau)}(x|w_\tau)\right) \cdot dx \\ &= \int_0^\tau \left(1 - F_{X|W(\tau)}(x|w_\tau)\right) \cdot dx. \end{aligned} \tag{2.72}$$

Finally, when $w_\tau \leq w_M$, given $W(\tau) = w_\tau$, the conditional mean of the variable $g(X)$ defined by the following transformation:

$$g(X) = \begin{cases} \tau + \Delta\tau - X, & X \leq \tau + \Delta\tau \\ 0, & X > \tau + \Delta\tau \end{cases} \tag{2.73}$$

where $\Delta\tau \geq 0$, can be computed as:

$$\begin{aligned} E\{g(X)|W(\tau) = w_\tau\} &= \int_\tau^{\tau+\Delta\tau} (\tau + \Delta\tau - x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot dx \\ &= \left[-(\tau + \Delta\tau - x) \cdot \left(1 - F_{X|W(\tau)}(x|w_\tau)\right)\right]_\tau^{\tau+\Delta\tau} - \int_\tau^{\tau+\Delta\tau} \left(1 - F_{X|W(\tau)}(x|w_\tau)\right) \cdot dx \\ &= \Delta\tau - \int_\tau^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot dx. \end{aligned} \tag{2.74}$$

2.7 The EM algorithm

The EM algorithm (Dempster, Laird, and Rubin (1977)) is a general approach broadly adopted to iteratively compute the ML estimates in the presence of missing values and/or

incomplete observations. It consists of a two-step sequence, an Expectation step (E-step) and a Maximization step (M-step), iterated until a given convergence condition is attained. To apply the procedure, it is necessary to select the values that one intends to treat as "observed data" and the ones that should be intended as "missing data". At the $(h+1)$th iteration, the E-step consists in evaluating the conditional mean $Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)})$:

$$Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)}) = E\{l(\boldsymbol{\xi}; \boldsymbol{Y}_{mis}, \boldsymbol{y}_{ava})|\boldsymbol{y}_{ava}, \boldsymbol{\xi}^{(h)}\} \tag{2.75}$$

of the complete log-likelihood $l(\boldsymbol{\xi}; \boldsymbol{Y}_{mis}, \boldsymbol{y}_{ava})$ (i.e., the logarithm of the likelihood formulated considering all data, available and missing) with respect to the missing data $\boldsymbol{Y}_{mis}$ given the available data $\boldsymbol{y}_{ava}$ and the current tentative estimate of model parameters (i.e., the estimate of the parameter vector available after the hth iteration, denoted by $\boldsymbol{\xi}^{(h)}$).

The corresponding M-step consists in maximizing the function $Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)})$ with respect to $\boldsymbol{\xi}$. The value of $\boldsymbol{\xi}$ that maximizes $Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)})$ is, by definition, the new tentative estimate $\boldsymbol{\xi}^{(h+1)}$.
The convergence condition is satisfied when the absolute difference:

$$\left| \frac{\ln(L(\boldsymbol{\xi}^{(h)}; \boldsymbol{z}))}{\ln(L(\boldsymbol{\xi}^{(h+1)}; \boldsymbol{z}))} - 1 \right| \tag{2.76}$$

drops below a predetermined, fixed value. If the convergence condition is attained at the hth iteration, then $\boldsymbol{\xi}^{(h+1)}$ is assumed to be the ML estimate of the parameter vector $\boldsymbol{\xi}$. Otherwise, the search continues with another two-step sequence. At the first iteration, the algorithm is initialized by assigning a tentative estimate, say $\boldsymbol{\xi}^{(0)}$.

In this manuscript, we adopted the EM algorithm to compute ML estimates under the three perturbed degradation processes presented above. In the following, the implementation of the EM will be particularized for the case of the perturbed gamma and inverse Gaussian processes in Chapter 2.7.1, and for the case of the perturbed gamma process with random effect in Chapter 2.7.2.

2.7.1 Perturbed gamma and inverse Gaussian processes

In the case of the perturbed gamma and inverse Gaussian processes, the realizations $\boldsymbol{z} = \{\boldsymbol{z}_1, ..., \boldsymbol{z}_m\}$ of the related perturbed measurements $\boldsymbol{Z} = \{\boldsymbol{Z}_1, ..., \boldsymbol{Z}_m\}$ (where $\boldsymbol{z}_i = \{z_{i,1}, ..., z_{i,n_i}\}$ is the realization of $\boldsymbol{Z}_i = \{Z_{i,1}, ..., Z_{i,n_i}\}$) will constitute the observed data whereas the missing data consist in the (unknown) values $\boldsymbol{w} = \{\boldsymbol{w}_1, ..., \boldsymbol{w}_m\}$ of the

(unobservable) true degradation levels $\boldsymbol{W} = \{\boldsymbol{W}_1, ..., \boldsymbol{W}_m\}$ of the m units at the measurement times, where $\boldsymbol{w}_i = \{w_{i,1}, ..., w_{i,n_i}\}$ is the realization of $\boldsymbol{W}_i = \{W_{i,1}, ..., W_{i,n_i}\}$. Therefore, Eq. (2.75) becomes:

$$Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)}) = E\{l(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{W})|\boldsymbol{Z} = \boldsymbol{z}, \boldsymbol{\xi}^{(h)}\}$$

where the complete log-likelihood can be formulated as:

$$L(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{w}) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) \cdot \prod_{i=1}^{m}\prod_{j=1}^{n_i} f_{\Delta W_{i,j}}(\Delta w_{i,j})$$

and the corresponding log-likelihood function as:

$$l(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{w}) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} \ln\left[f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j})\right] \cdot \prod_{i=1}^{m}\prod_{j=1}^{n_i} \ln\left[f_{\Delta W_{i,j}}(\Delta w_{i,j})\right].$$

In the case of the considered processes, the benefit of the EM procedure mainly lie in the M-step. Indeed, the conditional expectation $Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)}) = Q(a, b, \theta, \varphi, \nu|\boldsymbol{\xi}^{(h)})$ splits into the sum of the functions $Q_H\left(a, b, \theta|\boldsymbol{\xi}^{(h)}\right)$ and $Q_\varepsilon\left(\varphi, \nu|\boldsymbol{\xi}^{(h)}\right)$, where:

$$Q_H\left(a, b, \theta|\boldsymbol{\xi}^{(h)}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{ln\left(f_{\Delta W_{i,j}}(\Delta W_{i,j})\right)\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\},$$

given $\boldsymbol{\xi}^{(h)}$, depends only on the parameters of the hidden model, and $Q_\varepsilon\left(\varphi, \nu|\boldsymbol{\xi}^{(h)}\right)$

$$Q_\varepsilon\left(\varphi, \nu|\boldsymbol{\xi}^{(h)}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{ln\left(f_{Z_{i,j}|W_{i,j}}(z_{i,j}|W_{i,j})\right)\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\}$$

depends only on the parameters of the perturbing term.
Obviously, the functional form of $Q_H\left(a, b, \theta|\boldsymbol{\xi}^{(h)}\right)$ and $Q_\varepsilon\left(\varphi, \nu|\boldsymbol{\xi}^{(h)}\right)$ also depends on the model adopted to describe the error term and the hidden process, respectively.
In fact, when the error term is modeled by using the inverse gamma random variable described in Eq. (2.18), $Q_\varepsilon\left(\varphi, \nu|\boldsymbol{\xi}^{(h)}\right)$ reduces to (for the sake of clarity, here the functions

$\alpha(\cdot)$ and $\beta(\cdot)$ are denoted as $\alpha(\cdot, \varphi, \nu)$ and $\beta(\cdot, \varphi, \nu)$):

$$\begin{aligned} Q_{\varepsilon}^{1}\left(\varphi, \nu | \boldsymbol{\xi}^{(h)}\right) = & -\sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{E\left\{\alpha(W_{i,j};\varphi,\nu)|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\}}{z_{i,j}} \\ & + \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\beta(W_{i,j};\varphi,\nu)\cdot\ln[\alpha(W_{i,j};\varphi,\nu)]\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\} \\ & - \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\beta(W_{i,j};\varphi,\nu)\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\}\cdot\ln(z_{i,j}) \\ & - \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\ln\left(\Gamma\left(\beta(W_{i,j};\varphi,\nu)\right)\right)\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\}, \end{aligned} \tag{2.77}$$

whereas, when the error term is modeled by using the Gaussian distribution in Eq. (2.24), $Q_{\varepsilon}\left(\varphi, \nu | \boldsymbol{\xi}^{(h)}\right)$ reduces to:

$$\begin{aligned} Q_{\varepsilon}^{2}\left(\varphi, \nu | \boldsymbol{\xi}^{(h)}\right) = & -\frac{n_t}{2}\cdot\ln(2\pi) + \frac{n_t}{2}\cdot\ln(\varphi) - \frac{\nu}{2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\ln(W_{i,j})\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\} \\ & - \frac{\varphi}{2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\frac{(z_{i,j}-W_{i,j})^2}{W_{i,j}^{\nu}}\Big|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\right\} \end{aligned} \tag{2.78}$$

where the superscripts 1 and 2 in $Q_{\varepsilon}^{1}\left(\varphi, \nu | \boldsymbol{\xi}^{(h)}\right)$ and $Q_{\varepsilon}^{2}\left(\varphi, \nu | \boldsymbol{\xi}^{(h)}\right)$ indicate the option adopted to model the error term, in agreement with the symbols introduced in Chapter 2.3.

Similarly, from Eq. (2.11), under the gamma process, the function $Q_H\left(a, b, \theta | \boldsymbol{\xi}^{(h)}\right)$ reduces to:

$$\begin{aligned} Q_{H}^{G}\left(a, b, \theta | \boldsymbol{\xi}^{(h)}\right) = & -\frac{\sum_{i=1}^{m} E\{W_{i,n_i}|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\}}{\theta} \\ & + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left[\left(\frac{t_{i,j}}{a}\right)^{b} - \left(\frac{t_{i,j-1}}{a}\right)^{b}\right]\cdot E\{\ln(\Delta W_{i,j})|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)}\} \\ & - \ln(\theta)\cdot\sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^{b} - \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln\left[\Gamma\left(\left(\frac{t_{i,j}}{a}\right)^{b} - \left(\frac{t_{i,j-1}}{a}\right)^{b}\right)\right], \end{aligned} \tag{2.79}$$

whereas, from Eq. (2.12), under the inverse Gaussian process $Q_H\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)$ becomes:

$$\begin{aligned} Q_H^{IG}\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right) = &-\frac{n_t}{2}\cdot\ln(2\cdot\pi)+\frac{n_t}{2}\cdot\ln(\theta)-n_t\cdot b\cdot\ln(a)\\ &-\frac{3}{2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\{\ln(\Delta W_{i,j})|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}\\ &-\frac{1}{2\cdot\theta}\cdot\sum_{i=1}^{m}E\{W_{i,n_i}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}\\ &-\frac{\theta}{2\cdot a^{2\cdot b}}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\frac{(t_{i,j}^b-t_{i,j-1}^b)^2}{\Delta W_{i,j}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}\\ &+\frac{1}{a^b}\cdot\sum_{i=1}^{m}t_{i,n_i}^b+\sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln(t_{i,j}^b-t_{i,j-1}^b). \end{aligned} \tag{2.80}$$

where $n_t=\sum_{i=1}^{m}n_i$ and the superscripts G and IG in $Q_H^G\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)$ and $Q_H^{IG}\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)$ indicate that the hidden process used to compute $Q_H\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)$ are the gamma and the inverse Gaussian, respectively. In both Eqs. (2.79) and (2.80) the age function is modeled as $\eta(t)=(t/a)^b$.

For the sake of maximizing them in the M-step, the expressions in Eqs. (2.78), (2.79), and (2.80) can be further simplified. Indeed, from Eq. (2.78), by solving with respect to φ the equation:

$$\frac{\partial Q_\varepsilon^2\left(\varphi,\nu|\boldsymbol{\xi}^{(h)}\right)}{\partial\varphi}=0$$

the explicit form:

$$\tilde{\varphi}(\nu|\boldsymbol{\xi}^{(h)})=\frac{n_t}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\frac{(z_{i,j}-W_{i,j})^2}{W_{i,j}^{\nu}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}} \tag{2.81}$$

is obtained for the value $\tilde{\varphi}(\nu|\boldsymbol{\xi}^{(h)})$ that maximizes Eq. (2.81) with respect to φ when ν is set to the values indicated in the parentheses on the left side of the equation. By exploiting this result, $\nu^{(h+1)}$ can be obtained by numerically maximizing this 1-parameter function:

$$Q_\varepsilon^2\left(\varphi,\nu|\boldsymbol{\xi}^{(h)}\right)=-\frac{n_t}{2}\cdot\ln(2\pi)+\frac{n_t}{2}\cdot\ln(\tilde{\varphi})-\frac{\nu}{2}\cdot\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\ln(W_{i,j})\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}-\frac{n_t}{2}. \tag{2.82}$$

Then, $\varphi^{(h+1)}$ can be obtained from Eq. (2.81) by setting $\nu=\nu^{(h+1)}$.

Similarly, from Eq. (2.79), by solving with respect to θ the equation:

$$\frac{\partial Q_H^G\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)}{\partial\theta}=0$$

the explicit form:

$$\tilde{\theta}(a,b|\boldsymbol{\xi}^{(h)})=\frac{\sum_{i=1}^{m}E\left\{W_{i,n_i}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}}{\sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^b} \tag{2.83}$$

is obtained for the value $\tilde{\theta}(a,b)$ that maximizes Eq. (2.83) with respect to θ when a and b are set to the values indicated in the parentheses on the left side of the equation. By exploiting this result, $a^{(h+1)}$ and $b^{(h+1)}$ (i.e., the components of $\boldsymbol{\xi}^{(h+1)}$) can be obtained by numerically maximizing this two-parameter function:

$$\begin{aligned}Q_H^G(a,b|\tilde{\theta},\boldsymbol{\xi}^{(h)})=&-\sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln\left[\Gamma\left(\left(\frac{t_{i,j}}{a}\right)^b-\left(\frac{t_{i,j-1}}{a}\right)^b\right)\right]\\&+\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left[\left(\frac{t_{i,j}}{a}\right)^b-\left(\frac{t_{i,j-1}}{a}\right)^b\right]\cdot E\{\ln(\Delta W_{i,j})|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}\\&+\left[1+\ln\left(\frac{\sum_{i=1}^{m}E\{W_{i,n_i}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}}{\sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^b}\right)\right]\cdot\sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^b.\end{aligned} \tag{2.84}$$

Then, $\theta^{(h+1)}$ can be obtained from (2.83) by setting $a=a^{(h+1)}$ and $b=b^{(h+1)}$.

Conversely, in the case of the inverse Gaussian process, from Eq. (2.80), by solving with respect to θ and a the system of equations:

$$\begin{cases}\dfrac{\partial Q_H^{IG}\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)}{\partial\theta}=0\\[2ex]\dfrac{\partial Q_H^{IG}\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)}{\partial a}=0\end{cases}$$

the following explicit forms are obtained:

$$\tilde{\theta}(b|\boldsymbol{\xi}^{(h)}) = \frac{\left\{\frac{\sum_{i=1}^{m} E\{W_{i,n_i}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}}{\sum_{i=1}^{m} t_{i,n_i}^{b}}\right\}^2 \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\frac{\left(t_{i,j}^{b}-t_{i,j-1}^{b}\right)^2}{\Delta W_{i,j}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}}{n_t} - \frac{\sum_{i=1}^{m} E\{W_{i,n_i}|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{\xi}^{(h)}\}}{n_t} \tag{2.85}$$

$$\tilde{a}(b|\boldsymbol{\xi}^{(h)}) = \left[\frac{\left\{\frac{\sum_{i=1}^{m} E\{W_{i,n_i}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\}}{\sum_{i=1}^{m} t_{i,n_i}^{b}}\right\} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} E\left\{\frac{\left(t_{i,j}^{b}-t_{i,j-1}^{b}\right)^2}{\Delta W_{i,j}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\} - \sum_{i=1}^{m} t_{i,n_i}^{b}}{n_t}\right]^{\frac{1}{b}} \tag{2.86}$$

for the values $\tilde{\theta}(b)$ and $\tilde{a}(b)$ of θ and a that (jointly) maximize $Q_H^{IG}\left(a,b,\theta|\boldsymbol{\xi}^{(h)}\right)$ when the parameter b is set to the value reported in the parentheses. By exploiting these results, $b^{(h+1)}$ can be obtained by numerically maximizing the 1-parameter function that is obtained by setting θ and a in Eq. (2.80) to $\tilde{\theta}(b)$ and $\tilde{a}(b)$ (the resulting function is not reported here for the sake of readability). Then, $\theta^{(h+1)}$ and $a^{(h+1)}$ can be obtained by evaluating Eqs. (2.85) and (2.86) at $b = b^{(h+1)}$.
Therefore, when the error term is modeled according to the option 1), the M-step results in maximizing (numerically) two 2-parameter functions in the case of the gamma process, and one 2-parameter function and one 1-parameter function in the case of the inverse Gaussian. Whereas, when the error term is modeled by using the option 2), the M-step results in maximizing (numerically) one 2-parameter function and one 1-parameter function in the case of the gamma process, and two 1-parameter functions in the case of the inverse Gaussian process.

2.7.2 The perturbed gamma process with random effects

In the case of the perturbed gamma process with random effects, the observed data will still be the realizations $\boldsymbol{z} = \{\boldsymbol{z}_1, ..., \boldsymbol{z}_m\}$ of the related perturbed measurements $\boldsymbol{Z} = \{\boldsymbol{Z}_1, ..., \boldsymbol{Z}_m\}$ (where $\boldsymbol{z}_i = \{z_{i,1}, ..., z_{i,n_i}\}$ is the realization of $\boldsymbol{Z}_i = \{Z_{i,1}, ..., Z_{i,n_i}\}$), whereas the missing data are the (unknown) values $\boldsymbol{w} = \{\boldsymbol{w}_1, ..., \boldsymbol{w}_m\}$ of the (unobservable) true

degradation levels $\boldsymbol{W} = \{\boldsymbol{W}_1, ..., \boldsymbol{W}_m\}$ together with the set of unknown true values $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_m\}$ of the random rate parameters $\boldsymbol{\Lambda} = \{\Lambda_1, \ldots, \Lambda_m\}$ of the m units. Therefore, Eq. (2.75) becomes:

$$Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)}) = E\{l(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{W})|\boldsymbol{Z} = \boldsymbol{z}, \boldsymbol{\xi}^{(h)}\}$$

Therefore, considering also that, given $\boldsymbol{W} = \boldsymbol{w}$, $\boldsymbol{Z}$ is conditionally independent of $\boldsymbol{\Lambda}$, the complete likelihood (i.e., the one formulated considering all data, available and missing) can be formulated as:

$$L(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{\lambda}) = \prod_{i=1}^{m} f_{\boldsymbol{Z}_i|\boldsymbol{W}_i}(\boldsymbol{z}_i|\boldsymbol{w}_i) \cdot f_{\boldsymbol{W}_i|\Lambda_i}(\boldsymbol{w}_i|\lambda_i) \cdot g_{\Lambda_i}(\lambda_i),$$

where:

$$f_{\boldsymbol{Z}_i|\boldsymbol{W}_i}(\boldsymbol{z}_i|\boldsymbol{w}_i) = \prod_{j=1}^{n_i} f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) = \prod_{j=1}^{n_i} \frac{(\alpha(w_{i,j}))^{\beta(w_{i,j})} \cdot z_{i,j}^{-\beta(w_{i,j})-1}}{\Gamma(\beta(w_{i,j}))} \cdot e^{-\frac{\alpha(w_{i,j})}{z_{i,j}}},$$

$$f_{\boldsymbol{W}_i|\Lambda_i}(\boldsymbol{w}_i|\lambda_i) = \prod_{j=1}^{n_i} f_{\Delta W_{i,j}|\Lambda_i}(\Delta w_{i,j}|\lambda_i) = \prod_{j=1}^{n_i} \frac{\lambda_i^{\Delta\eta(t_{i,j-1},t_{i,j})} \cdot \Delta w_{i,j}^{\Delta\eta(t_{i,j-1},t_{i,j})-1}}{\Gamma\left[\Delta\eta(t_{i,j-1}, t_{i,j})\right]} \cdot e^{-\lambda_i \cdot \Delta w_{i,j}}, \tag{2.87}$$

and:

$$g_{\Lambda_i}(\lambda_i) = \frac{c^d \cdot \lambda_i^{d-1}}{\Gamma(d)} \cdot e^{-c\cdot\lambda_i}. \tag{2.88}$$

Thus, the complete likelihood function can be expressed as:

$$\begin{aligned} L(\boldsymbol{\xi}; \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{\lambda}) = \\ = \frac{c^{m\cdot d}}{[\Gamma(d)]^m} \cdot e^{-\sum_{i=1}^{m} \lambda_i \cdot (w_{i,n_i}+c)} \cdot \prod_{i=1}^{m} \lambda_i^{\eta(t_{i,n_i})+d-1} \cdot \prod_{i=1}^{m}\prod_{j=1}^{n_i} \frac{\Delta w_{i,j}^{\Delta\eta(t_{i,j-1},t_{i,j})-1}}{\Gamma\left[\Delta\eta(t_{i,j-1}, t_{i,j})\right]} \\ \times \prod_{i=1}^{m} f_{Z_{i,j}|W_{i,j}}(z_{i,j}|w_{i,j}) = \prod_{j=1}^{n_i} \frac{(\alpha(w_{i,j}))^{\beta(w_{i,j})} \cdot z_{i,j}^{-\beta(w_{i,j})-1}}{\Gamma(\beta(w_{i,j}))} \cdot e^{-\frac{\alpha(w_{i,j})}{z_{i,j}}} \end{aligned} \tag{2.89}$$

and the corresponding log-likelihood function can be written as:

$$\begin{aligned} l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{w},\boldsymbol{\lambda}) = m\cdot d\cdot\ln(c) - m\cdot\ln[\Gamma(d)] - \sum_{i=1}^{m}\lambda_i\cdot(w_{i,n_i}+c) + \sum_{i=1}^{m}[\eta(t_{i,n_i})+d-1]\cdot\ln(\lambda_i) \\ + \sum_{i=1}^{m}\sum_{j=1}^{n_i}[\Delta\eta(t_{i,j-1},t_{i,j})-1]\cdot\ln(\Delta w_{i,j}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln(\Gamma\left[\Delta\eta(t_{i,j-1},t_{i,j})\right]) \\ - \sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{\alpha(w_{i,j})}{z_{i,j}} + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\beta(w_{i,j})\cdot\ln(\alpha(w_{i,j})) \\ - \sum_{i=1}^{m}\sum_{j=1}^{n_i}[\beta(w_{i,j})+1]\cdot\ln(z_{i,j}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln\left[\Gamma(\beta(w_{i,j}))\right]. \end{aligned} \tag{2.90}$$

Under the considered perturbed gamma process with random effect, the conditional pdf of $\boldsymbol{W}$ given $\boldsymbol{Z}=\boldsymbol{z}$ is not available in closed form. However, the conditional pdf of $\boldsymbol{\Lambda}$ given $\boldsymbol{W}$ et $\boldsymbol{Z}=\boldsymbol{z}$ is available in closed form. Indeed, considered that $\boldsymbol{\Lambda}$ is conditionally independent of $\boldsymbol{Z}$ given $\boldsymbol{W}$ and that Λ_i is independent both of $\boldsymbol{W}_1,\ldots,\boldsymbol{W}_{i-1},\boldsymbol{W}_{i+1},\ldots,\boldsymbol{W}_m$ and $\Lambda_1,\ldots,\Lambda_{i-1},\Lambda_{i+1},\ldots,\Lambda_m$, and being:

$$g_{\Lambda_i|\boldsymbol{W}_i}(\lambda_i|\boldsymbol{w}_i) = \frac{f_{\boldsymbol{W}_i|\Lambda_i}(\boldsymbol{w}_i|\lambda_i)\cdot g_{\Lambda_i}(\lambda_i)}{f_{\boldsymbol{W}_i}(\boldsymbol{w}_i)}, \tag{2.91}$$

with (from Eqs. (2.87) and (2.88)):

$$f_{\boldsymbol{W}_i|\Lambda_i}(\boldsymbol{w}|\lambda_i)\cdot g_{\Lambda_i}(\lambda_i) = \frac{\lambda_i^{\eta(t_{i,n_i})+d-1}\cdot c^d}{\Gamma(d)}\cdot e^{-\lambda_i\cdot(w_{i,n_i}+c)}\cdot\prod_{j=1}^{n_i}\frac{\Delta w_{i,j}^{\Delta\eta(t_{i,j-1},t_{i,j})-1}}{\Gamma\left[\Delta\eta(t_{i,j-1},t_{i,j})\right]} \tag{2.92}$$

and (from Eq. (2.44)):

$$\begin{aligned} f_{\boldsymbol{W}_i}(\boldsymbol{w}_i) &= \prod_{j=1}^{n_i}\frac{1}{B\left[\Delta\eta(t_{i,j-1},t_{i,j}),\eta(t_{i,j-1})+d\right]}\cdot\frac{(w_{i,j-1}+c)^{\eta(t_{i,j-1})+d}\cdot\Delta w_{i,j}^{\Delta\eta(t_{i,j-1},t_{i,j})-1}}{(w_{i,j-1}+\Delta w_{i,j}+c)^{\eta(t_{i,j})+d}} \\ &= \frac{c^d\cdot\Gamma[\eta(t_{i,n_i})+d]}{(w_{i,n_i}+c)^{\eta(t_{i,n_i})+d}\cdot\Gamma(d)}\cdot\prod_{j=1}^{n_i}\frac{\Delta w_{i,j}^{\Delta\eta(t_{i,j-1},t_{i,j})-1}}{\Gamma\left[\Delta\eta(t_{i,j-1},t_{i,j})\right]}, \end{aligned} \tag{2.93}$$

it results:

$$g_{\Lambda_i|\boldsymbol{W}_i}(\lambda_i|\boldsymbol{w}_i) = g_{\Lambda_i|W_{i,n_i}}(\lambda_i|w_{i,n_i}) = \frac{(w_{i,n_i}+c)^{\eta(t_{i,n_i})+d}\cdot\lambda_i^{\eta(t_{i,n_i})+d-1}}{\Gamma[\eta(t_{i,n_i})+d]}\cdot e^{-\lambda_i\cdot(w_{i,n_i}+c)}. \tag{2.94}$$

Consequently, the conditional pdf of $\boldsymbol{\Lambda}$ given $\boldsymbol{W}$ and $\boldsymbol{Z} = \boldsymbol{z}$ is:

$$g_{\Lambda|\boldsymbol{Z},\boldsymbol{W}}(\lambda|\boldsymbol{z},\boldsymbol{w}) = \prod_{i=1}^{m} g_{\Lambda_i|W_{i,n_i}}(\lambda_i|w_{i,n_i}) = \prod_{i=1}^{m} \frac{(w_{i,n_i}+c)^{\eta(t_{i,n_i})+d} \cdot \lambda_i^{\eta(t_{i,n_i})+d-1}}{\Gamma[\eta(t_{i,n_i})+d]} \cdot e^{-\lambda_i \cdot (w_{i,n_i}+c)} \quad (2.95)$$

and, in particular:

$$E\{\Lambda_i|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W}\} = E\{\Lambda_i|W_{i,n_i}\} = \frac{\eta(t_{i,n_i})+d}{W_{i,n_i}+c} \quad (2.96)$$

and:

$$E\{\ln(\Lambda_i)|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W}\} = E\{\ln(\Lambda_i)|W_{i,n_i}\} = \psi[\eta(t_{i,n_i})+d] - \ln(W_{i,n_i}+c), \quad (2.97)$$

where $\psi(y) = d\ln[\Gamma(x)]/dx = \Gamma'(x)/\Gamma(x)$ is the digamma function. By exploiting these latter results, to perform the E-step it is firstly formulated the conditional mean of the log-likelihood with respect to $\boldsymbol{\Lambda}$ given $\boldsymbol{W}$ and $\boldsymbol{Z} = \boldsymbol{z}$, $E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W},\boldsymbol{\xi}^{(h)}\}$, which can be expressed in closed form. Then, we obtain $E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{\xi}^{(h)}\}$ by computing the conditional mean of $E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W},\boldsymbol{\xi}^{(h)}\}$ with respect to $\boldsymbol{W}$ given $\boldsymbol{Z} = \boldsymbol{z}$.

In fact, by the law of total expectation it is:

$$E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{\xi}^{(h)}\} = E\{E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W},\boldsymbol{\xi}^{(h)}\}|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{\xi}^{(h)}\}. \quad (2.98)$$

Under the proposed model, the expectation of $E\{l(\boldsymbol{\xi};\boldsymbol{z},\boldsymbol{W},\boldsymbol{\Lambda})|\boldsymbol{Z}=\boldsymbol{z},\boldsymbol{W},\boldsymbol{\xi}^{(h)}\}$ with respect to $\boldsymbol{W}$ given $\boldsymbol{Z} = \boldsymbol{z}$ cannot be computed in closed form and is performed via the particle filter described in Chapter 2.8.

The benefit of adopting the EM algorithm are most evident in the M-step, because the conditional mean $Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(h)})$ splits into the sum of the functions $Q_\lambda(c,d|\boldsymbol{\xi}^{(h)})$, $Q_\varepsilon(\varphi,\nu|\boldsymbol{\xi}^{(h)})$, and $Q_\eta(a,b|\boldsymbol{\xi}^{(h)})$, which (given $\boldsymbol{\xi}^{(h)}$) depend only on the parameters (c,d), (φ,ν), and (a,b), respectively. For the sake of clarity, here the age function is denoted as $\eta(\cdot;a,b)$, $\beta(\cdot)$ as

$\beta(\cdot;\varphi,\nu)$, and $\alpha(\cdot)$ as $\alpha(\cdot;\varphi,\nu)$:

$$\begin{aligned} Q_{\varepsilon}\left(\varphi,\nu\middle|\boldsymbol{\xi}^{(h)}\right) = &-\sum_{i=1}^{m}\sum_{j=1}^{n_i}\frac{E\left\{\alpha(W_{i,j};\varphi,\nu)|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}}{z_{i,j}} \\ &+\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\beta(W_{i,j};\varphi,\nu)\cdot\ln[\alpha(W_{i,j};\varphi,\nu)]\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\} \\ &-\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\beta(W_{i,j};\varphi,\nu)\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}\cdot\ln(z_{i,j}) \\ &-\sum_{i=1}^{m}\sum_{j=1}^{n_i}E\left\{\ln\left(\Gamma\left(\beta(W_{i,j};\varphi,\nu)\right)\right)\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}, \end{aligned} \tag{2.99}$$

$$\begin{aligned} &Q_{\eta}\left(a,b\middle|\boldsymbol{\xi}^{(h)}\right) = \\ &\sum_{i=1}^{m}\eta(t_{i,n_i};a,b)\cdot\left\{\psi\left[\eta(t_{i,n_i};a^{(h)},b^{(h)})+d^{(h)}\right]-E\left\{\ln\left(W_{i,n_i}+c^{(h)}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right)\right\}\right\} \\ &+\sum_{i=1}^{m}\sum_{j=1}^{n_i}[\Delta\eta(t_{i,j-1},t_{i,j}-1)]\cdot E\left\{\ln(\Delta W_{i,j}|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)})\right\} \\ &-\sum_{i=1}^{m}\sum_{j=1}^{n_i}\ln(\Gamma(\Delta\eta(t_{i,j-1},t_{i,j};a,b))), \end{aligned} \tag{2.100}$$

$$\begin{aligned} &Q_{\lambda}\left(c,d|\boldsymbol{\xi}^{(h)}\right) = m\cdot d\cdot\ln(c)-m\cdot\ln(\Gamma(d)) \\ &-\sum_{i=1}^{m}E\left\{\frac{\eta\left(t_{i,n_i};a^{(h)},b^{(h)}\right)+d^{(h)}}{W_{i,n_i}+c^{(h)}}\cdot w_{i,n_i}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\} \\ &+(d-1)\cdot\sum_{i=1}^{m}\left\{\psi\left[\eta(t_{i,n_i};a^{(h)},b^{(h)})+d^{(h)}\right]-E\left\{\ln\left(W_{i,n_i}+c^{(h)}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right)\right\}\right\} \\ &-c\cdot\sum_{i=1}^{m}E\left\{\frac{\eta\left(t_{i,n_i};a^{(h)},b^{(h)}\right)+d^{(h)}}{W_{i,n_i}+c^{(h)}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}, \end{aligned} \tag{2.101}$$

The maximization of Eq. (2.101) can be further simplified. Indeed, being:

$$\frac{\partial}{\partial c}Q_{\lambda}\left(c,d\middle|\boldsymbol{\xi}^{(h)}\right) = \frac{m\cdot d}{c}-\sum_{i=1}^{m}E\left\{\frac{\eta\left(t_{i,n_i};a^{(h)},b^{(h)}\right)+d^{(h)}}{W_{i,n_i}+c^{(h)}}\middle|\boldsymbol{Z}_i=\boldsymbol{z}_i,\boldsymbol{\xi}^{(h)}\right\}, \tag{2.102}$$

equating this to 0 and solving with respect to c it is possible to express c as a function of

d and $\boldsymbol{\xi}^{(h)}$:

$$\tilde{c}(d|\boldsymbol{\xi}^{(h)}) = \frac{m \cdot d}{\sum_{i=1}^{m} E\left\{ \frac{\eta\left(t_{i,n_i}; a^{(h)}, b^{(h)}\right) + d^{(h)}}{W_{i,n_i} + c^{(h)}} \middle| \boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)} \right\}}. \tag{2.103}$$

Hence, from Eq. (2.101), by replacing c with the expression on the right side of Eq. (2.103), the following function of d and $\boldsymbol{\xi}^{(h)}$ is obtained:

$$\begin{aligned} Q_{\lambda}^{*}\left(d|\boldsymbol{\xi}^{(h)}\right) = & \\ & m \cdot d \cdot \left(\ln(m \cdot d) - \ln\left(\sum_{i=1}^{m} E\left\{ \frac{\eta\left(t_{i,n_i}; a^{(h)}, b^{(h)}\right) + d^{(h)}}{W_{i,n_i} + c^{(h)}} \middle| \boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)} \right\} \right) \right) \\ & - m \cdot \ln(\Gamma(d)) \\ & + (d-1) \cdot \sum_{i=1}^{m} \left\{ \psi\left[\eta(t_{i,n_i}; a^{(h)}, b^{(h)}) + d^{(h)} \right] - E\left\{ \ln\left(W_{i,n_i} + c^{(h)} \middle| \boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)} \right) \right\} \right\} \\ & - \sum_{i=1}^{m} E\left\{ \frac{\eta\left(t_{i,n_i}; a^{(h)}, b^{(h)}\right) + d^{(h)}}{W_{i,n_i} + c^{(h)}} \middle| \boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}^{(h)} \right\}, \end{aligned} \tag{2.104}$$

Therefore, denoting by $d^{(h+1)}$ the value of d that maximizes Eq. (2.104), the value $c^{(h+1)}$ that, given $d^{(h+1)}$, maximizes Eq. (2.101) is readily obtained by evaluating Eq. (2.103) at $d = d^{(h+1)}$.

2.8 The particle filter

In this manuscript, the particle filter algorithm (Doucet, Johansen, et al. (2009)) is used to generate random samples from the (joint) conditional distribution of $\boldsymbol{W}$ given $\boldsymbol{Z} = \boldsymbol{z}$. The data generated by using this procedure are then used to compute (empirically) the likelihood function in Eq. (2.63), the distribution of the RUL, and the conditional expectations requested in the M-step of the EM algorithm described in Chapter 2.7.
The procedure described below allows to obtain N pseudorandom realizations of $\boldsymbol{W}_i$ given $\boldsymbol{Z}_i = \boldsymbol{z}_i$. Thus, to generate pseudorandom realizations of $\boldsymbol{W}_i$ given $\boldsymbol{Z}_i = \boldsymbol{z}_i$ it is necessary to replicate its use for any $i = 1, \ldots, m$.
The method consists of a two-step sequence: a prediction step and an update step, that is iterated n_i times. To apply the procedure, it is necessary to assign a value to the parameter vector $\boldsymbol{\xi}$.

The particle filter is a very flexible approach that is able to handle all the perturbed processes adopted in this manuscript with only minor changes. Specifically, the conditional pdf of $\Delta W(t_{i,j-1}, t_{i,j})$ given $W(t_{i,j-1})$, in case the stochastic process under analysis is the gamma process with random effect described in Chapter 2.4 is the pdf in Eq. (2.49), whereas in the case of the gamma process and the inverse Gaussian process illustrated in Chapter 2.3 (given that both have independent increments) coincides with the (unconditional) one in Eq. (2.11) or in Eq. (2.12), respectively.
Similarly, the conditional pdf $f_{Z_{i,j}|W_{i,j}}(\cdot|\cdot)$ is the one in Eq. (2.21) if the model adopted for the error term is the inverse gamma or the one in Eq. (2.24) if it is the Gaussian (i.e., option 1) or 2) using the notation introduced in Chapter 2.3).

— Step 1 (prediction step), jth iteration:
for any $k = 1, \ldots, N$, generate a pseudorandom realization ${}_k\Delta w_{i,j}$ of $\Delta W(t_{i,j-1}, t_{i,j})$ given $W(t_{i,j-1}) = {}^{j-1}_{k}w_{i,j-1}$, where ${}^{j-1}_{k}w_{i,j-1}$ is the kth pseudorandom realization of $W(t_{i,j-1})$ given $\boldsymbol{Z}_{i,j-1} = \boldsymbol{z}_{i,j-1}$ generated at the $(j-1)$th iteration of this algorithm. Then, compute the term ${}^{j-1}_{k}w_{i,j} = {}_k\Delta w_{i,j} + {}^{j-1}_{k}w_{i,j-1}$ and append it to the particle vector ${}^{j-1}_{k}w_{i,1}, \ldots, {}^{j-1}_{k}w_{i,j-1}$ defined at the $(j-1)$th iteration.
The output of this prediction step is a set of N vectors:

$$\begin{matrix} {}^{j-1}_{1}w_{i,1}, \ldots, {}^{j-1}_{1}w_{i,j-1}, {}^{j-1}_{1}w_{i,j} \\ \vdots \\ {}^{j-1}_{N}w_{i,1}, \ldots, {}^{j-1}_{N}w_{i,j-1}, {}^{j-1}_{N}w_{i,j} \end{matrix}$$

that we will refer to as particles.

— Step 2 (update step), jth iteration:
for any $k = 1, \ldots, N$, compute the importance weight of the kth particle as:

$${}_kq_{i,j} = \frac{f_{Z_{i,j}|W_{i,j}}\left(z_{i,j}|\,{}^{j-1}_{k}w_{i,j}\right)}{\sum_{k=1}^{N} f_{Z_{i,j}|W_{i,j}}\left(z_{i,j}|\,{}^{j-1}_{k}w_{i,j}\right)}.$$

Hence, resample the set of particles:

$$
\begin{gathered}
{}^{j-1}_{1}w_{i,1}, \ldots, {}^{j-1}_{1}w_{i,j-1}, {}^{j-1}_{1}w_{i,j} \\
\vdots \\
{}^{j-1}_{N}w_{i,1}, \ldots, {}^{j-1}_{N}w_{i,j-1}, {}^{j-1}_{N}w_{i,j}
\end{gathered}
$$

according to their importance weights and rename the new particles (i.e., the vectors) as:

$$
\begin{gathered}
{}^{j}_{1}w_{i,1}, \ldots, {}^{j}_{1}w_{i,j-1}, {}^{j}_{1}w_{i,j} \\
\vdots \\
{}^{j}_{N}w_{i,1}, \ldots, {}^{j}_{N}w_{i,j-1}, {}^{j}_{N}w_{i,j}.
\end{gathered}
$$

For $j = 1$, to initialize the algorithm, in the first prediction step draw a pseudorandom sample of size N from $W(t_{i,1})$, denote its elements as ${}_{1}w_{i,1}, \ldots, {}_{N}w_{i,1}$ and define the particles as:

$$
\begin{gathered}
{}^{0}_{1}w_{i,1} = {}_{1}w_{i,1} \\
\vdots \\
{}^{0}_{N}w_{i,1} = {}_{N}w_{i,1}.
\end{gathered}
$$

The particle ${}^{j-1}_{k}w_{i,1}, \ldots, {}^{j-1}_{k}w_{i,j-1}, {}^{j-1}_{k}w_{i,j}$ should be intended as a pseudorandom realization of $\boldsymbol{W}_{i,j-1}$ given $\boldsymbol{Z}_{i,j-1} = \boldsymbol{z}_{i,j-1}$, and ${}^{j}_{k}w_{i,1}, \ldots, {}^{j}_{k}w_{i,j-1}, {}^{j}_{k}w_{i,j}$ should be intended as a pseudorandom realization of $\boldsymbol{W}_{i,j}$ given $\boldsymbol{Z}_{i,j} = \boldsymbol{z}_{i,j}$, where $\boldsymbol{W}_{i,j} = \{W(t_{i,1}), \ldots, W(t_{i,j})\}$, $\boldsymbol{Z}_{i,j} = \{Z(t_{i,1}), \ldots, Z(t_{i,j})\}$, and $\boldsymbol{z}_{i,j} = \{z(t_{i,1}), \ldots, z(t_{i,j})\}$ is the realization of $\boldsymbol{Z}_{i,j}$.
The conditional pdfs that are needed to compute the likelihood function in Eq. (2.63) can be approximated as:

$$
f_{Z_{i,j}|\boldsymbol{Z}_{i,j-1}}(z_{i,j}|\boldsymbol{z}_{i,j-1}) \cong \frac{\sum_{k=1}^{N} f_{Z_{i,j}|W_{i,j}}\left(z_{i,j}|\,{}^{j-1}_{k}w_{i,j}\right)}{N}
$$

where ${}^{j-1}_{k}w_{i,j}$ is the last component of the particle vector ${}^{j-1}_{k}w_{i,1}, \ldots, {}^{j-1}_{k}w_{i,j-1}, {}^{j-1}_{k}w_{i,j}$ generated at the jth prediction step. Likewise, for example, the conditional mean of a

function $g(\boldsymbol{W}_i)$ of $\boldsymbol{W}_i$ given $\boldsymbol{Z}_i = \boldsymbol{z}_i$ and $\boldsymbol{\xi}$, can be computed as:

$$E\{g(\boldsymbol{W}_i)|\boldsymbol{Z}_i = \boldsymbol{z}_i, \boldsymbol{\xi}\} \cong \frac{\sum_{k=1}^{N} g\left({}_{k}^{n_i}\boldsymbol{w}_i\right)}{N}$$

where ${}_{k}^{n_i}\boldsymbol{w}_i$ is the particle ${}_{k}^{n_i}w_{i,1}, \ldots, {}_{k}^{n_i}w_{i,j-1}, {}_{k}^{n_i}w_{i,j}$ generated at the n_ith update step, under a perturbed model whose parameter vector is set to $\boldsymbol{\xi}$. Obviously, the quality of these approximations improves with N.
This particle filter algorithm is also used to compute the ML estimate of the cdf of $RUL(t)$ $F_{RUL(t)|\boldsymbol{Z}(t)}\left(\tau|\boldsymbol{z}(t)\right)$ (2.67) and the ML estimate of the $MRUL(t)$ (2.69). In fact, more specifically, by using the notations introduced in Chapter 2.6, if $t_l \leq t < t_{l+1}$, so that the set $\boldsymbol{z}(t) = \{Z(t_j); j \geq z, t_j \leq t\}$ contains l perturbed measurements of the degradation level of a certain (selected) unit, given a pseudorandom sample of size N from $W(t_l)|\boldsymbol{Z}(t) = \boldsymbol{z}(t)$, say ${}_{1}^{l}w_l, \ldots, {}_{N}^{l}w_l$, the ML estimate of the cdf $F_{RUL(t)}(\tau|\boldsymbol{z}(t))$ (2.67) can be computed as:

$$F_{RUL(t)}(\tau|\boldsymbol{z}(t)) \cong 1 - \frac{\sum_{k=1}^{N} F_{W(t+\tau)|W(t_l)}(w_M|\,{}_{k}^{l}w_l)}{N},$$

where both the parameters of the perturbed model used to generate the particles and the parameters of the conditional cdf $F_{W(t+\tau)|W(t_l)}(\cdot|\cdot)$ are set to their ML estimates. Similarly, the ML estimate of the $MRUL(t)$ (2.69) is computed as:

$$MRUL(t) \cong \frac{\sum_{k=1}^{N}\int_0^{\infty} F_{W(t+\tau)|W(t_l)}(w_M|\,{}_{k}^{l}w_l)\cdot d\tau}{N} = \int_0^{\infty}\frac{\sum_{k=1}^{N} F_{W(t+\tau)|W(t_l)}(w_M|\,{}_{k}^{l}w_l)}{N}\cdot d\tau,$$

where the integral is calculated numerically.

2.9 Conclusions

Stochastic processes can be a powerful tool for the successful implementation of advanced maintenance strategies. In this chapter, three stochastic process-based degradation models have been introduced. Specifically, the perturbed gamma process (formulated as in Giorgio, Mele, and Pulcini (2019)), and the new perturbed inverse Gaussian process and perturbed gamma process with random effect. Given that all of these processes can account for measurement error, ad-hoc formulations of the likelihood function and of the remaining useful life, based on noisy data alone, have been provided.

Obviously, the presence of measurement error complicates the already difficult endeavor of calibrating a degradation model. To address this issue, a new EM algorithm (with the corresponding implementation for all considered processes) has been proposed. All of these algorithm (i.e., formulating the likelihood, the remaining useful life, and using the EM algorithm) rely on an effective filtering procedure for their successful implementation. To this end, the adopted implementation of the particle filter has been described.

In the following chapters, the effectiveness of these stochastic processes as degradation models will be demonstrated. Moreover, maintenance policies that take advantage of their use will be presented.

Chapter 3

Reliability application of the new gamma degradation process with random effect and state-dependent measurement error

3.1 Introduction

The gamma process has shown its effectiveness in describing a wide range of degradation phenomena. It is commonly used when degradation progresses gradually over time through small non-negative increments. However, it is widely acknowledged that the basic form of the gamma process only accounts for the inherent variability associated with the temporal progression of degradation in a specific unit, referred to as temporal variability. In practical settings, degradation paths of technological units often exhibit other forms of variability. For instance, degradation data collected through in-service and non-destructive inspection methods are frequently affected by measurement errors, which hinder the accurate observation of the actual degradation level. Additionally, even nominally identical units may display degradation patterns that differ due to the influence of several factors, both exogenous and endogenous, with values that can vary across units. In this chapter, we focus on experimental situations where both measurement errors and unit-to-unit variability are present. Specifically, we consider the case where the heterogeneity between the observed degradation paths of different units cannot be attributed to any specific cause. This form of variability is generally known as random effect and is typically modeled by assuming that one or more parameters of the underlying degradation process vary

randomly from unit to unit.

In order to address the presence of measurement errors many perturbed degradation models have been proposed in the literature. However, the existing gamma-based perturbed degradation models found in the literature are often not mathematically tractable. Computationally intensive methods are required to both estimate model parameters from perturbed data and perform predictions based on the observed data.

The main goal of this chapter is to illustrate the results of some examples of application of the gamma-based perturbed degradation model presented in Chapter 2.4. Obviously, perturbed models that incorporate random effects are inherently more complex than those that only account for temporal variability. Therefore, it is crucial to strike a balance between model accuracy and computational complexity in order to facilitate their practical use. Indeed, the driving idea of this chapter is to develop this extension without significantly increasing the computational burden of the original model or imposing additional restrictive constraints.
To describe the underlying degradation process in the new model, we adopt the mathematically tractable gamma process with random effect, initially introduced by Lawless and Crowder (2004) which has proven to be flexible and effective in previous studies (e.g., see Xiao Wang (2008), Elsayed and Liao (2004), Tsai, Tseng, and Balakrishnan (2012), Y.-J. Yang et al. (2014), and H. Hao, Su, C. Li, et al. (2015)). A key characteristic of the proposed model, that is shared with the one suggested in Giorgio, Mele, and Pulcini (2019), is that the model construction assumes that the measurement error is stochastically dependent on the actual degradation level and that, given the actual degradation level, it is conditionally distributed as a three-parameter inverse gamma random variable. This particular setup, extensively discussed and justified in Giorgio, Mele, and Pulcini (2019), distinguishes the proposed model from other perturbed degradation models found in the literature, which typically assume normally distributed errors and/or treat them as statistically independent of the underlying degradation process (see Lu, Pandey, and W.-C. Xie (2013), Kallen and Van Noortwijk (2005), Le Son, Fouladirad, and Barros (2016), Bordes, Paroissin, and Salami (2016), Whitmore (1995), C.-Y. Peng and Tseng (2009), R. Zhou, Gebraeel, and Serban (2012), Si et al. (2014), S. Hao, J. Yang, and Bérenguer (2019a), S. Hao, J. Yang, and Bérenguer (2019b)).

To demonstrate the affordability of the proposed approach, two practical examples are developed to showcase the effectiveness of the proposed model, based on perturbed degradation measurements of carbon-film resistors (the first) and on fuel cell membranes

(the second).

For the sake of clarity, the remainder of this chapter only reports the results of the mentioned application examples. Indeed, the detailed mathematical description of the model is presented in Chapter 2.4, the formulation of the likelihood function is discussed in Chapter 2.5, the cumulative distribution function of the remaining useful life is presented in Chapter 2.6, while the EM algorithm and the particle filter are presented in Chapters 2.7 and 2.8, respectively.

3.2 Applicative example I

In this section, the proposed model is applied to the degradation data of carbon-film resistors tested at 173° , given in Meeker and Escobar (1998). These data consist of $m = 10$ degradation paths, each one with $n_i = 4$ measurements, for a total of 40 measurements. Measurement times are the same for all units ($t_{i,j} = t_j,\ \forall i = 1, \ldots, m$). The degradation measures the percent increase in resistance (with respect to the initial value). When the resistance of a resistor increases by more than 8% (i.e., the degradation level passes the threshold value $w_M = 8$) then it is declared as failed.
The data are reported in Table 3.1, and present a direct clue of the presence of measurement error under the form of a negative increment (specifically, in the case of path #7). This prevents from using a gamma process without measurement error, as well as any other degradation process with intrinsically non-negative increments.
Moreover, the data also show empirical evidence of positive correlation between observed increments (specifically, five out of six empirical estimates of the correlation coefficient between increments are positive), which might suggest the presence of random effect.

The MLEs of the six parameters of the proposed perturbed process (here referred to as model M1) are reported in the first column of Table 3.2, together with the corresponding value of the log-likelihood function and the value of the Akaike Information Criterion (AIC) (see Akaike (1974)). For the sake of comparison, the second column of the same table reports the MLEs of the parameters, the value of the log-likelihood function at the MLEs of model parameter, and the value of the AIC index obtained under the five-parameter perturbed (fixed effect) gamma process proposed in Giorgio, Mele, and Pulcini (2019) (here referred to as model M2).

The values of the AIC index obtained under the competing models indicate that the proposed gamma process with random effect (M1) fits the considered carbon-film data

better than the (fixed effect) gamma model (M2). Actually, given that in the absence of random effect model M1 reduces to model M2, as remarked in Giorgio, Guida, and Pulcini (2018), model M2 can be considered as asymptotically nested into model M1. Hence, the null hypothesis (H_0), that the considered data follow the (simpler) model M2, can also be (more formally) checked against the alternative hypothesis (H_1) that they follow model M1, by using the likelihood ratio test. The test statistic is defined as $LR = -2 \cdot [\ln(\widehat{L}_0) - \ln(\widehat{L}_1)]$, where $\widehat{L}_0$ and $\widehat{L}_1$ denote the MLEs of the log-likelihood functions obtained under models M2 and M1, respectively. The null hypothesis is rejected for large values of LR. In the considered case, when H_0 holds, the test statistic is approximately distributed as a χ^2 random variable with 1 degree of freedom. From Table 3.2, it results $LR = 20.52$. Thus, by comparing the observed value 20.52 with the quantiles of the χ^2 distribution with one degree of freedom, it is apparent that the null hypothesis can be rejected at any plausible significance level.

Resistor number	Measurement times (h)			
	452	1030	4341	8084
1	0.87	1.29	2.62	4.44
2	1.25	1.88	3.54	5.23
3	2.64	3.78	7.01	11.12
4	0.98	1.36	2.66	4.42
5	1.62	2.34	3.82	6.14
6	1.59	2.41	3.46	6.75
7	2.29	2.24	6.30	8.34
8	0.98	1.37	2.47	3.74
9	1.04	1.54	2.77	4.16
10	1.19	1.59	3.03	4.52

Table 3.1 – Degradation data of carbon-film resistors at 173°.

Figure 3.1 shows the (perturbed) degradation paths of the considered carbon-film resistors and the empirical estimates of the mean of the perturbed degradation process at the measurement times together with the MLE of $E\{Z(t)\}$ (i.e., the mean function of the perturbed process evaluated at the MLEs of the model parameters). The same figure shows the MLEs of the 0.90 probability bands of $Z(t)$. Note that the MLE of $E\{Z(t)\}$ coincides with the MLE of $E\{W(t)\}$ (see Eq. (2.60)).

Figure 3.2 shows the empirical estimates of the variance of $Z(t)$ at the measurement times and the MLEs of the variance of $Z(t)$ and $W(t)$ obtained under models M1 and

	Model	
	$M1$	$M2$
MLE		
$\widehat{a}$	0.0636	2.78
$\widehat{b}$	0.466	0.455
$\widehat{\lambda}$	/	0.136
$\widehat{c}$	0.255	/
$\widehat{d}$	11.1	/
$\widehat{\varphi}$	542	842
$\widehat{\nu}$	3.52	4.07
AIC	36.02	44.27
$\ln\left(L\left(\widehat{\boldsymbol{\xi}}, \boldsymbol{z}\right)\right)$	-24.01	-34.27

Table 3.2 – MLEs of the parameters, AIC index, and log-likelihood under models M1 and M2.

M2.

Figures 3.1 and 3.2 show that model M1 fits the considered data satisfactorily. In fact, both the MLEs of $E\{Z(t)\}$ and $V\{Z(t)\}$ obtained under model M1 are very close to the corresponding empirical estimates. Likewise, the probability bands contain a fraction of experimental points that is coherent with their nominal probability content (indeed, by chance, it is $36/40 = 0.9$). The figures also give evidence that model M1 outperforms model M2 in terms of fitting ability, as it was already concluded by adopting the AIC criterion and the likelihood ratio test.

In order to investigate the influence of the presence of random effect on the estimates of the $RUL(t)$, as well as to evaluate the effect of neglecting its presence, let us consider the MLEs of the ccdf of the RUL (i.e., of the residual reliability function, the complementary of the cdf in Eq. (2.67)) reported in Figure 3.3. The curves represented in the figure are the MLEs of the ccdf of $RUL(8084)$ of resistors #5, #7, and #8. These ccdf are computed as in equation (2.67), by setting the model parameters at the corresponding MLEs. In particular, from (2.67) and related comments, the ccdf of $RUL(8084)$ of the resistor $\#i$ can be denoted by $\bar{F}_{RUL(8084)|\boldsymbol{Z}_{i,4}}(\tau|\boldsymbol{z}_{i,4}) = 1 - F_{RUL(8084)|\boldsymbol{Z}_{i,4}}(\tau|\boldsymbol{z}_{i,4})$.

The paths of the considered resistors are represented in Figure 3.4, where the solid dots indicate the empirical estimates of the mean function at the measurement times. From Figure 3.3, it is apparent that the MLE of the ccdf $\bar{F}_{RUL(8084)|\boldsymbol{Z}_{7,4}}(\tau|\boldsymbol{z}_{7,4})$ obtained under model M2 for the resistor #7, whose path (see Figure 3.4) is entirely above the

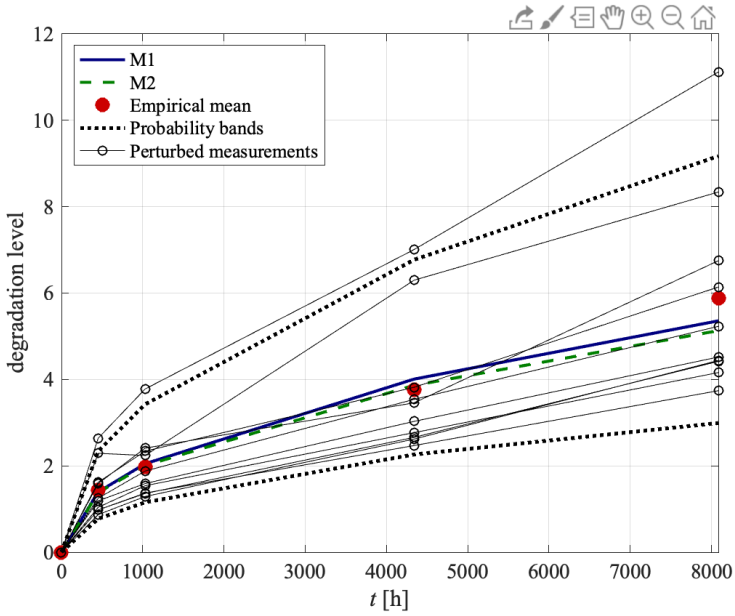

Figure 3.1 – Observed degradation paths of the carbon-film resistors (empty dots connected by thin solid lines), empirical estimate of $E\{Z(t)\}$ at measurement times (solid dots), MLEs of $E\{W(t)\} = E\{Z(t)\}$ obtained under model M1 (solid line) and M2 (dashed line), and MLE of 0.90 probability bands of $Z(t)$ obtained under model M1 (dotted lines).

empirical mean, is much more optimistic than the corresponding estimate obtained under (the more appropriate) model M1. In a specular way, it is also apparent that the MLE of the ccdf $\bar{F}_{RUL(8084)|\boldsymbol{Z}_{8,4}}(\tau|\boldsymbol{z}_{8,4})$ obtained under model M2 for resistor #8, whose path is below the empirical mean, is much more pessimistic than the one obtained under (the more appropriate) model M1. Indeed, the difference between the estimates provided by models M1 and M2 for a given resistor increases with the distance existing between its (observed) degradation path and the empirical mean. In fact, for example, Figure 3.3 also shows that the MLEs of the ccdf $\bar{F}_{RUL(8084)|\boldsymbol{Z}_{5,4}}(\tau|\boldsymbol{z}_{5,4})$ of the resistor #5 (see Figure 3.4) obtained under models M1 and M2 are relatively close one to the other.

This behavior is confirmed by the MLEs of the $MRUL(8084)$ of resistors reported in Table 3.3, that are computed as in Eq. (2.69), by setting the model parameters at the corresponding MLEs. In fact, from the table, it is apparent that the MLEs of the mean remaining useful life obtained under model M2 are larger than those obtained under model M1 in the case of resistors whose degradation progresses faster than the mean, being smaller in the case of those whose degradation progresses slower. At the same time,

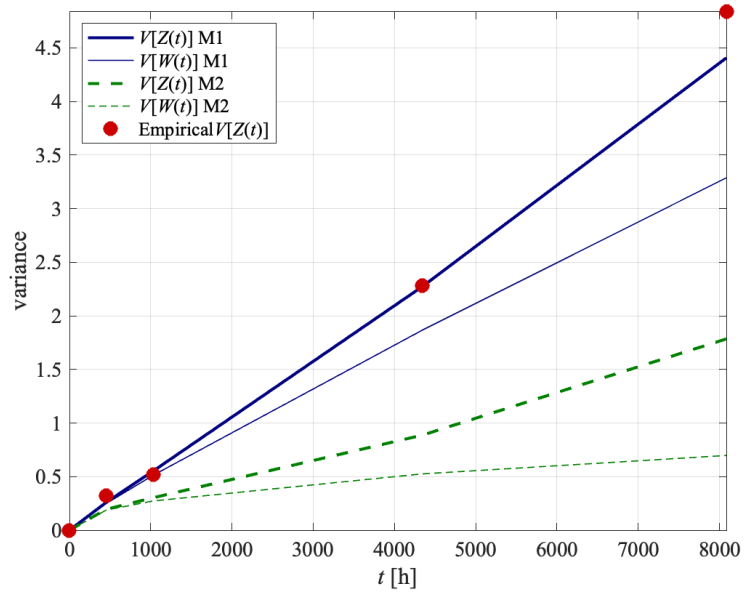

Figure 3.2 – Empirical estimates of $V\{Z(t)\}$ at measurement times (solid dots) together with the MLEs of $V\{Z(t)\}$ and $V\{W(t)\}$ obtained under model M1 (bold and thin solid lines) and M2 (bold and thin dotted lines).

	Resistor #									
	1	2	3	4	5	6	7	8	9	10
M1	32,165	14,861	3.85	29,774	7,927	7,904	1,419	35,867	27,465	22,207
M2	19,698	14,613	2,815	19,537	11,909	11,535	6,672	2,209	19,719	17,858

Table 3.3 – MLEs of the MRUL of the carbon-film resistors evaluated under model M1 and M2 at $t_4 = 8084$.

the difference between the estimates obtained under the considered models for a specified resistor is relatively small if its degradation path is close to the mean. From the values reported in Table 3.3, it is also apparent that the MRUL estimates obtained under model M1 vary from resistor to resistor more than those obtained under model M2.

Indeed, the difference existing among the MLEs of MRUL (and/or among those of the ccdf of RUL) obtained for different resistors under a given model, depends both on the actual state of the resistors (i.e., on the gap existing between their true degradation level and the threshold limit) and on the rapidity with which their degradation level is expected to progress over future time. Under the considered perturbed models, all the information available at time t_4 about the current true degradation state of the resistor

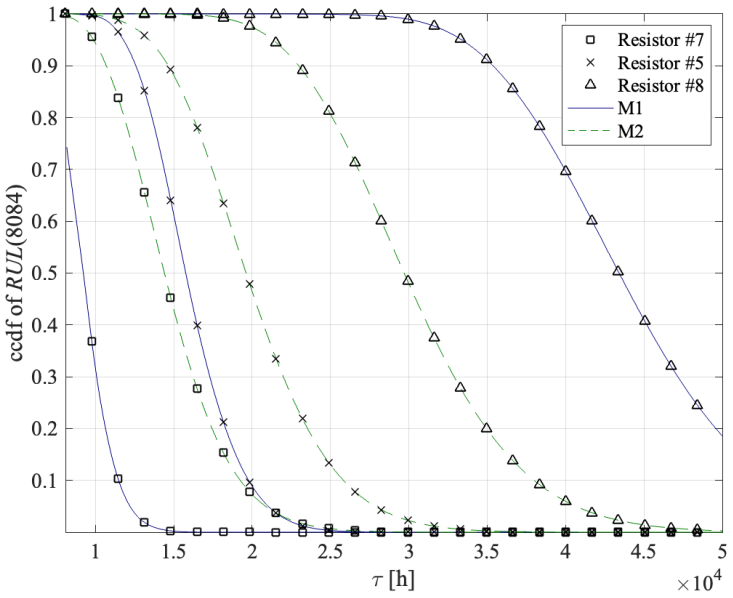

Figure 3.3 – MLEs of the conditional ccdf of $RUL(8084)$ of carbon-film resistors #5, #7, and #8 under models M1 (solid lines) and M2 (dashed lines), given the past perturbed measurements available for each resistor at time $t_4 = 8084$.

$\#i$ and its future degradation growth is contained in the conditional pdfs $f_{W(t_4)|\boldsymbol{Z}_{i,4}}(w|\boldsymbol{z}_{i,4})$ and $f_{\Delta W(t_4,t_4+\Delta t)|\boldsymbol{Z}_{i,4}}(\Delta w|\boldsymbol{z}_{i,4})$, respectively. Hence, to better understand how the presence of random effect impacts on the MLEs of the mean and ccdf of the remaining useful life, it is worth to see how it influences the MLEs of these pdfs. Figure 3.5 shows the MLEs of the conditional pdfs $f_{W(t_4)|\boldsymbol{Z}_{5,4}}(w|\boldsymbol{z}_{5,4})$, $f_{W(t_4)|\boldsymbol{Z}_{7,4}}(w|\boldsymbol{z}_{7,4})$, and $f_{W(t_4)|\boldsymbol{Z}_{8,4}}(w|\boldsymbol{z}_{8,4})$ of the degradation state of resistors #5, #7, and #8 at $t_4 = 8084$, obtained under models M1 and M2. The figure shows that the MLEs of these pdfs obtained under the model M1 differ (one from the others) more than those obtained under model M2. This result was expected because the (fixed effect) model M2 is less prone than model M1 to tolerate differences among the (hidden) degradation levels of the resistors. In fact, as it is shown in Figure 3.2, the MLE of $V\{W(t)\}$ obtained under model M2 is much smaller than the one obtained under model M1. Furthermore, it also results that the pdfs estimated under model M2 are more dispersed (in terms of coefficient of variation) than the corresponding ones obtained under model M1. This latter result can be intuitively understood by considering that, to explain the substantial difference existing between the empirical estimate of $V\{Z(t)\}$ and the MLE of $V\{W(t)\}$ (see Figure 3.2), model M2 overestimates the variance of the measurement error.

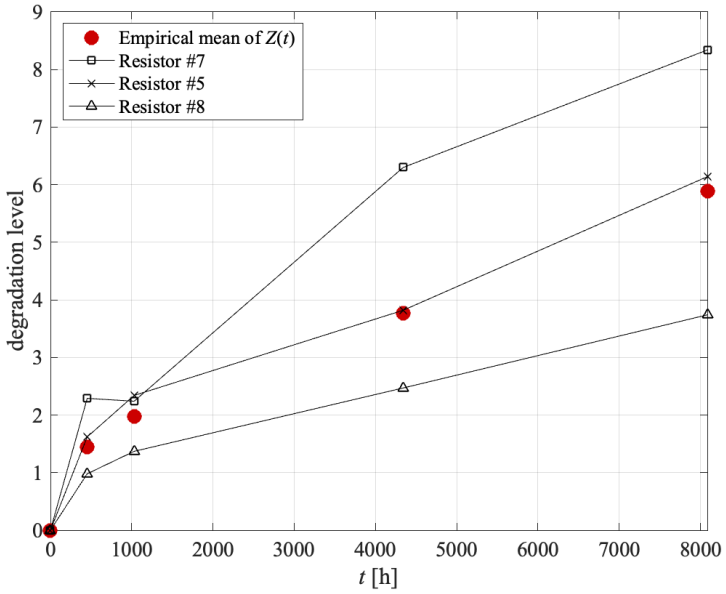

Figure 3.4 – Degradation paths of resistors #5, #7, and #8 and empirical estimates of $E\{W(t)\}$ at the measurement times.

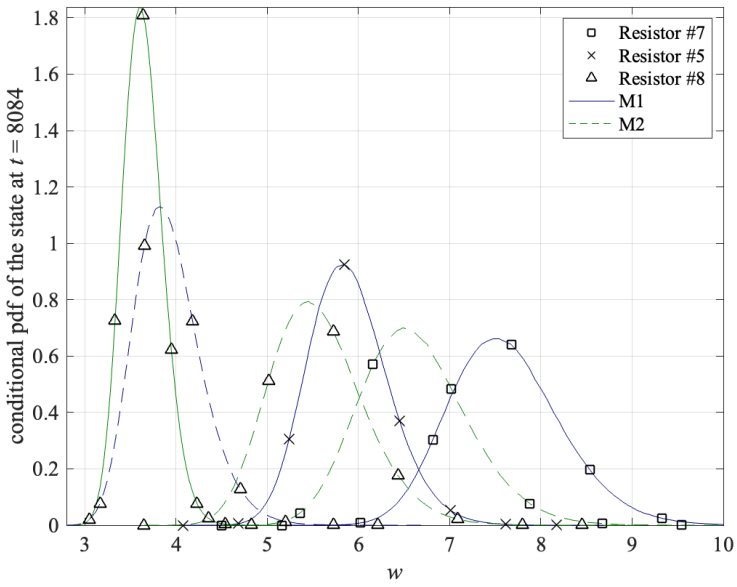

Figure 3.5 – MLEs of the conditional pdfs of the degradation state of the resistors #5, #7, and #8 at $t_4 = 8084$, given the perturbed measurements available at t_4 obtained under models M1 (solid lines) and M2 (dashed lines).

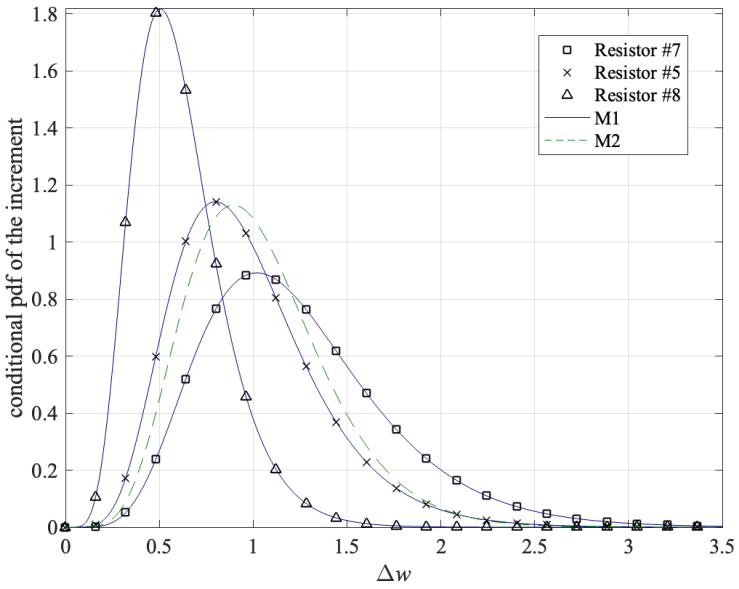

Figure 3.6 – MLEs of the conditional pdfs of the degradation increments $W(8084, 8084 + 4000)$ of the resistors #5, #7, and #8, obtained under models M1 (solid lines) and M2 (dashed line), given the perturbed measurements available at $t_4 = 8084$.

Figure 3.6 shows the MLEs of the conditional pdfs $f_{\Delta W(t_4,t_4+\Delta t)|\boldsymbol{Z}_{5,4}}(\Delta w|\boldsymbol{z}_{5,4})$, $f_{\Delta W(t_4,t_4+\Delta t)|\boldsymbol{Z}_{7,4}}(\Delta w|\boldsymbol{z}_{7,4})$, and $f_{\Delta W(t_4,t_4+\Delta t)|\boldsymbol{Z}_{8,4}}(\Delta w|\boldsymbol{z}_{8,4})$ of the degradation increment $\Delta W(t_4, t_4+\Delta t)$ of the resistors #5, #7, and #8, over the future time interval $(t_4, t_4+\Delta t)$, where $t_4 = 8084$ and (without loss of generality) $\Delta t = 4000$. Note that, under model M2, given that the hidden process has independent increments, the estimated pdfs of different resistors (dashed line) coincide with each other.

The figure confirms that the increments predicted under model M1 are larger or smaller with respect to those predicted under model M2 depending on the resistor. In fact, in the presence of unit to unit variability, the degradation process of resistor #7, whose observed path is above the empirical mean, is expected to progress faster than the degradation process of resistor #8, whose path is below the mean. Furthermore, it is also worth to note that the conditional pdf of the increment obtained for resistor #5 under the model M2 is very similar to the one obtained for the same resistor under the model M1. In fact, the path of the resistor #5 is close to the empirical mean.

3.3 Applicative example 2

The second example is developed by applying the proposed model to the microcrack growth data reported in Table 3.4. This dataset, given in Elsayed (2021), consists of the microcrack propagation paths of 10 fuel cell membranes. Measurements are taken at six different measurement times. The microcrack size is expressed in 10^{-6} *in* (inches). The measurement unit of time is not specified. The membranes are considered failed when the size of the microcrack exceeds the threshold $10,000 \times 10^{-6}$ *in*. By following Elsayed (2021), we assume that the microcrack propagation is monotone increasing and that data are affected by measurement error.

Fuel cell membrane #	Measurement times (h)					
	5	10	15	20	25	30
1	200	647	1507	1190	1651	1980
2	270	605	933	2738	4091	2444
3	137	566	1641	2332	1518	1499
4	179	1508	1127	719	1630	1080
5	282	640	799	2585	1424	1570
6	147	873	831	2520	1730	3125
7	19	1308	744	677	1767	1090
8	322	952	846	1662	4724	5727
9	286	1669	755	490	2537	1779
10	137	799	917	1856	2485	1566

Table 3.4 – Microcrack growth of fuel cell membranes.

The MLEs of the parameters of the proposed model (M1) and for the perturbed gamma model without random effect (M2) are reported in the first and second column of Table 3.5, respectively, together with the corresponding values of the log-likelihood function and the values of the AIC index.

The values of the AIC index reported in Table 3.5 shows that in this case the Akaike information criterion leads to prefer model M2. This conclusion is confirmed by the likelihood ratio test. In fact, from the results reported in the same table, it is easy to verify that in this case the test statistic $LR = -2 \cdot [\ln(\widehat{L}_0) - \ln(\widehat{L}_1)]$, where $\widehat{L}_0$ and $\widehat{L}_1$ denote the MLEs of the log-likelihood functions obtained under models M2 and M1, is equal to 0.49, a value that does not allow to reject the null hypothesis (that the considered data follow model M2) at any plausible level of significance. Figures 3.7 and 3.8 make apparent that both models M1 and M2 fit the considered data satisfactorily and also that, in this

	Model	
	$M1$	$M2$
MLE		
$\widehat{a}$	1.27	1.32
$\widehat{b}$	1.16	1.16
$\widehat{\lambda}$	/	67.3
$\widehat{c}$	2274	/
$\widehat{d}$	35.89	/
$\widehat{\varphi}$	0.0509	0.0731
$\widehat{\nu}$	1.42	1.47
AIC	935.76	933.26
$\ln\left(L\left(\widehat{\boldsymbol{\xi}}\right)\right)$	-461.89	-461.63

Table 3.5 – MLEs of the parameters, AIC index, and log-likelihood under the considered competing models.

case, their fitting abilities are comparable, confirming that model M1 provides a better trade-off between fitting ability and model complexity than model M2, for these data.

From Figure 3.10 and Table 3.6, it is apparent that the MLEs of the ccdf and mean of $RUL(30)$ obtained under models M1 and M2 for membranes #1, #4, and #8 (see Figure 3.10) are much more similar to each other than the MLEs of the ccdf of $RUL(8084)$ and $MRUL(8084)$ of resistors #5, #7, and #8 shown in Figure 3.3 and Table 3.3, respectively. Indeed, the estimates obtained for the resistors vary both from unit to unit (given the model) and from one model to the other (given the unit) much more than those obtained for the membranes. Similarly, Figures 3.11 and 3.12 show that, in this case, also

	Resistor #									
	1	2	3	4	5	6	7	8	9	10
M1	69.1	59.2	71.5	79.5	71.9	63.1	83.4	55.6	76.5	68.6
M2	68.9	66.8	69.4	71.6	69.8	67.7	73.0	65.1	70.8	68.9

Table 3.6 – MLEs of the MRUL of the fuel cell membranes evaluated under model M1 and M2 at $t_6 = 30$.

the MLEs of the conditional pdfs $f_{W(t_6)|\boldsymbol{Z}_{1,6}}(w|\boldsymbol{z}_{1,6})$, $f_{W(t_6)|\boldsymbol{Z}_{4,6}}(w|\boldsymbol{z}_{4,6})$, $f_{W(t_6)|\boldsymbol{Z}_{8,6}}(w|\boldsymbol{z}_{8,6})$, $f_{\Delta W(t_6,t_6+\Delta t)|\boldsymbol{Z}_{1,6}}(\Delta w|\boldsymbol{z}_{1,6})$, $f_{\Delta W(t_6,t_6+\Delta t)|\boldsymbol{Z}_{4,6}}(\Delta w|\boldsymbol{z}_{4,6})$, and $f_{\Delta W(t_6,t_6+\Delta t)|\boldsymbol{Z}_{8,6}}(\Delta w|\boldsymbol{z}_{8,6})$, obtained under model M1 and M2 are much more similar to each other than the corresponding estimates represented in Figures 3.5 and 3.6.

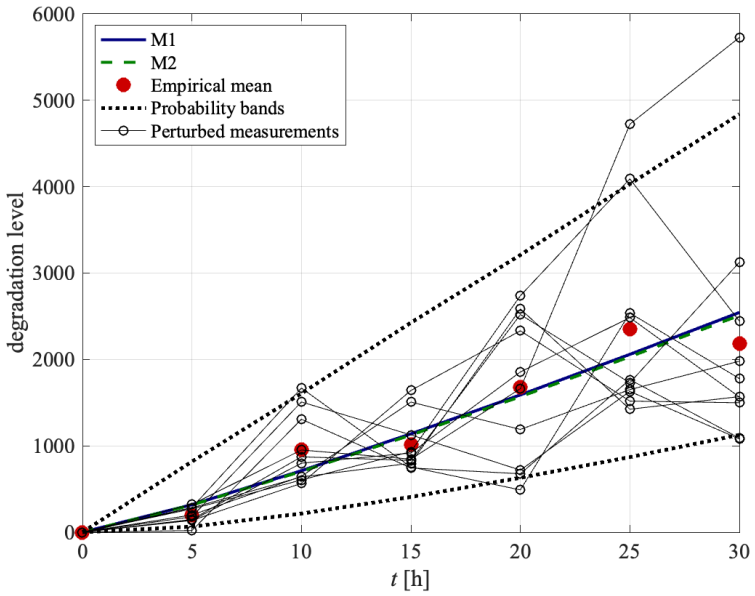

Figure 3.7 – Observed paths of fuel cell membranes (empty dots connected by thin solid lines), empirical estimate of $E\{Z(t)\}$ at measurement times (solid dots), MLEs of $E\{W(t)\} = E\{Z(t)\}$ obtained under model M1 (solid line) and M2 (dashed line), and MLE of 0.90 probability bands of $Z(t)$ obtained under model M1 (dotted lines).

Obviously, this result is also due to the circumstance that the paths of membranes #1, #4, and #8 represented in Figure 10 are more similar (one to the others) than the paths of resistors #5, #7, and #8, represented in Figure 3.4. On the other hand, as it is apparent from Figure 3.9 in this application there are no paths that are completely above or completely below the mean function. In fact, this is exactly the reason why according to the Akaike information criterion and the likelihood ratio test, in this application, there is not a significant presence of random effect.

3.4 Conclusions

In this chapter, the new degradation model described in Chapter 2.4, that accounts for the joint presence of measurement error and random effect, has been applied. Under this model, the degradation path of a specific unit is modeled as a gamma process. The presence of random effect is accounted for by assuming that the rate parameter of the unit specific gamma process varies randomly from unit to unit. In fact, the rate parameters of different units are assumed to be (unobservable) realizations of independent and identi-

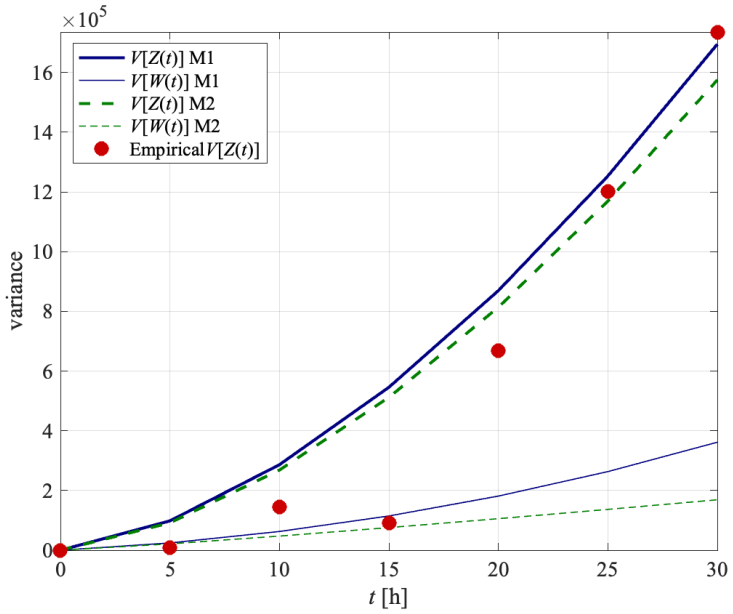

Figure 3.8 – Empirical estimates of $V\{Z(t)\}$ at measurement times (solid dots) together with the MLEs of $V\{Z(t)\}$ and $V\{W(t)\}$ obtained under model M1 (bold and thin solid lines) and M2 (bold and thin dotted lines).

cally distributed gamma random variables. Finally, to complete the model, it is assumed that the measurement error depends on the hidden measured degradation level, and that, given the actual degradation level, it is conditionally distributed as a three-parameter inverse gamma random variable.

Under this setting, directly maximizing the likelihood function has proven to be a challenging task (especially in terms of computational time) and therefore a procedure that combines a particle filter and an expectation-maximization algorithm has been suggested, that allows to solve this numerical issues. Moreover, a simple algorithm based on the same particle filter has been also adopted to compute (perturbed measurement-based) estimates of functions of the model parameters, such as the cumulative distribution function of the remaining useful life and the conditional (perturbed measurement-based) probability density functions of the true (hidden) degradation level. To demonstrate the affordability of the proposed estimation procedure, two applicative examples have been developed on the basis of two sets of degradation measurements of carbon-film resistors and fuel cell membranes, respectively.

In the first example, the presence of a random effect was found to be statistically significant, whereas in the second example it was not. The results obtained indicate that

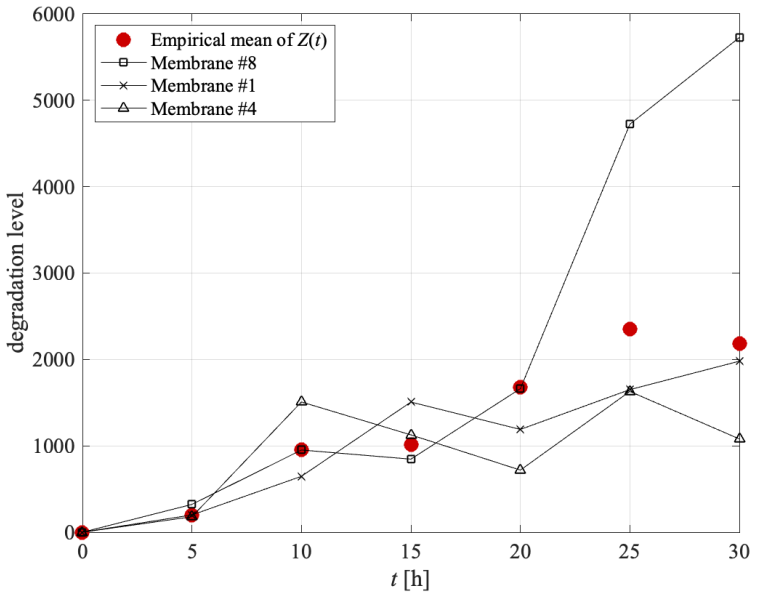

Figure 3.9 – Degradation paths of membranes #1, #4, and #8 and empirical estimates of $E\{W(t)\}$ at the measurement times.

the proposed model fits the first dataset significantly better than the perturbed gamma process suggested in Giorgio, Mele, and Pulcini (2019). On the other hand, both models adequately fit the second dataset, with the fixed effect model offering a better balance between fitting ability and complexity. The fitting ability of the two competing models was evaluated using the Akaike information criterion and likelihood ratio test. By comparing the estimates of cumulative distribution function and mean of the remaining useful life under the two models, we observed that while the estimates provided by the competing models were quite similar for the second dataset, the fixed effect model yielded noticeably poorer estimates than the proposed perturbed gamma model with random effect for the first dataset.

These examples show that neglecting the presence of random effect in the cases where it is significant can lead to significantly poorer prediction performances. In the context of maintenance decision-making then, intuition would suggest that when random effect is thought to be present, developing an ad-hoc maintenance strategy that directly takes it into account can lead to better results. Some examples of such strategies will be illustrated in Chapter 5.

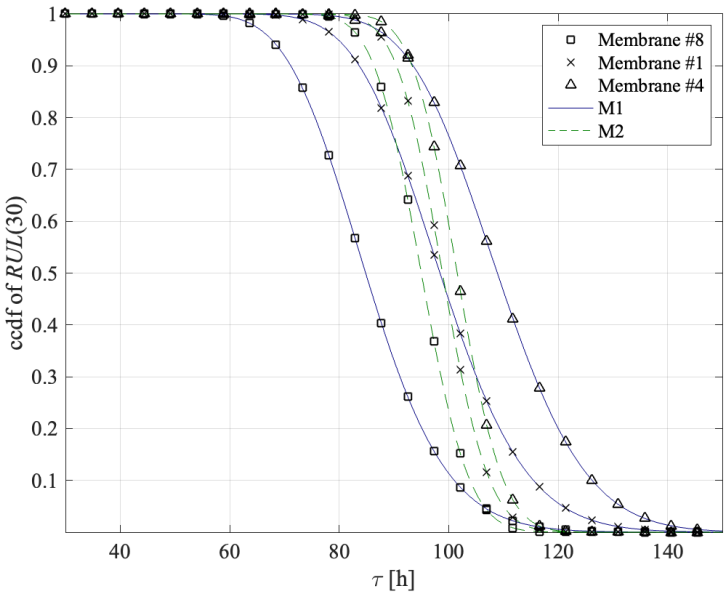

Figure 3.10 – MLEs of the conditional ccdf of $RUL(30)$ of fuel cell membranes #1, #4, and #8 under models M1 (solid lines) and M2 (dashed lines), given the past perturbed measurements available for each resistor at time $t_6 = 30$.

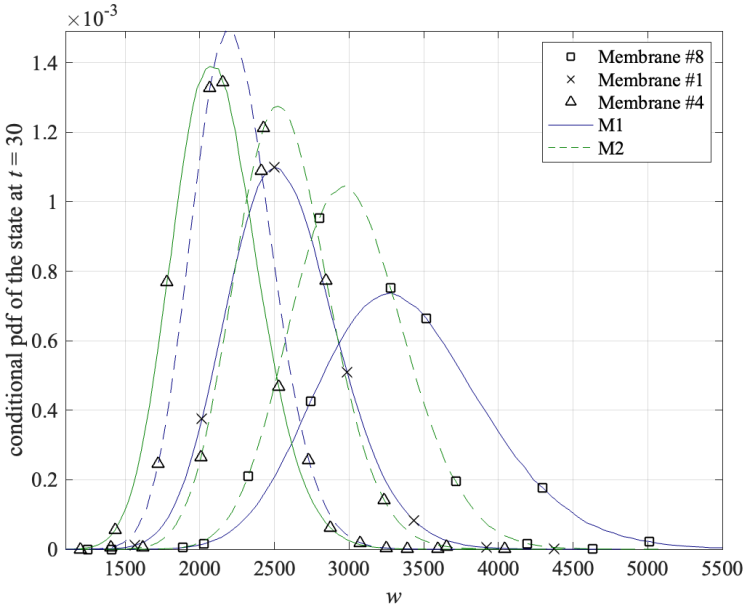

Figure 3.11 – MLEs of the conditional pdfs of the degradation state of the membranes #1, #4, and #8at $t_6 = 30$,obtained under models M1 (solid lines) and M2 (dashed lines).

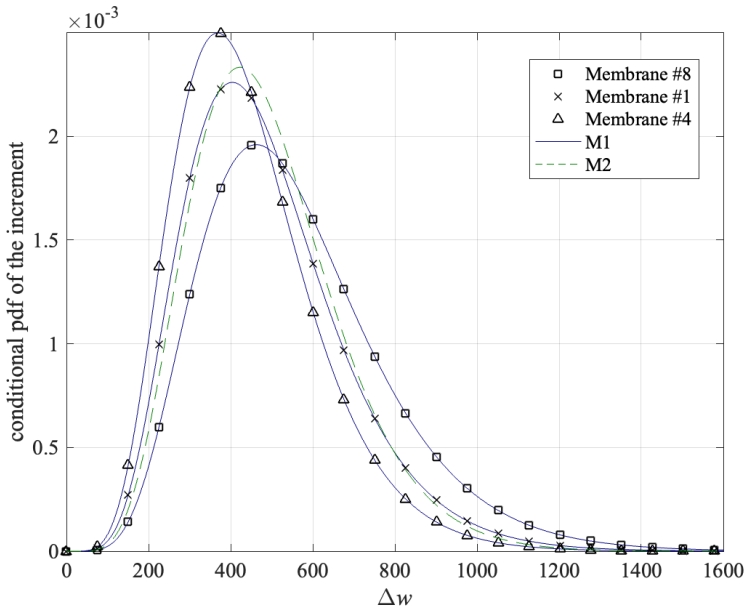

Figure 3.12 – MLEs of the conditional pdfs of the degradation increments $W(30, 30+5)$ of the membranes #5, #7, and #8, obtained under models M1 (solid lines) and M2 (dashed line), given the perturbed measurements available at $t_4 = 8084$.

Chapter 4

Misspecification analysis of gamma- and inverse Gaussian-based perturbed degradation processes

4.1 Introduction

In the previous chapters, we tackled the problem of the joint presence of temporal variability, unit-to-unit variability and measurement errors, which can all be seen as forms of aleatory uncertainty. In this chapter, we will focus our attention on what could be seen as epistemic uncertainty. In fact, the goal of this chapter is to investigate the misspecification issue of two stochastic processes that present very similar statistical features, the gamma and inverse Gaussian process. Both are widely applied in engineering and reliability, offering natural choices for modeling monotonic degradation phenomena (e.g., see Abdel-Hameed (1975), Van Noortwijk (2009), Wu et al. (2018), Xiaofei Wang et al. (2021), Wasan (1968), Xiao Wang and Xu (2010), Ye and N. Chen (2014), C.-Y. Peng (2015), L. A. Rodríguez-Picón, A. P. Rodríguez-Picón, and Alvarado-Iniesta (2019), Morita et al. (2021), Ye and M. Xie (2015), Kahle, Mercier, and Paroissin (2016)). Despite not being fully equivalent, they are often treated as such.

This situation makes the model misspecification issue of gamma and inverse Gaussian processes interesting and important. However, distinguishing between these models can be challenging, especially in practical applications where only noisy data can be gathered (as it often is when in-service and non-destructive inspections methods are implemented, e.g., Lu, Pandey, and W.-C. Xie (2013)).

Indeed, when measurement error is present, it can hinder the direct evaluation of a model's ability to describe the underlying degradation process, crucial in reliability and mainte-

nance applications. However, despite its importance, this misspecification problem has received limited attention in the literature, particularly when addressing situations with measurement errors. The basic issue (i.e., without measurement error) has been addressed in Zhang and Revie (2016) and Tseng and Yao (2017), but, to the best of our knowledge, no papers are available in the literature that tackle the misspecification issue in the presence of measurement error.

To address this gap, we conducted a comprehensive Monte Carlo study focusing on how the presence of measurement error impacts the misspecification issue of gamma and inverse Gaussian processes. The study has been carried out considered as competing models the perturbed Gamma process proposed by Giorgio, Mele, and Pulcini (2019) and a new perturbed inverse Gaussian process. Both competing models share the same modeling solution for the measurement error. Moreover, in order to facilitate the comparative analysis, the inverse Gaussian process has been reparameterized such that the considered competing models share the same parameters and the same functional forms of the mean and variance functions.

As in Giorgio, Mele, and Pulcini (2019), the error term is supposed to depend (in stochastic sense) on the hidden degradation level and, conditionally to the degradation level, is modeled as a 3 parameter inverse gamma random variable. This modeling solution distinguishes the considered models from other perturbed degradation models suggested in the literature, where the error is modeled by using a Gaussian distribution and/or is assumed to be stochastically independent of the hidden degradation process (e.g., Lu, Pandey, and W.-C. Xie (2013), Le Son, Fouladirad, and Barros (2016), Bordes, Paroissin, and Salami (2016), S. Hao, J. Yang, and Bérenguer (2019a), X. Chen et al. (2019)). Moreover, it ensures that the perturbed measurement is non negative, a result that is not guaranteed in the case where the error term is described by using the Gaussian model, especially when the magnitude of the standard deviation of the error term is comparable to that of the measured degradation level. Nonetheless, for the sake of generality, as an alternative modeling solution, we have also examined the case where the error term is modeled by using a Gaussian distribution. Also in this second case, by following Pulcini (2016), we have assumed that the measurement error depends in stochastic sense on the measured degradation level. Yet, as a special case under the proposed model, we have also considered the classical assumption where the error is independent of the hidden degradation process. For the sake of readability, the detailed formulation of these processes is not illustrated in this Chapter, but is available in Chapter 2.3.

Both the misspecification of a perturbed gamma process with a perturbed inverse Gaussian one and the symmetric case of the misspecification of a perturbed inverse Gaussian process with a perturbed gamma one are considered. Model parameters are estimated from perturbed data by using the maximum likelihood (ML) method.
The fitting ability of the considered competing perturbed models is evaluated by using the Akaike information criterion (AIC) (see, e.g., Akaike (1974)). The risk of incurring in a misspecification is evaluated as percentage of times the AIC leads to select the wrong model. The severity of a misspecification is evaluated in terms of its impact on ML estimates of reliability and mean remaining useful life. The impact of the presence of measurement error on risk and consequences of incurring in a misspecification is evaluated by comparing the obtained results with those obtained by carrying out the same misspecification analysis in the absence of measurement errors.

Unfortunately, as it also occurs in the case of other gamma and inverse Gaussian based perturbed degradation models, computing the likelihood functions, which are not available in closed form, requires intensive numerical methods that, at the same time, increase the computational burden and exacerbate convergence issues of numerical optimization algorithms used to retrieve the ML estimates. Indeed, mainly due to numerical problems and long computational times, the performance of ML estimators of parameters of the gamma and inverse Gaussian based perturbed degradation processes, and/or functions thereof, are typically investigated by using a relatively small number of synthetic datasets, a situation that does not allow to obtain results characterized by an adequate degree of accuracy.

In order to overcome this limitation, we suggest and adopt a new sequential Monte Carlo Expectation Maximization (EM) algorithm, which allows to drastically simplify the estimation task and (consequently) to find a better tradeoff between precision and computational affordability of the Monte Carlo study (again for the sake of clarity, the mathematical details about this procedure is not illustrated here but can be found in Chapter 2.7). In fact, the use of this algorithm allows us to perform a Monte Carlo study where each index of interest is evaluated on the basis of $2,000$ simulated datasets.
Nevertheless, the misspecification study is still a very time-consuming task, even when using this new algorithm. This is mainly because to correctly apply the Akaike information criterion a very accurate evaluation of the ML estimates of the 5-parameter log-likelihood function is required. In the considered misspecification analysis, we observed that in about 6% of cases the likelihood function of the competing models differ only in the fourth

significant figure and more rarely (in less than 1% of cases) only in the fifth significant figure. Obtaining this accuracy demands (in the considered setup) a very high number of particles (say, 500,000 or more), which significantly affects the computational burden.

4.2 Misspecification analysis

We address two misspecification issues: namely, the misspecification of a perturbed gamma process (PGP) with a perturbed inverse Gaussian process (PIGP) and the symmetric case of the misspecification of a PIGP with a PGP. To this aim, we have developed a large Monte Carlo study where three realistic experimental scenarios are simulated by using the setups described in Table 4.1. As already mentioned in Chapter 2.3, the age function is modeled by using the widely adopted power law function $\eta(t) = (t/a)^b$.

Setup	a	b	θ	φ	ν
$\mathcal{A}$	1.25	1	1.25	8	1
$\mathcal{B}$	1	1	1	10	1
$\mathcal{C}$	0.5	1	0.5	20	1

Table 4.1 – Setups $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ used to generate the datasets.

In particular, this specific choice of the parameters φ and ν allows (see Eq. (2.36)) to calibrate the error term so that the ratio between the variance of the perturbed and hidden processes (i.e., $V\{Z(t)\}/V\{W(t)\}$) is time independent and equal to $1+1/(\theta\cdot\varphi) = 1.10$, under all the considered setups. This means that the variance of the error term depends on the magnitude of the measurement degradation level and that, due to the presence of measurement errors, ($\forall t > 0$) the variance of $Z(t)$ is 10% higher than the variance of $W(t)$.

We have firstly investigated the case where the error term is modeled by using option 1). Under each setup, we have used the true model (i.e., either the PGP or the PIGP, depending on the misspecification issue of concern) to generate $N_t = 2000$ synthetic datasets. Each dataset consists of $m = 6$ degrading paths, which simulate the evolution of the perturbed degradation levels of as many degrading units. Each path consists of $n_i = 6$ perturbed measurements, taken at equally spaced inspection times ($t_1 = 1$, $t_2 = 2$, $t_3 = 3$, $t_4 = 4$, $t_5 = 5$, and $t_6 = 6$, expressed in time units) that are the same for all the units (that is: $t_{i,j} = t_j = j$, $\forall\, i, j$, $i = 1, \ldots, m$ and $j = 1, \ldots, 6$). Together with each perturbed measurement we also kept the value of the measured (hidden) degradation level, which is

generated (as an intermediate result) by the algorithm that we have adopted to generate the perturbed data. The true values of the hidden (i.e., measured) degradation levels are used to perform comparative analyses.

Note that, under all the considered settings the hidden processes are homogeneous (i.e., the parameter b is always set to 1). Hence, given that they also have independent increments, and that measurement times are equally spaced, the hidden increments $W(t_1)$, $\Delta W(t_1, t_2), \ldots, \Delta W(t_5, t_6)$ (where $W(t_1) \equiv \Delta W(t_0, t_1)$) of any considered degradation path are always independent and identically distributed. Moreover, as remarked in Chapter 2.3, given the setup, the mean and variance of these increments do not depend on the considered (i.e., either gamma or inverse Gaussian) hidden model. These convenient settings allow to create, without loss of generality, a direct and easily interpretable link between the risk of misspecification and the parameter a of the age function.

In fact, considered that, both in the case of the PGP and the PIGP, as $\Delta\eta(t_{j-1}, t_j)$ increases, due to the central limit theorem, the increment $\Delta W(t_{j-1}, t_j)$ tends in distribution to a Gaussian random variable, it is reasonable to expect that, under the considered settings, the similarity between the distributions of the mentioned increments, and consequently the risk of incurring in a misspecification, increases with $\Delta\eta(t_{j-1}, t_j)$ and thus, being $\Delta\eta(t_{j-1}, t_j) = a^{-1}$, decreases with a. Indeed, $\Delta\eta(t_{j-1}, t_j)$ is equal to 0.8 under the setup $\mathcal{A}$, to 1 under the setup $\mathcal{B}$, and to 2 under the setup $\mathcal{C}$. This situation is clearly highlighted in Figure 4.1, where the solid lines (in red) are the pdfs of the mentioned increments under the PGP and the dashed lines (in blue) are the corresponding pdfs under the PIGP. From Figure 4.1 it is evident that the similarity between the two increments increases moving from setup $\mathcal{A}$ to setup $\mathcal{C}$.
In fact, the setups in Table 4.1 have been specially designed to simulate scenarios with increasing (i.e., from $\mathcal{A}$ to $\mathcal{C}$) risk of incurring in a model misspecification. The values of θ have been calibrated so that the mean function of the PGP and PIGP (as well as the mean function of the corresponding hidden models) is the same under all considered setups. The variance of $Z(t)$ and $W(t)$, for any given $t > 0$, decreases moving from setup $\mathcal{A}$ to setup $\mathcal{C}$. Each dataset has been used to perform ML estimates of the parameters $(a, b, \theta, \varphi, \nu)$ of both the PGP and PIGP. Then, dataset by dataset, the AIC has been used to select the model that provides the best fit for that dataset. A misspecification is assumed to occur when the AIC leads to select the wrong model (that is, the PIGP in the case of a dataset generated by using the PGP and the PGP in the case of a dataset generated by using the PIGP). Under each considered setup, results obtained by performing these analyses have

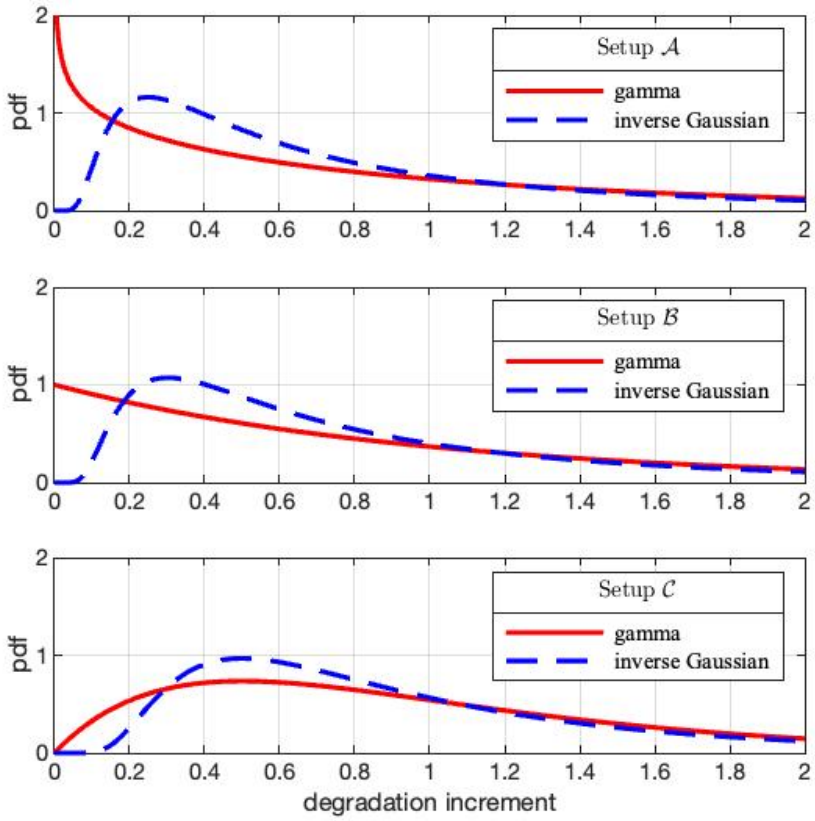

Figure 4.1 – Pdf of the increments $\Delta W(t_{j-1}, t_j)$ of the considered hidden competing processes under the adopted setups.

been used to evaluate the percent risk of misspecification as:

$$r_m\% = \left(\frac{N_m}{N_t}\right) \cdot 100. \tag{4.1}$$

where N_m indicates the number of datasets where a misspecification occurred and N_t is the total number of datasets (i.e., 2000) used to conduct the analyses. For the index r_m we have computed six values, three values from the datasets generated under the PGP process and three from those generated under the PIGP, which provide (under each setup) the risk of incurring in a misspecification of a PGP with a PIGP and the risk of incurring in the misspecification of a PIGP with a PGP, respectively.

The consequences of incurring in a misspecification have been evaluated in terms of its impact on the estimate of the remaining useful life measured by the root mean square

error (RMSE) of the ML estimators of the $MRUL(6)$ as:

$$RMSE_{M,d} = \sqrt{\sum_{k=1}^{N_d}\sum_{i=1}^{6}\frac{\left(\widehat{MRUL}(6)_{M,k,i} - MRUL(6)_{k,i}\right)}{6 \cdot N_d}}. \tag{4.2}$$

where:

— The subscript d of $RMSE_{M,d}$ specifies which are the N_d datasets used to compute the index. In particular, $d = t$ indicates that $RMSE_{M,d}$ is calculated by using all the N_t datasets, $d = \overline{m}$ indicates that it is calculated by using only the $N_{\overline{m}}$ datasets that did not lead to a misspecification, and $d = m$ indicates that it is calculated by considering only the N_m datasets that led to a misspecification.

— The subscript M of $RMSE_{M,d}$ specifies which is the model used to estimate the $MRUL(6)$. In particular, $M = PGP$ indicated that the estimates are obtained by using the PGP while $M = PIGP$ indicates that the estimates are obtained by using the $PIGP$.

— $MRUL(6)_{k,i}$ is the true value of the $MRUL$ of the unit whose degradation path, up to $t_6 = 6$, is described by the ith path ($i = 1, \ldots, m$) of the kth dataset ($k = 1, \ldots, N_d$). In fact, $MRUL(6)_{k,i}$ is computed, path by path, under the true model (i.e., the same model used to simulate the data) as:

$$\begin{aligned} MRUL(6)_{k,i} &= MRUL\left(t|W(6)_{k,i} = w(6)_{k,i}\right) \\ &= \int_0^\infty F_{W(t+\tau)|W(6)_{k,i}}\left(D|W(6)_{k,i} = w(6)_{k,i}\right) \cdot d\tau \\ &= \int_0^\infty F_{\Delta W(t,t+\tau)}\left(D - w(6)_{k,i}\right) \cdot d\tau \end{aligned} \tag{4.3}$$

where $F_{\Delta W(t,t+\tau)}(\cdot)$ is either the (2.16) or the (2.17) depending on the true model (i.e., PGP or PIGP, respectively), with parameters set to the values reported in Table 4.1 (according to the considered setup), and $w(6)_{k,i}$ is the true (hidden) value of the degradation level of the ith unit of the kth dataset at $t_6 = 6$ (i.e., the true value of $W(6)_{k,i}$), which in this simulation study is known. From each dataset six values of $MRUL(6)_{k,i}$;

— $\widehat{MRUL}(6)_{M,k,i}$ is the ML estimate of the $MRUL$ of the unit whose observed degradation history, up to $t_6 = 6$, is described by the ith perturbed path of the kth dataset, here denoted as $\boldsymbol{Z}(6)_{k,i}$. $\widehat{MRUL}(6)_{M,k,i}$ is computed as in (2.69), by us-

ing the model M (i.e., either $M = PGP$ or $M = PIGP$) with parameters set to the corresponding ML estimates obtained, under the same model M, from the kth dataset. In this case the distribution of the RUL used to compute the MRUL is conditional to $\boldsymbol{Z}(6)_{k,i} = \boldsymbol{z}(6)_{k,i}$. This estimate depends on the dataset, on the model M, and on the (whole) perturbed degradation path of the considered unit (note that both the PGP and PIGP are non-Markovian). From each dataset are obtained twelve values of $\widehat{MRUL}(6)_{M,k,i}$ six under the PGP and six under the PIGP. Hence, $\widehat{MRUL}(6)_{M,k,i}$ differs from the true value $MRUL(6)_{k,i}$ because it is conditional to the perturbed measurements instead than on the true (hidden) degradation level and because model parameters are estimated. Moreover, it is worth to underline that, when $\widehat{MRUL}(6)_{M,k,i}$ is computed by using the wrong model, the ML estimates of model parameters are obtained by using the wrong model;

Next, we have repeated the same analyses by modeling the error term according to option 2).

Furthermore, under both the error models 1) and 2), we have also considered the case where $\nu = 0$. In this latter case, the simulated data used to conduct the Monte Carlo study have been generated by using the setups described in Table 4.2, which differ from those given in Table 4.1 for the parameters of the error term only. In particular, differently

Setup	a	b	θ	φ	ν
$\mathcal{A}$	1.25	1	1.25	8/3	0
$\mathcal{B}$	1	1	1	10/3	0
$\mathcal{C}$	0.5	1	0.5	20/3	0

Table 4.2 – Setups $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$used to generate the datasets when $\nu = 0$.

than under the setups reported in Table 4.1, in this case the ratio between the variance of the perturbed and hidden processes (i.e., $V\{Z(t)\}/V\{W(t)\}$) is not time independent. In fact, under each setup (i.e., $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, respectively), here the value of the parameter φ has been selected to set the variance of the error term (that is equal to $V\{\varepsilon(t)\} = 1/\varphi$) to the value that $V\{\varepsilon(t)\}$ assumes at $t = 3$, when φ and ν are those given in Table 4.1. Note that, under the setups given in Table 4.1, $V\{\varepsilon(t)\}$ increases linearly from $t = 0$, where it is null, to $t = 6$, where it is equal to $6/\varphi$. Thus, adopting the setups given in Table 4.2 allows to set $V\{\varepsilon(t)\}$ to the value that the same variance assumes in mean over the time interval $(0, 6)$ under the setups given in Table 4.1. Accordingly, under the setups given in Table 4.2, given that $V\{\varepsilon(t)\}$ does not depend on t, the ratio $V\{Z(t)\}/V\{W(t)\}$

decreases at t increases.
It is worth to remark again that, under option 2) when $\nu = 0$, the adopted modeling solution reduces to the classical case where the error term is Gaussian distributed and independent of the measured degradation level.
Here, for the convenience of the readers we report a scheme of the step-by-step procedure we have used to evaluate the risk and consequences of model misspecification under a given setup and a given true model:

1. Select the true model (i.e., either PGP or the PIGP with a given error model) and a setup (i.e., either $\mathcal{A}$, $\mathcal{B}$, or $\mathcal{C}$) from Table 4.1 or Table 4.2, depending on ν;
2. Repeat steps 3-6 $N_t = 2000$ times;
3. Under the true model, simulate a synthetic dataset. At this step the true hidden degradation values, generated at an intermediate step of the simulation, are also kept;
4. Use the EM algorithm and the particle filter described in Chapters 2.7 and 2.8 to compute the ML estimates of the parameters of both the PGP and the PIGP. This step defines the "estimated processes" (i.e., the perturbed gamma and perturbed inverse Gaussian processes calibrated by using the ML estimates of model parameters);
5. Use the AIC to select the best model. In case of incorrect diagnosis (i.e., if the selected model is not the one used to generate the dataset) a misspecification is assumed to have occurred;
6. Path by path (i.e., for any $k = 1, 2, \ldots, 2000$ and $i = 1, 2, \ldots, 6$) compute $MRUL(6)_{k,i}$ under the model used to generate the data and the $\widehat{MRUL}(6)_{M,k,i}$ both under the (estimated) PGP and the PIGP;
7. Use the results obtained by these 2000 iterations to compute the indices (4.1) and (4.2).

For the sake of comparison, the same analysis has also been performed in the absence of measurement errors (i.e., by assuming that measurements provide exact values of measured degradation levels). In this latter case the competing models are the gamma process (GP) and the inverse Gaussian process (IGP) (i.e., the hidden processes) and the datasets used to perform the Monte Carlo study have been generated under the hidden models with parameters (i.e., a, b, and θ) set to the values given in Table 4.1 (which, as mentioned

above, coincide with those given in Table 4.2). The results of the mentioned analyses are reported in Chapters 4.2.1 and 4.2.2.

4.2.1 Risk of incurring in a misspecification

Table 4.3 reports the values of the percent risk of incurring in a misspecification both when the true model is the PGP and when the true model is the PIGP (i.e., in the symmetric case). The index $r_m(\%)$ is evaluated as in Eq. (4.1). This Table refers to the case where the error term is modeled by using option 1). The setups used to generate the data are those given in Table 4.1. As mentioned above, N_t, $N_{\overline{m}}$, and N_m indicate the total number of datasets used to perform the analysis, the number of datasets that did not lead to a misspecification, and the number of datasets that led to a misspecification, respectively.
Tables 4.4 reports, by using the same notation, the results of the same analyses in the absence of measurement errors. Table 4.3 shows that, under all setups, the risk of incurring

True process	Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
PGP	$\mathcal{A}$	2000	1362	638	31.9
	$\mathcal{B}$	2000	1252	748	37.4
	$\mathcal{C}$	2000	1065	935	46.8
PIGP	$\mathcal{A}$	2000	1484	516	25.8
	$\mathcal{B}$	2000	1470	530	26.5
	$\mathcal{C}$	2000	1368	632	31.6

Table 4.3 – Risk of misspecification when the error term is modeled by using option 1).

True process	Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
GP	$\mathcal{A}$	2000	1888	112	5.6
	$\mathcal{B}$	2000	1806	194	9.7
	$\mathcal{C}$	2000	1541	459	22.9
IGP	$\mathcal{A}$	2000	1843	157	7.85
	$\mathcal{B}$	2000	1814	186	9.30
	$\mathcal{C}$	2000	1665	335	16.7

Table 4.4 – Risk of misspecification in the absence of measurement error.

in a misspecification of a PGP with a PIGP is about $5 - 10\%$ higher than the one of

misspecifying a PIGP with a PGP. As expected, in both cases the risk increases moving from setup $\mathcal{A}$ to setup $\mathcal{C}$. In the worst-case scenario, represented by the setup $\mathcal{C}$, the risk of a wrong diagnosis when the true model is the PGP is close to 50%, whereas in case it is the PIGP it is close to 30%.

Results reported in Table 4.3 also show that the risk of misspecification depends on the setup more when the true model is the PGP than when the true model is the PIGP. An intuitive explanation for this is that passing from the setup $\mathcal{A}$ to the setup $\mathcal{C}$ the shape of the pdf of the increments of the hidden inverse Gaussian process depicted in Figure 4.1 changes less than the shape of the pdf of the increments of the gamma process. Moreover, it seems that, under every setup, when the true process is the PIGP, it is often possible to find a PGP that fits the simulated data in an acceptable manner, while the PIGP more rarely allows to adequately fit PGP data generated under the setups $\mathcal{A}$ and $\mathcal{B}$, where the increments $W(t_1)$, $\Delta W(t_1, t_2), \ldots, \Delta W(t_5, t_6)$ of the hidden process are gamma distributed with shape parameter (i.e., $\Delta\eta(t_1), \Delta\eta(t_1, t_2), \ldots, \Delta\eta(t_5, t_6)$) smaller than 1. In fact, from Figure 4.1, it is apparent that while the pdf of the increment of the gamma process obtained under the setup $\mathcal{C}$ is relatively similar to the pdfs of the increment of the inverse Gaussian process obtained under the setups $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, none of pdfs obtained under the inverse Gaussian process is similar to the pdf of the increment of the gamma process obtained under the setups $\mathcal{A}$ and $\mathcal{B}$).
The comparison with the results reported in Table 4.4 shows that, under all the setups, the presence of measurement errors increases the risk of incurring in a misspecification. Table 4.4 also shows that in the absence of measurement errors the risk of misspecifying a GP with an IGP is close to the one of misspecifying an IGP with a GP. Nonetheless, it seems again that the risk of misspecifying an IGP with a GP depends on the setup less than the risk of misspecifying GP with an IGP. Table 4.5 reports the value of the percent

True process	Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
PGP	$\mathcal{A}$	2000	1376	624	31.2
	$\mathcal{B}$	2000	1256	744	37.2
	$\mathcal{C}$	2000	1067	933	46.7
PIGP	$\mathcal{A}$	2000	1296	704	35.2
	$\mathcal{B}$	2000	1292	708	35.4
	$\mathcal{C}$	2000	1334	666	33.3

Table 4.5 – Risk of misspecification when the error term is modeled by using option 2).

risk of incurring in a misspecification in the case the error term is modeled according to option 2). The index $r_m(\%)$ is evaluated as in Eq. (4.1). The setups used to generate the data are again those given in Table 4.1.

Results reported in Tables 4.5 and 4.3 show that the risk of misspecifying a PGP with a PIGP does not significantly depend on the option adopted to model the error term. Moreover, they also show that, when the true model is the PIGP, results obtained by adopting the error term 2) differ from, and are closer to each other than, those obtained by adopting the error term 1). In fact, it seems that modeling the error term according to option 2) increases the risk of misspecifying a PIGP with a PGP and reduces the sensitivity of the mentioned risk on the setup. Indeed, somewhat surprisingly, Table 4.5 also shows that the risk of misspecifying a PIGP with a PGP under the setup $\mathcal{C}$ is very close to (and even slightly smaller than) the corresponding risk computed under the setups $\mathcal{A}$ and $\mathcal{B}$. We have carefully checked and confirmed these latter results in various ways, yet we do not have an intuitive explanation for them. In fact, we cannot exclude that the value obtained for $r_m(\%)$ under the setup $\mathcal{C}$ is smaller than those obtained under the other setups only for a matter of numerical accuracy. Indeed, the log-likelihood functions computed under the PIGP and PGP (i.e., those used to calculate the AIC index) in about 6% of cases differ only in the fourth significant figure and in less than 1% of cases only in the fifth significant figure, which the (time-demanding) numerical procedure used to compute the log-likelihood does not always allow to calculate in a sufficiently accurate manner.

As we have already mentioned above, the error term 2) allows the perturbed measurement to assume negative values, a result that in many applications can be unrealistic. About this point, we note that, in the case of the datasets used to produce the results reported in Table 4.5, when the true model is the PGP, we have obtained 989 out of $72,000$ negative perturbed measurements under Setup $\mathcal{A}$, 607 out of $72,000$ negative perturbed measurements under Setup $\mathcal{B}$, and 69 out of $72,000$ negative perturbed measurements under Setup $\mathcal{C}$. While, when the true model is the PIGP, we have obtained 443 negative perturbed measurements under Setup $\mathcal{A}$, 227 under Setup $\mathcal{B}$, and 17 under Setup $\mathcal{C}$.

Finally, Tables 4.6 and 4.7 report the values of the percent risk of incurring in a misspecification in the cases where the error term is modeled by using the setups reported in Table 2 (i.e., in the case $\nu = 0$). In particular, Table 4.6 refers to the cases where the error term is modeled by using option 1), while Table 4.7 reports the results obtained by using option 2).

Here, in the case of the datasets used to produce the results reported in Table 4.7 we have

True process	Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
PGP	$\mathcal{A}$	2000	1307	693	34.6
	$\mathcal{B}$	2000	1278	722	36.1
	$\mathcal{C}$	2000	1157	843	42.2
PIGP	$\mathcal{A}$	2000	1459	541	27.1
	$\mathcal{B}$	2000	1439	561	28.1
	$\mathcal{C}$	2000	1326	674	33.7

Table 4.6 – Risk of misspecification when the error term is modeled by using option 1) with $\nu = 0$.

True process	Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
PGP	$\mathcal{A}$	2000	1247	753	37.7
	$\mathcal{B}$	2000	1242	758	37.9
	$\mathcal{C}$	2000	1139	861	43.1
PIGP	$\mathcal{A}$	2000	1216	784	39.2
	$\mathcal{B}$	2000	1195	805	40.3
	$\mathcal{C}$	2000	1185	815	40.7

Table 4.7 – Risk of misspecification when the true processes is the PGP and the error term is modeled by using option 2) with $\nu = 0$.

obtained 3390 negative perturbed measurements under Setup $\mathcal{A}$, 2687 negative perturbed measurements under Setup $\mathcal{B}$, and 990 negative perturbed measurements under Setup $\mathcal{C}$ when the true model is the PGP. Similarly, when the true model is the PIGP, we have obtained 2643 negative perturbed measurements under Setup $\mathcal{A}$, 2016 under Setup $\mathcal{B}$, and 718 under Setup $\mathcal{C}$.

In general, by comparing Tables 4.3 and 4.5 with Tables 4.6 and 4.7, it seems that, when $\nu = 0$ the risk of misspecifying a PGP with a PIGP is similar to the one computed when $\nu = 1$. Yet it also seems that, when $\nu = 0$ the mentioned risk depends on the setup less than when $\nu = 1$, especially when the error term is modeled by using option 2). The same effect is also observed when the true model is the PIGP. However, from the same tables, it also seems that the risk of misspecifiying the PIGP with a PGP under the setups given in Table 4.2 is higher than under those given in Table 4.1.

It should be emphasized that, strictly speaking, the results obtained by setting $\nu = 0$ and $\nu = 1$ are not entirely comparable to each other because, as mentioned before, while

under the setups reported in Table 4.1 the variance of the error term, $\varepsilon(t)$, is proportional to the measured degradation level $W(t)$ under the setups given in Table 4.2 the variance of $\varepsilon(t)$ does not depend on $W(t)$. However, a possible intuitive explanation of the difference existing between the results obtained when $\nu = 0$ and when $\nu = 1$ can be given by focusing on the importance of the measurement obtained at the first measurement epoch. Indeed, as shown in Figure 4.1, the shapes of the pdfs of the increments of the hidden gamma and inverse Gaussian processes mainly differ in the left tail. The same figure also shows that the difference diminishes as $\eta(t)$ increases (i.e., moving from the setup $\mathcal{A}$ to the setup $\mathcal{C}$). For the same reasons the considered shapes become more and more similar as t increases, because $\eta(t)$ increases with t and both the gamma and the inverse Gaussian random variables as $\eta(t)$ increases tend in distribution to a Gaussian random variable (i.e., to the same Gaussian random variable, because the considered competing hidden processes have identical mean and variance functions). Based on this reasoning, it is reasonable to expect that the first perturbed measurements are the most useful to identify the hidden processes. Consequently, given that $V\{\varepsilon(t)\}$ at $t = 1$ is larger when $\nu = 0$ than when $\nu = 1$, it results that the perturbed measurements performed at $t = 1$ when $\nu = 0$ are less useful to identify the hidden model than the corresponding measurements performed when $\nu = 1$.

4.2.2 Consequences of incurring in a misspecification

Table 4.8 reports the value of the index $RMSE_{M,d}$ computed when the error term is modeled according to option 1). To assess the prognostic abilities of the competing models on a short, medium, and long timespan, we used three different values of the threshold: namely, $w_M = 7.5$, $w_M = 9$, and $w_M = 12$. For the sake of simplicity, this piece of information is not included in the notation of the RMSE, but is directly provided in the headings of the tables. The first column of Table 4.8 specifies which is the true model. The second columns indicates the setup under which the RMSE is computed. Details about these setups are given in Table 4.1. The third column specifies the model used to compute $\widehat{MRUL}(6)_{M,k,i}$, which is also the one used to obtain the MLEs of model parameters. Finally, the fourth column specifies the datasets used to compute the index. With the same notations as Tables 4.3-4.7, the subscripts t,m, and $\overline{m}$ indicate that the RMSE is evaluated by using all the N_t datasets, only the N_m datasets where a misspecification occurred, and only the $N_{\overline{m}}$ datasets where a misspecification did not occur, respectively. So, for example, the values reported in the fourth row of Table 4.8 are the RMSEs of the

ML estimator of $MRUL(6)$, computed as in Eq. (2.69), for all the considered thresholds, by using a PIGP with model parameters set to their ML estimates obtained from the N_m datasets generated unter a PGP (calibrated according to the setup $\mathcal{A}$) that led to a misspecification. Tables 4.8 shows that the prognostic abilities of the considered competing perturbed models are quite similar both when the true model is the PGP and when the true model is the PIGP. In addition, as expected, obtained results also show that the difference between the RMSEs obtained under the considered perturbed models decreases moving from the setup $\mathcal{A}$ to the setup $\mathcal{C}$. In fact, this is a direct consequence of the circumstance that (as shown by Figure 4.1) passing from the setup $\mathcal{A}$ to the setup $\mathcal{C}$ the difference between the hidden processes diminishes.

Table 4.9 reports the results obtained by performing the same analyses in the absence of measurement errors. The comparison between the results reported in Table 4.8 and those reported in Table 4.9 show how the presence of measurement errors impacts on the estimation of the $MRUL(6)$. In fact, it is evident that if by one side the presence of measurement errors negatively affects the performances of the ML estimator of $MRUL(6)$ constructed under the right perturbed model (i.e., for example, by using the PGP when the data are generated under a PGP), on the other side it also allows to mitigate the effect of using the wrong model. This is especially clear if one compares the results obtained in the first half of Table 4.8, where the true model is the PGP, whit those reported in the first half of Table 4.9, where the true model is the GP. In fact, for example, while the results reported in the fifth and sixth row of Table 4.9 (obtained under the setup $\mathcal{A}$ under the right GP) differ by several orders of magnitude from the RMSEs reported in the sixth row of the same table (obtained under the setup $\mathcal{A}$, by using the wrong IGP), the RMSEs reported in the fifth and sixth row of Table 4.8 (obtained under the right PGP and the wrong PIGP, respectively) are very close to each others. A similar situation is highlighted by the results reported in the eleventh and twelfth rows of the same tables (i.e., in the case of the setup $\mathcal{B}$).

Indeed, even more specifically, the results reported in Table 4.9 also show that, even in the absence of measurement error, the use of a good model selection criterion (such as the AIC) allows to greatly mitigate the consequences of adopting the wrong model. In fact, the results reported in the third and fourth row of Table 4.9 show that the ML estimators constructed under the wrong and right models in the case of the datasets, generated under the setup $\mathcal{A}$, that led to a misspecification (that is, those for which the AIC suggests the use of the wrong model) behave in a very similar manner. This conclusion is further

strengthened by the results reported in the first and second row of Table 4.9 (i.e., those obtained by using the datasets for which the AIC leads to select the right model) which as those reported in the fifth and sixth row of the same table, differ from each others by several orders of magnitude. In other words, from the results obtained under the setup $\mathcal{A}$, when the true model is the GP (i.e., in the absence of measurement error) it appears that the huge difference existing between the RMSEs computed on the basis of the estimates of the MRUL obtained from all the datasets by using the right GP and the wrong IGP (i.e., the difference existing between the results reported in the fifth row and in the sixth row of Table 4.9) is due to the (poor) estimates of the MRUL obtained under the wrong IGP from the datasets that do not cause a misspecification. The results reported in the seventh, eighth, ninth, and tenth rows of Table 4.9 show that the same situation unfolds also in the case of setup $\mathcal{B}$.

It is also interesting to note that, as shown by the results reported in the second half of Table 4.9, when the true model is the IGP the performances of the estimators constructed under the GP and the IGP seem to have very similar performances, regardless of whether a misspecification has occurred or not. As we have already remarked by commenting on the results in Tables 4.3, even in the absence of measurement error, the differences existing between the results obtained when the true process is the GP and when it is the IGP are probably due to the circumstance that when the true process is the IGP, under all the setups, it is possible to find a PGP that fits the simulated data in an acceptable manner, whereas the IGP more rarely allows to adequately fit PGP data generated under the setups $\mathcal{A}$ and $\mathcal{B}$. In fact, as shown in Figure 4.1, while the gamma pdf can assume shapes that are similar to those of all the inverse Gaussian pdf depicted in that figure, the inverse Gaussian pdf cannot assume shapes just as similar to the gamma pdfs corresponding to setups $\mathcal{A}$ and $\mathcal{B}$.

Table 4.10 reports the values of the index $RMSE_{M,d}$ computed when the error term is modeled according to option 2). The $RMSE_{M,d}$ computed in these cases are very similar to those reported in Tables 4.8. That is, it seems that the shape of the distribution of the error term does not significantly affect the properties of the ML estimator of the MRUL. Finally, Tables 4.11 and 4.12 report the values of the index $RMSE_{M,d}$ computed when $\nu = 0$ and the error term is modeled according to option 1) and option 2), respectively. The $RMSE_{M,d}$ computed in these cases are very similar to those reported in Tables 4.8 and 4.10. That is, it seems that also this specific feature of the distribution of the error term (i.e., setting $\nu = 0$ or $\nu = 1$) does not significantly affect the properties of the ML

estimator of the MRUL.

4.3 Conclusions

This chapter has investigated the risk and the consequences of the misspecification of a perturbed gamma process with a perturbed inverse Gaussian one and the symmetrical misspecification problem of a perturbed inverse Gaussian process with a perturbed gamma one. Two different models are used to describe the error term.
To facilitate the misspecification study, the (hidden) inverse Gaussian process is formulated by using a new parameterization that allows the considered competing models to share the same parameters and the same functional forms of mean and variance functions.

To conduct the analyses we have carried out a large Monte Carlo study where six realistic experimental scenarios, characterized by different misspecification risk, are simulated by using as many model setups. A setup defines the true model, which is either a perturbed gamma process or a perturbed inverse Gaussian process, depending on the misspecification issue of concern. Hence, with six setups, two hidden models, and two choices for the distribution of the error term, we have defined 24 true models, 12 perturbed gamma and 12 perturbed inverse Gaussian. Each true model has been used to generate 2000 synthetic datasets, each one consisting of the degradation path of six units. Together with each perturbed measurement we also kept the value of the measured (hidden) degradation level, which is generated at an intermediate step of the algorithm that we have adopted to generate the perturbed data. The true values of the hidden degradation level are used in the analysis for comparative purposes. From each dataset, we have estimated the parameters of both the competing perturbed processes by using the ML method. Hence, path by path we have estimated the mean remaining useful life by using the model estimated from the dataset which contains the considered path.

The ML estimation of the parameters of the competing perturbed models has been performed by using a new expectation maximization particle filter algorithm. Dataset by dataset the model selection is performed by using the Akaike information criterion. A misspecification has been assumed to occur when the Akaike information criterion has led to select the wrong model (i.e., for example, if it has led to select the perturbed gamma process on a dataset that had been generated under the perturbed inverse Gaussian process).
All the estimates have been used to compute the risk of incurring in a misspecification

and to evaluate its effect on the ML estimates of the mean remaining useful life. Finally, for the sake of comparison, the same study has also been repeated in the absence of measurement errors (that is, by considering as competing models a gamma process and an inverse Gaussian process).

Obtained results demonstrate that, when the error term is modeled by adopting option 1) (i.e., by using the 3-parameter inverse gamma distribution) the risk of incurring in a misspecification (as expected) is significantly influenced by the value of the shape parameter of the degradation increment of the hidden process (be it gamma or inverse Gaussian) defined between successive measurement times. Furthermore, and more specifically, it also seems that the risk of misspecifying a PGP with a PIGP depends on the setup more than the risk of misspecifying a PIGP with a PGP. In fact, while the PGP often allows to fit PIGP data in an acceptable manner independently of the setup, the PIGP more rarely allows to adequately fit PGP data generated under the setups $\mathcal{A}$ and $\mathcal{B}$ (i.e., cases where the increments of the hidden process between successive inspection epochs are gamma distributed with shape parameter smaller than or equal to 1). The results obtained by modeling the error term according to option 2) (i.e., by using a Gaussian distribution) indicate that the shape of the error term does not significantly influence the risk of misspecifying a PGP with a PIGP. Contrarily, it seems that modeling the error term according to option 2) increases the risk of misspecifying a PIGP with a PGP and reduces the sensitivity of the mentioned risk on the setup.
Finally, the results obtained by setting $\nu = 0$ indicate that also this modeling option does not significantly affect the risk of misspecifying a PGP with a PIGP. Yet, it also seems that, when $\nu = 0$ the mentioned risk depends on the setup less than when $\nu = 1$, especially when the error term is modeled by using option 2). The same effect is also observed when the true model is the PIGP. However, in this case, it also seems that when $\nu = 0$ the risk of misspecifiying the PIGP with a PGP is higher than when $\nu = 1$.

Obtained results also clearly show that the presence of measurement errors significantly increases the risk of selecting the wrong model. Obviously, this result was expected, since perturbed data do not allow to directly check whether the selected model is actually able to adequately fit the real (hidden) degradation process, being only useful to check the ability of the perturbed model to fit the perturbed measurements.
With respect to the consequences produced by a misspecification, obtained results show that the maximum likelihood estimators of the (perturbed measurement-based) remaining useful life constructed under the competing perturbed processes, irrespectively of the

model used to describe the error term, have very similar performances, regardless of whether a misspecification has occurred.
Moreover, the comparison of the results obtained in the presence of measurement errors with those obtained in its absence shows that if on one side the presence of measurement errors negatively affects the performances of the maximum likelihood estimator of the mean remaining useful life constructed under the right perturbed model, it also mitigates (with respect to the case where measurement errors are absent) the consequences of using the wrong model, especially when the true hidden process is the gamma. However, the results obtained in the absence of measurement errors also show that, when the true model is the gamma, even in experimental situations where the consequences of the use of the wrong model could be extremely severe, adopting an appropriate model selection procedure (such as the Akaike information criterion), allows to hugely mitigate the effect of a misspecification.

Finally, it is worth remarking that the analysis performed in this chapter, while being thorough and in-depth, is limited in its scope by only being focused on the consequences of a misspecification in terms of the remaining useful life. Indeed, in practical applications, it could also be interesting to assess the effects of a potential misspecification on maintenance costs. To this aim, in the next chapter (specifically, in Chapter 5.5), a second misspecification study, focused on its impact on maintenance costs, will be presented.

				$RMSE_{M,d}$		
True process	Setup	M	d	$w_M = 7.5$	$w_M = 9$	$w_M = 12$
PGP	$\mathcal{A}$	PGP	$\overline{m}$	1.67	2.34	3.99
		PIGP	$\overline{m}$	1.53	2.13	3.60
		PGP	m	1.89	2.64	4.44
		PIGP	m	1.94	2.69	4.46
		PGP	t	1.74	2.44	4.14
		PIGP	t	1.67	2.32	3.89
	$\mathcal{B}$	PGP	$\overline{m}$	1.38	1.92	3.21
		PIGP	$\overline{m}$	1.32	1.84	3.09
		PGP	m	1.37	1.91	3.20
		PIGP	m	1.33	1.85	3.07
		PGP	t	1.38	1.92	3.20
		PIGP	t	1.32	1.84	3.08
	$\mathcal{C}$	PGP	$\overline{m}$	0.76	1.09	1.89
		PIGP	$\overline{m}$	0.74	1.07	1.85
		PGP	m	0.78	1.12	1.92
		PIGP	m	0.78	1.11	1.91
		PGP	t	0.77	1.10	1.90
		PIGP	t	0.76	1.09	1.88
PIGP	$\mathcal{A}$	PGP	$\overline{m}$	1.56	2.20	3.72
		PIGP	$\overline{m}$	1.52	2.14	3.60
		PGP	m	1.54	2.17	3.64
		PIGP	m	1.38	1.91	3.14
		PGP	t	1.56	2.19	3.70
		PIGP	t	1.49	2.08	3.49
	$\mathcal{B}$	PGP	$\overline{m}$	1.26	1.77	2.98
		PIGP	$\overline{m}$	1.21	1.69	2.83
		PGP	m	1.29	1.84	3.15
		PIGP	m	1.20	1.69	2.87
		PGP	t	1.27	1.79	3.03
		PIGP	t	1.21	1.69	2.84
	$\mathcal{C}$	PGP	$\overline{m}$	0.74	1.04	1.77
		PIGP	$\overline{m}$	0.73	1.02	1.74
		PGP	m	0.77	1.10	1.90
		PIGP	m	0.75	1.06	1.83
		PGP	t	0.75	1.06	1.81
		PIGP	t	0.73	1.04	1.77

Table 4.8 – $RMSE_{M,d}$ of the ML estimators of $MRUL(6)$ when the error term is modeled by using option 1).

				$RMSE_{M,d}$		
True process	Setup	M	d	$w_M = 7.5$	$w_M = 9$	$w_M = 12$
GP	$\mathcal{A}$	PGP	$\overline{m}$	1.24	1.79	3.08
		PIGP	$\overline{m}$	$4.01 \cdot 10^{11}$	$1.70 \cdot 10^{12}$	$1.70 \cdot 10^{13}$
		PGP	m	0.94	1.36	2.34
		PIGP	m	2.03	2.75	4.38
		PGP	t	1.23	1.76	3.04
		PIGP	t	$3.90 \cdot 10^{11}$	$1.66 \cdot 10^{12}$	$1.62 \cdot 10^{13}$
	$\mathcal{B}$	PGP	$\overline{m}$	0.96	1.41	2.47
		PIGP	$\overline{m}$	$2.92 \cdot 10^{14}$	$2.10 \cdot 10^{15}$	$3.79 \cdot 10^{16}$
		PGP	m	0.75	1.09	1.90
		PIGP	m	1.22	1.73	2.92
		PGP	t	0.94	1.39	2.43
		PIGP	t	$2.77 \cdot 10^{14}$	$1.99 \cdot 10^{15}$	$3.60 \cdot 10^{16}$
	$\mathcal{C}$	PGP	$\overline{m}$	0.51	0.81	1.53
		PIGP	$\overline{m}$	1.63	2.54	4.98
		PGP	m	0.50	0.80	1.50
		PIGP	m	0.60	0.93	1.73
		PGP	t	0.51	0.81	1.53
		PIGP	t	1.46	2.28	4.45
IGP	$\mathcal{A}$	PGP	$\overline{m}$	0.96	1.47	2.67
		PIGP	$\overline{m}$	1.07	1.61	2.89
		PGP	m	0.81	1.21	2.12
		PIGP	m	0.85	1.25	2.17
		PGP	t	0.83	1.23	2.16
		PIGP	t	0.87	1.28	2.24
	$\mathcal{B}$	PGP	$\overline{m}$	0.71	1.12	2.08
		PIGP	$\overline{m}$	0.73	1.13	2.08
		PGP	m	0.70	1.06	1.91
		PIGP	m	1.73	1.09	1.95
		PGP	t	0.70	1.07	1.93
		PIGP	t	0.73	1.10	1.96
	$\mathcal{C}$	PGP	$\overline{m}$	0.37	0.62	1.20
		PIGP	$\overline{m}$	0.38	0.63	1.22
		PGP	m	0.43	0.69	1.32
		PIGP	m	0.43	0.69	1.30
		PGP	t	0.42	0.68	1.30
		PIGP	t	0.42	0.68	1.29

Table 4.9 – $RMSE_{M,d}$ of the ML estimators of $MRUL(6)$ in the absence of measurement error.

				$RMSE_{M,d}$		
True process	Setup	M	d	$w_M = 7.5$	$w_M = 9$	$w_M = 12$
PGP	$\mathcal{A}$	PGP	$\overline{m}$	1.73	2.44	4.15
		PIGP	$\overline{m}$	1.70	2.36	3.94
		PGP	m	2.04	2.85	4.81
		PIGP	m	2.02	2.79	4.64
		PGP	t	1.83	2.57	4.37
		PIGP	t	1.80	2.50	4.17
	$\mathcal{B}$	PGP	$\overline{m}$	1.33	1.85	3.12
		PIGP	$\overline{m}$	1.29	1.80	3.02
		PGP	m	1.50	2.14	3.67
		PIGP	m	1.51	2.14	3.68
		PGP	t	1.39	1.96	3.33
		PIGP	t	1.37	1.93	3.28
	$\mathcal{C}$	PGP	$\overline{m}$	0.76	1.08	1.86
		PIGP	$\overline{m}$	0.75	1.08	1.85
		PGP	m	0.75	1.06	1.82
		PIGP	m	0.75	1.07	1.84
		PGP	t	0.75	1.07	1.84
		PIGP	t	0.75	1.07	1.84
PIGP	$\mathcal{A}$	PGP	$\overline{m}$	1.59	2.21	3.70
		PIGP	$\overline{m}$	1.49	2.07	3.45
		PGP	m	1.51	2.14	3.69
		PIGP	m	1.46	2.05	3.50
		PGP	t	1.56	2.18	3.70
		PIGP	t	1.48	2.06	3.46
	$\mathcal{B}$	PGP	$\overline{m}$	1.27	1.78	2.99
		PIGP	$\overline{m}$	1.18	1.64	2.74
		PGP	m	1.32	1.88	3.26
		PIGP	m	1.26	1.78	3.03
		PGP	t	1.29	1.82	3.08
		PIGP	t	1.21	1.69	2.84
	$\mathcal{C}$	PGP	$\overline{m}$	0.76	1.08	1.86
		PIGP	$\overline{m}$	0.73	1.04	1.79
		PGP	m	0.77	1.10	1.92
		PIGP	m	0.74	1.06	1.83
		PGP	t	0.76	1.09	1.88
		PIGP	t	0.74	1.05	1.80

Table 4.10 – $RMSE_{M,d}$ of the ML estimators of $MRUL(6)$ when the error term is modeled by using option 2).

				$RMSE_{M,d}$		
True process	Setup	M	d	$w_M = 7.5$	$w_M = 9$	$w_M = 12$
PGP	$\mathcal{A}$	PGP	$\overline{m}$	1.42	1.97	3.29
		PIGP	$\overline{m}$	1.28	1.77	2.93
		PGP	m	1.64	2.31	3.95
		PIGP	m	1.71	2.42	4.17
		PGP	t	1.50	2.09	3.53
		PIGP	t	1.45	2.02	3.41
	$\mathcal{B}$	PGP	$\overline{m}$	1.27	1.80	3.07
		PIGP	$\overline{m}$	1.16	1.64	2.81
		PGP	m	1.34	1.88	3.16
		PIGP	m	1.21	1.68	2.81
		PGP	t	1.30	1.83	3.10
		PIGP	t	1.18	1.66	2.81
	$\mathcal{C}$	PGP	$\overline{m}$	0.73	1.07	1.92
		PIGP	$\overline{m}$	0.72	1.05	1.89
		PGP	m	0.74	1.07	1.89
		PIGP	m	0.73	1.07	1.88
		PGP	t	0.73	1.07	1.90
		PIGP	t	0.72	1.06	1.89
PIGP	$\mathcal{A}$	PGP	$\overline{m}$	1.36	1.90	3.21
		PIGP	$\overline{m}$	1.25	1.74	2.91
		PGP	m	1.83	2.64	4.61
		PIGP	m	1.61	2.31	3.99
		PGP	t	1.50	2.13	3.65
		PIGP	t	1.36	1.91	3.24
	$\mathcal{B}$	PGP	$\overline{m}$	1.11	1.56	2.66
		PIGP	$\overline{m}$	1.04	1.45	2.46
		PGP	m	1.40	2.04	3.64
		PIGP	m	1.24	1.79	3.13
		PGP	t	1.20	1.71	2.97
		PIGP	t	1.10	1.56	2.67
	$\mathcal{C}$	PGP	$\overline{m}$	0.66	0.96	1.70
		PIGP	$\overline{m}$	0.63	0.92	1.62
		PGP	m	0.72	1.07	1.94
		PIGP	m	0.69	1.02	1.82
		PGP	t	0.68	1.00	1.78
		PIGP	t	0.65	0.95	1.69

Table 4.11 – $RMSE_{M,d}$ of the ML estimators of $MRUL(6)$ when the error term is modeled by using option 1) with $\nu = 0$.

				$RMSE_{M,d}$		
Setup	M	d	$w_M = 7.5$	$w_M = 9$	$w_M = 12$	
PGP	$\mathcal{A}$	PGP	$\overline{m}$	1.76	2.48	4.23
		PIGP	$\overline{m}$	1.77	2.46	4.14
		PGP	m	2.03	2.95	5.27
		PIGP	m	2.13	3.08	5.49
		PGP	t	1.87	2.67	4.65
		PIGP	t	1.91	2.71	4.70
	$\mathcal{B}$	PGP	$\overline{m}$	1.36	1.93	3.29
		PIGP	$\overline{m}$	1.31	1.84	3.12
		PGP	m	1.39	1.99	3.49
		PIGP	m	1.38	1.98	3.46
		PGP	t	1.37	1.95	3.36
		PIGP	t	1.33	1.90	3.25
	$\mathcal{C}$	PGP	$\overline{m}$	0.71	1.03	1.82
		PIGP	$\overline{m}$	0.69	1.02	1.79
		PGP	m	0.65	0.95	1.70
		PIGP	m	0.64	1.94	1.69
		PGP	t	0.68	1.00	1.77
		PIGP	t	0.67	0.98	1.75
PIGP	$\mathcal{A}$	PGP	$\overline{m}$	1.59	2.24	3.85
		PIGP	$\overline{m}$	1.47	2.08	3.55
		PGP	m	1.61	2.31	4.07
		PIGP	m	1.54	2.21	3.86
		PGP	t	1.59	2.27	3.94
		PIGP	t	1.50	2.13	3.67
	$\mathcal{B}$	PGP	$\overline{m}$	1.21	1.71	2.94
		PIGP	$\overline{m}$	1.13	1.60	2.73
		PGP	m	1.33	1.92	3.39
		PIGP	m	1.26	1.82	3.20
		PGP	t	1.26	1.80	3.13
		PIGP	t	1.19	1.69	2.93
	$\mathcal{C}$	PGP	$\overline{m}$	0.68	0.99	1.79
		PIGP	$\overline{m}$	0.66	0.96	1.73
		PGP	m	0.71	1.05	1.87
		PIGP	m	0.68	1.00	1.79
		PGP	t	0.69	1.02	1.82
		PIGP	t	0.67	0.98	1.75

Table 4.12 – $RMSE_{M,d}$ of the ML estimators of $MRUL(6)$ when the error term is modeled by using option 2) with $\nu = 0$.

Chapter 5

Proposition of new maintenance policies in the presence of random effect and measurement error

5.1 Introduction

Maintenance activities for degrading units are typically scheduled by using either age-based or condition-based strategies (e.g., see Ahmad and Kamaruddin (2012) and Alaswad and Xiang (2017)). These approaches have typically the goal of optimizing some kind of objective function, such as the long-run average maintenance cost rate (H. Wang, Pham, et al. (2006), I. B. Gertsbakh (2000), Finkelstein, Shafiee, and Kotchap (2016), and Cha, Finkelstein, and Levitin (2017)).

Condition-based maintenance (CBM) strategies generally offer benefits over age-based maintenance (ABM) approaches, leading to reduced occurrence of failures and extended utilization of units throughout their operational lifespan. Jonge, R. Teunter, and Tinga (2017) provides an extensive discussion of the comparison between CBM and ABM.

In this chapter, attention is focused on cases where the degradation phenomenon under analysis can be modeled by a gamma process-based model. The performances of CBM approaches applied to gamma deteriorating units has been extensively investigated in the literature (reviews can be found in Van Noortwijk (2009) and Alaswad and Xiang (2017)). A potential limitation of several strategies proposed in the literature (e.g., see Grall, Bérenguer, and Dieulle (2002), Castanier, Bérenguer, and Grall (2003), Castanier, Grall, and Bérenguer (2005), Fouladirad and Grall (2011), Huynh, Barros, et al. (2011), Cholette et al. (2019)) is that it is usually assumed that inspections provide an exact measurement of the degradation level. However, especially when the degradation level is measured via in-service, non-destructive inspection methods (Lu, Pandey, and W.-C. Xie (2013)), the

measurements are often contaminated by random errors that can undermine the performances of CBM approaches. An example of a simple CBM policy for gamma deteriorating units that takes into account the presence of measurement error can be found in Kallen and Van Noortwijk (2005). More recently, some authors (Nguyen et al. (2019)) proposed an approach where the quality of the inspection procedure itself is a design parameter of the maintenance model. Similarly, S. Hao, J. Yang, and Bérenguer (2020) proposed a gamma process-based CBM policy where a perfect and an imperfect inspection procedures can be used in a combined manner.

An additional possible drawback found in numerous current CBM policies is their implicit connection between preventive replacements and inspection intervals. In fact, CBM policies often assume that such replacements are exclusively carried out during scheduled inspection epochs (see Y. Wang and Pham (2013) and Huynh, Grall, and Bérenguer (2019)). Nonetheless, considering the potentially hefty costs associated with inspections, it can be reasonable to presume that maintenance strategies enabling the avoidance of numerous inspections, especially when they are expected to give rise to foregone decisions, could yield economic benefits (e.g., see Fauriat and Zio (2020), Yuan, Higo, and Pandey (2021), and S. Kim, Choi, and N. H. Kim (2022)). Based on this general idea, Huynh, Grall, and Bérenguer (2019) introduced a novel maintenance strategy that separates decisions and scheduling for repairs/replacements and inspections, demonstrating its potential advantages in terms of long-term average maintenance cost rate. The degradation phenomenon, as described in their study, employs a standard homogeneous gamma process, which may pose limitations in practical applications. It is worth noting that this idea has been explored both before and after the work of Huynh, Grall, and Bérenguer (2019). Indeed, Crowder and Lawless (2007) had previously developed a maintenance scheme that utilizes degradation measurements to determine whether an immediate replacement is necessary or if scheduling a future replacement time is more suitable. Their work also discusses, for illustrative purposes, a scenario involving a single inspection in detail. Nonetheless, they also contemplate the option of enhancing the policy by incorporating a second inspection at the planned replacement time instead of opting for an automatic replacement. More recently, this general idea has also been exploited by Finkelstein, Cha, and Levitin (2020), Cha, Finkelstein, and Levitin (2022), and Cha, Finkelstein, and Levitin (2021), who proposed policies built on the stronger assumption that only one inspection is allowed over a maintenance cycle (i.e., between two consecutive repairs/replacements) and that the future replacement time is a design parameter

whose value is defined a priori, regardless of the measured degradation level. This particular maintenance approach is highly applicable to assets where continuous degradation monitoring is unfeasible. Examples include buried pipelines within the oil and gas sector, reformer tubes in ammonia plants, and drilling rigs (e.g., see Alaswad and Xiang (2017)). These assets necessitate costly and specialized inspections that often disrupt operations or require plant shutdowns for assessment.

With respect to the degradation modeling approaches, Crowder and Lawless (2007) adopt the gamma process with random effect firstly presented in Lawless and Crowder (2004). In Finkelstein, Cha, and Levitin (2020) the degradation process is modeled by using a non-homogeneous gamma process. In Cha, Finkelstein, and Levitin (2021) it is described by using a transformed gamma process (Giorgio, Guida, and Pulcini (2015)), whereas in Cha, Finkelstein, and Levitin (2022) it is considered the case of a mixed population composed of two different (i.e., strong and weak) gamma degrading items.

The main goal of this chapter is to illustrate three novel maintenance policies based on the driving ideas highlighted in this discussion, in Chapters 5.2, 5.3, and 5.4. Inspired by Finkelstein, Cha, and Levitin (2020), Cha, Finkelstein, and Levitin (2021), and Cha, Finkelstein, and Levitin (2022), all three policies are built on the assumption that only one inspection, performed at a predetermined time, is allowed during the life cycle of the unit. The outcome of this inspection will then inform subsequent decision-making.
Another common element between the three proposed policies, which also constitutes the main novelty with respect to Finkelstein, Cha, and Levitin (2020), Cha, Finkelstein, and Levitin (2021), and Cha, Finkelstein, and Levitin (2022), is that they assume that the degrading units under study are subjected to the joint presence of three forms of variability: temporal variability, unit-to-unit variability (random effect), and measurement error (see also Chapter 3.1). Indeed, a heterogeneous population of degrading units where the observed degradation process can be described by the perturbed process with random effect introduced in Chapter 2.4 and thoroughly discussed in Chapter 3 is considered. Maintenance activities refer to a generic unit randomly selected from this population.

Yet another common element, which also differentiates the three proposed policies from the aforementioned literature works, is that failures are assumed to be not self-announcing (see Chapter 1.2.1 and Chapter 2.6 for a detailed description of a not self-announcing failure behavior). Consequently, it is also assumed that failed units may continue operating past their failure point. The time elapsing between the failure of a unit and its eventual replacement is denoted as downtime, and additional costs due to reduced performance are

incurred during this time.
Lastly, for the sake of readability, some more assumptions, common across all three policies (and in the literature) are laid out here:

— inspections are instantaneous and non-destructive;

— any replacement (i.e., both corrective and preventive) restore the unit to an "as good as new" state. Consequently, the time between two successive replacements defines the cycle of a renewal process;

— the performance criterion adopted to optimize a policy is the long-run average maintenance cost rate, computed by adopting the renewal-reward theorem (e.g., see Ross (1983)).

The core concept of these policies involves utilizing information gathered during the inspection to promptly detect and replace units that have already failed or are on the verge of failure. Meanwhile, units with degradation levels comfortably below the failure threshold will have their replacement deferred to a later time. It is rational to believe that, particularly in scenarios where random effects is present, the inspection can differentiate between weak (i.e., units whose degradation progresses faster) and strong (i.e., units whose degradation progresses slower) units (see also Cha, Finkelstein, and Levitin (2022)) and schedule replacements accordingly.
Moreover, as remarked in Chapter 2.4, given that the adopted gamma process with random effect is an age and state dependent process, the probability that a failure occurs in a future time interval directly depends on the failure threshold, on the current degradation level of the unit, and on its current age. This feature also differentiates the considered scenario from the case of other failure threshold models built by using the basic homogeneous gamma process (e.g., Grall, Bérenguer, and Dieulle (2002), Castanier, Bérenguer, and Grall (2003), Castanier, Grall, and Bérenguer (2005), and Huynh, Grall, and Bérenguer (2019)), where the same probability only depends on the gap existing between the current degradation level of the unit and the failure threshold.
Therefore, the proposed approaches have the potential to outperform traditional condition-based and age-based policies by achieving a globally more advantageous trade-off among the costs of inspections, corrective replacements, and preventive replacements.

In addition to the three aforementioned policies, in Chapter 5.5 the results of a misspecification study focused on its impacts on maintenance costs computed under one of the proposed policies.

5.2 A hybrid maintenance policy for a deteriorating unit in the presence of random effect and measurement error

5.2.1 Description of the policy

This policy (see also Esposito, Mele, et al. (2021a)) consists in performing a single inspection at a predetermined time τ aimed at measuring the perturbed degradation level of the unit, and in adopting a condition-based rule to decide whether to immediately replace the unit or to delay its replacement of another predetermined interval of length $\Delta\tau$. In case of postponement, no further inspection will be carried out at time $\tau + \Delta\tau$ and the unit will be replaced no matter its state. Moreover, a second scenario, where a perfect measurement instead of the perturbed one is carried out, is also considered. Table 5.1 summarizes the condition-based rule adopted in the first scenario, where z_t denotes the observed value of the perturbed degradation level $Z(\tau)$ measured at the inspection time τ and L_Z is a maintenance threshold level. Table 5.2, where w_τ denotes the value of the measured degradation level $W(\tau)$ and L_W is a maintenance threshold level, reports the same rule in the second scenario. Tables 5.3 and 5.4 summarize all possible experimental

Measurement at τ	Decision
$z_\tau > L_Z$	Replacement at τ
$z_\tau \leq L_Z$	Replacement at $\tau + \Delta\tau$

Table 5.1 – Condition-based rule in the presence of measurement error.

Measurement at τ	Decision
$w_\tau > L_W$	Replacement at τ
$z_\tau \leq L_W$	Replacement at $\tau + \Delta\tau$

Table 5.2 – Condition-based rule in the absence of measurement error.

scenarios and, for each one of them, report the maintenance actions to be taken and the corresponding length of a maintenance cycle, in the presence and in the absence of measurement error, respectively. In Tables 5.3 and 5.4 $w_{\tau+\Delta\tau}$ denotes the true value of the degradation level $W(\tau + \Delta\tau)$ at time $\tau + \Delta\tau$, w_M is the failure threshold, and the length

of the maintenance cycle is denoted as ${}_1T(z_\tau, w_\tau)$ (when measurement error is present) and ${}_2T(w_\tau)$ (when it is not).

The vectors $\boldsymbol{\xi}_Z = \{L_Z, \tau, \Delta\tau\}$ and $\boldsymbol{\xi}_W = \{L_W, \tau, \Delta\tau\}$ represent the design parameters in the case with and without measurement error, respectively. It is worth to remark that ${}_1T(z_\tau, w_\tau)$ (${}_2T(w_\tau)$) functionally depends on the design parameter vector $\boldsymbol{\xi}_Z$ ($\boldsymbol{\xi}_W$), despite the adopted notation not highlighting it.

Experimental scenario	Maintenance action	Cycle length ${}_1T(z_\tau, w_\tau)$
$z_\tau > L_Z$ and $w_\tau \leq w_M$	Preventive replacement at τ	τ
$z_\tau > L_Z$ and $w_\tau > w_M$	Corrective replacement at τ	τ
$z_\tau \leq L_Z$ and $w_{\tau+\Delta\tau} \leq w_M$	Preventive replacement at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$z_\tau \leq L_Z, w_\tau \leq w_M$, and $w_{\tau+\Delta\tau} \leq w_M$	Corrective replacement at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$z_\tau \leq L_Z, w_\tau > w_M$, and $w_{\tau+\Delta\tau} > w_M$	Corrective replacement at $\tau + \Delta\tau$	$\tau + \Delta\tau$

Table 5.3 – Possible scenarios and corresponding maintenance actions and cycle lengths in the presence of measurement error.

5.2.2 The cost model

Maintenance costs are determined by assuming that the cost of a corrective replacement is c_c, the cost of a preventive replacement is c_p, the cost of an inspection is c_i, and the logistic cost (which is incurred each time a maintenance action, be it a repair or an inspection, is taken) is c_l. Moreover, a downtime cost is considered, computed as the product of a fixed cost rate c_d and the time spent in a failed state (i.e., the time elapsing between the possible failure of the unit and its (corrective) replacement). This cost accounts for the loss of utility (i.e, reduced performances and/or loss of production) resulting from operating the unit in a failed state.

Table 5.5 reports the maintenance cost ${}_1C(z_\tau, w_\tau, X)$ associated to each maintenance action described in Table 5.3, whereas Table 5.6 reports the maintenance cost ${}_2C(w_\tau, X)$ associated to the actions described in Table 5.4. In both tables, X denotes the lifetime

Experimental scenario	Maintenance action	Cycle length ${}_2T(w_\tau)$
$L_W < w_\tau \leq w_M$	Preventive replacement at τ	τ
$w_\tau \leq w_M$	Corrective replacement at τ	τ
$w_\tau \leq L_W$ and $w_{\tau+\Delta\tau} \leq w_M$	Preventive replacement at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$w_\tau \leq L_W$ and $w_{\tau+\Delta\tau} \leq w_M$	Corrective replacement at $\tau + \Delta\tau$	$\tau + \Delta\tau$

Table 5.4 – Possible scenarios and corresponding maintenance actions and cycle lengths in the absence of measurement error.

of the unit. Coherently with the assumption of not self-announcing failures, X is always denoted with the capital letter (even when z_τ of w_τ are known), to remark that, even in these special cases, the lifetime of the unit should be regarded as a random variable. Note that, as ${}_1T(z_\tau, w_\tau)$ and ${}_2T(w_\tau)$, the cost functions ${}_1C(z_\tau, w_\tau, X)$ and ${}_2C(w_\tau, X)$ functionally depend on the vectors of design parameters $\boldsymbol{\xi}_Z$ and $\boldsymbol{\xi}_W$. Yet again for simplicity, the adopted notation does not explicitly indicate this dependency.

Experimental scenario	Maintenance cost ${}_1C(z_\tau, w_\tau, X)$
$z_\tau > L_Z$ and $w_\tau \leq w_M$	$c_l + c_i + c_p$
$z_\tau > L_Z$ and $w_\tau > w_M$	$c_l + c_i + c_c + c_d \cdot (\tau - X)$
$z_\tau \leq L_Z$ and $w_{\tau+\Delta\tau} \leq w_M$	$2 \cdot c_l + c_i + c_p$
$z_\tau \leq L_Z, w_\tau \leq w_M,$ and $w_{\tau+\Delta\tau} > w_M$	$2 \cdot c_l + c_i + c_c + c_d \cdot (\tau + \Delta\tau - X)$
$z_\tau \leq L_Z, w_\tau > w_M,$ and $w_{\tau+\Delta\tau} > w_M$	$2 \cdot c_l + c_i + c_c + c_d \cdot (\tau + \Delta\tau - X)$

Table 5.5 – Possible scenarios and corresponding maintenance cost in the presence of measurement error.

Experimental scenario	Maintenance cost ${}_2C(w_\tau, X)$
$L_W < w_\tau \leq w_M$	$c_l + c_i + c_p$
$w_\tau > w_M$	$c_l + c_i + c_c + c_d \cdot (\tau - X)$
$w_\tau \leq L_W$ and $w_{\tau+\Delta\tau} \leq w_M$	$2 \cdot c_l + c_i + c_p$
$w_\tau \leq L_W$ and $w_{\tau+\Delta\tau} > w_M$	$2 \cdot c_l + c_i + c_c + c_d \cdot (\tau + \Delta\tau - X)$

Table 5.6 – Possible scenarios and corresponding maintenance cost in the absence of measurement error.

5.2.3 Formulation of the long-run average maintenance cost rate

In the presence of measurement error the long-run average maintenance cost rate ${}_1C_\infty(\boldsymbol{\xi}_Z)$ can be formulated as:

$$ {}_1C_\infty(\boldsymbol{\xi}_Z) = \frac{E\left\{{}_1C(Z(\tau), W(\tau), X)\right\}}{E\left\{{}_1T(Z(\tau), W(\tau))\right\}}, \tag{5.1} $$

where expectations have to be taken with respect to all the variables that are within the square brackets. Note that, in this case, the adopted notation explicitly shows that the long-run average maintenance cost rate depends on the design parameters vector. Similarly, in the absence of measurement error the long-run average maintenance cost rate ${}_2C_\infty(\boldsymbol{\xi}_W)$ can be expressed as:

$$ {}_2C_\infty(\boldsymbol{\xi}_W) = \frac{E\left\{{}_2C(W(\tau), X)\right\}}{E\left\{{}_2T(W(\tau))\right\}}. \tag{5.2} $$

The expected values in Eqs. (5.1) and (5.2) cannot be expressed in closed form. However, they can be numerically computed via Eqs. (5.3)-(5.6). Deriving these formulas is rather cumbersome, despite being conceptually simple. More details about their derivation can

be found in Appendix A.

$$
\begin{aligned}
E\{{}_1C(Z(\tau),W(\tau),X)\} &= c_l + c_i + c_p + [(c_c - c_p) + c_d \cdot \tau] \cdot [1 - F_{W(\tau)}(w_M)] \\
&+ (c_l + c_d \cdot \Delta\tau) \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_Z|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ (c_c - c_p) \cdot \int_0^{w_M} [1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)] \\
&\qquad \times F_{Z(\tau)|W(\tau)}(L_Z|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&- c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\
&- c_d \cdot \int_0^{w_M} \int_{\tau}^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot F_{Z(\tau)|W(\tau)}(L_Z|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau
\end{aligned} \tag{5.3}
$$

$$
E\{{}_1T(Z(\tau),W(\tau))\} = \tau + \Delta\tau \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_Z|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \tag{5.4}
$$

$$
\begin{aligned}
E\{{}_2C(W(\tau),X)\} &= c_l + c_i + c_p + (c_l + c_d \cdot \Delta\tau) \cdot F_{W(\tau)}(L_W) \\
&+ (c_c - c_p) \cdot \int_0^{L_W} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ [(c_c - c_p) + c_d \cdot \tau] \cdot [1 - F_{W(\tau)}(w_M)] \\
&- c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\
&- c_d \cdot \int_0^{L_W} \int_{\tau}^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau
\end{aligned} \tag{5.5}
$$

$$
E\{{}_2T(W(\tau))\} = \tau + \Delta\tau \cdot F_{W(\tau)}(L_W) \tag{5.6}
$$

where the expressions of the pdfs and cdfs included in these formulas are the ones given in Chapter 2.4.

5.2.4 Sensitivity analysis and comparison

This section provides the results of a small sensitivity analysis and a (simple) demonstrative example that shows how to select the best maintenance strategy when it is possible to choose between different (possibly imperfect) measurement procedures. Just to fix the ideas, we have considered an hypothetical application of the proposed approach to a pipeline subjected to corrosion. The unitary costs c_l, c_i, c_p, and c_c (expressed in monetary units) and the downtime cost rate c_d (expressed in monetary units/year) have been set as in Table 5.7, by using as rough reference the values provided in Dey (2004). The

values of the parameters used to calibrate the degradation process are given in Table 5.8. The considered settings have been loosely inspired by those given in Mahmoodian and Alani (2014) and Ellingwood and Mori (1993). Specifically, in this example, we assume that the shape parameter of the age function of the gamma process is equal to 1, a value that has been found appropriate for modeling the effect of corrosion. Nonetheless, based on Ellingwood and Mori (1993), other values of the parameter b can be possibly used to simulate degradation phenomena of different kinds (e.g., diffusion-controlled aging, reinforcement corrosion, etc.).

The pipelines are assumed to fail when their degradation level exceeds the threshold limit $w_M = 35\ mm$. The optimal values L_Z^*, τ^*, and $\Delta\tau^*$ of the design parameters and the

c_l	c_i	c_p	c_c	c_d
0.2	0.5	1	6	0.2

Table 5.7 – Parameters of the cost model.

$a\ [years]$	b	$c\ [mm]$	d	$\varphi\ [mm^{-1}]$	ν
5.88	1	2	4.7	1	1

Table 5.8 – Values used to calibrate the degradation process.

corresponding (minimum) long-run average maintenance cost rate ${}_1C_\infty(\boldsymbol{\xi}^*)$ determined by using the proposed maintenance policy, in the presence of measurement error, are given in Table 5.9. These values have been obtained by using Eqs. (5.1), (5.3), and (5.4).

Figure 5.1 shows the behavior of the long-run average maintenance cost rate ${}_1C_\infty(\boldsymbol{\xi}^*)$

${}_1C_\infty(\boldsymbol{\xi}^*)$	τ^*	$\Delta\tau^*$	L_Z^*
0.0092	134	190	10.8

Table 5.9 – Optimal values of the design parameters and corresponding (minimum) long-run average maintenance cost rate under the setup defined in Tables 5.7 and 5.8.

in the proximity of the optimum. The figure confirms that the vector $\boldsymbol{\xi}^*$, whose values are reported in Table 5.9, constitutes a minimum for ${}_1C_\infty(\boldsymbol{\xi}^*)$. Figure 5.2 displays the percent contribution of corrective and downtime costs to the (minimim) long-run average maintenance cost rate as a function of unitary corrective (c_c) and preventive (c_p) costs. As c_c and c_p vary, the other parameters have been kept to the values given in Tables 5.7 and

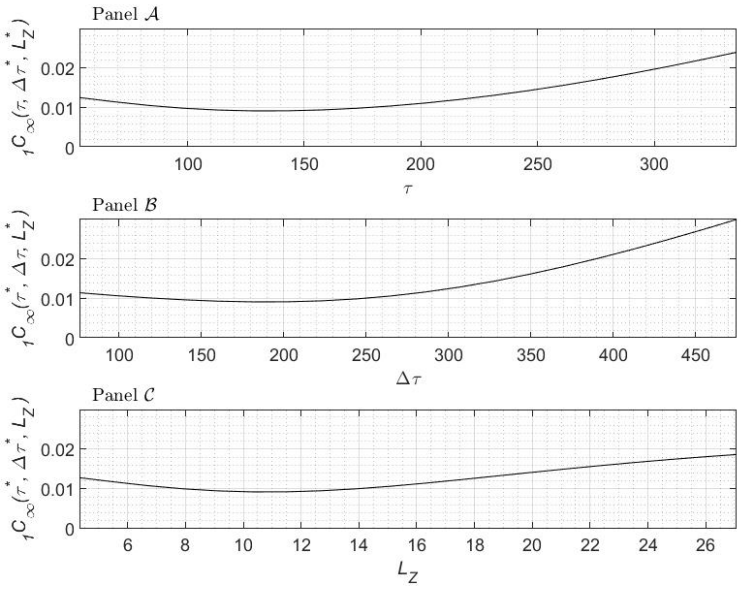

Figure 5.1 – Behavior of the long-run average maintenance cost rate as a function of τ (panel A), $\Delta\tau$ (panel B), and L_Z (panel C), given the optimal values of the other decision variables.

5.8. Results obtained in the absence of measurement error are qualitatevely analogous to those reported in Table 5.9, Figure 5.1, and Figure 5.2, with obvious difference in the optimal values of design parameters $\boldsymbol{\xi}_Z^* = \{\tau^*, \Delta\tau^*, L_W^*\}$, and long-run average maintenance cost rate ${}_2C_\infty(\boldsymbol{\xi}_W^*)$, that can be determined by using Eqs. (5.2), (5.5), and (5.6).

In fact, Figure 5.3 reports and compares the behavior of the (minimum) long-run average maintenance cost rate as a function of the inspection cost c_i in the presence (solid line) and in the absence (dashed line) of measurement error. For the sake of simplicity, the curves depicted in Figure 5.3 have been obtained by assuming that while c_i changes the other parameters of the model are kept to the values given in Tables 5.7 and 5.8.
As expected, the figure shows that (under the considered setup) the presence of measurement error increases the long-run average maintenance cost rate. Moreover, for example, the same figure shows that under the considered setting, if (as assumed in Table 5.7) the cost of the imperfect inspection is 0.5, resorting to a perfect measurement procedure (if this option exists) is convenient only if the corresponding inspection cost is smaller than 0.76.

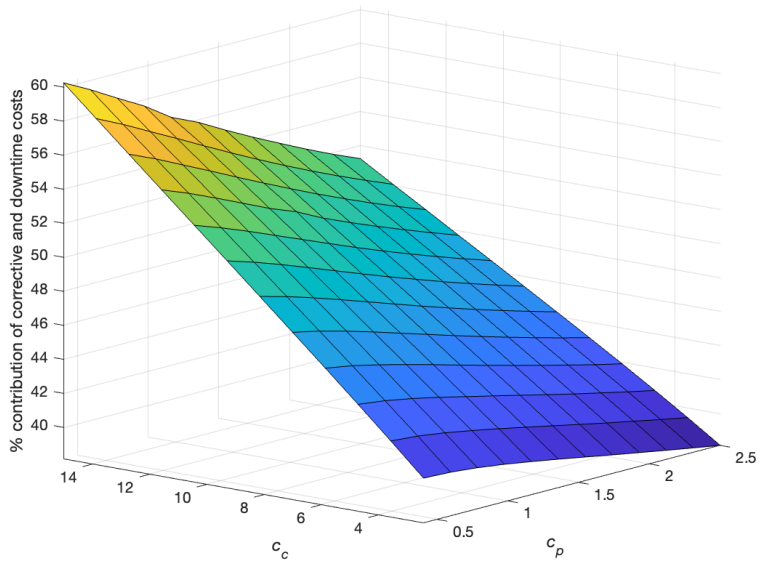

Figure 5.2 – Percent contribution of corrective and downtime costs to the long-run average maintenance cost rate as a function of c_c and c_p.

5.3 An adaptive hybrid maintenance policy for a gamma deteriorating unit in the presence of random effect

5.3.1 Description of the policy

We here suggest a new adaptive maintenance policy that generalizes the one proposed in Chapter 5.2, by allowing for the use of unit-specific condition-based replacement times (see also Esposito, Castanier, and Giorgio (2022b)). Indeed, here the future replacement time is defined by using a multilevel control limit decision rule that shares some features of the one adopted in Grall, Bérenguer, and Dieulle (2002).

Also in this case the policy is developed considering that a single inspection at a predetermined time τ is performed, and based on the outcome of this inspection a condition-based rule is adopted to decide whether to immediately replace the unit or to postpone its replacement to a later time $\tau + \Delta\tau$. Yet, differently than in Chapter 5.2, here the length of the interval $\Delta\tau$ will be determined based on the measured value of the degradation level at τ. Moreover, here we focus our attention only on the simplified case where measurement

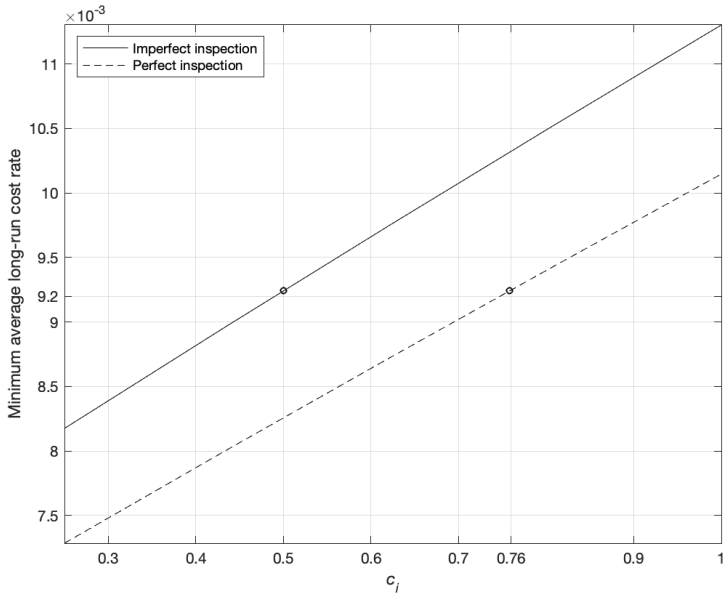

Figure 5.3 – Percent contribution of corrective and downtime costs to the long-run average maintenance cost rate as a function of c_c and c_p.

error is absent, although the proposed strategy can be adapted to also handle measurement error without excessive additional difficulties. Anyhow, the degradation process of the unit is still assumed to be modeled by the gamma process presented in Chapter 2.4, but neglecting the presence of measurement error (it is worth remarking that, in this case, this model reduces to the one firstly proposed by Lawless and Crowder (2004)).

Table 5.10 summarizes the adaptive condition-based rule, where $L_1 < L_2 < \cdots < L_k \leq w_M$ and $\Delta\tau_1 > \Delta\tau_2 > \cdots > \Delta\tau_k$.

Measurement at τ	Decision
$w_\tau > L_k$	Immediate replacement at τ
$L_{k-1} < w_\tau \leq L_k$	Postpone replacement to $\tau + \Delta\tau_k$
$\vdots$	$\vdots$
$L_1 < w_\tau \leq L_2$	Postpone replacement to $\tau + \Delta\tau_2$
$w_\tau \leq L_1$	Postpone replacement to $\tau + \Delta\tau_1$

Table 5.10 – Condition-based rule in the presence of measurement error.

The variables τ, $L_1 < L_2 < \cdots < L_k$ and $\Delta\tau_1 > \Delta\tau_2 > \cdots > \Delta\tau_k$

are treated as design parameters, and the set of design parameters is denoted by $\boldsymbol{\xi} = \{\tau, L_1, \ldots, L_k, \Delta\tau_1, \ldots, \Delta\tau_k\}$. The performance of the policy improves as k increases. On the other side, also the computational burden increases with k. To define the value of k, the minimization procedure is repeated by using increasing values of k, up to keep a satisfactory compromise between simplicity and effectiveness.

The optimal maintenance policy is defined as the one that (given the selected value of k) minimizes the long-run average maintenance cost rate, computed by using the renewal reward theorem. Note that this policy generalizes the one proposed in Chapter 5.2, which is obtained as a special case of this when $k = 1$. Table 5.11, where $w_{\tau+\Delta\tau_j}$ denotes the state

Experimental scenario	Maintenance action	Cycle length $T(w_\tau)$
$L_k < w_\tau \leq w_M$	Preventive replacement at τ	τ
$w_\tau > w_M$	Corrective replacement at τ	τ
$L_{k-1} < w_\tau \leq L_k$ and $w_{\tau+\Delta\tau_k} \leq w_M$	Preventive replacement at $\tau + \Delta\tau_k$	$\tau + \Delta\tau_k$
$L_{k-1} < w_\tau \leq L_k$ and $w_{\tau+\Delta\tau_k} > w_M$	Corrective replacement at $\tau + \Delta\tau_k$	$\tau + \Delta\tau_k$
$\vdots$	$\vdots$	$\vdots$
$w_\tau \leq L_1$ and $w_{\tau+\Delta\tau_1} \leq w_M$	Preventive replacement at $\tau + \Delta\tau_1$	$\tau + \Delta\tau_1$
$w_\tau \leq L_k$ and $w_{\tau+\Delta\tau_1} > w_M$	Corrective replacement at $\tau + \Delta\tau_1$	$\tau + \Delta\tau_1$

Table 5.11 – Possible scenarios and corresponding maintenance actions and cycle lengths.

at $\tau + \Delta\tau_j$ and w_M is the failure threshold, lists all the different experimental scenarios and reports for each one of them the maintenance action to be taken and the length of the maintenance cycle $T(w_\tau)$. Note that $T(w_\tau)$, despite the adopted notation not highlighting it, depends on $\boldsymbol{\xi}$.

Table 5.12 reports the maintenance cost $C(w_\tau, X)$ for each possible scenario, as a function of the measured degradation level w_τ and of the unit lifetime X. In this table the cost factors are denoted by using the same notation as in Chapter 5.2.

Note that, due to the assumption of not self-announcing failure, X cannot be directly

Experimental scenario	Maintenance cost $C(w_\tau, X)$
$L_k < w_\tau \leq w_M$	$c_l + c_i + c_p$
$w_\tau > w_M$	$c_l + c_i + c_c + c_d \cdot (\tau - X)$
$L_{k-1} < w_\tau \leq L_k$ and $w_{\tau+\Delta\tau_k} \leq w_M$	$2 \cdot c_l + c_i + c_p$
$L_{k-1} < w_\tau \leq L_k$ and $w_{\tau+\Delta\tau_k} > w_M$	$2 \cdot c_l + c_i + c_c + c_d \cdot (\tau + \Delta\tau_k - X)$
$\vdots$	$\vdots$
$w_\tau \leq L_1$ and $w_{\tau+\Delta\tau_1} \leq w_M$	$2 \cdot c_l + c_i + c_p$
$w_\tau \leq L_1$ and $w_{\tau+\Delta\tau_1} > w_M$	$2 \cdot c_l + c_c + c_c + c_d \cdot (\tau + \Delta\tau_1 - X)$

Table 5.12 – Possible scenarios and corresponding maintenance costs.

observed. Hence, in the cost function, it should be intended as a random variable even conditional to $W(\tau) = w_\tau$ (i.e., the state of the unit at the inspection time). To remark this aspect, the lifetime is denoted by using the capital letter. It is also worth to note that, as $T(w_\tau)$, the cost function $C(w_\tau, X)$ also depends on $\boldsymbol{\xi}$, although the adopted notation does not explicitly indicate this dependency.

5.3.2 Formulation of long-run average maintenance cost rate

By the renewal-reward theorem (e.g., see Ross (1983)), the long-run average maintenance cost rate $C_\infty(\boldsymbol{\xi})$ can be formulated as:

$$C_\infty(\boldsymbol{\xi}) = \frac{E\{C(W(\tau), X)\}}{E\{T(W(\tau))\}} \tag{5.7}$$

where expectations have to be taken with respect to all the variables that are within the parentheses. Note that, in this case, the adopted notation gives evidence that the long-run average maintenance cost rate depends on the set of design parameters $\boldsymbol{\xi}$.

The expected value $E\{T(W(\tau))\}$ at the denominator of Eq. (5.7) can be computed as:

$$\begin{aligned} E\{T(W(\tau))\} &= \int_0^{\infty} T(w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\ &= \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} T(w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau + \int_{L_k}^{\infty} T(w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \end{aligned}$$

which, from Table 5.11, can be rewritten as:

$$\begin{aligned} E\{T(W(\tau))\} &= \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} (\tau + \Delta\tau_j) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau + \int_{L_k}^{\infty} \tau \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\ &= \tau + \sum_{j=1}^{k} \Delta\tau_j \cdot [F_{W(\tau)}(L_j) - F_{W(\tau)}(L_{j-1})]. \end{aligned} \tag{5.8}$$

The expected value $E\{C(W(\tau), X)\}$ at the numerator of Eq. (5.7) can be computed as:

$$\begin{aligned} &E\{C(W(\tau), X)\} \\ &= \int_0^{\infty} \int_0^{\infty} C(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\ &= \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} \int_{\tau}^{\tau+\Delta\tau_j} C(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\ &= \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} \int_{\tau+\Delta\tau_j}^{\infty} C(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\ &+ \int_{L_k}^{w_M} \int_{\tau}^{\infty} C(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\ &+ \int_{w_M}^{\infty} \int_0^{\tau} C(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau. \end{aligned}$$

From Table 5.12, after a few simple (yet cumbersome) manipulations, $E\{C(W(\tau), X)\}$ becomes:

$$\begin{aligned} E\{C(W(\tau),X)\} &= c_l + c_i + c_p + c_l \cdot F_{W(\tau)}(L_k) + (c_c - c_p + c_d \cdot \tau) \cdot [1 - F_{W(\tau)}(w_M)] \\ &+ c_d \cdot \sum_{j=1}^{k} \Delta\tau_j \cdot [F_{W(\tau)}(L_j) - F_{W(\tau)}(L_{j-1})] \\ &+ (c_c - c_p) \cdot \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau_j)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\ &- c_d \cdot \sum_{j=1}^{k} \int_{L_{j-1}}^{L_j} \int_{\tau}^{\tau+\Delta\tau_j} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \\ &- c_d \cdot \int_{w_M}^{\infty} \int_{0}^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dx \cdot dw_\tau \end{aligned} \tag{5.9}$$

where the expressions of $f_{W(\tau)}(\cdot)$, $F_{W(\tau)}(\cdot)$, $F_{\Delta W(\tau,\tau+\Delta\tau_j)|W(\tau)}(\cdot|\cdot)$, and $F_{W(x)|W(\tau)}(\cdot|\cdot)$ are the ones given in Chapter 2.4. Under the considered degradation model, both the expected values in Eqs. (5.9) and (5.8) do not allow for a closed form solution, but can be easily computed by using numerical procedures.

5.3.3 Example of application

In order to demonstrate the effectiveness and affordability of the proposed policy, we consider an hypothetical application to the case of a corroding pipeline. Tables 5.13 and 5.14 report the values of the parameters used to calibrate the degradation process and the cost model, respectively. As in Chapter 5.2, these values have been set by using as rough reference the values provided by Mahmoodian and Alani (2014), Ellingwood and Mori (1993), and Dey (2004). The pipelines are assumed to fail when their degradation level exceeds the threshold limit $w_M = 35\ mm$. Let $C^*_{\infty,k}$ be the optimal long-run average

a $[years]$	b	c $[mm]$	d
5.88	1	50	21.25

Table 5.13 – Values used to calibrate the degradation process.

maintenance cost rate computed, under the proposed policy, where the number of classes has been set to k. Then, let $PRD(k)$ denote the percent relative difference between $C^*_{\infty,1}$

c_l	c_i	c_p	c_c	c_d
0.2	0.5	1	6	0.2

Table 5.14 – Parameters of the cost model.

and $C^*_{\infty,k}$, defined as:

$$PRD(k) = \frac{C^*_{\infty,1} - C^*_{\infty,k}}{C^*_{\infty,1}} \cdot 100. \tag{5.10}$$

It is worth to remark again that, as mentioned before, when $k = 1$ the proposed policy reduces to the one presented in Chapter 5.2. Hence, a positive value of the quantity in Eq. (5.10) should be seen as the percent reduction in the optimal cost computed with k classes compared with the optimal cost computed under the policy illustrated in Chapter 5.2.

k	$C^*_{\infty,k}$	$PRD(k)$
1	0.03806	0 %
10	0.03653	4.017 %
25	0.03650	4.098 %

Table 5.15 – Values of $C^*_{\infty,k}$ and $PRD(k)$ for $k = 1, 10, 25$.

Table 5.15 reports the values of $C^*_{\infty,k}$ and $PRD(k)$ for $k = 1$, $k = 10$, and $k = 25$, while Figure 5.4 depicts $PRD(k)$ as a function of the number of classes k.

Obtained results show that the proposed policy outperforms the one presented in Chapter 5.2. In fact, as k increases, $C^*_{\infty,k}$ and $PRD(k)$ monotonically decreases and increases, respectively, albeit less and less markedly, so that the values obtained when $k = 10$ are very close to those obtained for larger values of k. The step function represented (with the solid line) in Figure 5.5 shows the optimal value of $\Delta\tau$, denoted as $\Delta\tau^*$, varies as a function of the measured degradation level w_τ. This function has been obtained by setting $k = 25$. The dashed curve is the analog function obtained in the case $k = 1$. Note that, for the sake of facilitating the comparison, in this latter case $\Delta\tau^*$ has been determined by solving a constrained minimization problem where the inspection time τ is set to the value that is optimal for $k = 25$. These two figures clearly show the effectiveness of the proposed approach and give clear evidence of how the replacement times defined under the proposed adaptive maintenance policy could differ (depending on the unit specific degradation measure) from those planned under the less flexible policy

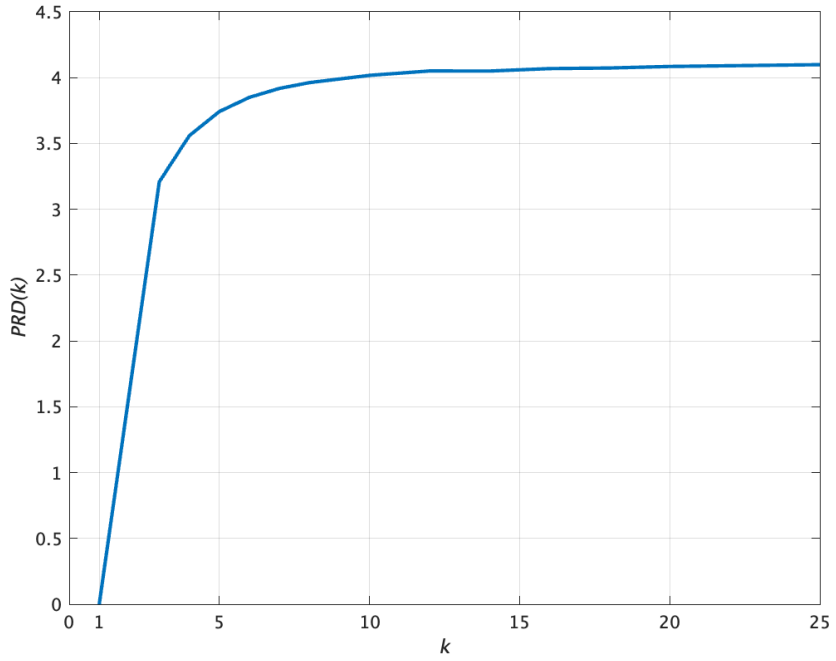

Figure 5.4 – Percent relative difference between the optimal cost computed under the current policy and the one computed as in chapter 5.2.

suggested in Chapter 5.2. Let ${}_cC^*_{\infty,k}$, ${}_dC^*_{\infty,k}$, ${}_iC^*_{\infty,k}$, ${}_lC^*_{\infty,k}$, and ${}_pC^*_{\infty,k}$ be the contribution to the optimal long-run average maintenance cost rate of corrective, downtime, inspection, logistic, and preventive costs, respectively, when the number of classes is set to k (i.e., it is $C^*_{\infty,k} = {}_cC^*_{\infty,k} + {}_dC^*_{\infty,k} + {}_iC^*_{\infty,k} + {}_lC^*_{\infty,k} + {}_pC^*_{\infty,k}$).
The bar chart in Figure 5.6 shows the values of the aforementioned contributions for the case of $k = 25$ (in blue) and $k = 1$ (in red).

While the contribution of the considered terms to $C^*_{\infty,k}$ is very similar, it is clear that the proposed generalized policy yields a smaller value for every cost factor, except for the logistic cost. This can be explained by the fact that when $k = 25$, as depicted in Figure 5.4, the policy is more likely to reschedule the replacement than when $k = 1$, and hence the last two scenarios of Table 5.12 (in which the logistic cost is sustained twice) are more likely.

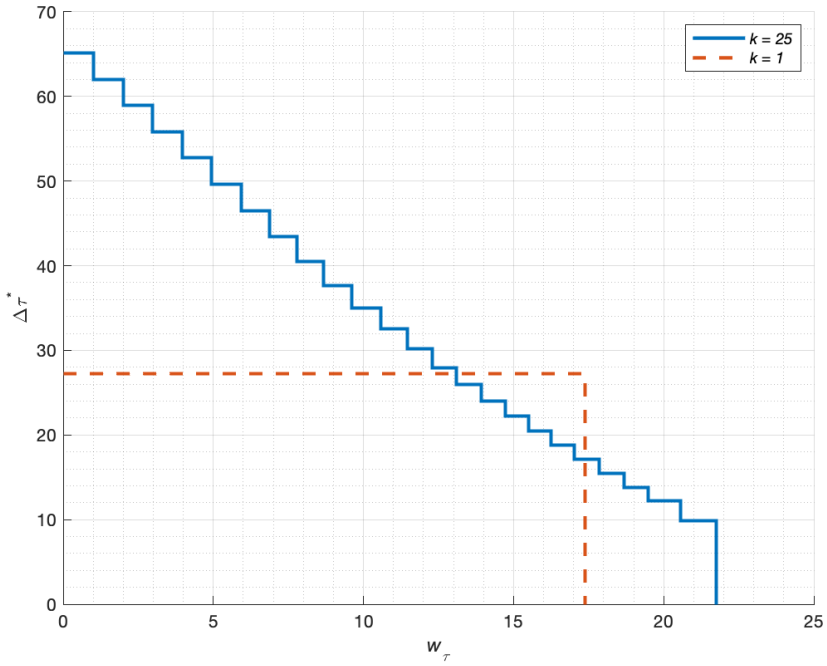

Figure 5.5 – Optimal value $\Delta\tau^*$ of $\Delta\tau$ as a function of the degradation level at the inspection w_τ for $k = 1$ (in red) and $k = 25$ (in blue).

5.4 A hybrid maintenance policy for a deteriorating unit in the presence of three forms of variability

5.4.1 Description of the policy

This policy (see also Esposito, Mele, et al. (2023a)) consists in performing a single inspection at a predetermined inspection time τ aimed at measuring the perturbed degradation level of the unit. Then, based on the outcome of this inspection, a condition-based rule is used to decide on whether to immediately replace the unit, to postpone the replacement to a later (also predetermined) time $\tau + \Delta\tau$, where the unit will be systematically replaced without any additional inspection, or to perform (at the same time τ) a second inspection, this time aimed at measuring the true degradation level. In case the second inspection is performed, the decision on the replacement will be taken based on the outcome of this perfect inspection.

Table 5.16, where z_τ denotes the observed value of the perturbed degradation measurement $Z(\tau)$ at the inspection time τ, w_τ is the true value of the actual degradation level

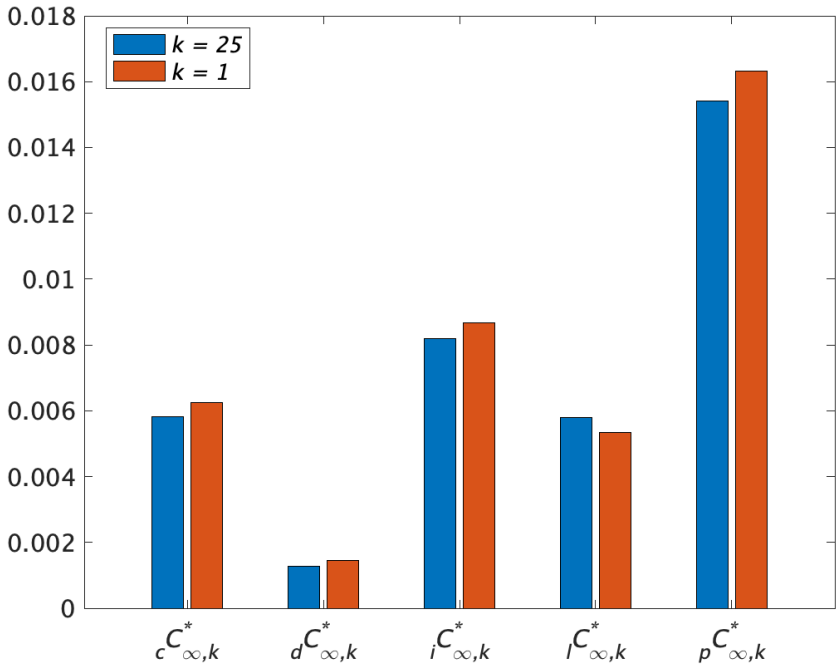

Figure 5.6 – Values of ${}_cC^*_{\infty,k}$, ${}_dC^*_{\infty,k}$, ${}_iC^*_{\infty,k}$, ${}_lC^*_{\infty,k}$, and ${}_pC^*_{\infty,k}$ obtained for $k = 25$ and $k = 1$.

$W(\tau)$ at time τ, and L_l, L_u, and L_W are (maintenance) threshold levels, summarizes the condition-based rule adopted to make the decision. The key idea of this policy is to exploit

Measurement at τ		
Perturbed	Perfect	Decision
$z_\tau > L_u$		Replacement at τ
$L_l < z_\tau \leq L_u$	$w_\tau > L_W$	Replacement at τ
$L_l < z_\tau \leq L_u$	$w_\tau \leq L_W$	Replacement at $\tau + \Delta\tau$
$z_\tau \leq L_l$		Replacement at $\tau + \Delta\tau$

Table 5.16 – Condition-based rule.

the information gathered via the perturbed inspection to timely identify and replaced the units which are already failed or close to failing, while postponing the replacement of units whose perturbed degradation level is safely below the threshold. In addition, in case it is believed that measuring also the exact degradation level can help decision-making, the perfect inspection can be additionally performed. Obviously, the presence of measurement error and the possibility of adopting two different measurement procedures complicates

decision-making. On the one hand, solely adopting the perturbed procedure saves maintenance costs, but on the other hand paying for the perfect inspection not only allows to reveal the potential failure of the unit, but also allows for better-informed decision-making.

Table 5.17 summarizes all possible experimental scenarios and reports for each of them the maintenance action to be taken and the resulting length of a maintenance cycle, denoted by $T(z_\tau, w_\tau)$. In this table, $w_{\tau+\Delta\tau}$ indicates the true value of the degradation level $W(\tau + \Delta\tau)$ at $\tau + \Delta\tau$ and w_M the failure threshold.

Measurement at τ				
Perturbed	Perfect	Lifetime	Maintenance action	Cycle length $T(z_\tau, w_\tau)$
$z_\tau > L_u$		$X < \tau$	Corrective at τ	τ
$z_\tau > L_u$		$X \geq \tau$	Preventive at τ	τ
$L_l < z_\tau \leq L_u$	$w_\tau > w_M$	$X < \tau$	Corrective at τ	τ
$L_l < z_\tau \leq L_u$	$L_W < w_\tau \leq w_M$		Preventive at τ	τ
$L_l < z_\tau \leq L_u$	$w_\tau \leq L_W$	$\tau \leq X < \tau + \Delta\tau$	Corrective at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$L_l < z_\tau \leq L_u$	$w_\tau \leq L_W$	$X \geq \tau + \Delta\tau$	Preventive at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$z_\tau < L_l$		$X \geq \tau + \Delta\tau$	Preventive at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$z_\tau < L_l$		$\tau \leq X < \tau + \Delta\tau$	Corrective at $\tau + \Delta\tau$	$\tau + \Delta\tau$
$z_\tau < L_l$		$X < \tau$	Corrective at $\tau + \Delta\tau$	$\tau + \Delta\tau$

Table 5.17 – Maintenance actions and corresponding cycle lengths

The parameter τ, $\Delta\tau$, L_l, L_u, and L_W should be intended as design parameters of the considered policy. Hereinafter, the vector of design parameters will be denoted by $\boldsymbol{\xi} = \{\tau, \Delta\tau, L_l, L_u, L_W\}$. The optimal value of $\boldsymbol{\xi}$, which defines the optimal policy is denoted as $\boldsymbol{\xi}^*$ and should be set a priori with the aim of minimizing the long-run average maintenance cost rate.

The cost model is summarized in Table 5.18 where the cost factors are denoted by using the same notation as in the previous chapters (with the only exception of the inspection cost that is not denoted by c_i anymore, but by c_z in the case of an imperfect inspection

and by c_w in the case of a perfect one). Finally, $C(z_\tau, w_\tau, X)$ is the maintenance cost per cycle formulated, scenario by scenario, for given values of the arguments z_τ, w_τ, and X.

Measurement at τ				
Perturbed	Perfect	State at τ	Lifetime	Maintenance cost
$z_\tau > L_u$		$w_\tau > w_M$	$X < \tau$	$c_l + c_z + c_c$ $+c_d \cdot (\tau - X)$
$z_\tau > L_u$		$w_\tau \leq w_M$		$c_l + c_z + c_p$
$L_l < z_\tau \leq L_u$	$w_\tau > w_M$		$X < \tau$	$c_l + c_z + c_w + c_c$ $+c_d \cdot (\tau - X)$
$L_l < z_\tau \leq L_u$	$L_W < w_\tau < w_M$			$c_l + c_z + c_w + c_p$
$L_l < z_\tau \leq L_u$	$w_\tau \leq L_W$		$\tau \leq X < \tau + \Delta\tau$	$2 \cdot c_l + c_z + c_w + c_c$ $+c_d \cdot (\tau + \Delta\tau - X)$
$L_l < z_\tau \leq L_u$	$w_\tau \leq L_W$		$X \geq \tau + \Delta\tau$	$2 \cdot c_l + c_z + c_w + c_p$
$z_\tau \leq L_l$			$X \geq \tau + \Delta\tau$	$2 \cdot c_l + c_z + c_p$
$z_\tau \leq L_l$		$w_\tau \leq w_M$	$\tau \leq X < \tau + \Delta\tau$	$2 \cdot c_l + c_z + c_c$ $+c_d \cdot (\tau + \Delta\tau - X)$
$z_\tau \leq L_l$		$w_\tau > w_M$	$X < \tau$	$2 \cdot c_l + c_z + c_c$ $+c_d \cdot (\tau + \Delta\tau - X)$

Table 5.18 – Maintenance costs

It is worth to emphasize that, coherently with the assumption of not self-announcing failure, in Table 5.18 the lifetime X is always denoted with the capital letter to intend that, even in the special case of z_τ and w_τ being known, X should be regarded as a random variable. It is also worth to remark that, although $T(z_\tau, w_\tau)$ and $C(z_\tau, w_\tau, X)$ functionally depend on $\boldsymbol{\xi}$, for economy of notation the adopted symbols do not highlight this dependency.

5.4.2 Formulation of the long-run average maintenance cost rate

By using the renewal/reward theorem (e.g., see Ross (1983)), the long-run average maintenance cost rate $C_\infty(\boldsymbol{\xi})$ can be formulated as:

$$C_\infty(\boldsymbol{\xi}) = \frac{E\{C(Z(\tau), W(\tau), X)\}}{E\{T(Z(\tau), W(\tau))\}} \tag{5.11}$$

where the expectations have been taken with respect to $Z(\tau)$, $W(\tau)$, and X. The expectations in Eq. (5.11) do not allow for closed form expressions, but can be numerically computed by using the (5.12) and (5.13), respectively.

$$
\begin{aligned}
E\{C(Z(\tau),W(\tau),X)\} &= c_l + c_z + c_p + (c_c - c_p + c_d \cdot \tau) \cdot \left[1 - F_{W(\tau)}(w_M)\right] \\
&+ c_w \cdot \int_0^{\infty} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ (c_l + c_d \cdot \Delta\tau) \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ (c_l + c_d \cdot \Delta\tau) \cdot \int_0^{L_W} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ (c_c - c_p) \cdot \int_0^{L_W} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ (c_c - c_p) \cdot \int_0^{w_M} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&- c_d \cdot \int_0^{w_M} \int_\tau^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx \\
&- c_d \cdot \int_0^{L_W} \int_\tau^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \\
&\qquad \times \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx \\
&- c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx,
\end{aligned}
\tag{5.12}
$$

$$
\begin{aligned}
E\{T(Z(\tau),W(\tau))\} &= \tau + \Delta\tau \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&+ \Delta\tau \cdot \int_0^{L_W} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau.
\end{aligned}
\tag{5.13}
$$

Details about the derivations of Eqs. (5.12) and (5.13) can be found in Appendix B.1

5.4.3 Application example

To demonstrate its effectiveness, we have applied the proposed policy to a real-world inspired case of a corroding pipeline. Pipelines are often buried underground, offshore, or in general in places that are difficult to access, so inspections are usually very costly and

failures (e.g., the considered soft failure consisting in excess of corrosion) and/or small leaks are difficult to detect unless specific inspections are performed.

The application has been developed by considering eight realistic experimental scenarios, which have been defined by adopting the setups described in Table 5.19. The cost parameters reported in the table have been inspired by Dey (2004). Without loss of generality, in this example, it is assumed that the unit-specific gamma processes are homogeneous with age function $\eta(t) = a \cdot t$, an assumption that was found to be appropriate for modeling the effect of corrosion (e.g., see Ellingwood and Mori (1993)). However, other modeling solutions can be adopted to describe degradation phenomena of different kinds (e.g., aging, reinforcement corrosion, etc.) without serious additional difficulties.

The parameter φ has been calibrated to define two groups of experimental scenarios where the measurement errors have different magnitude. Specifically, subscript 1 identifies the scenarios where the measurement errors are moderate (in this case it is $\varphi = 0.001$) while subscript 2 identifies those where the measurement errors are severe (in this case it is $\varphi = 0.005$). The setups, such as $\mathcal{A}_1$ and $\mathcal{A}_2$, that are indicated by the same letter with a different subscript only differ in the magnitude of the measurement errors. The logistic cost c_l is always equal to 0.2. The cost of the preventive replacement c_p is always equal to 2. The cost of the corrective replacement c_c is always equal to 6. The values of the other cost parameters vary according to the letter (i.e., $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$) used to identify the setup. The units are considered failed when their degradation level exceeds the value $w_M = 35\ mm$. The parameters of the hidden degradation process are set to the values

	c_l	c_z	c_w	c_p	c_c	c_d	$\varphi\ [mm^{-1}]$
Setup $\mathcal{A}_1$	0.2	0.2	1	2	6	0.6	0.01
Setup $\mathcal{A}_2$							0.005
Setup $\mathcal{B}_1$	0.2	0.2	0.8	2	6	0.6	0.01
Setup $\mathcal{B}_2$							0.005
Setup $\mathcal{C}_1$	0.2	0.2	0.4	2	6	0.6	0.01
Setup $\mathcal{C}_2$							0.005
Setup $\mathcal{D}_1$	0.2	0.8	1.2	2	6	0.1	0.01
Setup $\mathcal{D}_2$							0.005

Table 5.19 – Setups used to evaluate the performance of the proposed policy

reported in Table 5.20, which have been selected by using as rough reference those reported in Mahmoodian and Alani (2014). Without loss of generality, the value of the parameter ν is always set equal to 1. The results obtained by applying the suggested policy, hereinafter

a $[years]$	c $[mm]$	d	ν
0.17	5	2.125	1

Table 5.20 – Values of the parameters of the hidden degradation process.

indicated as $\mathcal{P}_0$, are compared to those obtained by adopting three alternative policies, indicated by $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$. The policies $\mathcal{P}_1$ and $\mathcal{P}_2$ are defined by assuming that, at the inspection time, only the perturbed and only the perfect measurement procedures can be adopted, respectively. The policy $\mathcal{P}_3$ is a classical (pure) age-based maintenance policy, which assumes that the unit is replaced at τ regardless of its state and without performing any inspection. It is worth to remark that the policies $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$ can be obtained as special cases from the policy $\mathcal{P}_0$ by imposing the following constraints:

— $\mathcal{P}_1$ (which uses only the imperfect measurement procedure) is obtained from $\mathcal{P}_0$ by setting $L_l = L_u$;

— $\mathcal{P}_2$ (which uses only the perfect measurement procedure) is obtained from $\mathcal{P}_0$ by setting $L_l = 0$, $L_u = +\infty$, and $c_z = 0$;

— $\mathcal{P}_3$ (which does not use any measurement procedure) is obtained from $\mathcal{P}_0$ by setting $L_l = L_u = 0$ and $c_z = 0$.

All the policies have been applied to determine, under each setup, the optimal set of design parameters and the corresponding optimal long-run average maintenance cost rate.

Setup by setup, to compare the performances of the policy $\mathcal{P}_0$ with those of the policies $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$ the percent relative difference $PRD(k)$ $(k = 1, 2, 3)$ has been evaluated:

$$PRD(k) = \frac{{}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*}) - {}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})}{{}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})} \cdot 100, \tag{5.14}$$

where ${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$ denotes the long-run average maintenance cost rate obtained by adopting the policy $\mathcal{P}_k$ $(k = 0, \ldots, 3)$ and ${}^{k}\boldsymbol{\xi}^{*}$ indicates the corresponding optimal set of design parameters, being ${}^{0}\boldsymbol{\xi} = \{{}^{0}\tau, {}^{0}\Delta\tau, {}^{0}L_l, {}^{0}L_u, {}^{0}L_W\}$, ${}^{1}\boldsymbol{\xi} = \{{}^{1}\tau, {}^{1}\Delta\tau, {}^{1}L_l = {}^{1}L_u\}$, ${}^{2}\boldsymbol{\xi} = \{{}^{2}\tau, {}^{2}\Delta\tau, {}^{2}L_W\}$, and ${}^{3}\boldsymbol{\xi} = \{{}^{3}\tau\}$. Thus, ${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$ is the long-run average maintenance cost rate given in Eq. (5.12), whereas the expressions for ${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$ under policies $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$ are provided in Appendix B.2.

Table 5.21 reports the optimal long-run average maintenance cost rate and the corresponding optimal values of the design parameters obtained under all the considered policies in the case of the setups $\mathcal{A}_1$ and $\mathcal{A}_2$. Tables 5.22, 5.23, and 5.24 report the same

quantities in the cases of the setups $\mathcal{B}_1$ and $\mathcal{B}_2$, $\mathcal{C}_1$ and $\mathcal{C}_2$, and $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively. Obviously, the results obtained under the policies $\mathcal{P}_2$ and $\mathcal{P}_3$, which do not use the imperfect measurement procedure, are not affected by the magnitude of the measurement errors. Thus, these results do not change passing from the left to the right side of the mentioned tables. For similar reasons, the results obtained under the policy $\mathcal{P}_1$ do not depend on c_w, those obtained under the policy $\mathcal{P}_2$ do not depend on c_z, and those obtained under the policy $\mathcal{P}_3$ depend neither on c_w nor on c_z. Table 5.21 shows that, under the

Setup		${}^kC_\infty({}^k\boldsymbol{\xi}^*)$	$PRD(k)$	${}^k\tau^*$	${}^k\Delta\tau^*$	${}^kL_l^*$	${}^kL_u^*$	${}^kL_W^*$
$\mathcal{A}_1$	$\mathcal{P}_0$	0.1005		22.2	43.9	3.76	7.81	10.3
	$\mathcal{P}_1$	0.1022	1.62 %	21.0	39.6	5.02	5.02	
	$\mathcal{P}_2$	0.1062	5.58 %	24.6	48.7		10.5	
	$\mathcal{P}_3$	0.1276	26.9 %	32.0				
$\mathcal{A}_2$	$\mathcal{P}_0$	0.1009		22.2	44.2	3.55	7.72	10.3
	$\mathcal{P}_1$	0.1027	1.83 %	20.9	39.7	4.81	4.81	
	$\mathcal{P}_2$	0.1062	5.20 %	24.6	48.7		10.5	
	$\mathcal{P}_3$	0.1276	26.4 %	32.0				

Table 5.21 – Optimal values of the design parameters, of the long-run average maintenance cost rate, and of $PRD(\cdot)$ under all the considered policies in the case of setups $\mathcal{A}_1$ and $\mathcal{A}_2$.

setups $\mathcal{A}_1$ and $\mathcal{A}_2$, the policy which yields the lowest long-run average maintenance cost rate (i.e., the best policy) is $\mathcal{P}_0$. It also shows that, under these setups, $\mathcal{P}_1$ outperforms $\mathcal{P}_2$, which in turn performs better than $\mathcal{P}_3$. In fact, for example, under the setup $\mathcal{A}_1$ the use of the policies $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$, in place of the policy $\mathcal{P}_0$, leads to percent increases of the long-run average maintenance cost rate equal to 1.62%, 5.58%, and 26.9%, respectively.

The setups $\mathcal{B}_1$ and $\mathcal{B}_2$ differ from $\mathcal{A}_1$ and $\mathcal{A}_2$ only for the value of c_w, which is 0.8 instead of 1. Table 5.22 shows that under the setups $\mathcal{B}_1$ and $\mathcal{B}_2$ the policies $\mathcal{P}_0$ and $\mathcal{P}_3$ are still the best and the worst one, respectively. On the other hand, due to the reduced value of c_w, which incentivizes the use of the perfect inspection, in these cases $\mathcal{P}_2$ outperforms $\mathcal{P}_1$. In fact, for example, in the case of the setup $\mathcal{B}_1$, using the policies $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$, instead of $\mathcal{P}_0$ produces a percent increase of the long-run average maintenance cost rate equal to 2.94%, 2.77%, and 28.5%, respectively.

The setups $\mathcal{C}_1$ and $\mathcal{C}_2$ differ from the setups $\mathcal{A}_1$, $\mathcal{B}_1$ and $\mathcal{A}_2$, $\mathcal{B}_2$ in the value of c_w only, which is smaller than in the other mentioned cases. These setups reflect an experimental scenario in which the perfect measurement procedure is only slightly more expensive than

Setup		${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$	$PRD(k)$	${}^{k}\tau^{*}$	${}^{k}\Delta\tau^{*}$	${}^{k}L_{l}^{*}$	${}^{k}L_{u}^{*}$	${}^{k}L_{W}^{*}$
$\mathcal{B}_1$	$\mathcal{P}_0$	0.0993		22.4	44.6	3.41	9.42	10.2
	$\mathcal{P}_1$	0.1022	2.94 %	21.0	39.6	5.02	5.02	
	$\mathcal{P}_2$	0.1020	2.77 %	24.0	47.9		10.3	
	$\mathcal{P}_3$	0.1276	28.5 %	32.0				
$\mathcal{B}_2$	$\mathcal{P}_0$	0.0996		22.4	44.8	3.23	9.35	10.2
	$\mathcal{P}_1$	0.1027	3.22 %	20.9	39.7	4.81	4.81	
	$\mathcal{P}_2$	0.1020	2.47 %	24.6	47.9		10.3	
	$\mathcal{P}_3$	0.1276	26.4 %	32.0				

Table 5.22 – Optimal values of the design parameters, of the long-run average maintenance cost rate, and of $PRD(\cdot)$ under all the considered policies in the case of setups $\mathcal{B}_1$ and $\mathcal{B}_2$.

Setup		${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$	$PRD(k)$	${}^{k}\tau^{*}$	${}^{k}\Delta\tau^{*}$	${}^{k}L_{l}^{*}$	${}^{k}L_{u}^{*}$	${}^{k}L_{W}^{*}$
$\mathcal{C}_1$	$\mathcal{P}_0$	0.0954		22.5	45.1	2.68	16.85	10.0
	$\mathcal{P}_1$	0.1022	7.09 %	21.0	39.6	5.02	5.02	
	$\mathcal{P}_2$	0.0934	-2.06 %	22.8	46.0			9.83
	$\mathcal{P}_3$	0.1276	33.7 %	32.0				
$\mathcal{C}_2$	$\mathcal{P}_0$	0.0956		22.5	45.2	2.55	17.0	9.96
	$\mathcal{P}_1$	0.1027	7.54 %	20.9	39.7	4.81	4.81	
	$\mathcal{P}_2$	0.0934	-2.19 %	22.8	46.0			9.83
	$\mathcal{P}_3$	0.1276	33.5 %	32.0				

Table 5.23 – Optimal values of the design parameters, of the long-run average maintenance cost rate, and of $PRD(\cdot)$ under all the considered policies in the case of setups $\mathcal{C}_1$ and $\mathcal{C}_2$.

the imperfect one. Table 5.23 shows that, under $\mathcal{C}_1$ and $\mathcal{C}_2$, the best policy is $\mathcal{P}_2$, which uses the perfect measurement procedure only. The same table also shows that $\mathcal{P}_0$, which allows for the optional use of the perfect measurement procedure, performs better than $\mathcal{P}_1$. As in the cases of the setups $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}_1$, and $\mathcal{B}_2$, the policies $\mathcal{P}_0$, $\mathcal{P}_1$, and $\mathcal{P}_2$ outperform the (pure) age-based policy $\mathcal{P}_3$. In fact, for example, under the setup $\mathcal{C}_1$ the use of $\mathcal{P}_2$ in place of $\mathcal{P}_0$ leads to a percent reduction of the long-run average maintenance cost rate equal to 2.06%, while $\mathcal{P}_1$ and $\mathcal{P}_3$ provide long-run average maintenance cost rates that are 7.09% and 33.7% higher than the one provided by $\mathcal{P}_0$. With respect to the other setups, $\mathcal{D}_1$ and $\mathcal{D}_2$ depict scenarios characterized by higher inspection costs c_w and c_z and smaller downtime cost rate c_d. Table 5.24 shows that, under these setups, the best policy is $\mathcal{P}_3$, followed by $\mathcal{P}_2$ and $\mathcal{P}_0$, which performs slightly better than $\mathcal{P}_1$.

Setup		${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})$	$PRD(k)$	${}^{k}\tau^{*}$	${}^{k}\Delta\tau^{*}$	${}^{k}L_{l}^{*}$	${}^{k}L_{u}^{*}$	${}^{k}L_{W}^{*}$
$\mathcal{D}_1$	$\mathcal{P}_0$	0.0809		47.5	93.5	7.73	10.57	15.9
	$\mathcal{P}_1$	0.0810	0.20 %	46.6	90.7	9.15	9.15	
	$\mathcal{P}_2$	0.0792	-2.06 %	44.5	92.1			15.2
	$\mathcal{P}_3$	0.0772	-4.55 %	75.7				
$\mathcal{D}_2$	$\mathcal{P}_0$	0.0812		48.0	94.3	7.33	10.3	16.0
	$\mathcal{P}_1$	0.0813	0.23 %	47.0	90.9	8.86	8.86	
	$\mathcal{P}_2$	0.0792	-2.40 %	44.5	92.1			15.2
	$\mathcal{P}_3$	0.0772	-4.87 %	75.7				

Table 5.24 – Optimal values of the design parameters, of the long-run average maintenance cost rate, and of $PRD(\cdot)$ under all the considered policies in the case of setups $\mathcal{D}_1$ and $\mathcal{D}_2$.

The results obtained under these latter setups show that, when the inspection costs are high with respect to other maintenance costs, performing inspections does not necessarily give economic advantages in the long run. Tables 5.21-5.24 also show that the magnitude of the measurement error (within the considered range) does not significantly influence the discussed results. Hence, the comments provided above apply both when measurements are moderately and severely perturbed. In fact, the value of φ only influences the long-run average maintenance cost rate under the policies $\mathcal{P}_0$ and $\mathcal{P}_1$, which in general slightly increases as φ decreases. Consequently, also the percent relative difference $PRD(\cdot)$ slightly changes passing from the case where the error is moderate to the one when it is severe.

The severity of the measurement error can be expressed in terms of the ratio $V\{Z(t)\}/V\{W(t)\}$. Unfortunately, both the variance of $W(t)$ and $Z(t)$ (see Eqs. (2.52) and (2.61)), as well as $V\{Z(t)\}/V\{W(t)\}$ depend on t. Hence, it is not possible to quantify the ratio $V\{Z(t)\}/V\{W(t)\}$ by a unique number. However, it can be useful to note that, for example, in the case of the setup $\mathcal{A}_1$ at ${}^{0}\tau^{*} = 22.2$ it is $V\{Z({}^{0}\tau^{*})\}/V\{W({}^{0}\tau^{*})\} = 1.57$ while under the setup $\mathcal{A}_2$ at ${}^{0}\tau^{*} = 22.2$ it is $V\{Z({}^{0}\tau^{*})\}/V\{W({}^{0}\tau^{*})\} = 2.15$. Similarly, under the setup $\mathcal{D}_1$ at ${}^{0}\tau^{*} = 47.5$ it is $V\{Z({}^{0}\tau^{*})\}/V\{W({}^{0}\tau^{*})\} = 1.30$ while under the setup $\mathcal{D}_2$ at ${}^{0}\tau^{*} = 48.0$ it is $V\{Z({}^{0}\tau^{*})\}/V\{W({}^{0}\tau^{*})\} = 1.61$.

To deepen and better understand how the results provided by the considered policies change with the setup, it is useful to investigate how the long-run average maintenance cost rate ${}^{k}C_{\infty}(\cdot)$ splits into the components ${}^{k}_{c}C_{\infty}(\cdot)$, ${}^{k}_{d}C_{\infty}(\cdot)$, ${}^{k}_{l}C_{\infty}(\cdot)$, ${}^{k}_{p}C_{\infty}(\cdot)$, ${}^{k}_{w}C_{\infty}(\cdot)$, and ${}^{k}_{z}C_{\infty}(\cdot)$, that denote the contributions to ${}^{k}C_{\infty}(\cdot)$ of corrective replacement, downtime, logistic, preventive replacement, perfect inspection, and perturbed inspection costs, re-

spectively. The expressions of these cost rates are given in Appendix B.3. Obviously, it is ${}^{k}C_{\infty}(\cdot) = {}^{k}_{c}C_{\infty}(\cdot)+{}^{k}_{d}C_{\infty}(\cdot)+{}^{k}_{l}C_{\infty}(\cdot)+{}^{k}_{p}C_{\infty}(\cdot)+{}^{k}_{w}C_{\infty}(\cdot)+{}^{k}_{z}C_{\infty}(\cdot)$. Figure 5.7 and Figure 5.8 show how in the case of setups $\mathcal{A}_1$ and $\mathcal{D}_1$, respectively, the (minimum) long-run average maintenance cost rate splits into the aforementioned contributions. The setup $\mathcal{D}_1$ differs from the setup $\mathcal{A}_1$ for the values of c_w, c_z, and c_d, which pass from 0.8, 0.2, and 0.6 under setup $\mathcal{A}_1$ to 1.2, 0.8, and 0.1 under setup $\mathcal{D}_1$, respectively.

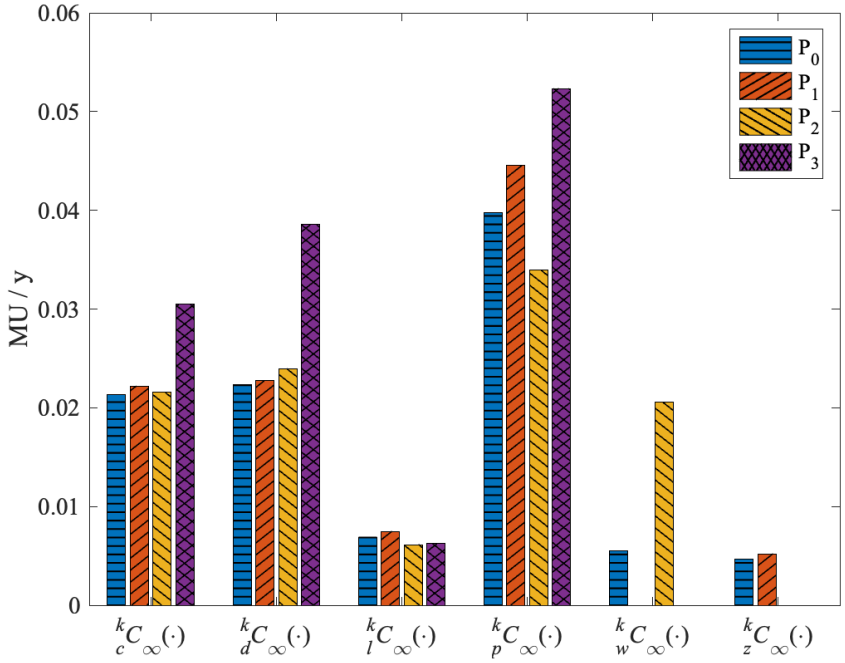

Figure 5.7 – Values of ${}^{k}_{c}C_{\infty}(\cdot)$, ${}^{k}_{d}C_{\infty}(\cdot)$, ${}^{k}_{l}C_{\infty}(\cdot)$, ${}^{k}_{p}C_{\infty}(\cdot)$, ${}^{k}_{w}C_{\infty}(\cdot)$, and ${}^{k}_{z}C_{\infty}(\cdot)$ under each of the considered policies in the case of the setup $\mathcal{A}_1$.

From Figure 5.7 it is apparent that in the case of the setup $\mathcal{A}_1$ the main contribution, under all the policies, is ${}^{k}_{p}C_{\infty}(\cdot)$, followed by ${}^{k}_{d}C_{\infty}(\cdot)$ and ${}^{k}_{c}C_{\infty}(\cdot)$. This shows that, under this setup, the maintenance cycle ends preferentially with a preventive replacement, hence units are generally replaced when they are not yet failed. In fact, replacing failed units implies not only a higher (compared to the preventive one) corrective replacement cost but also a downtime cost that, depending on the setup, can be quite hefty.
Conversely, in the case of the setup $\mathcal{D}_1$, where the downtime cost rate c_d is lower than under the other setups, the consequences of failures are milder. Hence, the setup $\mathcal{D}_1$, compared to the setup $\mathcal{A}_1$, leads all of the considered policies to incentivize more corrective replacements, by increasing the probability of failure in the maintenance cycle. This sit-

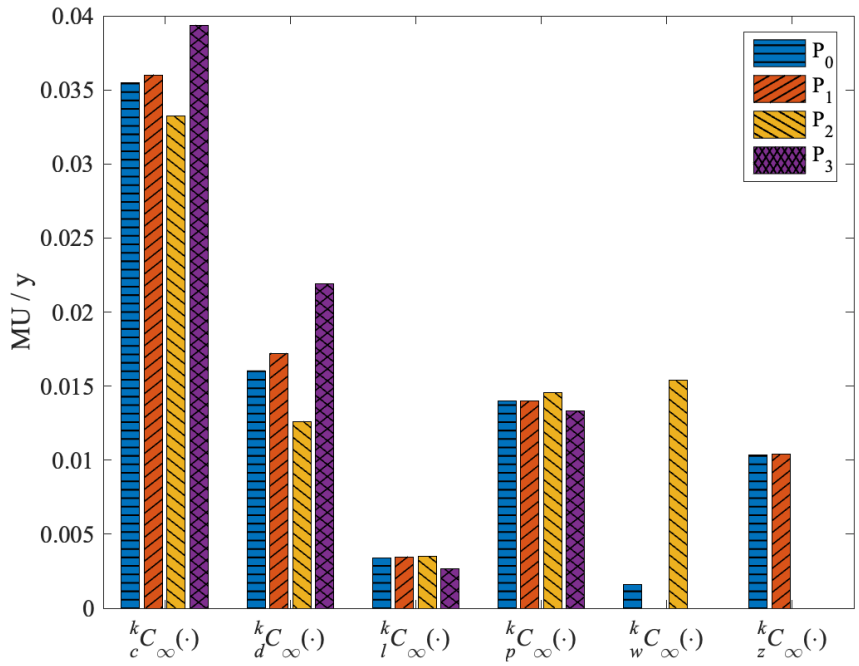

Figure 5.8 – Values of ${}^{k}_{c}C_{\infty}(\cdot)$, ${}^{k}_{d}C_{\infty}(\cdot)$, ${}^{k}_{l}C_{\infty}(\cdot)$, ${}^{k}_{p}C_{\infty}(\cdot)$, ${}^{k}_{w}C_{\infty}(\cdot)$, and ${}^{k}_{z}C_{\infty}(\cdot)$ under each of the considered policies in the case of the setup $\mathcal{D}_1$.

uation is clearly shown in Figure 5.8, where it is apparent that (under all the considered policies) the dominant contribution is now ${}^{k}_{c}C_{\infty}(\cdot)$. Tables 5.21 and 5.24 show that this result is primarily accomplished by lengthening the maintenance cycle. In fact, from these tables, it is immediate to verify that the values of ${}^{k}\tau^{*}$ and ${}^{k}\Delta\tau^{*}$ obtained under the setup $\mathcal{D}_1$ are all higher than the corresponding values obtained under the setup $\mathcal{A}_1$.

5.5 Impact on maintenance decision-making of misspecification of gamma with inverse Gaussian process

Expanding on the analysis performed in Chapter 4, here we investigate how the misspecification issue of gamma and inverse Gaussian processes impacts on maintenance costs (see also Esposito, Castanier, Giorgio, et al. (2022)). However, here we focus solely on the consequences of the misspecification of a gamma process with an inverse Gaussian process

(i.e., we consider neither measurement error nor the symmetric case of the misspecification of an inverse Gaussian process with a gamma one). The processes adopted in this analysis are still the ones presented in Chapter 2.3, though only the hidden processes should be considered here. Maintenance costs are computed under the simple condition-based policy presented in Chapter 5.2 (obviously, the case where measurement error is absent should be considered).

Similarly to Chapter 4, to perform the analysis we have generated 3 sets of 1000 synthetic datasets under 3 different gamma degradation processes. As first step, dataset by dataset, the gamma and inverse Gaussian models are fitted to data by means of ML estimation. Hence, all the estimated processes are used to optimize the considered maintenance policy. A misspecification is assumed to occur if the Akaike information criterion leads to prefer the inverse Gaussian process. The effect of a misspecification is evaluated in terms of long-run average maintenance cost rate.

Neglecting the presence of measurement error hugely simplifies the estimation task. Indeed, there is no need here for the EM algorithm and classical ML estimation can be performed. Details about the formulation of the likelihood function under the considered setup can be found in Chapter 5.5.1. Then, the results of the misspecification analysis are illustrated in Chapter 5.5.2.

5.5.1 Maximum likelihood estimation of model parameters

Let us consider m identical units operating under homogeneous conditions, suppose that the degradation level of the unit i $(i = 1, ..., m)$ is measured at times $t_{i,1}, ..., t_{i,n_i}$, and denote by $W_{i,j} = W(t_{i,j})$ the degradation level of the unit i at the measurement time $t_{i,j}$, by $w_{i,j}$ its realization, and by $\boldsymbol{\xi}$ the vector of model parameters. Then, under these setting, the likelihood function of the considered competing processes can be formulated as:

$$L(\boldsymbol{\xi}; \boldsymbol{w}) = \prod_{i=1}^{m} \prod_{j=1}^{n_i} f_{\Delta W_{i,j}}(\Delta w_{i,j}), \tag{5.15}$$

where $\boldsymbol{w} = \{w_{1,1}, ..., w_{1,n_1}, ..., w_{m,1}, ..., w_{m,n_m}\}$, $t_{i,0} = 0$, $w_{i,0} = 0$, $\Delta W_{i,j} = W_{i,j} - W_{i,j-1}$, $\Delta w_{i,j} = w_{i,j} - w_{i,j-1}$. Therefore, the log-likelihood function $l(\boldsymbol{\xi}; \boldsymbol{w})$ can be formulated as:

$$l(\boldsymbol{\xi}; \boldsymbol{w}) = ln[L(\boldsymbol{\xi}; \boldsymbol{w})] = \sum_{i=1}^{m} \sum_{j=1}^{n_i} ln \left[f_{\Delta W_{i,j}}(\Delta w_{i,j}) \right]. \tag{5.16}$$

The maximum likelihood estimate (MLE) of the parameter vector $\widehat{\boldsymbol{\xi}}$ is the value of $\boldsymbol{\xi}$ that maximizes $l(\boldsymbol{\xi}, \boldsymbol{w})$, given $\boldsymbol{w}$. In the case of the considered gamma process, $\boldsymbol{\xi} = \{a, b, \theta\}$ and, from Eq. (2.11), the log-likelihood function in Eq. (5.16) can be expressed as:

$$\begin{aligned} l(\boldsymbol{\xi}; \boldsymbol{w}) = &- \frac{\sum_{i=1}^{m} w_{i,n_i}}{\theta} + \sum_{i=1}^{m}\sum_{j=1}^{n_i}\left[\left(\frac{t_{i,j}}{a}\right)^b - \left(\frac{t_{i,j-1}}{a}\right)^b\right] \cdot \ln(\Delta w_{i,j}) \\ &- \ln(\theta) \cdot \sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^b - \sum_{i=1}^{m}\sum_{j=1}^{n_i} \ln\left[\Gamma\left(\left(\frac{t_{i,j}}{a}\right)^b - \left(\frac{t_{i,j-1}}{a}\right)^b\right)\right]. \end{aligned} \tag{5.17}$$

Then, by solving with respect to θ the equation:

$$\frac{\partial l(\boldsymbol{\xi}; \boldsymbol{w})}{\partial \theta} = \frac{\partial l(a, b, \theta; \boldsymbol{w})}{\partial \theta} = 0$$

the explicit form:

$$\tilde{\theta}(a, b) = \frac{\sum_{i=1}^{m} w_{i,n_i}}{\sum_{i=1}^{m}\left(\frac{t_{i,n_i}}{a}\right)^b} \tag{5.18}$$

is obtained for the value $\tilde{\theta}(a, b)$ that maximizes the log-likelihood function (5.17) with respect to θ when a and b are set to the values indicated in the parentheses. Exploiting this result, the MLEs $\hat{a}$ and $\hat{b}$ of a and b can be conveniently retrieved by numerically maximizing, with respect to a and b, the two parameter profile log-likelihood $l_p(a, b; \boldsymbol{w})$ that is obtained by replacing $\tilde{\theta}(a, b)$ to θ in the log-likelihood in Eq. (5.17). Once $\hat{a}$ and $\hat{b}$ are available, the MLE $\hat{\theta}$ of θ can be readily computed from Eq. (5.18) as $\hat{\theta} = \tilde{\theta}(\hat{a}, \hat{b})$.

Whereas in the case of the inverse Gaussian process, from Eq. (2.12), the log-likelihood function (5.16) becomes:

$$\begin{aligned} l(\boldsymbol{\xi}; \boldsymbol{w}) = &- \frac{n_t}{2} \cdot \ln(2 \cdot \pi) + \frac{n_t}{2} \cdot \ln(\theta) - n_t \cdot b \cdot \ln(a) - \frac{3}{2} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} \ln(\Delta w_{i,j}) \\ &- \frac{1}{2 \cdot \theta} \cdot \sum_{i=1}^{m} w_{i,n_i} - \frac{\theta}{2 \cdot a^{2 \cdot b}} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{\left(t_{i,j}^b - t_{i,j-1}^b\right)^2}{\Delta w_{i,j}} \\ &+ \frac{1}{a^b} \cdot \sum_{i=1}^{m} t_{i,n_i}^b + \sum_{i=1}^{m}\sum_{j=1}^{n_i} \ln\left(t_{i,j}^b - t_{i,j-1}^b\right), \end{aligned} \tag{5.19}$$

where $n_t = \sum_{i=1}^{m} n_i$. Then, from Eq. (5.17), by solving with respect to θ and a the following

system of equations:

$$\begin{cases} \frac{\partial l(\boldsymbol{\xi};\boldsymbol{w})}{\partial \theta} = \frac{\partial l(a,b,\theta;\boldsymbol{w})}{\partial \theta} = 0 \\ \frac{\partial l(\boldsymbol{\xi};\boldsymbol{w})}{\partial a} = \frac{\partial l(a,b,\theta;\boldsymbol{w})}{\partial a} = 0 \end{cases}$$

we get the explicit forms:

$$\tilde{\theta}(b) = \frac{\left(\frac{\sum_{i=1}^{m} w_{i,n_i}}{\sum_{i=1}^{m} t_{i,n_i}^b}\right)^2 \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{\left(t_{i,j}^b - t_{i,j-1}^b\right)^2}{\Delta w_{i,j}} - \sum_{i,j}^{m} w_{i,n_i}}{n_t}, \tag{5.20}$$

$$\tilde{a}(b) = \left[\frac{\frac{\sum_{i=1}^{m} w_{i,n_i}}{\sum_{i=1}^{m} t_{i,n_i}^b} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{\left(t_{i,j}^b - t_{i,j-1}^b\right)^2}{\Delta w_{i,j}} - \sum_{i=1}^{m} t_{i,n_i}^b}{n_t}\right]^{\frac{1}{b}} \tag{5.21}$$

for the values $\tilde{\theta}(b)$ and $\tilde{a}(b)$ of θ and a that (jointly) maximize the log-likelihood function (5.19) when the parameter b is set to the value reported in the parentheses.
By using these results, the MLE $\hat{b}$ of b can be conveniently retrieved by numerically maximizing the one-parameter profile log-likelihood $l_p(b;\boldsymbol{w})$, that is obtained by replacing $\tilde{\theta}(b)$ and $\tilde{a}(b)$ to θ and a in the log-likelihood (5.19). Once $\hat{b}$ is available the MLEs $\hat{\theta}$ and $\hat{a}$ of θ and a can be then readily obtained from Eqs. (5.20) and (5.21) as $\hat{\theta} = \tilde{\theta}(\hat{b})$ and $\hat{a} = \tilde{a}(\hat{b})$.

5.5.2 Results of the misspecification analysis

The misspecification analysis has been developed by considering a real-world inspired application where the degrading unit under study is a pipeline subjected to corrosion. The degradation process of the pipeline is described by using three different (alternative) gamma processes, hereinafter referred to as the true processes. The shape parameter b of the age function $\eta(t) = (t/a)^b$ of the true processes is always assumed to be equal to 1. The other parameters of the true gamma degradation processes (i.e., a and θ) have been calibrated by using as rough reference the values provided in Mahmoodian and Alani (2014). The complete description of the adopted setups (i.e., $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, respectively) is provided in Table 5.25.

The true gamma processes have been used to generate 3 sets of $N_t = 1000$ synthetic datasets (i.e., one for each setup). Each dataset consists of $m = 6$ degrading paths, which simulate the evolution of the degradation levels of as many units. Each path consists of $n_i = 6$ measurements, taken at equally spaced inspection times ($t_1 = 30$, $t_2 = 60$, $t_3 = 90$,

Setup	a $[years]$	b	θ $[mm]$
$\mathcal{A}$	40	1	8
$\mathcal{B}$	30	1	6
$\mathcal{C}$	15	1	3

Table 5.25 – Setups $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ used to generate the datasets.

$t_4 = 120$, $t_5 = 150$, and $t_6 = 180$, expressed in time units) that are the same for all the units (that is: $t_{i,j} = t_j$, $\forall\ i, j\ , i = 1, \ldots, m$ and $j = 1, \ldots, 6$).

These datasets have been used to estimate, dataset by dataset, the parameters of the gamma and inverse Gaussian processes via the maximum likelihood method described in Chapter 5.5.1. Dataset by dataset, the misspecification is assumed to occur if the Akaike information criterion (e.g., see Akaike (1974)) leads to prefer the (wrong) inverse Gaussian process instead of the (right) gamma process. The results of the misspecification analysis are reported in Table 5.26, where N_t is the total number of datasets used for the analysis, N_m is the number of times a misspecification has occurred, $N_{\overline{m}}$ is the number of times a misspecification has not occurred, and r_m is the percent risk of incurring in a misspecification, that is evaluated as:

$$r_m\% = \left(\frac{N_m}{N_t}\right) \cdot 100.$$

These results give evidence that the risk of incurring in a misspecification increases

Setup	N_t	$N_{\overline{m}}$	N_m	$r_m(\%)$
$\mathcal{A}$	1000	945	55	5.5
$\mathcal{B}$	1000	909	91	9.1
$\mathcal{C}$	1000	868	232	23.2

Table 5.26 – Results of the misspecification analysis.

moving from the setup $\mathcal{A}$ to the setup $\mathcal{C}$. Actually, as in Chapter 4, these setups have been specially conceived with the aim of representing three possible experimental scenarios that, while sharing the same mean function, are characterized by increasing risk (i.e., from $\mathcal{A}$ to $\mathcal{C}$) of incurring in a misspecification. For the sake of brevity, here we will not repeat the explanation as to why this is the case (see Chapter 4.2.1 for more details).

The impact of a misspecification on the performances of the considered maintenance policy is evaluated in terms of long-run average maintenance cost rate. The cost model is

formulated by setting the unitary costs c_c, c_p, c_i, and c_l (expressed in monetary units) and the downtime cost rate c_d (expressed in monetary units/year) to the values given in Table 5.27, that have been loosely inspired to the values provided in Dey (2004). The pipelines are assumed to fail when their degradation level exceeds the threshold limit $w_M = 35\ mm$.

c_c	c_p	c_i	c_l	c_d
6	1	0.5	0.2	0.2

Table 5.27 – Parameters of the cost model

The effect of a misspecification is evaluated, dataset by dataset, by computing the optimal value of the design parameter vector and the corresponding minimum long-run average maintenance cost rate under each estimated model. Moreover, the same design parameter vector (i.e., the one obtained under both the estimated models) is used to compute the corresponding true long-run average maintenance cost rate, by using the true gamma model. Finally, the true gamma model is used to determine the true optimal value of the design parameters and the corresponding true minimum long-run average maintenance cost rate. Obtained results have been used to compute the following indices:

$$MRD_{M,d}^{(1)} = \frac{\sum_{k=1}^{N_d} \frac{C_\infty(\widehat{\boldsymbol{\zeta}}_{M,k}^*) - C_\infty(\boldsymbol{\zeta}^*)}{C_\infty(\boldsymbol{\zeta}^*)}}{N_d}, \tag{5.22}$$

$$SDRD_{M,d}^{(1)} = \sqrt{\frac{\sum_{k=1}^{N_d} \left[\frac{C_\infty(\widehat{\boldsymbol{\zeta}}_{M,k}^*) - C_\infty(\boldsymbol{\zeta}^*)}{C_\infty(\boldsymbol{\zeta}^*)} - MRD_{M,d}^{(1)}\right]^2}{N_d}}, \tag{5.23}$$

where:

— $C_\infty(\cdot)$ is the true value of the long-run average maintenance cost rate. This is computed as in Eq. (5.2) by using the true gamma process);

— $\boldsymbol{\zeta}^* = (\tau^*, \Delta\tau^*, L^*)$ is the true optimal value of the vector of design parameters $\boldsymbol{\zeta} = (\tau, \Delta\tau, L)$ (i.e., the value of $\boldsymbol{\zeta}$ that minimizes $C_\infty(\boldsymbol{\zeta})$);

— $C_\infty(\boldsymbol{\zeta}^*)$ is the true minimum long-run average maintenance cost rate;

— $\widehat{C}_{\infty,M,k}(\cdot)$ is the long-run average maintenance cost rate computed by using the model M with a, b, and θ set at their MLEs obtained (under the model M) from the k-th dataset (the hat « ˆ » indicates that $\widehat{C}_{\infty,M,k}(\cdot)$ can be intended as an estimate of $C_\infty(\cdot)$);

— $\widehat{\boldsymbol{\zeta}}^*_{M,k} = (\widehat{\tau}^*_{M,k}, \widehat{\Delta\tau}^*_{M,k}, \widehat{L}^*_{M,k})$ is the value of $\boldsymbol{\zeta} = (\tau, \Delta\tau, L)$ that minimizes $\widehat{C}_{\infty,M,k}(\boldsymbol{\zeta})$ (the hat «ˆ» indicates that $\widehat{\boldsymbol{\zeta}}^*_{M,k}$ can be intended as an estimate of $\boldsymbol{\zeta}^*$);

— $\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k})$ is the (estimated) minimum long-run average maintenance cost rate computed by using the model M with a, b, and θ set at their MLE obtained (under the model M) from the k-th dataset;

— $C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})$ is the true long-run average maintenance cost rate obtained by setting $\boldsymbol{\zeta} = \widehat{\boldsymbol{\zeta}}^*_{M,k}$ (this cost is evaluated by using the true gamma process).

The subscript M on the left sides of Eqs. (5.22) and (5.23) serves to identify the model used to compute $\widehat{\boldsymbol{\zeta}}^*_{M,k}$, with $M = GP$ standing for the gamma process and $M = IGP$ standing for the inverse Gaussian process. The subscript d indicates the datasets used to compute the index. Specifically, $d = t$ indicates that the index is computed by using all the N_t datasets, $d = m$ indicates that it is computed by using only the N_m datasets that lead to a misspecification, and $d = \overline{m}$ indicates that it is computed by using only the datasets that do not lead to a misspecification. The index (5.22) is the mean of the relative

		$MRD^{(1)}_{M,d}$		
M	d	Setup $\mathcal{A}$	Setup $\mathcal{B}$	Setup $\mathcal{C}$
GP	$\overline{m}$	0.042	0.043	0.042
IGP	$\overline{m}$	0.43	0.20	0.046
GP	m	0.043	0.039	0.046
IGP	m	0.044	0.033	0.039
GP	t	0.042	0.043	0.043
IGP	t	0.41	0.18	0.044

Table 5.28 – Results of the study: index $MRD^{(1)}_{M,d}$.

difference $[C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\boldsymbol{\zeta}^*)]/C_\infty(\boldsymbol{\zeta}^*)$. Hence, for example, the value $MRD^{(1)}_{G,\overline{m}} = 0.043$, reported in the first row of Table 5.28 in the column relative to the setup $\mathcal{B}$, indicates that using the design parameter vector $\widehat{\boldsymbol{\zeta}}^*_{G,k}$ (i.e., the one determined by minimizing $\widehat{C}_{\infty,M,k}(\cdot)$) in place of $\boldsymbol{\zeta}^*$ (i.e., the one determined by minimizing $C_\infty(\cdot)$) results in a long-run average maintenance cost rate that (in mean) is 4.3% higher than the true minimum long-run average maintenance cost rate $C_\infty(\boldsymbol{\zeta}^*)$. Thus, roughly speaking, this means that under the setup $\mathcal{B}$, using the estimated gamma process in place of the true one in the cases where a misspecification has not occurred, causes (in mean) a 4.3% increase of the maintenance cost. This difference, when $M = GP$, is only due to the circumstance that the MLEs of

		$SDRD^{(1)}_{M,d}$		
M	d	Setup $\mathcal{A}$	Setup $\mathcal{B}$	Setup $\mathcal{C}$
GP	$\overline{m}$	0.069	0.076	0.067
IGP	$\overline{m}$	0.92	0.59	0.045
GP	m	0.058	0.061	0.074
IGP	m	0.047	0.036	0.049
GP	t	0.068	0.075	0.069
IGP	t	0.90	0.57	0.046

Table 5.29 – Results of the study: index $SDRD^{(1)}_{M;d}$.

the model parameters do not coincide with the true values. Differently, when $M = IGP$ the cost increase depends both on the use of a wrong model and on the fact that its parameters are estimated, with the aggravating circumstance that the estimates of the parameters are obtained under the wrong model.

Table 5.28 shows that, under all the setups, the results obtained when $M = IGP$ and $d = m$ are very similar to those obtained when $M = GP$ and $d = \overline{m}$. This means that using the inverse Gaussian process in place of the true gamma one, when its use is suggested by the AIC, determines a negligible effect on the long-run average maintenance cost rate. On the other side, the results obtained when $M = IGP$ and $d = \overline{m}$ give clear evidence that using the inverse Gaussian model when the AIC leads to prefer the gamma process can produce severe consequences, especially when the risk of incurring in a misspecification is low (i.e., setup $\mathcal{A}$).

The index (5.23) is the standard deviation of the relative difference $[C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\boldsymbol{\zeta}^*)]/C_\infty(\boldsymbol{\zeta}^*)$. This index serves to understand how the considered relative difference varies from dataset to dataset with respect to its mean value reported in Table 5.29. The observed differences are due to the circumstance that the estimates of model parameters (obviously) vary from dataset to dataset. Obtained results show that, in all the considered cases, the (quadratic) mean of the difference between the values obtained from the generic dataset and the mean reported in Table 5.28 is about 1.5 times greater than the mean.

Finally, the results of the Monte Carlo study have been also used to compute the

indices:

$$MRD^{(2)}_{M,d} = \frac{\sum_{k=1}^{N_d} \frac{\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})}{C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})}}{N_d}, \tag{5.24}$$

$$SDRD^{(2)}_{M,d} = \sqrt{\frac{\sum_{k=1}^{N_d} \left[\frac{\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})}{C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})} - MRD^{(2)}_{M,d}\right]^2}{N_d}}, \tag{5.25}$$

The results obtained for these indices are reported in Tables 5.30 and 5.31.

The indices (5.24) and (5.25) are the mean and the standard deviation of the relative

		$MRD^{(2)}_{M,d}$		
M	d	Setup $\mathcal{A}$	Setup $\mathcal{B}$	Setup $\mathcal{C}$
GP	$\overline{m}$	-0.0053	-0-056	-0.046
IGP	$\overline{m}$	-0.0076	0.18	0.17
GP	m	-0.11	-0.074	-0.066
IGP	m	0.12	0.14	0.066
GP	t	-0.056	-0.057	-0.050
IGP	t	-0.00079	0.17	0.15

Table 5.30 – Results of the study: index $MRD^{(2)}_{M,d}$.

		$SDRD^{(2)}_{M,d}$		
M	d	Setup $\mathcal{A}$	Setup $\mathcal{B}$	Setup $\mathcal{C}$
GP	$\overline{m}$	0.25	0.23	0.17
IGP	$\overline{m}$	0.34	0.29	0.21
GP	m	0.25	0.21	0.17
IGP	m	0.24	0.24	0.19
GP	t	0.25	0.23	0.17
IGP	t	0.34	0.29	0.21

Table 5.31 – Results of the study: index $SRDR^{(2)}_{M,d}$.

difference $[\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})]/C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})$. A positive (negative) value of the index $MRD^{(2)}_{M,d}$ indicates that the estimates of the long-run average maintenance cost rate $\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k})$ assume (dataset by dataset) values that in mean are greater (smaller) than the corresponding true long-run average maintenance cost rate $C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})$. More specif-

ically, for example, the value $MRD^{(2)}_{G,m}$ reported in the first row of Table 5.30, in the column relative to the setup $\mathcal{B}$, indicates that the long-run average maintenance cost rate calculated under the estimated gamma process, in the cases where the misspecification has not occurred, are in mean 5.6% smaller than the corresponding true values. It is interesting to note that the use of the gamma process leads to underestimate the true long-run average maintenance cost rate, while the inverse Gaussian process leads (in almost all cases) to overestimate it. However, in all cases, the values obtained for the $MRD^{(2)}_{M,d}$ are rather small. The values of the index $SDRD^{(2)}_{M_d}$ reported in Table 5.31 show how the estimates of the relative difference $[\widehat{C}_{\infty,M,k}(\widehat{\boldsymbol{\zeta}}^*_{M,k}) - C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})]/C_\infty(\widehat{\boldsymbol{\zeta}}^*_{M,k})$ obtained from the considered N_d synthetic datasets fluctuate around the corresponding mean reported in Table 5.31.

5.6 Conclusions

In this chapter, three new maintenance policies have been proposed for units whose degradation path is affected by unit-to-unit variability in the form of random effect. The proposed policies consists in performing, at a predefined epoch, an inspection aimed at measuring the degradation level of the unit and in using the outcome of this measurement to inform subsequent decision-making via a condition-based rule.

In Chapter 5.2 the basic decision-making scheme is presented, showing its effectiveness and how measurement error impacts on maintenance costs. The policy there illustrated takes advantage of the heterogeneity present in the population by using the inspection to timely distinguish between weak and strong units. Indeed, the inspection can be envisaged as a classifier, casting all units in a "weak" class (immediately replacing them) and a "strong" class (postponing their replacement).

The adaptive policy presented in Chapter 5.3 takes the same general idea one step further. In fact, whereas in case of the non-adaptive policy of Chapter 5.2 all units cast into the "strong" class are replaced at the same time, in this case the replacement time will be adaptively scheduled based on the measured degradation level at the inspection time. Indeed, here the inspection classifies units into several "strong" classes, and to each one of them assigns an ad-hoc replacement time. Obtained results show that, under the considered setup, noticeable savings can be achieved by adopting this adaptive strategy over the non-adaptive one.

On the other hand, the policy presented in Chapter 5.4 expands on the simple scheme

of Chapter 5.2 by integrating different measurement procedure at the inspection time. Indeed, in this case the measurement is initially performed by using an imperfect procedure that provides a perturbed measurement of the true degradation level. Based on the outcome of this first measurement procedure, a condition-based rule is then used to decide whether to immediately replace the unit, to postpone its replacement to a (also predefined) later time, or to use second perfect measurement procedure, which has an extra cost. In this latter case, based on the outcome of this second measurement procedure it is finally decided whether to immediately replace the unit or to postpone its replacement to a predefined future time.
The key idea of the policy, in this case, is to use the perturbed inspection to timely identify the units that are safely below the failure threshold and those that are close to failure or even already failed, for which the maintenance decision is unambiguous. Then, an additional (perfect) measurement can be performed to further assess the true state of the remaining unit and inform further decision-making.
The performances of this latter policy have been compared with those of three alternative policies obtained by assuming that only one (either the perturbed one or the perfect one, respectively) or none of the considered measurement procedures is adopted. The comparative study has been developed by considering eight setups that differ in the values of some of the cost parameters and/or in the magnitude of the measurement error. Obtained results have shown that, depending on the relative values of the cost parameters, adopting the more flexible scheme can be very advantageous.

Overall, comparing the three policies has shown that, when inspection costs are high, restricting the policy to a single inspection can be a viable strategy. However, it has also shown that, for some particular setting of the cost model (i.e., high inspection costs, low consequences of undetected failure) even a single inspection is already superfluous.

Finally, this chapter also presented the results of a misspecification study investigating the impact on maintenance costs of misspecifying a gamma process with an inverse Gaussian process. As it was already observed in the larger misspecification study presenter in Chapter 4, also in this case it seems that, in the experimental situations when the risk of a wrong diagnosis is higher, its consequences are milder. Conversely, when the risk of a misspecification is lower, the consequences can be severe. Nevertheless, if model selection is performed by using an appropriate statistical criterion (such as the adopted AIC) then the performances of the two models are very similar.

In this chapter, the proposed policies are built only considering inspections and re-

placements are available maintenance actions. In the next chapter, we will prioritize maintenance actions that we will refer to as "prescriptive", in the sense that the scope of these actions is not limited to the strictly maintenance domain, but tap also into the more general operational and economical aspects of the system. In conjunction with the more classical maintenance actions, the prescriptive actions might provide an overall better coordinated response between economical profitability and maintenance cost control.

Chapter 6

A new paradigm: prescriptive maintenance

6.1 Introduction

The main role of maintenance planning in industrial applications is to guarantee reliable and safe functioning of equipment. However, given that it often entails disruption to normal operations (for example, to perform in-depth inspections) it is typically perceived as a time-intensive and costly task. As touched upon also in some previous chapters, this situation often generates a conflict between the commercial incentive to maximum utilization of equipment and ensuring its safe operation through regular maintenance.
On the other hand, maintenance is also an activity that could create opportunities for improvement. In the search for maximum operational efficiency, modern maintenance strategies such as prescriptive maintenance (PsM) have been proposed as a potentially effective tool.

Indeed, a PsM framework does not see maintenance-related decisions and operational-related decisions are separated and conflicting, but rather considers them as two sides of the same coin (see Pinciroli, Baraldi, and Zio (2023)). Therefore, maintenance recommendations (i.e., the prescriptions) should be defined by taking into account all functionalities of a system (see Iung (2019)). These prescriptions go beyond describing what, when, and how to perform maintenance, but also provide precise operative instructions on how to adjust the system operating conditions to reach a desired outcome (see Longhitano et al. (2021a)).
Although an increasing number or papers dealing with this new paradigm have been presented in the recent literature, a formal definition of PsM is still missing. To the best of the author's knowledge, the most serious attempt at a definition has been made by Longhitano et al. (2021b). In that paper, the authors argue that a PsM approach should:

— use failure predictions, or data-driven degradation models, to quantitatively assess

the best course of action;

— as Iung (2019) suggests, account for all functionalities of a system;

— adopt a detailed model of the system under study, which should capture the effects of prescriptive actions;

— given that the prescriptions will also influence the future failure behavior of the system, said model should also have a closed-loop structure and be able to integrate real-time data to correct the course of action.

The same authors also propose an example of application of PsM in Longhitano et al. (2021b) and Longhitano et al. (2021a), where they model the degradation of a critical component, the brake pad, of a fleet of vehicles. The goal of their PsM solution is to provide the best mission schedule for the fleet and decide maintenance dates, considering the effect one has on the other. They aim to maximize the utilization of the component while minimizing workshop visits (especially the extra ones with respect to the prearranged calendar), assigning different missions to different vehicles in order to manage the degradation of their brake pads.

The idea of using prescriptive actions to influence the degradation and the operating lifetime, with the goal of conforming to a predetermined maintenance calendar, is a recurring theme in the PsM literature, also touched upon by other authors, such as Dirkes et al. (2023). In this manuscript, we will also explore a similar scenario.

Drawing inspiration from modern prescriptive maintenance ideas, in this chapter we propose two prescriptive policies inspired by the one presented in Chapter 5.2. In fact, both policies assume that, due to its high cost, a single inspection can be performed during the life cycle of the unit and use the information gathered by means of this inspection to decide, via a condition-based rule, whether to immediately replace the unit or to postpone its replacement to a future time. The main difference with respect to the policy of Chapter 5.2 is that, here, we suppose that it is possible to manipulate the usage rate of the unit according to economic considerations. Moreover, we also assume that the time at which maintenance operations can be carried out are supposed to be subjected to significant constraints (for example, limited maintenance personnel availability, contractual clauses, etc...). The driving idea behind these policies is to investigate whether introducing another degree of freedom (in the form of adjusting the usage rate), in a scenario where maintenance dates are constrained, can provide a better trade-off between preventive replacement, corrective replacement, inspection, and operational costs.

A change in the usage rate is supposed to affect both the future evolution of the degradation and operational costs. Indeed, one of the main challenge of prescriptive maintenance is accounting for and quantifying how prescriptive actions (such as manipulating the usage rate) impact the degradation of the unit, and even more developing a comprehensive cost model that captures their effects on maintenance and operational costs.

The prescriptive policies presented in this chapter are developed based on the following assumptions:

— a single degrading unit is considered;

— failure occurs when the degradation level of the unit passes an assigned threshold w_M;

— failures are not self announcing and the unit keeps operating past its failure point, but with reduced performance/additional costs;

— any replacement (i.e., both corrective and preventive) restores the unit to an "as good as new" state. Therefore, the time elapsing between two successive replacement times defines the cycle of a renewal process;

— at the start of each cycle the usage rate is set to u_0, where $u_{min} \leq u_0 \leq u_{max}$, u_{min} is the minimum usage rate, and u_{max} is the maximum usage rate;

— it is possible to change the usage rate only at the inspection time.

The performance measure adopted to define the optimal policy is the long-run average maintenance utility rate.

6.2 A prescriptive maintenance policy for a gamma deteriorating unit

6.2.1 Description of the policy

This prescriptive maintenance policy (see also Esposito, Castanier, and Giorgio (2022a)) is developed starting from the simple decision scheme proposed in Chapter 5.2. In fact, it consists in performing a single inspection at an age-based time τ and deciding whether to immediately replace the unit or to postpone its replacement to time $\tau + \Delta\tau$, by using a condition-based rule. The main novelty with respect to the policy illustrated in Chapter 5.2 is that here the condition-based rule includes the possibility of adjusting the usage rate of the unit.

The policy suits particularly well the cases in which the time intervals between successive inspections/replacements are subjected to constraints. In fact, for simplicity (yet without loss of generality), the proposed maintenance policy is formulated by assuming that times between successive inspections/replacements are constrained to be equal (i.e., it is assumed that $\Delta\tau = \tau$).

The condition based rule adopted to take the decision at τ is summarized in Table 6.1, where w_τ is the measured degradation level at time τ, L ($L \leq w_M$) is the preventive maintenance threshold, and $u_{min} \leq u_p \leq u_{max}$. The components of the vector $\boldsymbol{\xi} = \{\tau, L, u_0, u_p\}$

Measurement at τ	Decision
$w_\tau > L$	Replacement at τ
$w_\tau \leq L$	Replacement at 2τ, set usage rate to u_p

Table 6.1 – Condition-based rule adopted to decide the replacement time and the usage rate.

should be intended as design parameters. The value of $\boldsymbol{\xi}$ that maximizes the long-run average maintenance reward rate (and hence defines the optimal policy) is denoted by $\boldsymbol{\xi}^* = \{\tau^*, L^*, u_0^*, u_p^*\}$. Table 6.2, where $w_{2\tau}$ is the degradation level at time 2τ, lists all the possible scenarios, along with the corresponding maintenance actions to be taken and the length of a maintenance cycle $T(w_\tau)$. It is worth to note that, although the notation does

Experimental Scenario	Maintenance action	Cycle length $T(w_\tau)$
$L < w_\tau \leq w_M$	Preventive replacement at τ	τ
$w_\tau > w_M$	Corrective replacement at τ	τ
$w_\tau \leq L,\ w_{2\tau} \leq w_M$	Preventive replacement at τ	2τ
$w_\tau \leq L,\ w_{2\tau} > w_M$	Corrective replacement at τ	2τ

Table 6.2 – Possible scenarios and corresponding maintenance actions and cycle lengths.

not highlight it, $T(w_\tau)$ functionally depends on $\boldsymbol{\xi}$.

6.2.2 Degradation model

The gamma process $\{Y(t); t \geq 0\}$ is a monotonic increasing process with independent and gamma distributed increments. Thus, given an initial condition (here $Y(0) = 0$), it is fully defined by the (pdf) of its generic increment $\Delta Y(t_1, t_2)$, with $t_2 > t_1 \geq 0$:

$$f_{\Delta Y(t_1,t_2)}(\delta) = \frac{\delta^{\Delta\eta(t_1,t_2)} - 1}{\theta^{\Delta\eta(t_1,t_2)} \cdot \Gamma\left(\Delta\eta(t_1,t_2)\right)} \cdot e^{-\frac{\delta}{\theta}}$$

or by the corresponding cdf:

$$F_{\Delta Y(t_1,t_2)}(\delta) = \frac{\gamma\left(\Delta\eta(t_1,t_2), \frac{\delta}{\theta}\right)}{\Gamma\left(\Delta\eta(t_1,t_2)\right)}$$

where θ ($\theta > 0$) is the scale parameter, $\eta(t)$ is a non-negative, monotone increasing function (referred to as the age function), $\Delta\eta(t_1, t_2) = \eta(t_2) - \eta(t_1)$, $\Gamma(\cdot)$ is the complete gamma function, and $\gamma(\cdot)$ is the lower incomplete gamma function. In this chapter, it is assumed that, given the history of the usage rate $u_t = \{u(y), 0 \leq y \leq t\}$ up to and including the time t, the degradation process $\{W(t); t \geq 0; u_t\}$ of the generic unit can be described by using a gamma process, whose increment $\Delta W(t, t + dt)$ over the elementary time interval $(t, t + dt)$ is gamma distributed with conditional pdf:

$$f_{\Delta W(t,t+dt)}(\delta; u_t) = f_{\Delta W(t,t+dt)}\left(\delta; u(t)\right) = \frac{\delta^{\eta'(t;u(t))dt-1} \cdot e^{-\frac{\delta}{\theta}}}{\theta^{\eta(t)} \cdot \Gamma\left(\eta'\left(t; u(t)\right) dt\right)}. \tag{6.1}$$

Note that the pdf (6.1) also implies that given $u(t)$:

— the degradation increment $\Delta W(t, t + dt)$ is assumed to be independent of $W(t)$ and of the past history of the usage rate $u_{t^-} = \{u(y), 0 \leq y < t\}$;

— as in Tseng, Balakrishnan, and Tsai (2009), the pdf of $\Delta W(t, t + dt)$ is assumed to depend on $u(t)$ only through the value of its shape parameter $\eta'(t; u(t)) \cdot dt$ (in fact, the scale parameter θ is assumed to be independent of the usage rate).

The function $\eta'(t; u(t))$ must be positive and integrable with respect to t. In this chapter, the power-law expression $\eta'(t; u(t)) = a\left(u(t)\right) \cdot b \cdot t^{b-1}$ is adopted.

Therefore, under the proposed maintenance policy, the degradation level $W(t)$ at time t, for any $t \leq \tau$ and given u_0 (i.e., at any time before the inspection, given the value of the usage rate between $t = 0$ and $t = \tau$), is distributed as a gamma random variable with

pdf:

$$f_{W(t)}(\delta; u_0) = \frac{\delta^{\eta(t;u_0)-1}}{\theta^{\eta(t;u_0)} \cdot \Gamma\left(\eta(t;u_0)\right)} \cdot e^{-\frac{\delta}{\theta}} \tag{6.2}$$

and cdf:

$$F_{W(t)}(\delta; u_0) = \frac{\gamma\left(\eta(t;u_0); \frac{\delta}{\theta}\right)}{\Gamma\left(\eta(t;u_0)\right)} \tag{6.3}$$

where:

$$\eta(t;u_0) = \int_0^t \eta^{'}(y; u(y)) \cdot dy = \int_0^t a(u(y)) \cdot b \cdot y^{b-1} \cdot dy = a(u_0) \cdot t^b.$$

Similarly, the degradation increment $\Delta W(\tau, t)$ for any $t > \tau$ and given u_p (i.e., at any time after the inspection and given the value of the usage rate between $t = \tau$ and $t = 2\tau$) is gamma distributed with conditional pdf:

$$f_{\Delta W(\tau,t)}(\delta; u_p) = \frac{\delta^{\Delta\eta(\tau,t;u_p)-1}}{\theta^{\Delta\eta(\tau,t;u_p)} \cdot \Gamma\left(\Delta\eta(\tau,t;u_p)\right)} \cdot e^{-\frac{\delta}{\theta}} \tag{6.4}$$

and conditional cdf:

$$F_{\Delta W(\tau,t)}(\delta; u_p) = \frac{\gamma\left(\Delta\eta(\tau,t;u_p); \frac{\delta}{\theta}\right)}{\Gamma\left(\Delta\eta(\tau,t;u_p)\right)} \tag{6.5}$$

where:

$$\Delta\eta(\tau,t;u_p) = \int_\tau^t \eta^{'}(y; u(y)) \cdot dy = \int_\tau^t a(u(y)) \cdot b \cdot y^{b-1} \cdot dy = a(u_p) \cdot (t^b - \tau^b).$$

The adopted notations give evidence that the pdfs $f_{W(t)}(\delta; u_0)$ and $f_{\Delta W(\tau,2\tau)}(\delta; u_p)$ and the cdfs $F_{W(t)}(\delta; u_0)$ and $F_{\Delta W(\tau,2\tau)}(\delta; u_p)$ depend on the usage rate (i.e., on u_0 and u_p, according to the time window). Conversely, for the sake of notation simplicity, they do not explicitly highlight that $f_{\Delta W(\tau,t)}(\delta; u_p)$ and $F_{\Delta W(\tau,t)}(\delta; u_p)$ are conditional distributions (i.e., are formulated given u_t).

Finally, it is not hard to verify that, under the same assumptions, for $t \leq \tau$, the conditional cdf of $W(t)$ given $W(\tau)$ can be expressed as:

$$F_{W(t)|W(\tau)}(w_t|w_\tau; u_0) = \mathbb{B}\left(\frac{w_t}{w_\tau}; \eta(t;u_0), \Delta\eta(t,\tau;u_0)\right) \tag{6.6}$$

where $\Delta\eta(t,\tau;u_0) = \eta(\tau;u_0) - \eta(t;u_0)$, and $\mathbb{B}(z;\alpha,\beta)$ is the regularized beta function:

$$\mathbb{B}(z;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)} \cdot \int_0^z x^{\alpha-1} \cdot (1-x)^{\beta-1} \cdot dx.$$

6.2.3 The cost model and formulation of the long-run average maintenance reward rate

Under the proposed policy, the cost model is developed considering the cost of a preventive replacement c_p, the cost of a corrective replacement c_c ($c_c \leq c_p$), the inspection cost c_i, which is incurred only when an ad hoc inspection is done, and the logistic cost c_l, which is incurred each time an inspection or a replacement (even in the absence of an inspection) are performed.

The usage rate is assumed to impact the cost model through a reward term and a penalty term. Specifically, we assume that the reward earned by operating the unit at usage rate u can be computed as the product of a reward rate $r(u)$ and the operating time of the unit, while the penalty cost can be expressed as the product of a penalty rate $c_{pen}(u)$ and the downtime of the unit (i.e., the time elapsing from the potential failure of the unit until its eventual replacement). This penalty cost is sustained only in case of failure of the unit and is supposed to capture the effect of the reduced performances/additional costs resulting from operating the unit past its failure time.

Table 5.12 summarizes the utility $U(w_\tau, X)$ for each possible scenario, as a function of the degradation level w_τ and the lifetime of the unit X. As $T(w_\tau)$, also $U(w_\tau, X)$ functionally depends on $\boldsymbol{\xi}$, despite the adopted notation not highlighting it.

Experimental scenario	Utility $U(w_\tau, X)$
$L < w_\tau \leq w_M$	$-c_l - c_i - c_p + r(u_0) \cdot \tau$
$w_\tau > w_M$	$-c_l - c_i - c_c + r(u_0) \cdot \tau$ $-c_{pen}(u_0) \cdot (\tau - X)$
$w_\tau \leq L$, $w_{2\tau} \leq w_M$	$-2 \cdot c_l - c_i - c_p + r(u_0) \cdot \tau + r(u_p) \cdot \tau$
$w_\tau \leq L$, $w_{2\tau} > w_M$	$-2 \cdot c_l - c_i - c_c + r(u_0) \cdot \tau + r(u_p) \cdot \tau$ $-c_{pen}(u_p) \cdot (2\tau - X)$

Table 6.3 – Possible scenarios and corresponding utility.

By the renewal-reward theorem (see Ross (1983)), the long-run average maintenance reward rate function $U_\infty(\boldsymbol{\xi})$ is defined as:

$$U_\infty(\boldsymbol{\xi}) = \frac{E\{U(W(\tau), X)\}}{E\{T(W(\tau))\}} \tag{6.7}$$

where expectations have to be taken with respect to all the variables that are within the

parentheses.

The expected values in Eq. (6.7) can be computed via Eqs. (6.8)-(6.9).

$$
\begin{aligned}
&E\{U(W(\tau),X)\} = \\
&= \int_0^L \int_\tau^{2\tau} U(w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau;u_p)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau \\
&+ \int_0^L \int_{2\tau}^{\infty} U(w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau;u_p)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau \\
&+ \int_L^{w_M} \int_\tau^{\infty} U(w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau;u_p)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau \\
&+ \int_{w_M}^{\infty} \int_0^{\tau} U(w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau;u_0)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau \\
&= -c_l - c_i - c_p + r(u_0)\cdot\tau + [r(u_p) - c_{pen}(u_p)]\cdot\tau\cdot F_{W(\tau)}(L;u_0) \\
&- c_l\cdot F_{W(\tau)}(L;u_0) - [c_c - c_p + c_{pen}(u_0)\cdot\tau]\cdot[1 - F_{W(\tau)}(w_M;u_0)] \\
&- (c_c - c_p)\cdot\int_0^L [1 - F_{\Delta W(\tau,2\tau)}(w_M - w_\tau;u_p)]\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dw_\tau \\
&+ c_{pen}(u_p)\cdot\int_0^L\int_\tau^{2\tau} F_{\Delta W(\tau,x)}(w_M - w_\tau;u_p)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau \\
&+ c_{pen}(u_0)\cdot\int_{w_M}^{\infty}\int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau;u_0)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dx\cdot dw_\tau
\end{aligned}
\tag{6.8}
$$

$$
\begin{aligned}
&E\{T(W(\tau))\} = \int_0^L T(w_\tau)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dw_\tau + \int_L^{\infty} T(w_\tau)\cdot f_{W(\tau)}(w_\tau;u_0)\cdot dw_\tau \\
&= 2\tau\cdot\int_0^L f_{W(\tau)}(w_\tau;u_0)\cdot dw_\tau + \tau\cdot\int_L^{\infty} f_{W(\tau)}(w_\tau;u_0)\cdot dw_\tau = \tau\cdot[1 + F_{W(\tau)}(L;u_0)]
\end{aligned}
\tag{6.9}
$$

6.2.4 Example of application

We consider a hypothetical application of the proposed approach to a corroding pipeline, assuming that it is possible to control flow velocity according to convenience. In fact, pipelines are often buried underground, or offshore, and only periodic inspections can be performed (Alaswad and Xiang (2017)). Moreover, it has been observed that flow velocity can influence the rate of corrosion (Utanohara and Murase (2019), Yoneda et al. (2016)). Following Utanohara and Murase (2019) and Yoneda et al. (2016), we assume that the corrosion rate is affected by flow velocity via a power-law function. In particular,

we suppose that it affects the scale parameter of the age function:

$$a(u) = a \cdot \left(\frac{u}{u_{max}}\right)^d.$$

In this preliminary analysis, without loss of generality, it is assumed that the reward rate $r(u)$ depends linearly on flow velocity:

$$r(u) = r \cdot \frac{u}{u_{max}}$$

and that the penalty cost rate is a fixed fraction of the reward rate:

$$c_{pen}(u) = c_{pen} \cdot r(u), \tag{6.10}$$

where u_{max} is the maximum allowable flow velocity, and a, r, and c_{pen} are the corresponding maximum degradation rate, reward rate, and penalty cost rate, respectively (i.e., $a = a(u_{max})$, $r = r(u_{max})$, and $c_{pen} = c_{pen}(u_{max})$). Depending on the practical scenario under study, c_{pen} might be greater or smaller than r (that is, operating a failed unit may incur in a penalty that is greater than the corresponding reward). In this case, we set $c_{pen} = 0.8$, to investigate the case where the penalty is relevant but does not exceed the reward. The values of the parameters used to calibrate the degradation process have been inspired by those provided in Mahmoodian and Alani (2014) and are reported in Table 6.4. The same table also report the values of the unitary costs c_c, c_p, c_i, c_l, and c_{pen}, inspired by those found in Dey (2004). The pipelines are assumed to fail when their degradation level exceeds the threshold $w_M = 35\ mm$. The maximum and minimum allowable usage rates have been set to $u_{max} = 1$ and $u_{min} = 0$.

$a\ [years]$	b	$\theta\ [mm]$	c_p	c_c	c_i	c_l	c_{pen}
0.17	1	2.35	1	6	0.5	0.2	0.8

Table 6.4 – Parameters used to calibrate the degradation process and the cost model.

Let $\mathcal{P}_1$ be the proposed policy. For the sake of investigating the effect of changing the usage rate, we compare the performance of $\mathcal{P}_1$ with three special cases obtained by considering the following constraints:

— $\mathcal{P}_2$: $u_0 = u_M$;

— $\mathcal{P}_3$: $u_p = u_0$;

— $\mathcal{P}_4$: $u_0 = u_p = u_M$.

We consider four experimental scenarios that differ in the values of r and d, and are summarized in Table 6.5. Table 6.6 reports the optimal values of the design parameters

Scenario	d	r
$\mathcal{A}$	1.3	0.2
$\mathcal{B}$	2.3	0.2
$\mathcal{C}$	1.3	0.05
$\mathcal{D}$	2.3	0.05

Table 6.5 – Considered setups.

and the corresponding optimal long-run average maintenance reward rate computed under the four aforementioned policies, in the case of the scenario $\mathcal{A}$. Tables 6.7, 6.8, and 6.9 report the same results obtained under the scenarios $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$. Tables 6.6-6.9 give

	τ^*	L^*	u_0^*	u_p^*	U_∞^*
$\mathcal{P}_1$	34.2	16.1	1	1	0.1636
$\mathcal{P}_2$	34.2	16.1	1	1	0.1636
$\mathcal{P}_3$	34.2	16.1	1	1	0.1636
$\mathcal{P}_4$	34.2	16.1	1	1	0.1636

Table 6.6 – Optimal values of the design parameters and corresponding long-run average maintenance reward rate obtained under the scenario $\mathcal{A}$.

	τ^*	L^*	u_0^*	u_p^*	U_∞^*
$\mathcal{P}_1$	50.4	11.5	1	0.91	0.1643
$\mathcal{P}_2$	50.4	11.5	1	0.91	0.1643
$\mathcal{P}_3$	34.2	16.1	1	1	0.1636
$\mathcal{P}_4$	34.2	16.1	1	1	0.1636

Table 6.7 – Optimal values of the design parameters and corresponding long-run average maintenance reward rate obtained under the scenario $\mathcal{B}$.

evidence that, depending on the particular scenario, it is more or less advantageous to adopt $\mathcal{P}_1$ instead of the less flexible $\mathcal{P}_2$, $\mathcal{P}_3$, and $\mathcal{P}_4$.

The effect of changing the usage rate is especially apparent in the case of the scenarios $\mathcal{C}$ and $\mathcal{D}$. In fact, for example, in the case of scenario $\mathcal{D}$, $\mathcal{P}_1$ allows to obtain a long-run average maintenance reward rate that is 22% greater than the one obtained under $\mathcal{P}_3$ and

	τ^*	L^*	u_0^*	u_p^*	U_∞^*
$\mathcal{P}_1$	52.6	15.2	1	0.65	$2.027 \cdot 10^{-2}$
$\mathcal{P}_2$	52.6	15.2	1	0.65	$2.027 \cdot 10^{-2}$
$\mathcal{P}_3$	35.2	16.4	1	1	$1.432 \cdot 10^{-2}$
$\mathcal{P}_4$	35.2	16.4	1	1	$1.432 \cdot 10^{-2}$

Table 6.8 – Optimal values of the design parameters and corresponding long-run average maintenance reward rate obtained under the scenario $\mathcal{C}$.

	τ^*	L^*	u_0^*	u_p^*	U_∞^*
$\mathcal{P}_1$	61.2	16.2	0.95	0.71	$2.355 \cdot 10^{-2}$
$\mathcal{P}_2$	54.5	16.4	1	0.74	$2.348 \cdot 10^{-2}$
$\mathcal{P}_3$	83.7	16.4	0.68	0.68	$1.924 \cdot 10^{-2}$
$\mathcal{P}_4$	35.2	16.4	1	1	$1.423 \cdot 10^{-2}$

Table 6.9 – Optimal values of the design parameters and corresponding long-run average maintenance reward rate obtained under the scenario $\mathcal{D}$.

65% greater than the one obtained under $\mathcal{P}_4$, where $\mathcal{P}_3$ is the policy that does not allow for the possibility of modifying the usage rate on the basis of the degradation measurement performed at τ and $\mathcal{P}_4$ is the one that sets both u_0 and u_p to the maximum allowable value u_{max}.

6.3 An adaptive prescriptive maintenance policy for a gamma deteriorating unit

6.3.1 Description of the policy

This prescriptive policy (see also Esposito, Castanier, and Giorgio (2023)) extends the one presented in Chapter 6.2. Also in this case the policy consists in performing a single inspection at a predetermined inspection time τ whose outcome is a measurement of the current degradation level of the unit (hereinafter denoted as w_τ). This measurement is then used to decide, via a condition-based rule, whether to immediately replace the unit or to postpone its replacement to time 2τ (no additional inspection will be performed at 2τ). In case the replacement is postponed, if it is deemed economically convenient, the usage rate for the remainder of the operating life of the unit can be changed within predetermined limits, according to convenience, influencing both the future evolution of

the degradation process and the operating costs.

The main novelty of this approach with respect to the one suggested in Chapter 6.2 is that, here, in case the replacement is postponed the new usage rate is determined based on the adaptive rule described in Table 6.10, where $L_1 < L_2 < \cdots < L_k < w_M$, $u_1 > u_2 > \cdots > u_k$ denote the usage rates that (based on w_τ) are used in the time interval $(\tau, 2\tau)$, and L_k is the preventive replacement threshold.

Measurement at τ	Decision
$w_\tau > L_k$	Replacement at τ
$L_{k-1} < w_\tau \leq L_k$	Replacement at 2τ, set usage rate to u_k
$\vdots$	$\vdots$
$L_1 < w_\tau \leq L_2$	Replacement at 2τ, set usage rate to u_2
$w_\tau \leq L_1$	Replacement at 2τ, set usage rate to u_1

Table 6.10 – Adaptive condition-based rule adopted to perform decision-making.

At the beginning of the maintenance cycle the usage rate is set to u_0 and its value is also a design parameter which should should be set on the basis of economic considerations. The components of the vector $\boldsymbol{\xi} = \{L_1, \ldots, L_k, u_0, u_1, \ldots, u_k\}$ should be intended as design parameters. The value of $\boldsymbol{\xi}$ which maximizes the long-run average maintenance reward rate and defines the optimal policy is denoted by $\boldsymbol{\xi}^* = \{L_1^*, \ldots, L_k^*, u_0^*, u_1^*, \ldots, u_k^*\}$.

The number of classes k should be envisaged as a "hyperparameter" that must be set a priori with the aim of finding a satisfactory trade-off between performance of the policy and computational burden, which both increase with k.

It is worth remarking that, when $k = 1$, the proposed policy coincides with the one suggested in Chapter 6.2. Table 6.11, where $w_{2\tau}$ is the degradation level at time 2τ, lists all the possible scenarios along with the corresponding maintenance actions to be taken and the length of a maintenance cycle $T(w_\tau)$. Note that, despite not highlighting it, $T(w_\tau)$ functionally depends on $\boldsymbol{\xi}$.

6.3.2 The cost model and formulation of the long-run average maintenance reward rate

The cost model is developed under the same assumptions (and notation) of Chapter 6.2. Table 6.12 lists, for each possible scenario, the corresponding utility $U(w_\tau, X)$ as a

Experimental scenario	Maintenance action	Cycle length $T(w_\tau)$
$L_k < w_\tau \leq w_M$	Preventive replacement at τ	τ
$w_\tau > w_M$	Corrective replacement at τ	τ
$w_\tau \leq L_k$ and $w_{2\tau} \leq w_M$	Preventive replacement at 2τ	2τ
$w_\tau \leq L_k$ and $w_{2\tau} > w_M$	Corrective replacement at 2τ	2τ

Table 6.11 – Possible experimental scenarios and corresponding maintenance action and cycle length.

function of the degradation level at the inspection time w_τ and of the lifetime of the unit X.

The long-run average maintenance reward rate $U_\infty(\boldsymbol{\xi})$ is computed via the renewal-reward theorem (see Ross (1983)) as:

$$U_\infty(\boldsymbol{\xi}) = \frac{E\{U(W(\tau), X)\}}{\{T(W(\tau))\}} \tag{6.11}$$

where expectations have to be taken with respect to both $W(\tau)$ and X. The optimal value $U_\infty(\boldsymbol{\xi}^*)$ obtained when $\boldsymbol{\xi} = \boldsymbol{\xi}^*$ is denoted as U_∞^*.
The expected values in Eq. (6.11) are not available in closed form but can be computed via Eqs. (6.12)-(6.13).

$$\begin{aligned}
&E\{U(W(\tau), X)\} = \\
&= \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} \int_{\tau}^{2\tau} U(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau; u_h) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot dw_\tau \\
&+ \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} U(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau; u_h) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot dw_\tau \\
&+ \int_{L_k}^{w_M} \int_{\tau}^{\infty} U(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau; u_0) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot \cdot dw_\tau \\
&+ \int_{w_M}^{\infty} \int_{0}^{\tau} U(w_\tau, x) \cdot f_{X|W(\tau)}(x|w_\tau; u_0) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot dw_\tau \\
&= -c_l - c_i - c_p - c_l \cdot F_{W(\tau)}(L_k; u_0) + r(u_0) \cdot \tau \\
&- [c_c - c_p + c_{pen}(u_0) \cdot \tau] \cdot \left[1 - F_{W(\tau)}(w_M; u_0)\right]
\end{aligned}$$

Experimental scenario	Utility $U(w_\tau, X)$
$L_k < w_\tau \leq w_M$	$-2 \cdot c_l - c_i - c_p + r(u_0) \cdot \tau$
$w_\tau > w_M$	$-2 \cdot c_l - c_i - c_c + r(u_0) \cdot \tau$ $-c_{pen}(u_0) \cdot (\tau - X)$
$L_{k-1} < w_\tau \leq L_k$ and $w_{2\tau} \leq w_M$	$-2 \cdot c_l - c_i - c_p + r(u_0) \cdot \tau + r(u_k) \cdot \tau$
$L_{k-1} < w_\tau \leq L_k$ and $w_{2\tau} > w_M$	$-2 \cdot c_l - c_i - c_c + r(u_0) \cdot \tau + r(u_k) \cdot \tau$ $-c_{pen}(u_k) \cdot (2\tau - X)$
$\vdots$	$\vdots$
$L_1 < w_\tau \leq L_2$ and $w_{2\tau} \leq w_M$	$-2 \cdot c_l - c_i - c_p + r(u_0) \cdot \tau + r(u_2) \cdot \tau$
$L_1 < w_\tau \leq L_2$ and $w_{2\tau} > w_M$	$-2 \cdot c_l - c_i - c_c + r(u_0) \cdot \tau + r(u_2) \cdot \tau$ $-c_{pen}(u_2) \cdot (2\tau - X)$
$w_\tau \leq L_1$ and $w_{2\tau} \leq w_M$	$-2 \cdot c_l - c_i - c_p + r(u_0) \cdot \tau + r(u_1) \cdot \tau$
$w_\tau \leq L_1$ and $w_{2\tau} > w_M$	$-2 \cdot c_l - c_i - c_c + r(u_0) \cdot \tau + r(u_1) \cdot \tau$ $-c_{pen}(u_1) \cdot (2\tau - X)$

Table 6.12 – Possible scenarios and corresponding utility.

$$
\begin{aligned}
&+ \tau \cdot \sum_{h=1}^{k} \left[r(u_h) - c_{pen}(u_h) \right] \cdot \left[F_{W(\tau)}(L_h; u_0) - F_{W(\tau)}(L_{h-1}; u_0) \right] \\
&- (c_c - c_p) \cdot \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} \left[1 - F_{\Delta W(\tau,2\tau)}(w_M - w_\tau; u_h) \right] \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dw_\tau \\
&+ \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} \int_{\tau}^{2\tau} c_{pen}(u_h) \cdot F_{\Delta W(\tau,x)}(w_M - w_\tau; u_h) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot dw_\tau \\
&+ c_{pen}(u_0) \int_{w_M}^{\infty} \int_{0}^{\tau} F_{W(x)|W(\tau)}(w_M | w_\tau; u_0) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dx \cdot dw_\tau
\end{aligned} \tag{6.12}
$$

$$
\begin{aligned}
E\left\{T\left(W(\tau)\right)\right\} &= \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} T(w_\tau) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dw_\tau + \int_{L_k}^{\infty} T(w_\tau) \cdot f_{W(\tau)}(w_\tau; u_0) \cdot dw_\tau \\
&= 2\tau \cdot \sum_{h=1}^{k} \int_{L_{h-1}}^{L_h} f_{W(\tau)}(w_\tau; u_0) \cdot dw_\tau + \tau \cdot \int_{L_k}^{\infty} f_{W(\tau)}(w_\tau; u_0) \cdot dw_\tau \\
&= \tau \cdot \left[1 + F_{W(\tau)}(L_k; u_0) \right]
\end{aligned} \tag{6.13}
$$

Also under this policy, the degradation model is the gamma-based process described in

Chapter 6.2.2, where the expressions of the pdfs and cdf included in Eqs. (6.12) and (6.13) can be found.

6.3.3 Example of application

As in Chapter 6.2, the proposed adaptive policy is applied to a real-world inspired case of corroding pipelines. Under the same notation, we have adopted the following expressions for the age function, reward rate, and penalty cost rate:

$$a(u) = a \cdot \left(\frac{u}{u_{max}}\right)^d, \tag{6.14}$$

$$r(u) = r \cdot \frac{u}{u_{max}}, \tag{6.15}$$

$$c_{pen}(u) = c_{pen} \cdot r(u). \tag{6.16}$$

The parameters of the degradation model and of the cost model are reported in Tables 6.13. The failure threshold w_M has been set to 35 mm, while the maximum and minimum allowable usage rates have been set to $u_{max} = 1$ and $u_{min} = 0$.

a $[years]$	b	θ $[mm]$	d	c_p	c_c	c_i	c_l	c_{pen}	r
0.24	1	2.35	2	1	6	0.5	0.2	0.8	0.12

Table 6.13 – Parameters used to calibrate the degradation process and the cost model.

For the sake of comparison, let $\mathcal{P}_0(k)$ be the proposed policy and $\mathcal{P}_1$ be the special case of $\mathcal{P}_0(k)$ where $k = 1$ and $u_0 = u_1$. Essentially, $\mathcal{P}_1$ is a condition-based policy where the usage rate is set a priori to a fixed value that does not change during the maintenance cycle. Figure 6.1 depicts the optimal long-run average maintenance utility rate as a function of the inspection time τ under $\mathcal{P}_0(k)$ (for $k = 1, 3$, and 5) and under $\mathcal{P}_1$. This figure shows that, under the considered setup, adopting the more flexible policy $\mathcal{P}_0(k)$ can provide noticeable improvements in the long-run average maintenance utility rate. To understand how these performances are achieved, it is necessary to investigate how $\mathcal{P}_0(k)$ adaptively assigns the usage rate.

Figure 6.2 depicts the optimal values of the usage rates u assigned by the considered policies in case $\tau = 35$, as a function of w_τ. Here, the preventive replacement threshold L_k coincides with the smallest value of w_τ where the assigned usage rate is 0. Figure 6.2 shows that, under policy $\mathcal{P}_0$, to a unit that is barely degraded at $\tau = 35$ it will be

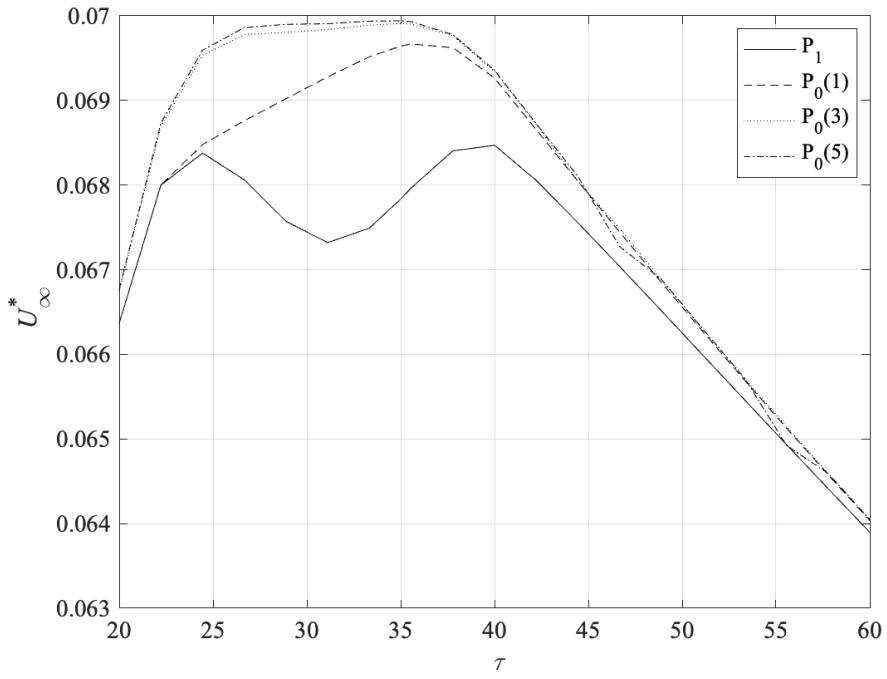

Figure 6.1 – Optimal long-run average maintenance reward rate as a function of the inspection time τ under the considered policies.

assigned a higher value of the usage rate, while to more degraded units it will be assigned a progressively lower usage rate. Moreover, the same figure also shows that, due to the lack of flexibility, policy $\mathcal{P}_1$ must be more conservative than $\mathcal{P}_0(k)$. Indeed, under $\mathcal{P}_0(k)$ the preventive replacement threshold is higher than under $\mathcal{P}_1$, increasing the probability that the replacement is postponed, which in turn prolongs the operating life of the unit. As the number of classes k increases, this effect is further accentuated.

Let ${}_cU^*_\infty$, ${}_iU^*_\infty$, ${}_lU^*_\infty$, ${}_pU^*_\infty$, and ${}_{pen}U^*_\infty$ be the contributions to the optimal long-run average maintenance reward rate U^*_∞ of corrective, inspection, logistic, preventive, and penalty costs, respectively, and ${}_rU^*_\infty$ be the contribution to U^*_∞ of the reward term (obviously, it is ${}_cU^*_\infty + {}_pU^*_\infty + {}_lU^*_\infty + {}_iU^*_\infty + {}_{pen}U^*_\infty + {}_rU^*_\infty = U^*_\infty$. The bar chart in Figure 6.3 shows (in black) the values of these contributions under policy $\mathcal{P}_1$ and (in grey) under policy $\mathcal{P}_0(5)$ (i.e., $\mathcal{P}_0$ with $k = 5$). The same values are also reported in Table 6.14, together with the total utility U^*_∞.

Figure 6.3 and Table 6.14 show that the reward earned under $\mathcal{P}_1$ is greater than the one earned under $\mathcal{P}_0(5)$ (i.e., the contribution of the reward term ${}_rU^*_\infty$ is higher under $\mathcal{P}_1$ than under $\mathcal{P}_0(5)$). However, this is compensated by the contributions of all the other

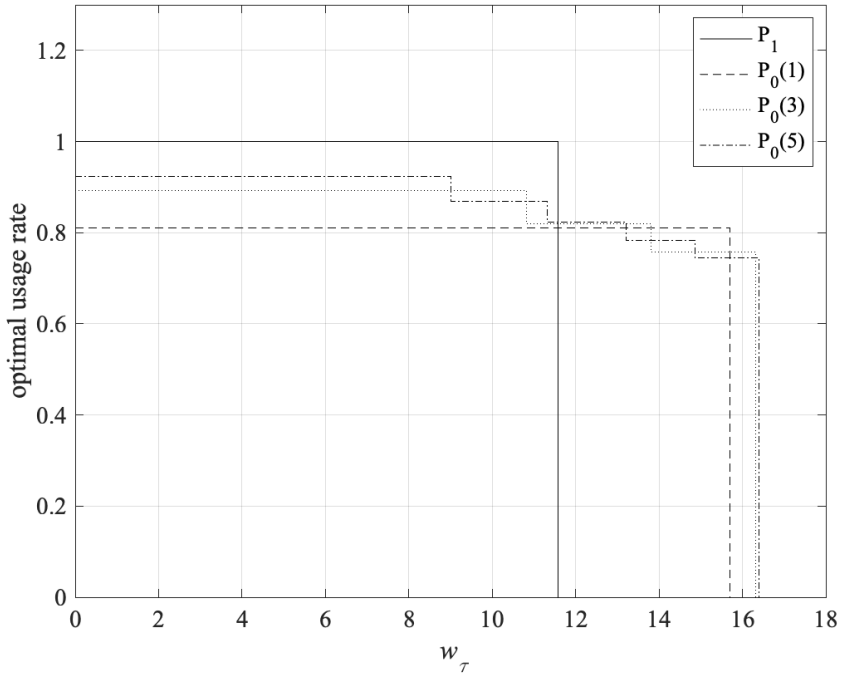

Figure 6.2 – Optimal values of the usage rate in $(\tau, 2\tau)$ as a function of w_τ at $\tau = 35$ under $\mathcal{P}_0(k)$ (with $k = 1, 3$, and 5) and $\mathcal{P}_1$.

(negative) cost factors, which are all smaller than under $\mathcal{P}_0(5)$. These results show that, by adaptively assigning a usage rate tailored to the actual degradation level of the unit, the policy $\mathcal{P}_0(5)$ is able to prolong the useful life of the unit while carefully managing the risk of failure.

6.4 Conclusions

In this chapter, two maintenance policy inspired by concepts of modern prescriptive maintenance have been proposed. Similarly to the ones discussed in Chapter 5, these policies rely on a single inspection performed at a predetermined time and, based on the outcome of this inspection, adopt a condition-based decision-making rule. Yet, with respect to the policies presented in Chapter 5, where the intervals between successive maintenance actions could be set freely by the policy based on economic considerations, here significant constraints regarding them are assumed to be in place, a situation that is commonly encountered in practice. Indeed, the driving idea behind the two prescrip-

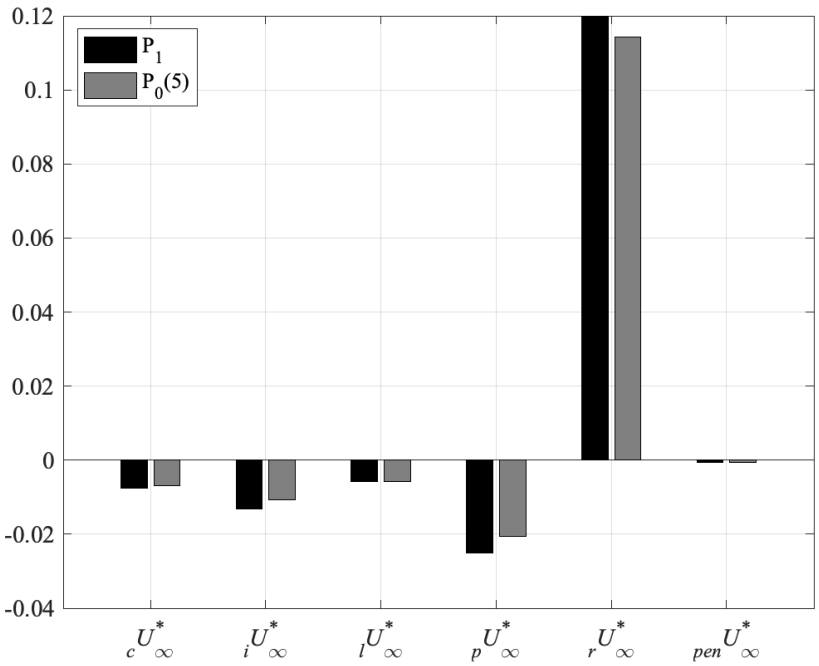

Figure 6.3 – Values of ${}_cU^*_\infty$, ${}_iU^*_\infty$, ${}_lU^*_\infty$, ${}_pU^*_\infty$, ${}_rU^*_\infty$, and ${}_{pen}U^*_\infty$

tive policies is that, especially in these cases where maintenance times are subjected to constraints, introducing another degree of freedom in the form of adjusting the usage rate can help in reaching a better global trade-off between inspection, replacement, and operational costs.

Obviously, changing the usage rate of a unit has an impact both on the degradation evolution and on operational costs. Therefore, a comprehensive cost model that can account for these effects should be developed. Indeed, this is currently one of the main challenges of the implementation of prescriptive maintenance. In Chapter 6.2 the basic decision-making scheme is presented. Similarly to the policy of Chapter 5.2, at the inspection time it is decided whether to immediately replace the unit or to postpone its replacement and, if deemed economically convenient, change its usage rate.
Then, the adaptive policy presented in Chapter 6.3 expands on the aforementioned basic scheme by assuming that, in case the replacement is postponed, an ad-hoc usage rate can be assigned to different units based on the measured degradation level at the inspection time.

Obtained results show that, depending on the relative values of parameters of the

	$\mathcal{P}_0(5)$	$\mathcal{P}_1$
${}_cU^*_\infty$	-0.0068	-0.0076
${}_iU^*_\infty$	-0.0108	-0.0076
${}_lU^*_\infty$	-0.0057	-0.0076
${}_pU^*_\infty$	-0.0205	-0.0076
${}_rU^*_\infty$	-0.1144	-0.0076
${}_{pen}U^*_\infty$	$-5.49 \cdot 10^{-4}$	$-6.64 \cdot 10^{-4}$
U^*_∞	-0.0068	-0.0076

Table 6.14 – Values of ${}_cU^*_\infty$, ${}_iU^*_\infty$, ${}_lU^*_\infty$, ${}_pU^*_\infty$, ${}_rU^*_\infty$, ${}_{pen}U^*_\infty$, and U^*_∞ obtained under $\mathcal{P}_1$ and under $\mathcal{P}_0(5)$.

degradation and cost model, adjusting the usage rate can be an economically advantageous strategy. Indeed, comparing the performances of the proposed prescriptive policy with simpler condition-based ones (obtained from the suggested policy by constraining the usage rate), manipulating the usage rate seems to allow the prescriptive policy to prolong and better exploit the operational life of the unit while carefully managing the risk of failure.

Chapter 7

A preliminary analysis of a prescriptive block replacement policy for a degrading production system

7.1 Introduction

In industrial production processes, maintenance and production planning play a crucial role in ensuring smooth operation of the system. Maintenance is responsible for overseeing the reliability and availability of equipment, while production planning focuses on maximizing the efficiency of the production process.

Even though the underlying goal for both is the maximization of production, their requirements can often be conflicting with each other. Indeed, production planning prioritizes the maximization of output, which may lead to a push for increased utilization of equipment, deferring or neglecting crucial maintenance actions, which might potentially result in unexpected breakdowns and overall lower production efficiency. On the other hand, maintenance is focused on ensuring the reliability and longevity of the equipment regardless of production constraints, which may require frequent interruption of operations for inspections, repairs, and replacement, in order to minimize downtime due to potential failures.

For these reasons, maintenance is typically considered a capacity-consuming activity and often results in compromising between minimizing disruption to production while guaranteeing reliable performance at the lowest cost (e.g., see Bajestani, Banjevic, and Beck (2014), Hajej and Rezg (2020), Lai, Z. Chen, and Bidanda (2019), Nasr, Salameh, and Moussawi-Haidar (2017)).

Taking advantage of the recent improvements in sensors and monitoring technologies,

maintenance strategies that incorporate the real-time information about the state of the equipment, such as Condition-Based Maintenance (CBM) have been extensively studied in the literature (e.g., see Van Noortwijk (2009) and Alaswad and Xiang (2017)). Within the context of production processes, Jafari and Makis (2015) developed a maintenance policy where the production facility is regularly inspected and the decision about whether to perform maintenance or to continue production is made based on the outcome of the inspection. By using a semi-Markov decision process, the authors solve the problem of the joint optimization of lot sizing and preventive maintenance and show that this approach can be significantly advantageous with respect to the traditional age-based policy in terms of long-run average maintenance cost rate.

The same authors also deal with a similar problem, but in the case of a partially observable facility (using a hidden Markov process) in Jafari and Makis (2016b). They extend the analysis also in the case of a two-unit system where the condition of most expensive unit is monitored and only age information is available for the other one in Jafari and Makis (2016a). Other papers studying the effects of CBM on production planning can be found in H. Peng and Houtum (2016), Cheng, B. H. Zhou, and L. Li (2017), Cheng, B. H. Zhou, and L. Li (2018), Khatab et al. (2019), and Zheng et al. (2021). From these papers we can remark that the main advantage of CBM lies in the fact that maintenance decisions can be taken on the basis on the degradation measures collected on the system under analysis during its operating life.

Another interesting area of research is the study of how production decisions, specifically modifications in the production rate, impact on the risk of failure. In some papers, such as Martinelli (2010), or Tan (2019), it is assumed that higher production rates entail higher failure risks but have no permanent effect on the degradation of the system. Other articles, such as Ruifeng and Subramaniam (2012), Ayed, Sofiene, and Nidhal (2012), Cheng, B. H. Zhou, and L. Li (2016), Zied, Sofiene, and Nidhal (2011), and Majdouline et al. (2022), assume that the degradation is dependent on the production rate, but do not take into account direct condition monitoring (instead, they use other tools, such as, for example, virtual age).

Some other authors have also considered the possibility of changing the production rate of a system based on its real-time state with the aim of influencing its degradation rate. Specifically, Broek, R. H. Teunter, Jonge, et al. (2020) analyze a single-unit production system subjected to regular condition monitoring where the gathered degradation measures are used to dynamically adjust the production rate with the aim of maximizing

production revenues and minimizing maintenance costs. A similar analysis, but in the case of a two-unit system with economic dependencies, is performed in Broek, R. H. Teunter, De Jonge, et al. (2021b).
Broek, R. H. Teunter, De Jonge, et al. (2021a) compare the performances of condition-based maintenance policies (i.e., policies where maintenance decisions are taken based on the measurements) against those of condition-based production policies (i.e., policies where production decisions are taken based on the measurements).
On the other hand, new maintenance paradigms, such as prescriptive maintenance, are getting increased attention in the very recent literature (see also Chapter 6) as a potentially useful tool to solve the conflict between maintenance and production planning.

In this chapter, we aim to apply concepts of modern prescriptive maintenance to develop a prescriptive block replacement maintenance policy for a production system undergoing stochastic degradation. The proposed policy consists in performing regular inspections aimed at measuring the degradation level of the production system. Based on the outcome of the inspections, it is decided whether to immediately replace the system or to continue operation. If deemed economically convenient, the production rate can be adjusted (either increased or decreased). Any change in the production rate should have an effect both on the future evolution of the degradation of the system and on its operating costs. Moreover, the policy must also take into account a production requirement, with the global objective being to maximize production revenue while minimizing downtime.

It is worth to emphasize that, the fact that in a prescriptive maintenance context, maintenance decisions and degradation influence each other generates serious mathematical complexity, both in the degradation modeling and (even more) in the optimization process. This work should therefore be seen as exploratory in this domain.

7.2 Description of the policy

The proposed policy relies on the following assumptions:

— A single production system, producing one-type items is considered;

— The processing time (i.e., the time it takes for the production system to produce a single item) is deterministic;

— The production horizon is divided in equally spaced Production Periods (PP) of length τ;

— Let $N_k(\tau)$ be the number of items produced in a PP of length τ. In each PP the system is expected to meet a production requirement of $N_k(\tau) \geq n_\tau$ items;

— The system is subjected to a gradual deterioration process that can be described via a gamma-based process;

— Failure is defined by the first passage time of the degradation process to a fixed threshold, say w_M. When failure occurs, production is halted immediately, and the system is inoperative until maintenance is performed;

— Both preventive and corrective replacements restore the system to an "as good as new" state and are performed in negligible time.

Based on this assumptions, the policy can be seen as a combination of three decision rules:

— A block replacement rule: preventive and corrective replacements, as well as inspections, can only be performed at predefined, periodic inspection times. Specifically, here it is supposed that inspections and replacements can only be performed at the end of each PP;

— A condition-based rule: based on the outcome of the inspection, the preventive and corrective replacement actions might be triggered. Specifically, if at the end of the kth PP ($k = 1, \ldots, \infty$), the inspection reveals that the system is failed it is correctively replaced, whereas if the system is still in a working state but its degradation level has already passed a preventive threshold, say L, it is deemed too degraded and is preventively replaced. Otherwise, production continues in the $(k+1)$th PP;

— A prescriptive rule: it is possible, only once within each PP, to change the Production Rate (PR) of the system. The PR can be either increased or decreased and can be changed at any time within the PP, with the only constraint that it cannot be changed while an item is being produced. However, it can be changed from its nominal value (denoted by u_{nom}) after q_k items are produced in the kth PP to a new value (denoted by u_k) which will be kept for the remainder of the PP. The PR will be reset to the nominal value at the start of the next PP. The decision about if, how, and when to change the PR during the kth PP will be taken based on the degradation level measured at the end of the $(k-1)$th PP. A change in the PR will have an effect both on the future degradation evolution of the system and on operating costs. In the following, a pair (q, u) will be referred to as a "prescriptive decision".

Therefore, the decision parameters of the proposed policy are the preventive threshold L, and the vectors $\boldsymbol{u}_k$ and $\boldsymbol{q}_k$, where $\boldsymbol{u}_k = \{u_1, u_2, u_3, \ldots\}$ and $\boldsymbol{q}_k = \{q_1, q_2, q_3, \ldots\}$.

7.3 The degradation process

The gamma process $\{Y(t); t \geq 0\}$ is a monotonic increasing process with independent and gamma distributed increments. Hence, it can be completely defined by an initial condition (here $Y(0) = 0$) and by either the probability density function (pdf) or the cumulative distribution function (cdf) of its generic increment $\Delta Y(t, t + \Delta t) = Y(t + \Delta t) - Y(t)$, given by, respectively:

$$f_{\Delta Y(t,t+\Delta t)}(y) = \frac{y^{\Delta\eta(t,t+\Delta t)-1}}{\theta^{\Delta\eta(t,t+\Delta t)} \cdot \Gamma\left(\Delta\eta(t, t+\Delta t)\right)} \cdot e^{-\frac{y}{\theta}} \tag{7.1}$$

and:

$$F_{\Delta Y(t,t+\Delta t)}(y) = \frac{\gamma\left(\Delta\eta(t, t+\Delta t), y/\theta\right)}{\Gamma\left(\Delta\eta(t, t+\Delta t)\right)} \tag{7.2}$$

where θ ($\theta > 0$) is the scale parameter, $\Gamma(\cdot)$ and $\gamma(\cdot)$ are the complete and lower incomplete gamma functions, respectively, $\Delta\eta(t, t + \Delta t) = \eta(t + \Delta t) - \eta(t)$, and $\eta(t) = a \cdot t^b$ is the age function. For the sake of simplicity, yet without loss of generality, in this chapter we assume that the considered gamma process is homogeneous (i.e., $b = 1$).

Following Tseng, Balakrishnan, and Tsai (2009), we assume that the PR influences the degradation process only through the age function (i.e., the scale parameter θ is independent of the PR). Specifically, we suppose that increasing the production rate will reduce the processing time of an item (e.g., items will be produced faster), but at the same time it also increases the mean degradation rate by increasing the parameter a of the age function. Under these assumptions, a system producing one item at PR u will sustain a degradation increment $\Delta W(0, \delta(u))$ that has the following pdf:

$$f_{\Delta W(0,\delta(u))}(\Delta w) = \frac{\Delta w^{a(u)\cdot\delta(u)-1}}{\theta^{a(u)\cdot\delta(u)} \cdot \Gamma\left(a(u) \cdot \delta(u)\right)} \cdot e^{-\frac{\Delta w}{\theta}}, \tag{7.3}$$

where $\delta(u)$ is a monotonic decreasing function of u that measures the processing time of one item at PR u, and $a(u)$ is a monotonic increasing function of u that captures the effect of the PR on the mean degradation rate.

Hereinafter, to highlight the influence of the prescriptive decisions on the degradation

evolution of the system, the degradation increment within the kth PP will be denoted by $\Delta W_k(z)$ (where $0 < z < \tau$), whereas the cumulated degradation level after k PP will be denoted by $W(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k)$ (where $\boldsymbol{q}_k = \{q_i, i = 1, \ldots, k\}$ and $\boldsymbol{u}_k = \{u_i, i = 1, \ldots, k\}$ keep track of the entire history of prescriptive decisions up to the kth PP). The pdf of $\Delta W_k(z)$ can be expressed as:

$$f_{\Delta W_k(z)}(\Delta w) = \frac{\Delta w^{\eta_k(z)-1}}{\theta^{\eta_k(z)} \cdot \Gamma\left(\eta_k(z)\right)} \cdot e^{-\frac{\Delta w}{\theta}} \tag{7.4}$$

where:

$$\eta_k(z) = \begin{cases} z \cdot a(u_{nom}) & z \leq q_k \cdot \delta(u_{nom}) \\ q_k \cdot \delta(u_{nom}) \cdot a(u_{nom}) + (z - q_k \cdot \delta(u_{nom})) \cdot a(u_k) & z > q_k \cdot \delta(u_{nom}) \end{cases}, \tag{7.5}$$

then, being $W(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k) = \sum_{i=1}^{k} \Delta W_i(\tau)$, thanks to the reproductive property of the gamma random variable, the pdf of $W(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k)$ can be expressed as:

$$f_{W(k\tau;\boldsymbol{q}_k,\boldsymbol{u}_k)}(w_{k\tau}; \boldsymbol{q}_k, \boldsymbol{u}_k) = \frac{w_{k\tau}^{\eta(k\tau;\boldsymbol{q}_k,\boldsymbol{u}_k)-1}}{\theta^{\eta(k\tau;\boldsymbol{q}_k,\boldsymbol{u}_k)} \cdot \Gamma\left(\eta(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k)\right)} \cdot e^{-\frac{w_{k\tau}}{\theta}} \tag{7.6}$$

and its cdf:

$$F_{W(k\tau;\boldsymbol{q}_k,\boldsymbol{u}_k)}(w_{k\tau}; \boldsymbol{q}_k, \boldsymbol{u}_k) = \frac{\gamma\left(\eta(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k), w_{k\tau}/\theta\right)}{\Gamma\left(\eta(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k)\right)} \tag{7.7}$$

where $\eta(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k) = \sum_{i=1}^{k} \eta_k(\tau)$ captures the effect of the entire history of prescriptive decisions up to the kth PP.

Eqs. (7.4) and (7.5) give clear evidence of the fact that the degradation increment within the kth PP only depends on the prescriptive decision made in the kth PP itself (i.e., the pair (q_k, u_k)). Conversely, Eqs. (7.6) and (7.7) show that the degradation level after k PPs is affected by the entire history of prescriptive decisions $(\boldsymbol{q}_k, \boldsymbol{u}_k)$. As already mentioned, the production system is assumed to fail when its degradation process passes for the first time a fixed threshold w_M. Given that inspections will be regularly performed at the end of each PP, we can suppose that the degradation value at the end of the $(k-1)$th PP, $w_{(k-1)\tau}$, will always be known before starting the kth PP (note that this also implies that a failed system would be systematically detected and will not start production already failed). Consequently, we can express the "local" (i.e., measured in the kth PP) lifetime X_k as:

$$X_k = inf\{x_k : \Delta W_k(x) > w_M - w_{(k-1)\tau}\}. \tag{7.8}$$

That is, x_k $(0 < x_k \leq \tau)$ is the first passage time, measured between 0 and τ, when the degradation process exceeds the failure threshold in the kth PP (obviously, $x_k \neq tau$ in the sole PP when failure occurs). In this sense, the "calendar" (i.e., measured from the start of the life cycle) lifetime will be:

$$X = \sum_{k=1}^{\infty} X_k$$

7.4 Cost formulation

The cost model is developed by considering the cost of an inspection c_i, the cost of a preventive replacement c_p, and the cost of a corrective replacement c_c. Moreover, we assume that a reward r_n is earned for each item produced in the kth PP until the production requirement n_τ is met. Any additional item produced in the same period beyond the requirement yields a reward $r_s < r_n$ (due to storage costs, for example). Conversely, if the requirement is not met and fewer than n_τ items will be produced, then a penalty cost c_{lp} for each non-produced item is incurred.
The reward earned in the kth PP, scenario by scenario, is summarized in Table 7.1, where $w_{k\tau}$ is the degradation level of the system at the kth inspection (i.e., at the end of the kth PP), and N_k is the number of items produced in the kth PP.

Measured degradation level	N_k	Reward $R_k(w_{k\tau}, x_k)$
$w_{k\tau} > w_M$	$N_k \geq n_\tau$	$n_\tau \cdot r_n + (N_k - n_\tau) \cdot r_s - c_c$
$w_{k\tau} > w_M$	$N_k < n_\tau$	$N_k(\tau) \cdot r_n - (n_\tau - N_k) \cdot c_{lp} - c_c$
$L < w_{k\tau} \leq w_M$	$N_k \geq n_\tau$	$n_\tau \cdot r_n + (N_k - n_\tau) \cdot r_s - c_p$
$L < w_{k\tau} \leq w_M$	$N_k < n_\tau$	$N_k \cdot r_n - (n_\tau - N_k) \cdot c_{lp} - c_p$
$w_{k\tau} \leq L$	$N_k \geq n_\tau$	$n_\tau \cdot r_n + (N_k - n_\tau) \cdot r_s$
$w_{k\tau} \leq L$	$N_k < n_\tau$	$N_k \cdot r_n - (n_\tau - N_k) \cdot c_{lp}$

Table 7.1 – Reward earned in the kth PP.

The number of produced items $N_k(\tau)$ in the kth PP can be computed as:

$$N_k = \begin{cases} x_k/\delta(u_{nom}) & x_k < q_k \cdot \delta(u_{nom}) \\ q_k + (x_k - q_k \cdot \delta(u_{nom}))/\delta(u_k) & x_k \geq q_k \cdot \delta(u_{nom}) \end{cases} . \tag{7.9}$$

From Eq. (7.9) it is clear that, given the prescriptive decision in the kth interval (q_k, u_k) the only other relevant quantity to compute N_k is the value of the local lifetime x_k.

The adopted performance measure is the long-run average maintenance reward rate, formulated via the renewal/reward theorem (e.g., see Ross (1983)):

$$R_\infty(L, \boldsymbol{q}, \boldsymbol{u}) = \frac{E\left\{\sum_{k=0}^{\infty} R_k(w_{k\tau}, x_k)\right\}}{MTBR}, \tag{7.10}$$

where $MTBR$ is the mean time between two replacements (i.e., the length of the maintenance cycle) given by:

$$MTBR = \tau \cdot \sum_{k=0}^{\infty} P(W(k\tau; \boldsymbol{q}_k, \boldsymbol{u}_k) < L) = \tau \cdot \sum_{k=0}^{\infty} F_{W(k\tau;\boldsymbol{q}_k,\boldsymbol{u}_k)}(L). \tag{7.11}$$

The optimal values of L, $\boldsymbol{q}$, and $\boldsymbol{u}$, which define the optimal policy, are the values L^*, $\boldsymbol{q}^*$, and $\boldsymbol{u}^*$ that maximize Eq. (7.10). The corresponding optimal long-run average maintenance reward rate is denoted by R^*_∞.

For simplicity of notation, Eq. (7.10) does not highlight that both the expected reward per cycle (at the numerator) and the $MTBR$ depend on the design variables L, $\boldsymbol{q}$, and $\boldsymbol{u}$. Conversely, it gives evidence that the expected reward explicitly depends on the whole sequence of measured degradation values at the end of every PP (i.e., $w_{k\tau}$ $\forall k = 1, \ldots, \infty$). This, combined with the fact that (in general) the production horizon is infinite, and that the maintenance decision taken in each PP will influence the degradation level at its end, which in turn influences the prescriptive decisions in the following PPs, leads to complexity in determining the optimal values $\boldsymbol{q}^*$ and $\boldsymbol{u}^*$ over an infinite horizon. Here, the long-run average maintenance reward rate is computed by means of Monte Carlo simulations. It is worth to remark that the considered degradation process is still Markovian and therefore only the degradation level at the start of a PP is relevant to decide if, when and how to change the PR during the PP.

7.5 Application example

We here illustrate an example of application of the proposed policy. The values of the parameters of the degradation model and of the cost model are reported in Table 7.2. For the sake of simplicity, we assume that the production rate can be selected from a set of three values: the nominal value u_{nom}, the maximum value u_{max} (that represents an accelerated PR) and the minimum value u_{min} (which represents a reduced PR). Table 7.3 reports the values of $\delta(u)$ and $a(u)$ for the considered production rates. Let Pol be the

w_M	θ	τ	r_n	r_s	c_{lp}	c_c	c_p	n_τ
10	0.2	100	5	3	4	2000	500	90

Table 7.2 – Values of the parameter of the degradation model and of the cost model.

	u_{min}	u_{nom}	u_{max}
$\delta(u)$	0.6	1	1.4
$a(u)$	1.2	1	0.6

Table 7.3 – Effect of the production rate on $\delta(u)$ and $a(u)$.

proposed policy. Then, let $Pol(u_{nom})$, $Pol(u_{min})$, and $Pol(u_{max})$ be the special cases of Pol where the PR is constrained to $u = u_{nom}$, $u = u_{min}$, and $u = u_{max}$, respectively. Table 7.4 reports the results obtained by optimizing all of the considered policies. It is worth to remark that the parameter L here is not optimized, but it is set to $L = 8.5$. Obtained

Policy	$MTBR$	$E\{N_k\}$ per cycle	R^*_∞
Pol	564	166.67	5.50
$Pol(u_{min})$	754	89.95	3.76
$Pol(u_{nom})$	668	99.05	3.49
$Pol(u_{max})$	575	164.24	5.30

Table 7.4 – $MTBR$, expected number of produced items per cycle, and long-run average maintenance reward rate under all of the considered policies.

results show that the more flexible Pol provides the best performance in terms of long-run average maintenance reward rate. Table 7.5 shows the optimal sequence of decisions on a randomly selected cycle (this is obtained by means of Monte Carlo simulation), under policy Pol. We can observe that, during the first 5 PPs, when the degradation is relatively low, the optimal decision is to set the PR to u_{max} for the entire PP. Then, when the degradation level gets closer to the preventive threshold, in the 6th period, the optimal decision shifts to slowing down the PR after having produced $q_6 = 65$ items (note that, even with the slowdown, in this PP the production target is still attained, as $N_6 = 90 = n_\tau$).

Subsequently, in the 7th PP, when the measured degradation at its start (i.e., $w_{6\tau} = 8.39$) is only slightly smaller than the preventive threshold (i.e.,$L = 8.5$), the PR is slowed down relatively quickly, after only $q_7 = 10$ items. In this period only $N_7 = 74$ items are produced, which yields a gross reward of $N_7 \cdot r_n = 370$. After accounting for the penalty

$(n_\tau - N_7) \cdot c_{lp} = 64$ and the cost of a preventive replacement $c_p = 500$, this period yields a (negative) reward $R_7 = -194$.
Therefore, we can conclude that, in this particular cycle, *Pol* opted to accelerate the PR as much as possible at the start, then slow down in the 6th PP (while still respecting the production target) so that the preventive replacement can be deferred to the 7th PP. In this latter period, production is slowed down considerably to minimize the probability of failure. This shows that, in this example, the prescriptive policy *Pol* is able to prolong the operational lifetime of the system while still managing the probability of failure.

PP (k)	$w_{(k-1)\tau}$	u_k	q_k	N_k	R_k
1	0	u_{max}	0	166	678
2	1.5	u_{max}	0	166	678
3	3.06	u_{max}	0	166	678
4	4.91	u_{max}	0	166	678
5	5.59	u_{max}	0	166	678
6	7.88	u_{min}	65	90	450
7	8.39	u_{min}	10	74	-194

Table 7.5 – A one-cycle production example ending with a preventive replacement.

For the sake of understanding how *Pol* achieves this performances, let us now focus on a single PP, with an initial degradation level relatively close to the threshold. Indeed, as observed in Table 7.5, it is in these cases that we can expect the policy to carefully manipulate the PR. Moreover, since the considered degradation process is state-independent, it will behave in the same way regardless of the initial degradation level. To perform this analysis, we conducted two constrained optimization problems, under policy *Pol*, where the prescriptive decision is constrained to (u_{max}, q) and (u_{min}, q), respectively. In other words, we will analyze the evolution of a single prescriptive decision, in a single PP, as a function of the measured degradation level at the start of the interval.

Figure 7.1 reports the result of this analysis, in red for the case of an acceleration and in blue for the case of a deceleration. Specifically, the x-axis reports the degradation level at the start of the PP, denoted by w, whereas the y-axis reports the value of q (i.e., the number of items produced before a change in the PR).
Figure 7.1 shows that, until $w = 7.8$, it is more convenient to accelerate the PR right away. As w increases further, an acceleration becomes less and less advantageous, and after $w = 8.1$ the acceleration is relegated to the end of the PP (let us remind that in this specific example the optimization is constrained, hence the acceleration is forced even

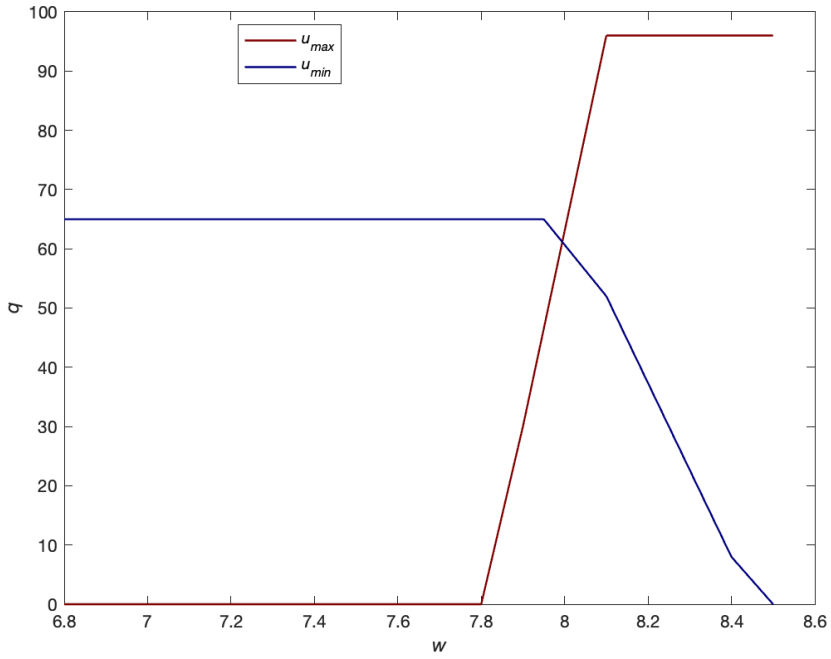

Figure 7.1 – Evolution of the optimal number of produced items before an acceleration (red line) or a deceleration (blue line) of the PR as a function of the initial degradation.

though it may not be the overall optimal decision).
Figure 7.1 also shows that, until a value of $w = 7.9$, it is only convenient to slow down the PR after 65 items have been produced. We can suspect that the specific value of 65 has been selected because, under the considered setup, producing 65 items at the nominal PR and subsequently decelerating guarantees that $N_k = n_\tau = 90$ items will be produced in the PP, hence respecting the production target. Then, as w increases further, it is more convenient to decelerate earlier and earlier.

7.6 Conclusions

The question of the integration of maintenance in the production planning process remains a largely open subject. Due to its inherent random behavior, maintenance strategies are often optimized by using long-term average criterion, which can usually lead to very conservative recommendations that are sometimes difficult to accept from the operational point of view.

In this chapter, we carried out a preliminary analysis of the use of prescriptive maintenance on a simplified production problem. To this aim, we proposed a maintenance policy that can be seen as a combination of three decision rule: an age-based rule (i.e., maintenance interventions can be carried out only at regular and predetermined intervals), a condition-based rule (i.e., replacements are triggered by direct degradation measurements), and a prescriptive rule (i.e., the production rate can be adjusted based on the gathered measurements and economic considerations).

The rationale behind this policy is that manipulating the production rate can offer more flexibility and simultaneously maximize production while managing the probability of unexpected failures.

Despite the simplifying assumptions we made, the fact that prescriptive decisions influence the future evolution of the degradation, and vice versa, quickly gives rise to mathematical complexity in optimizing the considered strategy. Nevertheless, we can retain from this chapter that a prescriptive policy might be an effective approach to maintaining a production system.

General conclusions

In this manuscript, we embarked on a journey to unravel the complexities of degradation modeling and maintenance optimization, seeking to address practical challenges in this field. In this concluding chapter, we aim to summarize this research journey, synthesize its main contributions, and provide some ideas about future perspectives.

The main contributions of this PhD program have been oriented along two main axes: (i) studying, developing, proposing, and comparing stochastic processes that can be used to model the evolution of the degradation of real systems, and (ii) developing maintenance strategies that can take advantage of the use of these degradation models.

With respect to the first axis, attention has been focused on developing stochastic processes that can account for different sources of variability. Specifically, a new perturbed gamma degradation process that can simultaneously take into account temporal variability, unit-to-unit variability and measurement error has been proposed. As it is commonly observed in the literature, introducing more model complexity often leads to mathematically intractable solutions, especially when it comes to calibrate a model and perform predictions with it. To overcome this issues, a new expectation-maximization (EM) algorithm, which also takes advantage of a particle filter, has been developed that allows to efficiently estimate model parameters. The same particle filter is then used also to perform predictions.

The same EM algorithm has then also been used to carry out a large misspecification study, investigating the model selection issue of perturbed gamma and perturbed inverse Gaussian degradation models. In fact, these two models exhibit very similar characteristics and are often treated as equal to each other. Due mainly to the long computational times involved in the parameter estimation procedures, the misspecification studies presented in the current literature have always been limited to a relatively small number of datasets (which hinders the generalizability of the results) and have not considered the impacts of measurement error. By taking advantage of the aforementioned EM algorithm, we managed to reduce the computational times enough so that a large study, considering different measurement error distribution and involving several thousands of estimations could be carried out.

Then, a small exploratory extension of the misspecification study has also been carried out to evaluate its impact on maintenance costs, a matter that has seldom received attention in the literature.
Obtained results have shown that, in the absence of measurement error, the consequences of using a wrong model selection can be severe, both in terms of remaining useful lifetime and of maintenance costs. On the other hand, if by one side the presence of measurement error massively complicates the estimation procedure, it also seems to equally reduce the consequences of a wrong diagnosis. However, regardless of the presence of measurement error, if model selection is performed by adopting a rigorous statistical criterion, then the performances of the considered models are quite similar.

With respect to the second axis, focus has mainly been directed to developing condition-based maintenance policies that can exploit the newly developed perturbed gamma process with random effects. Indeed, we proposed three maintenance policies based on the assumption that a single inspection can be performed during the life cycle of the system under study, with the core issue being how to better exploit the information gathered at a single time point in the presence of a heterogeneous population. After proposing a basic decision-making scheme, we investigated how an adaptive rule and considering inspections of different quality can influence the performance of the policy.
The comparative analysis carried out showed that, when inspection costs are high, a policy based on a single inspection can be a fruitful strategy. Moreover, in some particular scenarios where inspection costs and the consequences of an undetected failure are particularly high and mild, respectively, it can even be more economically advantageous to resort to simpler purely systematic maintenance approaches.

The same basic maintenance decision-making scheme, with the restriction on the number of inspections, has also been extended by integrating elements from modern prescriptive maintenance. Prescriptive maintenance is a new concept that has been gaining increased attention from researcher in the field of maintenance optimization. By simultaneously considering all aspects of a system in a holistic framework, including commercial and operational aspects, prescriptive maintenance might help alleviate the structural conflict in maintenance optimization between maximum commercial utilization of a system and ensuring its safe operation.
Specifically, we investigated how, in an experimental scenario where maintenance interventions cannot be scheduled according to convenience, adding one more degree of freedom (in the form of manipulating the usage rate of the system, or the production rate of a

production system) affects maintenance and operational costs.

Perspectives

Based on the insights gathered during this research journey, we believe that this activity opens these research directions:

— all of the stochastic processes adopted in this manuscript assume a monotonic increasing degradation rate function (here intended as the derivative of the mean function), as it is also commonly done in most of the literature (e.g., see Abdel-Hameed (1975), Whitmore (1995), Kahle, Mercier, and Paroissin (2016)). However, real-world degradation processes often exhibit bathtub-shaped, three-phases behavior (I. Gertsbakh and Kordonskiy (1969)), corresponding to a S-shaped degradation rate function: an accommodation phase, a steady-state phase, and a degenerative phase. Models with monotonic increasing degradation rate can only account for two of these three phases, but a degradation process that can model all three phases might provide better prognostic performances. We already took some preliminary steps in this direction (this has not been detailed in the manuscript). Specifically, in Piscopo et al. (2023), we proposed a gamma-based process built around a bathtub-shaped degradation rate which also integrates random effect and measurement error. The main potential of this new model is its ability to perform accurate lifetime prediction based on early data alone (i.e., predict the degradation behavior in the second and third phase by only observing data from the first phase). This latter feature can have interesting properties in maintenance optimization: indeed, in some applications different phases may need different actions to be performed (e.g., see El Hajj et al. (2023));

— Both policies presented in Chapter 6 investigate how to manipulate the usage rate in an experimental scenario where maintenance intervention dates are subjected to constraints, in the case of a single unit. With the same reasoning, it is easy to imagine how, in a scenario with multiple units working together (such as a pumping station with several pumps), adjusting some usage rate-like measure (such as the volume of water being pumped by a specific pump) can be used to control the probability of failure and synchronize the optimal maintenance dates of all units to a common calendar in a block maintenance scheme. This idea, also known as "load sharing" is not a new concept in the maintenance literature (e.g., see Broek, R. H. Teunter,

De Jonge, et al. (2021b)) and has also received some attention in the prescriptive literature (e.g., see Longhitano et al. (2021a)). A scheme like this would of course introduce strong dependency between the degradation processes of the considered units, increasing the complexity of the required modeling solution;

— Again with respect to the prescriptive policies, one of the cornerstones of their conception is the fact that inspections are supposed to be very expensive due to difficulty in accessing the equipment and/or requiring interruption of service and specialized maintenance personnel. Similarly, it is also supposed that carrying out any other maintenance action (such as a replacement, regardless of whether an inspection is performed) entails a logistic cost. However, a prescriptive action, such as changing the usage rate, is much less invasive than other action and might not even require on-site personnel to be performed. Similarly, sensors producing cheap low quality measurements can be installed and remotely operated. In this scenario, we can imagine an extension of the decision-making scheme proposed in Chapter 6 where, in between the already scheduled maintenance times, some low-quality measurements can be carried out, which would then inform a potential adjustment in the usage rate, if deemed convenient.

Appendix A

Eq. (5.3) can be derived as follows:

$$
\begin{aligned}
&E\left\{{}_1C(Z(\tau),W(\tau),X)\right\}\\
&\quad=\int_0^{\infty}\int_0^{\infty}\int_0^{\infty}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_0^{w_M}\int_0^{L_Z}\int_{\tau}^{\tau+\Delta\tau}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_0^{w_M}\int_0^{L_Z}\int_{\tau+\Delta\tau}^{\infty}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_{w_M}^{\infty}\int_0^{L_Z}\int_0^{\tau}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_0^{w_M}\int_{L_Z}^{\infty}\int_{\tau}^{\infty}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_{W_M}^{\infty}\int_{L_Z}^{\infty}\int_0^{\tau}{}_1C(z_\tau,w_\tau,X)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx
\end{aligned}
$$

Each of these five integrals represents one of the scenarios in Table 5.3. Thus, substituting from the same Table:

$$
\begin{aligned}
&E\left\{{}_1C(Z(\tau),W(\tau),X)\right\}\\
&\quad=\int_0^{w_M}\int_0^{L_Z}\int_{\tau}^{\tau+\Delta\tau}[2\cdot c_l+c_i+c_c+c_d\cdot(\tau+\Delta\tau-x)]\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{w_M}\int_0^{L_Z}\int_{\tau+\Delta\tau}^{\infty}(2\cdot c_l+c_i+c_p)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{w_M}^{\infty}\int_0^{L_Z}\int_0^{\tau}[2\cdot c_l+c_i+c_c+c_d\cdot(\tau+\Delta\tau-x)]\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{w_M}\int_{L_Z}^{\infty}\int_{\tau}^{\infty}(c_l+c_i+c_p)\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{w_M}^{\infty}\int_{L_Z}^{\infty}\int_0^{\tau}[c_l+c_i+c_c+c_d\cdot(\tau-x)]\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=c_l+c_i+c_p+c_l\cdot\int_0^{\infty}\int_0^{L_Z}f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dz_\tau\cdot dw_\tau\\
&\quad+(c_c-c_p)\cdot\int_0^{w_M}\int_0^{L_Z}\int_{\tau}^{\tau+\Delta\tau}f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dx\cdot dz_\tau\cdot dw_\tau\\
&\quad+c_d\cdot\int_0^{w_M}\int_0^{L_Z}\left[\int_{\tau}^{\tau+\Delta\tau}(\tau+\Delta\tau-x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot dx\right]\\
&\qquad\qquad\times f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dz_\tau\cdot dw_\tau\\
&\quad+c_d\cdot\int_{w_M}^{\infty}\int_0^{L_Z}\left[\tau+\Delta\tau-\int_0^{\tau}x\cdot f_{X|W(\tau)}(x|w_\tau)\cdot dx\right]\\
&\qquad\qquad\times\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dz_\tau\cdot dw_\tau\\
&\quad+c_d\cdot\int_{w_M}^{\infty}\int_{L_Z}^{\infty}\left[\tau-\int_0^{\tau}x\cdot f_{X|W(\tau)}(x|w_\tau)\cdot dx\right]\\
&\qquad\qquad\times\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dz_\tau\cdot dw_\tau
\end{aligned}
$$

Which, by using results given in Chapters 2.4 and 2.6, and after a few additional manipulations, provides the expression in Eq. (5.3). Eqs. (5.4)-(5.6) can be derived using a similar approach.

B.1

Eq. (5.12) can be derived starting as follows:

$$
\begin{aligned}
&E\{C(Z(\tau),W(\tau),X)\}\\
&\quad=\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad=\int_0^{w_M}\int_{L_u}^{\infty}\int_{\tau}^{\infty} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{w_M}^{\infty}\int_{L_u}^{\infty}\int_0^{\tau} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{L_W}\int_{L_l}^{L_u}\int_{\tau}^{\tau+\Delta\tau} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{L_W}\int_{L_l}^{L_u}\int_{\tau+\Delta\tau}^{\infty} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{L_W}^{w_M}\int_{L_l}^{L_u}\int_{\tau}^{\infty} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{w_M}^{\infty}\int_{L_l}^{L_u}\int_0^{\tau} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{w_M}\int_0^{L_l}\int_{\tau}^{\tau+\Delta\tau} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_0^{w_M}\int_0^{L_l}\int_{\tau+\Delta\tau}^{\infty} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx\\
&\quad+\int_{w_M}^{\infty}\int_0^{L_l}\int_0^{\tau} C(z_\tau,w_\tau,x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx
\end{aligned}
\tag{B.1}
$$

Hence, from Eq. (B.1), replacing $C(z_\tau, w_\tau, x)$ (integral by integral) with the appropriate expression specified in Table 5.18, Eq. (B.2) is readily obtained:

$$
\begin{aligned}
&E\{C(Z(\tau), W(\tau), X)\} \\
&= \int_0^{w_M} \int_{L_u}^{\infty} \int_{\tau}^{\infty} (c_l + c_z + c_p) \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_{w_M}^{\infty} \int_{L_u}^{\infty} \int_{0}^{\tau} [c_l + c_i + c_c + c_d \cdot (\tau - x)] \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_0^{L_W} \int_{L_l}^{L_u} \int_{\tau}^{\tau+\Delta\tau} [2 \cdot c_l + c_z + c_w + c_c + c_d \cdot (\tau + \Delta\tau - x)] \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_0^{L_W} \int_{L_l}^{L_u} \int_{\tau+\Delta\tau}^{\infty} (2 \cdot c_l + c_z + c_w + c_p) \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_{L_W}^{w_M} \int_{L_l}^{L_u} \int_{\tau}^{\infty} (c_l + c_z + c_w + c_p) \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_{w_M}^{\infty} \int_{L_l}^{L_u} \int_{0}^{\tau} [c_l + c_z + c_w + c_c + c_d \cdot (\tau - x)] \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_0^{w_M} \int_{0}^{L_l} \int_{\tau}^{\tau+\Delta\tau} [2 \cdot c_l + c_z + c_c + c_d \cdot (\tau + \Delta\tau - x)] \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_0^{w_M} \int_{0}^{L_l} \int_{\tau+\Delta\tau}^{\infty} (2 \cdot c_l + c_i + c_p) \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx \\
&+ \int_{w_M}^{\infty} \int_{0}^{L_l} \int_{0}^{\tau} [2 \cdot c_l + c_z + c_c + c_d \cdot (\tau + \Delta\tau - x)] \\
&\qquad \times f_{X|W(\tau)}(x|w_\tau) \cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dz_\tau \cdot dx
\end{aligned}
\tag{B.2}
$$

where:

$$\int_0^{w_M}\int_{L_u}^{\infty}\int_{\tau}^{\infty}(c_l+c_z+c_p)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx$$
$$=\int_0^{w_M}\int_{L_u}^{\infty}(c_l+c_z+c_p)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau$$

and:

$$\int_{L_W}^{w_M}\int_{L_l}^{L_u}\int_{\tau}^{\infty}(c_l+c_z+c_w+c_p)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\cdot dx$$
$$=\int_{L_W}^{w_M}\int_{L_l}^{L_u}(c_l+c_z+c_w+c_p)\cdot f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau$$

Eq. (B.2) can be then rewritten as:

$$\begin{aligned}
&E\{C(Z(\tau),W(\tau),X)\}\\
&\quad=c_l+c_z+c_p+c_w\cdot[F_{Z(\tau)}(L_u)-F_{Z(\tau)}(L_l)]+(c_c+c_p+c_d\cdot\tau)\cdot[1-F_{W(\tau)}(w_M)]\\
&\quad+c_l\cdot F_{Z(\tau)}(L_l)+c_l\cdot\int_0^{L_W}\left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau)-F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right]\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\\
&\quad+(c_c-c_p)\cdot\int_0^{L_W}\int_{\tau}^{\tau+\Delta\tau}\left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau)-F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right]\\
&\qquad\qquad\times f_{X|W(\tau)}(x|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx\\
&\quad+(c_c-c_p)\cdot\int_0^{w_M}\int_{\tau}^{\tau+\Delta\tau}f_{X|W(\tau)}(x|w_\tau)\cdot F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx\\
&\quad+c_d\cdot\Delta\tau\cdot\int_{w_M}^{\infty}F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\\
&\quad+c_d\cdot\int_0^{w_M}\int_{\tau}^{\tau+\Delta\tau}\cdot f_{X|W(\tau)}(x|w_\tau)\cdot F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx\\
&\quad+c_d\cdot\int_0^{L_W}\left[\int_{\tau}^{\tau+\Delta\tau}(\tau+\Delta\tau-x)\cdot f_{X|W(\tau)}(x|w_\tau)\cdot dx\right]\\
&\qquad\qquad\times\left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau)-F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right]\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\\
&\quad-c_d\cdot\int_{w_M}^{\infty}\int_0^{\tau}x\cdot f_{X|W(\tau)}(x|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx
\end{aligned}\tag{B.3}$$

Hence, from Eq. (B.3), by exploiting the results given in Eqs. (2.70), (2.72), and (2.74), the expression in Eq. (5.12) is obtained.

Eq. (5.13) can be more easily derived by following the scheme given in Table 5.17:

$$\begin{aligned}
E\{T(Z(\tau),W(\tau))\} &= \tau\cdot[1-F_{Z(\tau)}(L_u)]+(\tau+\Delta\tau)\cdot F_{Z(\tau)}(L_l)\\
&+\tau\cdot\int_{L_W}^{\infty}\int_{L_l}^{L_u} f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau\\
&+(\tau+\Delta\tau)\cdot\int_{0}^{L_W}\int_{L_l}^{L_u} f_{Z(\tau)|W(\tau)}(z_\tau|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dz_\tau \qquad \text{(B.4)}\\
&=\tau+\Delta\tau\cdot F_{Z(\tau)}(L_l)\\
&\quad+\Delta\tau\cdot\int_{0}^{L_W}\left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau)-F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right]\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau
\end{aligned}$$

B.2

Under the policy $\mathcal{P}_k$ $(k=1,2,3)$ the long-run average maintenance cost rate is computed as:

$${}^{k}C_{\infty}({}^{k}\boldsymbol{\xi}^{*})=\frac{E\left\{{}^{k}C(Z(\tau),W(\tau),X)\right\}}{E\left\{{}^{k}T(Z(\tau),W(\tau))\right\}}$$

where $E\left\{{}^{k}C(Z(\tau),W(\tau),X)\right\}$ and $E\left\{{}^{k}T(Z(\tau),W(\tau))\right\}$ are obtained as special cases from Eqs. (5.12) and (5.13), respectively, by imposing the constraint relative to policy $\mathcal{P}_k$. Thus, in the case of policy $\mathcal{P}_1$, by imposing the constraint $L_l=L_u$, Eq. (5.12) simplifies to:

$$\begin{aligned}
E\left\{{}^{1}C(Z(\tau),W(\tau),X)\right\} &= c_l+c_z+c_p+(c_c-c_p+c_d\cdot\tau)\cdot\left[1-F_{W(\tau)}(w_M)\right]\\
&+(c_l+c_d\cdot\Delta\tau)\cdot\int_{0}^{\infty}F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\\
&+(c_c-c_p)\cdot\int_{0}^{w_M}\left[1-F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M-w_\tau|w_\tau)\right]\\
&\qquad\times F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\\
&-c_d\cdot\int_{0}^{w_M}\int_{\tau}^{\tau+\Delta\tau}F_{\Delta W(\tau,x)|W(\tau)}(w_M-w_\tau|w_\tau)\cdot F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx\\
&-c_d\cdot\int_{w_M}^{\infty}\int_{0}^{\tau}F_{W(x)|W(\tau)}(w_M|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau\cdot dx,
\end{aligned}\qquad \text{(B.5)}$$

and Eq. (5.13) reduces to:

$$E\left\{{}^{1}T(Z(\tau),W(\tau))\right\}=\tau+\Delta\tau\cdot\int_{0}^{\infty}F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\cdot f_{W(\tau)}(w_\tau)\cdot dw_\tau. \qquad \text{(B.6)}$$

Similarly, in the case of policy $\mathcal{P}_2$, by imposing $L_l = 0$, $L_u = +\infty$, and $c_z = 0$, it is:

$$\begin{aligned} E\left\{{}^2C(Z(\tau), W(\tau), X)\right\} &= c_l + c_w + c_p + (c_c - c_p + c_d \cdot \tau) \cdot \left[1 - F_{W(\tau)}(w_M)\right] \\ &+ (c_l + c_d \cdot \Delta\tau) \cdot \int_0^{L_W} f_{W(\tau)}(w_\tau) \cdot dw_\tau \\ &+ (c_c - c_p) \cdot \int_0^{L_W} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\ &- c_d \cdot \int_0^{L_W} \int_\tau^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx \\ &- c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx, \end{aligned} \tag{B.7}$$

$$E\left\{{}^2T(Z(\tau), W(\tau))\right\} = \tau + \Delta\tau \cdot \int_0^{\infty} f_{W(\tau)}(w_\tau) \cdot dw_\tau. \tag{B.8}$$

Finally, in che case of policy $\mathcal{P}_3$, by imposing $L_l = L_u = 0$, and $c_z = 0$ it results:

$$\begin{aligned} E\left\{{}^3C(Z(\tau), W(\tau), X)\right\} &= c_l + c_p + (c_c - c_p + c_d \cdot \tau) \cdot \left[1 - F_{W(\tau)}(w_M)\right] \\ &- c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx, \end{aligned} \tag{B.9}$$

$$E\left\{{}^3T(Z(\tau), W(\tau))\right\} = \tau. \tag{B.10}$$

B.3

The long-run average maintenance cost rate ${}^kC_\infty(\cdot)$ can be intended as the sum of the contributions of corrective replacement costs, preventive replacement costs, perturbed inspection costs, perfect inspection costs, logistic costs, and downtime costs. These contributions can be expressed as:

$${}^k_cC_\infty(\cdot) = \frac{{}^k_cC(\cdot)}{E\left\{{}^kT(Z(\tau), W(\tau))\right\}}, \ldots {}^k_zC_\infty(\cdot) = \frac{{}^k_zC(\cdot)}{E\left\{{}^kT(Z(\tau), W(\tau))\right\}},$$

where $E\left\{{}^{k}T(Z(\tau),W(\tau))\right\}$ is the expected length of a maintenance cycle under policy $\mathcal{P}_k$ and the numerators, from Eq. (5.12), can be expressed as:

$$
\begin{aligned}
{}^{k}_{c}C(\cdot) &= c_c \cdot \left[1 - F_{W(\tau)}(w_M)\right] \\
&\quad + c_c \cdot \int_0^{L_W} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&\quad + c_c \cdot \int_0^{w_M} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau
\end{aligned}
\tag{B.11}
$$

$$
\begin{aligned}
{}^{k}_{d}C(\cdot) &= c_d \cdot \tau \cdot \left[1 - F_{W(\tau)}(w_M)\right] \\
&\quad + c_d \cdot \Delta\tau \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&\quad + c_d \cdot \Delta\tau \cdot \int_0^{L_W} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&\quad - c_d \cdot \int_0^{w_M} \int_{\tau}^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \\
&\qquad \times F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx \\
&\quad - c_d \cdot \int_0^{L_W} \int_{\tau}^{\tau+\Delta\tau} F_{\Delta W(\tau,x)|W(\tau)}(w_M - w_\tau|w_\tau) \\
&\qquad \times \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx \\
&\quad - c_d \cdot \int_{w_M}^{\infty} \int_0^{\tau} F_{W(x)|W(\tau)}(w_M|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \cdot dx,
\end{aligned}
\tag{B.12}
$$

$$
\begin{aligned}
{}^{k}_{p}C(\cdot) &= c_p - c_p \cdot \int_0^{L_W} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&\quad - c_p \cdot \int_0^{w_M} \left[1 - F_{\Delta W(\tau,\tau+\Delta\tau)|W(\tau)}(w_M - w_\tau|w_\tau)\right] \\
&\qquad \times F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau
\end{aligned}
\tag{B.13}
$$

$$
\begin{aligned}
{}^{k}_{p}C(\cdot) &= c_l + c_l \cdot \int_0^{\infty} F_{Z(\tau)|W(\tau)}(L_l|w_\tau) \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \\
&\quad + c_l \cdot \int_0^{L_W} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau
\end{aligned}
\tag{B.14}
$$

$$
{}^{k}_{w}C(\cdot) = c_w \cdot \int_0^{\infty} \left[F_{Z(\tau)|W(\tau)}(L_u|w_\tau) - F_{Z(\tau)|W(\tau)}(L_l|w_\tau)\right] \cdot f_{W(\tau)}(w_\tau) \cdot dw_\tau \tag{B.15}
$$

$$
{}^{k}_{z}C(\cdot) = c_z. \tag{B.16}
$$

As remarked in Chapter 5.4.3, under $\mathcal{P}_1$ it is $L_l = l_u$ (hence it results ${}^{1}_{w}C(\cdot) = 0$), under $\mathcal{P}_2$ it is $L_l = 0$, $L_u = +\infty$, and $c_z = 0$ (hence it results ${}^{2}_{z}C(\cdot) = 0$, and under $\mathcal{P}_3$ it is $L_l = L_u = 0$ and $c_z = c_w = 0$ (hence it results ${}^{3}_{w}C(\cdot) = {}^{3}_{z}C(\cdot) = 0$).

BIBLIOGRAPHY

Abdel-Hameed, Mohamed (1975), « A gamma wear process », *in*: *IEEE transactions on Reliability* 24(*2*), pp. 152–153.

Ahmad, Rosmaini and Shahrul Kamaruddin (2012), « An overview of time-based and condition-based maintenance in industrial application », *in*: *Computers & industrial engineering* 63(*1*), pp. 135–149.

Akaike, Hirotugu (1974), « A new look at the statistical model identification », *in*: *IEEE transactions on automatic control* 19(*6*), pp. 716–723.

Alaswad, Suzan and Yisha Xiang (2017), « A review on condition-based maintenance optimization models for stochastically deteriorating system », *in*: *Reliability engineering & system safety* 157, pp. 54–63.

Ayed, Souheil, Dellagi Sofiene, and Rezg Nidhal (2012), « Joint optimisation of maintenance and production policies considering random demand and variable production rate », *in*: *International Journal of Production Research* 50(*23*), pp. 6870–6885.

Bajestani, Maliheh Aramon, Dragan Banjevic, and J Christopher Beck (2014), « Integrated maintenance planning and production scheduling with Markovian deteriorating machine conditions », *in*: *International Journal of Production Research* 52(*24*), pp. 7377–7400.

Bautista, Lucía, Inma T Castro, and Luis Landesa (2022), « Condition-based maintenance for a system subject to multiple degradation processes with stochastic arrival intensity », *in*: *European Journal of Operational Research* 302(*2*), pp. 560–574.

Bismut, Elizabeth, Mahesh D Pandey, and Daniel Straub (2022), « Reliability-based inspection and maintenance planning of a nuclear feeder piping system », *in*: *Reliability Engineering & System Safety* 224, p. 108521.

Bordes, Laurent, Christian Paroissin, and Ali Salami (2016), « Parametric inference in a perturbed gamma degradation process », *in*: *Communications in Statistics-Theory and Methods* 45(*9*), pp. 2730–2747.

Broek, Michiel AJ uit het, Ruud H Teunter, Bram De Jonge, et al. (2021a), « Joint condition-based maintenance and condition-based production optimization », *in*: *Reliability Engineering & System Safety* 214, p. 107743.

Broek, Michiel AJ uit het, Ruud H Teunter, Bram De Jonge, et al. (2021b), « Joint condition-based maintenance and load-sharing optimization for two-unit systems with economic dependency », *in*: *European Journal of Operational Research* 295(*3*), pp. 1119–1131.

Broek, Michiel AJ uit het, Ruud H Teunter, Bram de Jonge, et al. (2020), « Condition-based production planning: Adjusting production rates to balance output and failure risk », *in*: *Manufacturing & Service Operations Management* 22(*4*), pp. 792–811.

Castanier, Bruno, Christophe Bérenguer, and Antoine Grall (2003), « A sequential condition-based repair/replacement policy with non-periodic inspections for a system subject to continuous wear », *in*: *Applied stochastic models in business and industry* 19(*4*), pp. 327–347.

Castanier, Bruno, Nicola Esposito, Massimiliano Giorgio, et al. (2020), « Misspecification Analysis of a Gamma-with an Inverse Gaussian-Based Degradation Model in the presence of Measurement Error », *in*: *e-proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*, Research Publishing Services, pp. 1–8.

Castanier, Bruno, Nicola Esposito, Massimiliano Giorgio, Agostino Mele, et al. (2020), « A Perturbed Gamma Process with Random Effect and State-Dependent Error », *in*: *E-proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*, Research Publishing Services, pp. 1–8.

Castanier, Bruno, Antoine Grall, and Christophe Bérenguer (2005), « A condition-based maintenance policy with non-periodic inspections for a two-unit series system », *in*: *Reliability Engineering & System Safety* 87(*1*), pp. 109–120.

Cha, Ji Hwan, Maxim Finkelstein, and Gregory Levitin (2017), « On preventive maintenance of systems with lifetimes dependent on a random shock process », *in*: *Reliability Engineering & System Safety* 168, pp. 90–97.

Cha, Ji Hwan, Maxim Finkelstein, and Gregory Levitin (2021), « Age-replacement policy for items described by stochastic degradation with dependent increments », *in*: *IMA Journal of Management Mathematics* 33(*2*), pp. 273–287.

Cha, Ji Hwan, Maxim Finkelstein, and Gregory Levitin (2022), « Replacement Policy for Heterogeneous Items Subject to Gamma Degradation Processes », *in*: *Methodology and Computing in Applied Probability* 24(*3*), pp. 1323–1340.

Chen, Xudan et al. (2019), « Inverse Gaussian–based model with measurement errors for degradation analysis », *in*: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 233(*6*), pp. 1086–1098.

Cheng, Guo Qing, Bing Hai Zhou, and Ling Li (2016), « Joint optimisation of production rate and preventive maintenance in machining systems », *in*: *International Journal of Production Research* 54(*21*), pp. 6378–6394.

Cheng, Guo Qing, Bing Hai Zhou, and Ling Li (2017), « Joint optimization of lot sizing and condition-based maintenance for multi-component production systems », *in*: *Computers & Industrial Engineering* 110, pp. 538–549.

Cheng, Guo Qing, Bing Hai Zhou, and Ling Li (2018), « Integrated production, quality control and condition-based maintenance for imperfect production systems », *in*: *Reliability Engineering & System Safety* 175, pp. 251–264.

Cholette, Michael E et al. (2019), « Degradation modeling and condition-based maintenance of boiler heat exchangers using gamma processes », *in*: *Reliability Engineering & System Safety* 183, pp. 184–196.

Crowder, Martin and Jerald Lawless (2007), « On a scheme for predictive maintenance », *in*: *European Journal of Operational Research* 176(*3*), pp. 1713–1722.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin (1977), « Maximum likelihood from incomplete data via the EM algorithm », *in*: *Journal of the royal statistical society: series B (methodological)* 39(*1*), pp. 1–22.

Dey, Prasanta Kumar (2004), « Decision support system for inspection and maintenance: a case study of oil pipelines », *in*: *IEEE transactions on engineering management* 51(*1*), pp. 47–56.

Dirkes, Steffen et al. (2023), « Prescriptive Lifetime Management for PEM fuel cell systems in transportation applications, Part I: State of the art and conceptual design », *in*: *Energy Conversion and Management* 277, p. 116598.

Doucet, Arnaud, Adam M Johansen, et al. (2009), « A tutorial on particle filtering and smoothing: Fifteen years later », *in*: *Handbook of nonlinear filtering* 12(*656-704*), p. 3.

El Hajj, Boutros et al. (2023), « Stochastic Multiphasic Multivariate State-Based Degradation and Maintenance Meta-Models for RC Structures Subject to Chloride Ingress », *in*: *Infrastructures* 8(*2*), p. 36.

Ellingwood, Bruce R and Yasuhiro Mori (1993), « Probabilistic methods for condition assessment and life prediction of concrete structures in nuclear power plants », *in*: *Nuclear engineering and design* 142(*2-3*), pp. 155–166.

Elsayed, EA (2021), *Reliability engineering*, John Wiley & Sons.

Elsayed, EA and HT Liao (2004), « A geometric Brownian motion model for field degradation data », *in*: *International Journal of Materials and Product Technology* 20(*1-3*), pp. 51–72.

Esposito, Nicola, Bruno Castanier, and Massimiliano Giorgio (2022a), « A Prescriptive Maintenance Policy for a Gamma Deteriorating Unit », *in*: *32nd European Safety and Reliability Conference (ESREL 2022)*.

Esposito, Nicola, Bruno Castanier, and Massimiliano Giorgio (2022b), « An Adaptive Hybrid Maintenance Policy for a Gamma Deteriorating Unit in The Presence of Random Effect », *in*: *32nd European Safety and Reliability Conference (ESREL 2022)*.

Esposito, Nicola, Bruno Castanier, Massimiliano Giorgio, et al. (2022), « Impact on performances of a condition-based maintenance policy of misspecification of gamma with inverse Gaussian degradation process », *in*: *Proc. of the 8th Intl. Symp. on Reliability Engineering and Risk Management (ISRERM 2022)*, Research Publishing, pp. 800–807.

Esposito, Nicola, Bruno Castanier, and Massimiliano Giorgio (2022c), « Maintenance conditionnelle prescriptive de type bloc pour des systàmes de production sousmis à dégradation », *in*: *32nd European Safety and Reliability Conference (ESREL 2022)*.

Esposito, Nicola, Bruno Castanier, and Massimiliano Giorgio (2022d), « Prescriptive block replacement policy for production degrading systems », *in*: *IFAC-PapersOnLine* 55(*10*), pp. 1974–1979.

Esposito, Nicola, Bruno Castanier, and Massimiliano Giorgio (2023), « An adaptive Prescriptive Maintenance Policy for a Gamma Deteriorating Unit », *in*: *33nd European Safety and Reliability Conference (ESREL 2023)*.

Esposito, Nicola, Agostino Mele, et al. (2021a), « A hybrid maintenance policy for a deteriorating unit in the presence of random effect and measurement error », *in*: *E-proceedings of the 31th European Safety and Reliability Conference*, Research Publishing, pp. 116–123.

Esposito, Nicola, Agostino Mele, et al. (2021b), « Misspecification Analysis of a Gamma-with an Inverse Gaussian-Based Perturbed Degradation Model by Using a New Expectation Maximization Particle Filter Algorithm », *in*: *Proceedings of the 31st European Safety and Reliability Conference*, Research Publishing Services, pp. 681–688.

Esposito, Nicola, Agostino Mele, et al. (2022), « A new gamma degradation process with random effect and state-dependent measurement error », *in*: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X211067299.

Esposito, Nicola, Agostino Mele, et al. (2023a), « A hybrid maintenance policy for a deteriorating unit in the presence of three forms of variability », *in*: *Reliability Engineering & System Safety* 237, p. 109320.

Esposito, Nicola, Agostino Mele, et al. (2023b), « Misspecification analysis of gamma- and inverse Gaussian-based perturbed degradation processes », *in*: *Applied Stochastic Models in Business and Industry*.

Fauriat, William and Enrico Zio (2020), « Optimization of an aperiodic sequential inspection and condition-based maintenance policy driven by value of information », *in*: *Reliability Engineering & System Safety* 204, p. 107133.

Finkelstein, Maxim, Ji Hwan Cha, and Gregory Levitin (2020), « On a new age-replacement policy for items with observed stochastic degradation », *in*: *Quality and Reliability Engineering International* 36(*3*), pp. 1132–1143.

Finkelstein, Maxim, Mahmood Shafiee, and Anselme N Kotchap (2016), « Classical optimal replacement strategies revisited », *in*: *IEEE Transactions on Reliability* 65(*2*), pp. 540–546.

Fouladirad, Mitra and Antoine Grall (2011), « Condition-based maintenance for a system subject to a non-homogeneous wear process with a wear rate transition », *in*: *Reliability Engineering & System Safety* 96(*6*), pp. 611–618.

Gertsbakh, Ilia Borukhovich (2000), *Reliability theory: with applications to preventive maintenance*, Springer Science & Business Media.

Gertsbakh, Ilya and Kh B Kordonskiy (1969), *Models of failure*, Springer-Verlag.

Giorgio, Massimiliano, Maurizio Guida, and Gianpaolo Pulcini (2015), « A new class of Markovian processes for deteriorating units with state dependent increments and covariates », *in*: *IEEE Transactions on Reliability* 64(*2*), pp. 562–578.

Giorgio, Massimiliano, Maurizio Guida, and Gianpaolo Pulcini (2018), « The transformed gamma process for degradation phenomena in presence of unexplained forms of unit-to-unit variability », *in*: *Quality and Reliability Engineering International* 34(*4*), pp. 543–562.

Giorgio, Massimiliano, Agostino Mele, and Gianpaolo Pulcini (2019), « A perturbed gamma degradation process with degradation dependent non-Gaussian measurement errors », *in*: *Applied Stochastic Models in Business and Industry* 35(*2*), pp. 198–210.

Giorgio, Massimiliano and Gianpaolo Pulcini (2018), « A new state-dependent degradation process and related model misidentification problems », *in*: *European Journal of Operational Research* 267(*3*), pp. 1027–1038.

Grall, Antoine, Christophe Bérenguer, and Laurence Dieulle (2002), « A condition-based maintenance policy for stochastically deteriorating systems », *in*: *Reliability Engineering & System Safety* 76(*2*), pp. 167–180.

Hajej, Zied and Nidhal Rezg (2020), « An optimal integrated lot sizing and maintenance strategy for multi-machines system with energy consumption », *in*: *International Journal of Production Research* 58(*14*), pp. 4450–4470.

Hao, Huibing, Chun Su, Chunping Li, et al. (2015), « LED lighting system reliability modeling and inference via random effects gamma process and copula function », *in*: *International Journal of Photoenergy* 2015.

Hao, Songhua, Jun Yang, and Christophe Bérenguer (2019a), « A perturbed inverse Gaussian process model with time varying variance-to-mean ratio », *in*: *ESREL 2019-29th European Safety and Reliability Conference*, pp. 739–745.

Hao, Songhua, Jun Yang, and Christophe Bérenguer (2019b), « Degradation analysis based on an extended inverse Gaussian process model with skew-normal random effects and measurement errors », *in*: *Reliability Engineering & System Safety* 189, pp. 261–270.

Hao, Songhua, Jun Yang, and Christophe Bérenguer (2020), « Condition-based maintenance with imperfect inspections for continuous degradation processes », *in*: *Applied Mathematical Modelling* 86, pp. 311–334.

Hoyle, Edward, Lane P Hughston, and Andrea Macrina (2011), « Lévy random bridges and the modelling of financial information », *in*: *Stochastic Processes and their Applications* 121(*4*), pp. 856–884.

Hu, Jiawen, Jingyuan Shen, and Lijuan Shen (2020), « Opportunistic maintenance for two-component series systems subject to dependent degradation and shock », *in*: *Reliability Engineering & System Safety* 201, p. 106995.

Huynh, Khac Tuan, Anne Barros, et al. (2011), « A periodic inspection and replacement policy for systems subject to competing failure modes due to degradation and traumatic events », *in*: *Reliability Engineering & System Safety* 96(*4*), pp. 497–508.

Huynh, Khac Tuan, Antoine Grall, and Christophe Bérenguer (2019), « A parametric predictive maintenance decision-making framework considering improved system health prognosis precision », *in*: *IEEE Transactions on Reliability* 68(*1*), pp. 375–396.

Iung, Benoît (2019), « De la maintenance prédictive à la maintenance prescriptive: Une évolution nécessaire pour l'industrie du futur », *in*: *Conference on Complexity Analysis of Industrial Systems and Advanced Modeling, CAISAM 2019.*

Jafari, Leila and Viliam Makis (2015), « Joint optimal lot sizing and preventive maintenance policy for a production facility subject to condition monitoring », *in*: *International Journal of Production Economics* 169, pp. 156–168.

Jafari, Leila and Viliam Makis (2016a), « Joint optimization of lot-sizing and maintenance policy for a partially observable two-unit system », *in*: *The International Journal of Advanced Manufacturing Technology* 87, pp. 1621–1639.

Jafari, Leila and Viliam Makis (2016b), « Optimal lot-sizing and maintenance policy for a partially observable production system », *in*: *Computers & Industrial Engineering* 93, pp. 88–98.

Jonge, Bram de, Ruud Teunter, and Tiedo Tinga (2017), « The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance », *in*: *Reliability engineering & system safety* 158, pp. 21–30.

Kahle, Waltraud, Sophie Mercier, and Christian Paroissin (2016), *Degradation processes in reliability*, John Wiley & Sons.

Kallen, Maarten-Jan and Jan M Van Noortwijk (2005), « Optimal maintenance decisions under imperfect inspection », *in*: *Reliability engineering & system safety* 90(*2-3*), pp. 177–185.

Khatab, Abdelhakim et al. (2019), « Integrated production quality and condition-based maintenance optimisation for a stochastically deteriorating manufacturing system », *in*: *International Journal of Production Research* 57(*8*), pp. 2480–2497.

Kim, Seokgoo, Joo-Ho Choi, and Nam Ho Kim (2022), « Inspection schedule for prognostics with uncertainty management », *in*: *Reliability Engineering & System Safety* 222, p. 108391.

Lai, Xinfeng, Zhixiang Chen, and Bopaya Bidanda (2019), « Optimal decision of an economic production quantity model for imperfect manufacturing under hybrid maintenance policy with shortages and partial backlogging », *in*: *International Journal of Production Research* 57(*19*), pp. 6061–6085.

Lawless, Jerald and Martin Crowder (2004), « Covariates and random effects in a gamma process model with application to degradation and failure », *in*: *Lifetime data analysis* 10, pp. 213–227.

Le Son, Khanh, Mitra Fouladirad, and Anne Barros (2016), « Remaining useful lifetime estimation and noisy gamma deterioration process », *in*: *Reliability engineering & system safety* 149, pp. 76–87.

Longhitano, Pedro Dias et al. (2021a), « A closed-loop prescriptive maintenance approach for an usage dependent deteriorating item-Application to a critical vehicle component », *in*: *ESREL 2021-31st European Safety and Reliability Conference*, Research Publishing Services, pp. 2465–2472.

Longhitano, Pedro Dias et al. (2021b), « Proposition of a Generic Decision Framework for Prescriptive Maintenance », *in*: *World Congress on Engineering Asset Management*, Springer, pp. 263–273.

Lu, Dongliang, Mahesh D Pandey, and Wei-Chau Xie (2013), « An efficient method for the estimation of parameters of stochastic gamma process from noisy degradation measurements », *in*: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 227(*4*), pp. 425–433.

Mahmoodian, Mojtaba and Amir Alani (2014), « Modeling deterioration in concrete pipes as a stochastic gamma process for time-dependent reliability analysis », *in*: *Journal of pipeline systems engineering and practice* 5(*1*), p. 04013008.

Majdouline, I et al. (2022), « Integrated production-maintenance strategy considering quality constraints in dry machining », *in*: *International Journal of Production Research* 60(*9*), pp. 2850–2864.

Martinelli, Francesco (2010), « Manufacturing systems with a production dependent failure rate: structure of optimality », *in*: *IEEE Transactions on Automatic Control* 55(*10*), pp. 2401–2406.

Meeker, William Q and Luis A Escobar (1998), « Statistical methods for reliability data », *in*: *A. Wiley Interscience Publications*.

Morita, Lia HM et al. (2021), « Optimal burn-in policy based on a set of cutoff points using mixture inverse Gaussian degradation process and copulas », *in*: *Applied Stochastic Models in Business and Industry* 37(*3*), pp. 612–627.

Nasr, Walid W, Moueen Salameh, and Lama Moussawi-Haidar (2017), « Economic production quantity with maintenance interruptions under random and correlated yields », *in*: *International Journal of Production Research* 55(*16*), pp. 4544–4556.

Nguyen, Khanh TP et al. (2019), « Joint optimization of monitoring quality and replacement decisions in condition-based maintenance », *in*: *Reliability Engineering & System Safety* 189, pp. 177–195.

Oumouni, Mestapha and Franck Schoefs (2021), « A Perturbed Markovian process with state-dependent increments and measurement uncertainty in degradation modeling », *in*: *Computer-Aided Civil and Infrastructure Engineering* 36(*8*), pp. 978–995.

Peng, Chien-Yu (2015), « Inverse Gaussian processes with random effects and explanatory variables for degradation data », *in*: *Technometrics* 57(*1*), pp. 100–111.

Peng, Chien-Yu and Sheng-Tsaing Tseng (2009), « Mis-specification analysis of linear degradation models », *in*: *IEEE Transactions on Reliability* 58(*3*), pp. 444–455.

Peng, Hao and Geert-Jan van Houtum (2016), « Joint optimization of condition-based maintenance and production lot-sizing », *in*: *European Journal of Operational Research* 253(*1*), pp. 94–107.

Pinciroli, Luca, Piero Baraldi, and Enrico Zio (2023), « Maintenance optimization in Industry 4.0 », *in*: *Reliability Engineering & System Safety*, p. 109204.

Piscopo, Antonio et al. (2023), « Remaining useful life estimation of gamma degrading units characterized by a bathtub-shaped degradation rate in the presence of random effect and measurement error », *in*: *ESREL 2023-33rd European Safety and Reliability Conference*, Research Publishing Services.

Pulcini, Gianpaolo (2016), « A perturbed gamma process with statistically dependent measurement errors », *in*: *Reliability Engineering & System Safety* 152, pp. 296–306.

Rodríguez-Picón, Luis Alberto, Anna Patricia Rodríguez-Picón, and Alejandro Alvarado-Iniesta (2019), « Degradation modeling of 2 fatigue-crack growth characteristics based on inverse Gaussian processes: A case study », *in*: *Applied Stochastic Models in Business and Industry* 35(*3*), pp. 504–521.

Ross, Sheldon (1983), *Introduction to Stochastic processes*.

Ruifeng, Chen and Velusamy Subramaniam (2012), « Increasing production rate in Kanban controlled assembly lines through preventive maintenance », *in*: *International Journal of Production Research* 50(*4*), pp. 991–1008.

Si, Xiao-Sheng et al. (2014), « Estimating remaining useful life with three-source variability in degradation modeling », *in*: *IEEE Transactions on Reliability* 63(*1*), pp. 167–190.

Stuart, A and K Ord (1991), « Kendall's advanced theory of statistics, 5th edn., ed », *in*: *E. Arnold* 2.

Sun, Bo et al. (2021), « An improved inverse Gaussian process with random effects and measurement errors for RUL prediction of hydraulic piston pump », *in*: *Measurement* 173, p. 108604.

Tan, Barış (2019), « Production control with price, cost, and demand uncertainty », *in*: *OR Spectrum* 41(*4*), pp. 1057–1085.

Tsai, Chih-Chun, Sheng-Tsaing Tseng, and Narayanaswamy Balakrishnan (2012), « Optimal design for degradation tests based on gamma processes with random effects », *in*: *IEEE Transactions on Reliability* 61(*2*), pp. 604–613.

Tseng, Sheng-Tsaing, Narayanaswamy Balakrishnan, and Chih-Chun Tsai (2009), « Optimal step-stress accelerated degradation test plan for gamma degradation processes », *in*: *IEEE Transactions on Reliability* 58(*4*), pp. 611–618.

Tseng, Sheng-Tsaing and Yu-Cheng Yao (2017), « Misspecification analysis of Gamma with inverse Gaussian degradation processes », *in*: *Statistical Modeling for Degradation Data*, pp. 193–208.

Utanohara, Yoichi and Michio Murase (2019), « Influence of flow velocity and temperature on flow accelerated corrosion rate at an elbow pipe », *in*: *Nuclear Engineering and Design* 342, pp. 20–28.

Van Noortwijk, Jan M (2009), « A survey of the application of gamma processes in maintenance », *in*: *Reliability Engineering & System Safety* 94(*1*), pp. 2–21.

Wang, Hongzhou, Hoang Pham, et al. (2006), *Reliability and optimal maintenance*, vol. 14197, Springer.

Wang, Xiao (2008), « A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects », *in*: *Statistica Sinica*, pp. 1153–1163.

Wang, Xiao and Dihua Xu (2010), « An inverse Gaussian process model for degradation data », *in*: *Technometrics* 52(*2*), pp. 188–197.

Wang, Xiaofei et al. (2021), « Degradation data analysis based on gamma process with random effects », *in*: *European Journal of Operational Research* 292(*3*), pp. 1200–1208.

Wang, Yaping and Hoang Pham (2013), « Maintenance modeling and policies », *in*: *Stochastic Reliability and Maintenance Modeling: Essays in Honor of Professor Shunji Osaki on his 70th Birthday*, Springer, pp. 141–158.

Wasan, MT (1968), « On an inverse Gaussian process », *in*: *Scandinavian Actuarial Journal* 1968(*1-2*), pp. 69–96.

Whitmore, GA (1995), « Estimating degradation by a Wiener diffusion process subject to measurement error », *in*: *Lifetime data analysis* 1, pp. 307–319.

Wu, Zeyu et al. (2018), « A gamma process-based prognostics method for CCT shift of high-power white LEDs », *in*: *IEEE Transactions on Electron Devices* 65(*7*), pp. 2909–2916.

Yang, Yuan-Jian et al. (2014), « Reliability analysis of direct drive electrohydraulic servo valves based on a wear degradation process and individual differences », *in*: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 228(*6*), pp. 621–630.

Ye, Zhi-Sheng and Nan Chen (2014), « The inverse Gaussian process as a degradation model », *in*: *Technometrics* 56(*3*), pp. 302–311.

Ye, Zhi-Sheng and Min Xie (2015), « Stochastic modelling and analysis of degradation for highly reliable products », *in*: *Applied Stochastic Models in Business and Industry* 31(*1*), pp. 16–32.

Yoneda, Kimitoshi et al. (2016), « Development of flow-accelerated corrosion prediction method (1) Acquisition of basic experimental data including low temperature condition », *in*: *Mechanical Engineering Journal* 3(*1*), pp. 15–00232.

Yuan, Xian-Xun, Eishiro Higo, and Mahesh D Pandey (2021), « Estimation of the value of an inspection and maintenance program: A Bayesian gamma process model », *in*: *Reliability Engineering & System Safety* 216, p. 107912.

Zhai, Qingqing and Zhi-Sheng Ye (2017), « Robust degradation analysis with non-Gaussian measurement errors », *in*: *IEEE Transactions on Instrumentation and Measurement* 66(*11*), pp. 2803–2812.

Zhang, Mimi and Matthew Revie (2016), « Model selection with application to gamma process and inverse Gaussian process », *in*.

Zheng, Rui et al. (2021), « Joint optimization of lot sizing and condition-based maintenance for a production system using the proportional hazards model », *in*: *Computers & Industrial Engineering* 154, p. 107157.

Zhou, Rensheng, Nagi Gebraeel, and Nicoleta Serban (2012), « Degradation modeling and monitoring of truncated degradation signals », *in*: *IIE Transactions* 44(*9*), pp. 793–803.

Zied, Hajej, Dellagi Sofiene, and Rezg Nidhal (2011), « Optimal integrated maintenance/production policy for randomly failing systems with variable failure rate », *in*: *International Journal of Production Research* 49(*19*), pp. 5695–5712.

COLLEGE SCIENCES
DOCTORAL DE L'INGENIERIE
PAYS DE LA LOIRE ET DES SYSTEMES

université angers

Titre : Extension du processus gamma et son application en maintenance prescriptive et conditionnelle

Mot clés : processus gamma, erreur de mesure, effect random, maintenance prescriptive, maintenance conditionnelle

Résumé : La maintenance est cruciale dans divers aspects de la vie, car elle permet d'éviter des dommages coûteux et de garantir un fonctionnement optimal. La recherche academique s'efforce d'améliorer l'efficacité de la maintenance, en particulier dans les systèmes complexes. Les modèles traditionnels s'appuient sur des distributions de durée de vie, mais celles-ci peuvent être limitées par la rareté ou l'imprécision des données relatives aux défaillances. Des approches alternatives basées sur des processus stochastiques promettent de surmonter ces limitations. Les processus stochastiques offrent des avantages tels que l'intégration d'informations technologiques et l'utilisation de données de dégradation historiques et en temps réel. Cette activité de recherche vise à proposer des modèles de dégradation stochastiques précis et à les appliquer à des stratégies de maintenance développées ad hoc, en tenant compte d'incertitudes telles que les aléas environnementaux et les erreurs de mesure. En outre, elle se penche sur des paradigmes émergents tels que la maintenance prescriptive, qui équilibre l'utilisation du système et le risque pour une approche holistique de l'optimisation de la maintenance.

Title: Extension of the gamma process and its application in prescriptive and condition-based maintenance under imperfect degradation information.

Keywords: gamma process, measurement error, random effect, prescriptive maintenance, condition-based maintenance

Abstract: Maintenance is crucial in various aspects of life, preventing costly damage and ensuring optimal function. Academic research has strived to enhance maintenance efficiency, especially in complex systems. Traditional models rely on lifetime distributions, but these can be limited by scarce or inaccurate failure data. Alternative approaches based on stochastic processes promise to overcome these limitations. Stochastic processes offer advantages like integrating technological information and utilizing historical and real-time degradation data. This research activity aims to propose accurate stochastic degradation models and apply them to ad-hoc developed maintenance strategies, considering uncertainties like environmental randomness and measurement errors. Additionally, it delves into emerging paradigms like prescriptive maintenance, which balances system utilization and risk for a holistic approach to maintenance optimization.