Evaluating maintenance policies by quantitative modeling and analysis

Enrico Zio^{1,2,*}, Michele Compare²

¹ Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-Electricite' de France, Ecole Centrale Paris and Supelec

² Politecnico di Milano, Italy

* Corresponding author: enrico.zio@ecp.fr, enrico.zio@supelec.fr, enrico.zio@polimi.it

Abstract: The growing importance of maintenance in the evolving industrial scenario and the technological advancements of the recent years have yielded the development of modern maintenance strategies such as the Condition-Based Maintenance (CBM) and the Predictive Maintenance (PrM). In practice, assessing whether these strategies really improve the maintenance performance becomes a fundamental issue. In the present work, this is addressed with reference to an example concerning the stochastic crack growth of a generic mechanical component subject to fatigue degradation. It is shown that modeling and analysis provide information useful for setting a maintenance policy.

1 Introduction

In the last decades, the fast evolution of the industrial scenario has boosted the economic relevance of maintenance in all sectors of industry; the main reasons are:

- the extensive mechanization of industry has reduced the number of production personnel and expanded the capital inventories; this has led to an increment of the portion of the employees working in maintenance and of the maintenance costs; for example, in refineries the maintenance and operations departments are usually the largest (Dekker 1996) and in various sectors maintenance costs constitute a portion of 15% to 70% of the total production costs (Bevilacqua and Braglia 2000).
- The enhancement of the functionality requirements of the systems, linked to the just-in-time production philosophies (which require high availability of the equipment), to the market demand for products of high quality (which calls for production systems maintained and calibrated so to meet the strict tolerance ranges of the products), etc. (Dekker 1996, Dekker and Scarf 1998)

- The outsourcing of maintenance, which has required the clear specification of the maintenance activities beyond the day-by-day routine (Dekker 1996, Dekker and Scarf 1998).
- The increased complexity of the systems and the rising costs of material and labor; i.e., systems are made up of a number of components larger than in the past; these components are more expensive, need to be maintained, and the maintenance actions are also more costly (Pham and Wang 1996).
- The tightening of health and safety legislations in some industries (e.g., air traffic management, aircrafts, nuclear power plants, hospital patient monitoring systems, etc.), which call for maintenance policies capable of guaranteeing that systems fulfill the applicable safety levels during the whole lifetime (Cooke 2003).
- The opening of the energy market, which has forced the producers to be more competitive by reacting promptly and reliably to the demand/offer dynamics, while avoiding the penalties related to the occurrence of service black-out also through more efficient and effective maintenance (Zio 2009a).

Interest in maintenance can be expected to continue increasing in the next future, as the industrial scenario continues to evolve. An example is the development of non-fossil-fuel energy production plants (nuclear, solar, wind, etc.), which is receiving worldwide attention in the last decades (e.g., Amato et al. 2011): maintenance represents a major portion of the total production cost of such technologies, and its optimization can play a role for their competitiveness with respect to fossil-fuel energy production plants.

Given the dimension, complexity and economic relevance of the problem, maintenance must be supported by modeling. In this respect, a huge amount of approaches to maintenance modeling, optimization and management have been propounded in the literature to cope with the maintenance problem, in the evolving technological context. Usually these approaches are divided into two main groups: corrective maintenance (CM) and scheduled maintenance.

Under the CM strategy, the components are operated until failure (events F in Figure 1, top); then, repair or renovation actions are performed (events R in Figure 1, top). This is the oldest approach to maintenance and is nowadays still adopted in some industries, especially for equipment which is neither safety-critical nor crucial for the production performance of the plant, and whose spare parts are easily available and not expensive (Zio & Compare 2011).

Scheduled maintenance policies can be further divided into three groups: Preventive Maintenance (PM), Condition-Based Maintenance (CBM) and Predictive Maintenance (PrM).

Preventive Maintenance (PM) encompasses all actions performed in an attempt to retain an item in specified conditions by providing systematic inspection, detection and prevention of incipient failures (MIL-STD-721C).

The first scientific approaches to PM date back to the 1960's (McCall 1965 and Barlow and Proschan 1965). Since then, a huge number of PM models and optimization methods have been introduced with the aim of reducing failures, for safety reasons, and unplanned downtime, for economic reasons (see Wang 2002 for a survey). For example, the so-called 'age-replacement' models (a very well-known class of PM models, see for example Barlow and Hunter, 1960) consider that a component is preventively maintained at some predetermined age A or repaired at failure, whichever comes first (Figure 1, middle).

In recent years, the relative affordability of on-line monitoring technology has led to a growing interest in new maintenance paradigms such as the CBM and PrM (e.g., Wang 2009, 2011, 2012; Wang et al. 2011, Wang et al. 2012). These are founded on the possibility of monitoring the system to obtain information on its conditions, which is then used to both identify problems at an early stage and predict their evolution in the future. On this basis, a decision is taken on the next maintenance action. This allows a dynamic approach to maintenance based on failure anticipation, aimed at optimizing the equipment lifetime usage. Figure 1, bottom shows that in case of CBM and PrM the maintenance actions are performed upon either the dynamic event D or failure, where D represents the achievement of the safety threshold in case of CBM or the prognosticated failure time in case of PrM.

Figure 1: Synoptic of the maintenance policies

Any company interested in pursuing such maintenance strategies must consider the risks related to the lack of experience and the capital expenditures needed to purchase the necessary instrumentation and software. This requires an evaluation of the opportunity of adopting such advanced maintenance policies founded on specialized knowledge and modern technology. In this sense, evaluating under which conditions and to which extent the CBM and PrM settings can improve the plant performance becomes a fundamental issue. Such evaluation must be made in comparison to the performance of the 'traditional' CM and PM policies.

In the present work, this evaluation analysis is carried out by way of a reference example concerning the stochastic crack growth of a mechanical component subject to fatigue degradation. Different maintenance approaches are applied to show decision makers a way to go for gaining full understanding of the characteristics of the different maintenance policies, and of their benefits.

The remainder of the paper is organized as follows: Section 2 introduces the reference example; the performance of the CM policy is assessed in Section 3, and is compared to that of the PM policy in Section 4. Section 5 and Section 6 apply the CBM and PrM policies, respectively, to the considered component and compare the resulting performances with those of the PM and CM schemes. In particular, the prognostic method which the PrM approach relies on is the Particle Filtering (PF) technique (e.g., Arulampalam et al. 2002). A general discussion is proposed in Section 7, whereas Section 8 concludes the work.

2 Reference example

In this Section, the example concerning the fatigue degradation process of a mechanical component is presented. It constitutes the workbench for the comparison of the different maintenance policies, aimed at identifying the conditions under which one is to be preferred over another.

The reference mission time T does not end with the failure of the component; it is set to $T=10000\text{h}$. As it will be explained below, this entails that the performances of the maintenance strategies are assessed with regards to the transient period, rather than to the steady state.

The crack growth mechanism due to fatigue is a complex physical phenomenon, which has been widely investigated in the literature in several contexts (for example, see Marquis & Solin 1999 and Shigley et al. 2004 for surveys). It initiates from micro defects such as the foundry flaws, subject to oscillating loads that locally strain the component. This leads to the creation of small cracks. Once initiated, the degradation process can propagate under stressful operating conditions, and the depth of the crack can grow up to limits threatening the component structural integrity. In the present case, it is assumed that the component fails when the crack

depth reaches the entire thickness d of the mechanical component. In particular, it is supposed that the value of d is affected by uncertainty, being this situation very common in industrial practice (Baraldi et al. 2011). Such uncertainty is described by a uniform probability distribution of failure crack depth values between 100mm and 110mm.

Several models have been developed to describe the evolution of the fatigue degradation mechanism (e.g., gamma processes (van Noortwijk 2009), Weibull-based discrete processes (e.g., Wormsen and Härkegård 2004), etc. In this work, we adopt the randomized Paris – Erdogan fatigue crack growth model, a popular approach propounded in the literature (e.g., Provan 1987, Kozin and Bogdanoff 1989). The justification of this choice is threefold: on one side, such model has been shown to suitably approximate the degradation evolution of mechanical components subject to fatigue loads (Provan 1987, Kozin and Bogdanoff 1989); on another side, it has been successfully applied in applications to prognostics (e.g., Cadini et al. 2009); finally, the precise modeling of the crack growth process is not the central focus of the present work and, in this view, the simplicity of the model adopted is appealing for the purposes of the paper. Its most widely used expressions is:

$$x_k = x_{k-1} + e^{\xi_k} \cdot C \cdot (\beta \sqrt{x_{k-1}})^n \Delta t \quad (1)$$

where x_k is the crack depth at the (discretized) k^{th} time instant, β , C and n are constant parameters, which can be estimated from measured crack growth data (Myötyri et al. 2006); $\xi_k \approx N(0, \sigma_\xi^2)$, $k = 1, 2, \dots$, are independent and identically distributed random variables, whereas Δt is a sufficiently small time interval (for simplicity, $\Delta t = 1h$).

The model (1) makes x_k , $k = 1, 2, \dots$, a Markov process with independent increments, which are log-normally distributed (Myötyri et al. 2006). The values of the model parameters are taken from (Cadini et al. 2009) and reported in Table 1.

Table 1: parameters of the Paris-Erdogan model.

Model Parameters	Values
C	0.005
β	1
n	1.3
Δt	1 cycle time step

Notice that in the considered setting, the standard deviation (std) σ_ξ of the Gaussian noise is the only variable that determines the amount of uncertainty on the component failure time, being the noise a zero-mean variable. This entails that the performances of the maintenance policies may strongly depend on the value of σ_ξ . In recognition of its importance, σ_ξ is included in the set of the decision parameters. Notice that the decision parameters are those characteristics influencing the final decision about which is the most suitable maintenance policy for the system of interest, independently on whether the decision maker can actually modify their values. On the contrary, the decision variables are those the decision makers usually consider to optimize the maintenance policy (e.g., the inspection interval).

2.1 Structure of the maintenance strategies

This Section presents the maintenance decision framework. First, the aspects related to detection are discussed; then, the durations of the replacement actions and the corresponding costs are detailed. This allows to identify additional ‘decision parameters’, which the performance of the maintenance policies is expected to be sensitive to.

2.1.1 Detection

In this work, it is assumed that the crack depth can be observed only through inspections. For the sake of simplicity, these are performed periodically, being the non-periodic inspection strategies of difficult implementation in an industrial context (Deloux et al. 2009). The length dt of the time interval between two successive inspections is a decision variable, which, in general, may range from $dt \rightarrow 0$ to ∞ ; the former case is generally referred to as continuous monitoring, the latter corresponds to non-monitored components.

Generally speaking, the observation z_k of the crack depth that is acquired at the component inspection is just an estimation of its true value x_k , since errors and imprecision always affect measurements. In this work, the uncertainty about the observations is described by a white Gaussian noise $v \approx N(0, \sigma_v^2)$, which enters the physical law linking z_k to the depth x_k , leading to the following conditional probability density function (pdf, for further details see Cadini et al. 2009):

$$p(z_k | x_k) = \frac{1}{\sqrt{2\pi\sigma_v^2}} \cdot e^{-\frac{\left(\ln\left(\frac{z_k}{d-z_k}\right) - \mu_k\right)^2}{2\sigma_v^2}} \cdot \frac{d}{(d-z_k)z_k} \quad (2)$$

where:

$$\mu_k = \beta_0 + \beta_1 \cdot \ln\left(\frac{x_k}{d - x_k}\right) \quad (3)$$

where β_0 and β_1 are parameters to be estimated from experimental data. In the reference example here investigated, they take the values 0.06 and 1.25, respectively.

Finally, notice that the variance σ_v^2 of the Gaussian noise may influence the performances of the maintenance policies; thus, it is considered among the decision parameters.

2.1.2 Cost and duration of the maintenance actions

The actions performed by the maintenance operators are generally divided into two groups: preventive and corrective, depending on whether they are performed before or after the component failure, respectively. In the present reference example, it is assumed that the mechanical component cannot be fixed, but just replaced. Therefore, only two types of actions are possible:

- Preventive replacement of the component, for which the degraded component is substituted by a new one, before its failure. Such action is supposed to last 100h.
- Corrective replacement, performed after the component failure. Due to the fact that the failure event is unscheduled, this action brings an additional duration of 100h, leading to a total duration of 200h. In particular, the additional time may be caused by the supplementary time needed for performing the procedure of replacement after failure or to the time elapsed between the occurrence of the failure and the start of the replacement actions.

In regard to the costs of the maintenance actions, these are assumed to be proportional to the unavailability of the component. This situation is typical of the plants where the main maintenance costs are related to the business interruption (e.g., solar energy production plants, Amato et al. 2011, Parker 1993). In this setting, the identification of the best maintenance policy, which is in general a multi-objective optimization problem, comes down to the issue of identifying the policy that minimizes the value of the unavailability.

Finally, the identified decision parameters and variables are listed in Table 2 jointly with their best estimates (taken from Cadini et al. 2009), which define the ‘nominal’ condition, i.e., the condition in which the performances of the different strategies are initially assessed. Then, a sensitivity analysis is carried out for

every maintenance strategy, in order to assess the impact of variations of the decision variables and parameters on the component unavailability. In particular, the sensitivity analyses are performed in a crude and local approach (Zio 2009b); that is, the decision variables and parameters are varied one at a time around their nominal values. The sensitivity analyses allow providing the maintenance decision makers with some insights in support of their choices.

Notice that the value of the time interval needed to preventively replace the component does not belong to the set of decision variables. In fact, we are interested in investigating the sensitivity of the performance of the maintenance strategies to the ratio between the durations of corrective and preventive replacement actions, rather than to these single variables. Thus, when sensitivity analyses are performed, the length of the preventive maintenance actions remains fixed, while that of the corrective actions varies.

Table 2: decision variables and parameters considered in the reference example.

Parameters	Nominal Values
Std of the Gaussian noise σ_{ξ}	1.7
Inspection Interval (I)	20h
Std of Measurement Noise σ_{ν}	0.47
Duration of corrective replacement	200h

3 Corrective maintenance

In this Section, CM is applied to the considered reference example. Notice that, there are industries (e.g., nuclear) in which the safety and regulatory requirements do not allow to adopt the CM policy, which invalidates the analysis of its performance. This is not the case of the present case study, which assumes that there is no constraint on the applicability of the CM policy.

Figure 2 (left) shows the values of the unavailability of the component over time, with the related 68.3% confidence intervals, obtained by applying the Monte Carlo (MC) method with 10^3 trials (Marseguerra & Zio 2002); the time step Dt of the bins partitioning the time horizon is 1h. Figure 2 (right) zooms in on the first part of the mission time, and reports the instantaneous unavailability only one out of every ten time instants, in order to improve the understanding of the Figure.

Such behavior of the mean unavailability in the bins can be explained by referring to a population of identical components which are operated until failure, and never undergo any preventive maintenance action. The first

peak of unavailability (about 0.8), in correspondence of $t=850h$, tells us that about 80% of these components fail between $t=650h$ and $t=850h$. In fact, those failed within this interval turn out to be unavailable at $t=850h$, since the duration of the replacement actions is 200h. Components continue to fail in the immediately subsequent part of the mission time, but the corresponding contribution to the unavailability is marginal and heavily counterbalanced by the increase due to the components that are reset into operation after the maintenance actions. This leads to the decreasing behavior observed in the interval $[850h, 1250h]$. In particular, the unavailability reaches zero in correspondence of about $t=1250h$, which means that no component fails after $t=1050h$. A second peak occurs at $t=1790h$, and can be justified analogously to the first one. The fact that its value is smaller (0.65) stems from the different starting situation: although also in this case all components are available at $t=1250h$, they are working since different time instants (i.e., their failure and repair times are different). In other words, they form a less homogeneous population, because their lives are shifted one another, and this leads to more and more smoothed peaks. These considerations also explain the subsequent oscillating behavior of the unavailability.

Figure 2: unavailability over time in case of components that are operated until failure, and then replaced.

On the other side, decision makers may be more interested in the mean value of the unavailability over the mission time. In the nominal case, this falls in the interval $[0.20-1.98e-4, 0.20+1.98e-4]$ with a confidence level of 68.3%.

Figure 3 shows the sensitivity of the mean unavailability to the duration of the replacement actions. As expected, the larger the duration the larger the unavailability. Notice that, this relation is almost linear in the range analyzed. Notice also that the 68.3% confidence intervals are so narrow that they reduce to points in Figure 3. This is due to the combination of small variability of the mean unavailability in this study and the large number (10^3) of MC simulations performed. Thus, the actual value of the mean unavailability corresponding to the different durations of the corrective replacement actions is affected by a small amount of uncertainty. This can be assumed as a general consideration, also valid for the subsequent Figures of the paper.

Figure 3: sensitivity of the CM policy to the duration of the replacement actions.

4 Preventive maintenance

A number of approaches have been propounded in the literature to cope with the issue of setting a PM policy (for example, see Zio & Compare 2012 and Wang 2002 for surveys), among which there is the ‘age-dependent’ model (Wang 2002) presented in the Introduction, which is applied to the reference example by reason of its simplicity.

On the other side, the key issue in any PM model is to determine the optimal maintenance intervals (i.e., the time spans between two successive maintenance actions), which balance the costs of maintenance actions with the production and safety benefits obtained from repairing/substituting a component before it experiences a failure.

In the present work, the optimization of the PM policy, i.e., the identification of the age A that minimizes the unavailability, is performed on a statistical basis. Namely, it is here supposed that a large collection of components' failure times is available, which allows to estimate the corresponding probability density function (pdf, Figure 4), or Cumulative Distribution Function (cdf) and the associated percentiles of interest. In practice, the dataset of failure times has been built by simulating a large number (i.e., 10^4) of stochastic evolutions of the component's crack growth and comparing them with samples of the failure threshold d , drawn from the uniform distribution in $[100\text{mm}, 110\text{mm}]$.

The performance of the PM policy is assessed in correspondence of the distribution percentiles (Figure 5). For example, the 10th percentile, which corresponds to $A=627\text{h}$, defines the time interval between two preventive actions whose length is longer than the lives of 10% of components. The minimum of the mean unavailability ($0.143 \pm 3.30\text{e-}4$) is achieved just in correspondence of the tenth percentile. Shorter intervals steer the PM policy in the direction of inefficiency of component usage, whereas longer intervals push towards non-effectiveness of maintenance. Namely, in the first case the components are poorly exploited, i.e., removed when their crack depths are small; this entails that the number of components needed to cover the mission time is larger, and thus larger is the number of times the components are maintained. In the second case, the preventive replacement actions are belated and therefore rarely performed, since the components fail. In this respect, the last row of Table 3 confirms that the nominal setting of the PM policy is effective in avoiding the components to fail (on average $1.24/(11.76+1.24) \approx 10\%$ of the components fail, as expected), but the components are poorly exploited, since they are removed when the depth of their crack is on average 62.45 mm, still far from the failure threshold.

Notice also that the mean unavailability corresponding to the 90th percentile in Figure 5 approaches the value corresponding to the nominal CM scheme; this proves that if the preventive maintenance actions are rarely performed (every 844h), then they are not capable of preventing the components to fail.

Figure 4: pdf of the components' failure times.

Figure 5: mean unavailability corresponding to different percentiles of the distribution of the failure times.

Table 3: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of the standard deviation of the Gaussian noise entering the degradation process.

	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
0.32	0.25 ± 0.016	2.75 ± 0.016	100.5 ± 0.034
1	0.46 ± 0.021	4.54 ± 0.021	91.74 ± 0.11
1.7	1.24 ± 0.034	11.76 ± 0.034	62.45 ± 0.17

The mean unavailability corresponding to the three different values of σ_ξ (i.e., 0.32, 1 and 1.7) is shown in Figure 6, which refers to the nominal value of t (i.e., the tenth percentile of failure times=627h). Its large reduction in correspondence of smaller uncertainties is the sum of two contributions:

1. Smaller values of the stds lead to larger failure times of the components and thus to a smaller number of maintenance actions performed within the mission time, and consequently to a smaller unavailability. Table 3, columns 2 and 3, show that the number of maintenance actions varies from 3 in correspondence of $\sigma_\xi = 0.32$ to 13, in correspondence of $\sigma_\xi = 1.7$.
2. Smaller uncertainty on the failure time makes the estimation of the generic percentile, and in particular the tenth, closer to the actual value of the failure time (to stress the concept, in a deterministic setting any percentile coincides with the value of the failure time, which is exactly known). Table 3, column 4, shows that the mean crack depths of the components replaced before failure ranges from around 100 in correspondence of $\sigma_\xi = 0.32$ and around 62, in correspondence of $\sigma_\xi = 1.7$.

This second contribution proves that the amount of uncertainty on the failure times influences the performance of the maintenance policy; thus, this aspect should be taken into account when making the final choice of the maintenance scheme.

Figure 6: distributions of the failure times for different values of the std of the noise entering the degradation process (a), and the corresponding values of the mean unavailability.

Figure 7 shows that the mean unavailability over the mission time is quite insensitive to the length of the time interval needed to perform the corrective maintenance actions. In fact, these latter are rarely performed; thus, the resulting increase in unavailability has a small impact on the value of the mean unavailability over the entire mission time.

Figure 7: sensitivity of the PM policy to the duration of the replacements performed after failure.

Finally, Figure 8(left) shows the behavior of the unavailability over the mission time in the nominal PM setting. For visualization, a zoom in the first part of the mission time is reported in Figure 8(right). Notice that the first peak, at 627h, is very close to 1; that is, in the nominal PM setting almost surely the component will be unavailable at $t=627h$. To explain such behavior we can refer, again, to a population of similar components. Only the components, very few (Figure 4), that fail before $t=627h-200h=427h$ will be available at the replacement time, since they are reset into operation before this time instant. The other components either fail after 427h and then are under repair at $t=627h$ or are replaced at such time instant. The behavior of the mean unavailability in the succeeding part of the time axis can be explained analogously to that of Figure 2: the scattering of the crowd due to the shift of the components' lives leads to more smoothed peaks.

Notice that the almost sure unavailability at some time instants in the initial part of the mission time entails that in a plant with a large number of similar components, these become unavailable at the same time if put together into operation. This may be a serious constraint on the applicability of the PM policy in some industrial fields (e.g., energy distribution), where the companies have to guarantee the continuity of the service and fulfill the safety and warranty requirements.

Figure 8: component unavailability over the mission time, in the nominal setting (i.e., $l_l=627h$, $\sigma_\xi=1.7$).

5 Condition-Based Maintenance

From the previous Section, it emerges that the main drawback of the PM policy lies in its tendency to replace components before a complete exploitation of their useful lives, which entails also the danger of imposing actions when unnecessary, interrupting operation and possibly introducing malfunctions due to errors of maintenance operators.

In this Section, the CBM policy is applied to the reference example. In particular, the decision on whether performing a preventive maintenance replacement is based on the comparison of the current degradation level with a safety threshold S_T . Finally, measures of the crack depth are periodically acquired by inspections, and suffer from the uncertainty described in Section 2.

From this, it clearly appears that the decision parameters to be taken into account when assessing the performance of the CBM policy are:

1. the safety threshold S_T . This is initially set to 10% smaller than 100mm, i.e., 90mm.
2. The magnitude of the measurement error ($\sigma_v=0.47$ in the nominal case).

Whereas the decision variable is:

3. The length of inspection interval h ($=20h$ in the nominal case).

Figure 9 shows the behavior of the mean unavailability vs the safety threshold S_T , for different values of the h (i.e., $h=5h$, top left, $h=20h$, top right, $h=50h$, bottom left, and $h=100h$, bottom right), and for different values of the std σ_v of the Gaussian noise entering the measurement equation. Some considerations arise by looking at this Figure:

- Figure 9, top right, shows that the mean unavailability achievable by applying the CBM policy to the mechanical component in the nominal condition is 0.127 ± 2.37 , i.e., smaller than those of the PM and CM policies.
- In the nominal case, the mean unavailability reaches its minimum in correspondence of $S_T=90\text{mm}$. Smaller values of the threshold prevent the policy to fully exploit the components' useful lives, whereas larger values undermine the capability of avoiding failures.
- The value of S_T at which the mean unavailability reaches its minimum decreases as the length of h increases. It is 95mm when $h=5h$, 90mm when $h=20h$, 75mm when $h=50h$ and 60mm when $h=100h$. This is an intuitive result: if the inspections are more rarely performed, then the threshold needs to be set at smaller values to avoid failures, which heavily influences the mean unavailability over the mission time.
- In general, the more frequent the inspections, the larger the promptness in detecting the achievement of the safety threshold, and the smaller the mean unavailability. However, this reasoning does not hold in two cases: for large values of σ_v (e.g., $\sigma_v = 2.23$) and for small values of the safety threshold (e.g., $S_T=60\text{mm}$).
 - In the first case, the measurement errors are very large and lead to wrong decisions. Namely, Figure 10 shows the pdfs of the results of the measurements corresponding to different values of σ_v , when the actual value of the crack depth is 90mm : the larger the value of σ_v , the larger the interval of values that the measurements can take (even the entire range $[0, 100]$ when $\sigma_v = 2.23$). Obviously, underestimations of the crack length result in missed alarms, whereas overestimations lead to false alarms. Then, less frequent inspections on one hand prevent the decision makers from making the mistakes of removing healthy components (false alarms); on the other hand, the components are more prone to fail (missed alarms). In this respect, notice that large values of σ_v lead to an increment of the number of failures only in correspondence of large values of h (Table 4), for any value of S_T . That is, an increment in the

number of measurements performed in the time interval between the achievement of the threshold and the failure instants results in the decrease of the probability of having missed alarms. Notice that for brevity the 68.3% confidence intervals have not been reported in Table 4, this information being irrelevant for the considerations made.

- In the second case, i.e., S_T set to very low values (e.g., 60mm), the delay in measurement acquisition due to large inspection intervals allows a better exploitation of the component useful life.
- Let us consider the cases in which inspections are not frequent (e.g., $I=50h$ or $100h$) and the thresholds are large (e.g., 90-100mm). Then, Figure 9 shows that the larger the value of σ_v (i.e., the poorer the measurement precision), the better the performance (i.e., the smaller the mean unavailability). This counterintuitive result is the consequence of the beneficial effect that the overestimations of the crack depths produce in this particular situation, which is not fully counterbalanced by the negative effect of the underestimations (i.e., failures of the components). In fact, the closeness of the threshold to the failure limit guarantees that underestimations of the cracks' depths do not lead to replacement of under-exploited components, whereas overestimation errors prevent the components to fail.
- There are a number of combinations of the values of the three considered decision variables/parameters that lead the CBM performance to be worse than those of PM and CM. Also in this case, modeling and analysis are shown to be useful for determining whether the application of a given maintenance policy results in an improvement of its performance.

Figure 9: performance of the CBM policy (ordinate, Y) in correspondence of different values of the abscissa $X=S_T$, for different settings of II and σ_v , with the corresponding 68.3% confidence intervals [$Y-L, Y+U$].

Table 4: mean number of failures for different pairs of II and safety threshold values, when $\sigma_v = 2.23$.

II	60mm	65mm	70mm	75mm	80mm	85mm	90mm	95mm	100mm
5h	0	0	0	0	0	0.0030	0.0040	0.0100	0.2330
20h	0,0120	0,0100	0,0200	0,0130	0,0350	0,0690	0,1160	0,3050	1,5670
50h	0,1730	0,2560	0,3720	0,4800	0,6400	0,8960	1,4290	2,1520	4,2360
100h	1,4140	1,6110	1,9560	2,2430	2,6640	3,1770	3,7980	4,8610	6,5390

Concerning the sensitivity of the performance of the CBM policy to the duration of the corrective actions, this analysis is unnecessary. In fact, notice that the mean unavailability corresponding to the CBM policy is smaller than that of the PM policy. This entails that also the number of performed corrective actions is smaller.

Therefore, the previous considerations about the not relevant sensitivity of the maintenance policy to the corrective maintenance durations hold even more in this case.

With regards to the sensitivity of the CBM performance to the standard deviation σ_{ξ} , a decrease of its value reduces the number of maintenance actions performed during the mission time, and improves the capability of preventing failures (Table 5). This second aspect suggests that the amount of uncertainty on the failure times influences the performance of the CBM strategy as well as that of the PM (in particular, the mean unavailability associated to these policies are similar, Figure 11). This is a counterintuitive result, being the CBM performance expected to be dependent on the length of the II and on the uncertainty on the measurements, but not on the uncertainty on the failure times. Indeed, the justification of this result is suggested by the values of the mean crack depths at the preventive replacement times: large times to reach the failure threshold make the CBM policy more sensitive to the measurement errors. In fact, the slower growth of the crack depth allows to perform a larger number of measurements when the crack is approaching the safety threshold. This increases the probability of having false alarms (see also Figure 10), and certainly of replacing the components before their complete exploitation, but, on the other hand, saves the extra-time related to the corrective actions (as it emerges by comparing the first columns of Table 3 and Table 5).

Figure 10: distributions of the results of the measurements for different values of σ_v , when the current value of the crack depth is 90mm.

Table 5: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of the standard deviation of the Gaussian noise entering the degradation process.

σ_ξ	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
0.32	0	$3.01 \pm 3.13\text{e-}3$	78.63 ± 0.32
1	0	$5.04 \pm 6.01\text{e-}3$	80.93 ± 0.26
1.7	0.511 ± 0.021	11.62 ± 0.026	86.38 ± 0.22

Figure 11: sensitivity of the performance of the CBM policy for different values of σ_{ξ} , and its comparison with the PM policy.

Finally, Figure 12(left) shows the behavior of the unavailability over the mission time; this is characterized by the short time needed to achieve the steady state, which is due to the quickness of the ‘cycles’; for example, the unavailability starts to decrease after just 335h, whereas in the CM this happens at $t=850h$. Such rapidity is the consequence of both the tendency for the CBM to fast replace the components and avoid failures, and the consequent small periods after which the components are reset into operation (see Table 5). This also justifies the relatively small values of the peaks. For visualization, the behavior in first part of the mission time is reported in Figure 12(right).

Figure 12: unavailability over the mission time.

6 Predictive maintenance

PrM can be regarded as an improvement of CBM: the knowledge of the current degradation state of the component (diagnostics) is complemented with a prediction of its future behavior and thus of its Remaining Useful Life (RUL) (e.g., Fan et al. 2011, Grall et al. 2002, Lu et al. 2007, You et al. 2010). The accurate estimation of the RUL provides time to opportunely plan and prepare the repair or the replacement of the component, e.g., by delaying the maintenance to the next planned plant outage, by provisioning with spare parts only at time of necessity, by optimizing staff utilization, while remaining acceptably confident that the system will not fail before maintenance. On this basis, a decision is taken on the time of the next maintenance action, in order to optimize the equipment lifetime usage.

The possibility of acquisition and capability of handling system and process information in real time for estimating the component RUL is a necessary condition for performing a predictive maintenance. Then, equipping the system of interest with a prognostic tool with these characteristics is a fundamental issue. In this respect, a number of prognostics approaches are emerging in the literature. These can be divided into two

main categories (Zio 2010): Model-Based approaches (statistical distributions of failure times, Markov models and degradation models), in which a model of the degradation process leading to failure is built and used to predict the evolution of the equipment state in the future and infer its probability of failure and time to failure, and Data-Driven approaches (Pattern Similarity, Artificial Neural Network, Support Vector Machines, etc.), which are developed on the basis of historical data of component behavior.

In the present work, the Particle Filtering technique has been adopted to perform prognostics, since it has proven capable of robustly predicting the future behavior of the distribution that describes the uncertainty on the actual depth of the crack (Cadini et al. 2009). The characteristics of such technique can be found in (Arulampalam et al. 2002). Briefly, PF estimates the pdf of the degradation state of the component (i.e., its crack depth) at any time instant with a set of weighted particles, which constitute a probability mass function (pmf). When a measure of the crack depth is acquired, such pmf is adjusted in a Bayesian perspective (Arulampalam et al. 2002).

In the present work, the prognostic model, based on PF, is embedded within a maintenance scheme as follows:

- Cracks are periodically measured, with period Δ (in the nominal case $\Delta=20\text{h}$).
- When a measure is acquired, the estimation of the distribution of the current degradation state is updated, and the future evolution of such distribution is estimated. This allows to identify the time instant t_{PT} at which the probability of exceeding the safety threshold S_T reaches the value P_T . In analogy with the maintenance schemes previously investigated, P_T is set to 10%.
- The preventive replacement action is performed either when t_{PT} is achieved or at the acquisition of a measurement if the updated distribution leads the component degradation state to exceed S_T with a probability larger than P_T .

For example, Figure 13 shows the evolution of the estimations of $t_{0.1}$ when the failure time is 753h (dashed line) and $\Delta=20\text{h}$. It can be seen that the estimations become more and more accurate with time, and the final value is 744h. At this time instant, the component undertakes a preventive replacement action; thus, failure is avoided, and the component life is widely exploited. Notice also that, the first estimation of $t_{0.1}$, performed at $t=20\text{h}$, is equal to 624h, i.e., very close to the value of the 10th percentile (627h) of the statistics of failure times. As a matter of fact, the first estimation of the time instant at which the probability of having a failure exceeds 0.1, which is not sensitive to the successive Bayesian updates, is expected to be coincident with the time instant before which only 10% of components fail. This also proves that the prediction of the distribution of the failure times made by PF is reasonably precise.

Figure 13: evolution of the estimation of $t_{0,1}$.

In analogy with the previous Sections, the performance of the PrM policy is estimated in correspondence of three different values of σ_ξ . Figure 14 allows to compare the performance of PrM with those of PM and CBM. The first result which stands out, is the closeness of the performances of PrM and PM when the uncertainty on the evolution of the degradation process is small. This becomes even clearer by comparing the first two rows of Table 3 with the corresponding ones of Table 6, which reports the same parameters of Table 3 with the values they take in the PrM setting. Namely, the two strategies lead the component to have the same number of corrective and preventive actions. This is due to the precise knowledge of the behavior of the degradation process, which is available when σ_ξ is small. To further clarify, notice that in case of deterministic evolution of the crack growth, the maintenance time A of the PM scheme and the time $t_{0,1}$ provided by PF would be exactly the same, with consequent equal performances of the maintenance policies. From this, a remarkable conclusion can be drawn on the case study analyzed: if the degradation behavior of the component does not suffer from large variability, then a PrM policy does not lead to a significant improvement over the performance of a PM policy.

Further, notice the difference of the mean crack depths of the preventively replaced components, which reveals that PrM policy allows a fuller exploitation of the components' lives, which work for longer time intervals. This additional working time has not an appreciable effect on the mean unavailability over the considered mission time, since the speed of the crack growth is very high when the crack depth is large; that is,

the components take few hours (in the order of 10h) to reach 101mm from 91mm or 104mm from 100mm, and this shift does not decrease the number of maintenance actions required in the mission time.

Note that in the case considered, CBM is superior to PrM. By a detailed analysis of the results, we see that this is indeed true for some conditions, such as the nominal one, where PrM indeed allows exploiting the components longer (columns 3 of Table 5 and Table 6) but this is paid with a larger number of failures due to missed alarms (columns 1 of Table 5 and Table 6). On the other hand, there are also a number of conditions (e.g., if $l=50h$ or $100h$ and the other decision parameters are set to their nominal values) in which the performance of PrM is superior to that of CBM (compare Figure 15 and Figure 9). Furthermore, in general we conjecture that the full potential of PrM would be exploited in a multi-component system when combined with opportunistic maintenance: in such situation, the knowledge of the RULs of the components can play a fundamental role in the effectiveness of the maintenance policy.

Figure 14: sensitivity of the performance of the PrM policy in correspondence of different values of σ_ξ , and its comparison with both CBM and PM policies.

Table 6: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of σ_{ξ} .

σ_{ξ}	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
0.32	0.32±0.016	2.68±0.016	104.4±0.05
1	0.54±0.021	4.45±0.021	103±0.048
1.7	2.21±0.041	9.22 ±0.048	98±0.047

Figure 15 shows the results of the sensitivity analysis with respect to the updating interval of the prediction of t_{PT} . The performance of the policy worsens in correspondence of more frequent updating. In fact, in such setting the PF allows to deeply exploit the component usage, which is paid with a large number of failures (Table 7). The mean unavailability corresponding to the other values of $//$ remains almost constant, which is the consequence of the reliability of the PF predictions.

Figure 15: sensitivity of PrM to //

Table 7: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of //

//	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
5	4.33±0.050	6.71±0.054	103.33±0.039

20	2.21±0.041	9.22 ±0.048	98±0.047
50	1.68±0.038	10.06±0.042	93.2±0.067
100	1.57±0.036	10.38±0.039	88.1±0.094

Figure 16 shows the sensitivity of the performance of the PrM to the values of the percentiles P_T , which define the time instants t_{pT} at which the components are preventively replaced. It is found that the larger the percentile the larger the corresponding mean unavailability of the component. This behavior is different from that corresponding to the PM policy (Figure 5), where a decreasing behavior of the mean unavailability can be observed in correspondence of small percentiles, followed by an increasing behavior in correspondence of large percentiles. The difference in the first part of the P_T axis can be explained by looking at Table 8; also in the PrM setting, small values of P_T improve the capability of the maintenance policy in avoiding failures, paying the price of a worse exploitation of the components lives. However, this latter aspect is attenuated by the application of the prognostic tool, even for very small values of P_T (i.e., 5); thus, the increase of the mean unavailability in their correspondence is avoided.

Figure 16: sensitivity of PrM to P_T

Table 8: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of P_T .

P_T	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
5	1.34±0.03	10.41±0.0384	94.66±0.05
10	2.21±0.041	9.22 ±0.048	98±0.047
20	4.14±0.049	6.97±0.052	101.54±0.041

50	8.41±0.0395	2.14±0.042	105.06±0.051
----	-------------	------------	--------------

Notice that in Table 8 the mean number of failures constitutes a percentage of the total number of maintenance actions which is larger than P_T . For example, with regards to the first row, $1.34=11.4\%$ of $1.34+10.41$. Indeed, such percentage (e.g., 11%) represents the frequency at which an overestimation of t_{pT} occurs, given that its prediction is performed every 20h. This is not the probability that at t_{pT} the safety threshold S_T is reached, which is equal to P_T .

Finally, Figure 17 and Table 9 show the results of the analysis of the sensitivity of the PrM performance to the uncertainty on measurements; this is quite insensitive to σ_v , in agreement with our expectation about the robustness of the Particle Filtering technique. The small reduction in mean unavailability that can be observed in correspondence of $\sigma_v = 1$ stems from the overestimations caused by measurement errors, which lead to remove the components before failures; this saves the extra-time related to the corrective maintenance actions (Table 9, first column), while allowing a loss in terms of component exploitation (Table 9, last column) which is so small that it does not entail a larger number of preventive maintenance actions performed within the mission time.

Figure 17: sensitivity of PrM to σ_v

Table 9: mean number of corrective and preventive actions, and mean crack depth at the preventive replacement times for different values of σ_v .

σ_v	Mean number of corrective actions	Mean number of preventive actions	Mean crack depth at preventive replacement time
0.22	2.56±0.047	8.81±0.042	99.7±0.039
0.47	2.21±0.041	9.22 ±0.048	98±0.047
1	1.51±0.039	10.24±0.035	94.2±0.067
2.22	2.55±0.0471	9.14±0.043	90.4±0.094

7 General discussion

From the above, it emerges that the choice about the maintenance strategy to be set may depend on the quantity and quality of the required information as well as the complexity of the (software and hardware) tools to be employed. CM policy does not require any information, algorithm and instrumentation; on the contrary, a database of the components' failure times is needed for the application of the PM policy, which also does not involve the adoption of any instrument; the CBM approach is founded on component monitoring and requires more or less refined software and hardware tools, depending on the particular case study; PrM relies on prognostics, which may be performed by means of more or less refined algorithms and technologies, whose development/employment necessarily requires a certain knowledge of the degradation process. In this respect, the task of deriving a model of the degradation process is usually more complicated than just statistically describing the binary transition from a good state to a failed state (Grall et al 2002). Thus, when setting a maintenance strategy the first step is the identification of those policies which are really applicable, given the information and resources available.

In this work, we have compared the different strategies just on the basis of the a-priori evaluation of their performances, not considering the problem of budgetary constraints that could prevent from adopting the advanced maintenance policies and the problem of the information available to set and optimize the maintenance strategy.

The results obtained on the considered reference example, can be summarized as follows:

- The CM performance is better than that corresponding to the other considered strategies when the ratio between the duration of the corrective and preventive maintenance actions is not far from 1, and the other decision parameters take their nominal values.
- PM is to be preferred over the other policies when a small amount of uncertainty affects the degradation process. Indeed, under this condition its performance is similar to those of CBM and PrM, with the advantage of avoiding their complexity.

- In nominal conditions, CBM is more performing than the other policies.
- Finally, the PrM policy based on Particle Filtering is robust and little sensitive to a number of decision parameters. Notice that the full potential of this setting, which is related to the possibility of exploiting maintenance interventions on near components offered by the knowledge of the RUL, has not been investigated in this work.

8 Conclusions

In this work, various maintenance schemes have been considered in their fundamental aspects and thoroughly analyzed in their application to a generic mechanical component subject to fatigue degradation. Table 10 reports the values of the maintenance performance indicator considered, i.e., the component mean unavailability over the mission time with the corresponding 68.3% confidence intervals, in the nominal case (see Table 2) and would lead to the conclusion that the more advanced CBM and PrM policies have better performance. However, we have seen that this is not always true, since in a number of situations the traditional CM and PM are more performing than CBM and PrM. This confirms the fact that modeling and analysis play a key role in supporting maintenance decision makers in the task of identifying the best maintenance policy for the specific component.

Table 10: performance of the different maintenance policies, in the nominal case.

Policy	Mean Unavailability
Corrective Maintenance	0.20±1.98e-4
Preventive Maintenance	0.143 ±3.30e-4
Condition-Based Maintenance	0.127±2.37e-4
Predictive Maintenance	0.137±4.03e-4

9 Reference

- Amato, A., Compare, M., Gallisto, M., Maccari, A., Paganelli, M., Zio, E. 2011. Business interruption and loss of assets risk assessment in support of the design of an innovative concentrating solar power plant. Renewable Energy. Vol. 36 (5), pp. 1558-1568.
- Arulampalam, M. S., Maskell, S., Gordon, N. and Clapp, T. 2002. A tutorial on particle filters for online nonlinear/ non-Gaussian Bayesian tracking. IEEE Trans. Signal Processing. Vol. 50, no. 2, pp. 174-188.

Baraldi, P., Balestrero, A., Compare, M., Zio, E., Benetrix, L., Despujols, A. 2011. A new modeling framework of component degradation. Proceedings of European Safety and Reliability Conference ESREL 2011, 18-22 September, Troyes, France.

Barlow, R.E. and Proschan, F. 1965. Mathematical Theory of Reliability, John Wiley, New York.

Bevilacqua, M. and Braglia, M. 2000. The analytic hierarchy process applied to maintenance strategy selection. *Reliability Engineering and System Safety*. Vol. 70, pp. 71-83.

Cadini, F., Zio, E. Avram, D. 2009. Model-based Monte Carlo state estimation for condition-based component replacement. *Reliability Engineering and System Safety*. Vol. 94 (3), pp. 752-758.

Cooke, F.L. 2003. Plant maintenance strategy: evidence from four British manufacturing firms. *Journal of Quality in Maintenance Engineering*. Vol. 9 (3), pp. 239-249.

Dekker, R. 1996. Application of maintenance optimization: a review and analysis. *Reliability Engineering and System Safety*. Vol. 51, pp. 229-240.

Dekker, R. and Scarf, P.A. 1998. On the impact of optimization models in maintenance decision making: the state of the art. *Reliability Engineering and System Safety*. Vol. 60, pp. 111-119.

Deloux, E., Castanier, B., Bérenguer, C. 2009. Predictive maintenance Policy for a gradually deteriorating system subject to stress. *Reliability Engineering and System Safety*, Vol. 94, pp. 418-431.

Fan, H., Hu, C., Chen, M., Zhou, D. 2011. Cooperative Predictive Maintenance of Repairable Systems With Dependent Failure Modes and Resource Constraint. *IEEE Transactions on Reliability*, Vol. 60 (1), pp. 144-157.

Grall, A., Dieulle, L., Bérenguer, C., Roussignol, M. 2002. Continuous-Time Predictive-Maintenance Scheduling for a Deteriorating System. *IEEE Transactions on Reliability*, Vol. 51 (2), pp. 141-150.

Kozin, F., Bogdanoff, J.L. 1989. Probabilistic models of fatigue crack growth: results and speculations. *Nuclear Engineering and Design*. Vol. 115, pp.143–171.

Lu, S., Tu, Y.C., Lu, H. 2007. Predictive Condition-based Maintenance for Continuously Deteriorating Systems. *Quality and Reliability Engineering International*. Vol. 23, pp. 71-81.

Marquis, V., Solin J. 1999. Fatigue design and Reliability. ESIS, Elsevier.

Marseguerra, M. and Zio, E. 2002. Basics of Monte Carlo method with Application to System Reliability. Life Long Learning – LiLoLe-Verlag.

McCall, J.J. 1965. Maintenance policies for stochastically failing equipment: a survey. *Management Science*, Vol. 11, pp. 493-524.

MIL-STD 721C. 1981. Definitions of terms for Reliability and maintainability. United States - Department of Defense.

Myötyri, E., Pulkkinen, U., Simola, K. 2006. Application of stochastic filtering for lifetime prediction. *Reliability Engineering and System Safety*. Vol. 91 (2), pp. 200-208.

Orchard, M., Kacprzynski, G., Goebel, K., Saha, B. and Vachtsevanos, G. 2008. Advances in Uncertainty Representation and Management for Particle Filtering Applied to Prognostics. International Conference on Prognostics and Health Management.

Park, K.S. 1993. Condition-Based Predictive Maintenance by Multiple Logistic Function. *IEEE Transactions on Reliability*, Vol. 42 (4), pp. 556-560.

Pham, H. and Wang, H. 1996. Imperfect Maintenance. *European Journal of Operational Research*. Vol. 94, pp. 425-438.

Provan, J.W., editor. 1987. Probabilistic fracture mechanics and reliability. Martinus Nijhoff Publishers.

Shigley, J. E., Mischke, C. R., Brown, T. H. 2004. Standard Handbook of Machine Design. 3rd Edition. Mc Graw-Hill.

van Noortwijk, J.M. 2009. A survey of the application of gamma processes in maintenance. *Reliability Engineering and System Safety*. Vol. 94, pp. 2-21.

Wang, H. 2002. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*. Vol. 139, pp. 469-489.

Wang, W. 2009. An inspection model for a process with two types of inspections and repairs. *Reliability Engineering and System Safety*. Vol. 94, pp. 526-533.

Wang, W. 2011, An overview of a semi-stochastic filtering approach for residual life estimation with applications in condition based maintenance. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 225 (2), pp. 188-198.

Wang, W. 2012. A simulation-based multivariate Bayesian control chart for real time condition-based maintenance of complex systems, *European Journal of Operational Research*, Vol. 218, pp. 726-734

Wang, W., Carr, M., Xu W. and Kobbacy, K.A.H. 2011. A model for residual life prediction based on Brownian motion with an adaptive drift. *Microelectronics Reliability*, Vol. 51, pp. 285-293.

Wang, W., Hussin B. and Jefferis, T. 2012. A case study of condition based maintenance modelling based upon the oil analysis data of marine diesel engines, *International Journal of Production Economics*, Vol. 136 (1), pp. 84-92.

Wormsen, A. Härkegård, G. 2004. A statistical investigation of fatigue behaviour according to Weibull's weakest link theory. Proceedings of the 15th European Conference on Fracture, August 11-13, 2004, Stockholm, Sweden. (2004)

You, M.Y., Liu, F., Wang, W., Meng, G. 2010. Statistically planned and Individually Improved Predictive Maintenance Management for Continuously Monitored Degrading Systems. *IEEE Transactions on Reliability*, Vol. 59 (4), pp. 744-753.

Zio, E. 2009a. Reliability Engineering: Old problems and New Challenges. *Reliability Engineering and System Safety*. Vol. 94-2, pp. 125-141.

Zio, E. 2009b. Computational methods for reliability and risk analysis. World Scientific. Singapore.

Zio, E. 2010. Prognostics and health management of industrial equipment. Accepted on *Reliability Engineering and System Safety*.

Zio, E., Compare, M. 2012. A snapshot on maintenance modeling and applications. *Marine Technology and Engineering*. Eds. Guedes Soares, Garbatov, Fonseca, Teixeira. Volume 2, pp. 1413-1425. CRC Press.