# A hybrid condition-based maintenance policy for continuously monitored components with two degradation thresholds*

Joeri Poppe[1], Robert N. Boute[1,2], and Marc R. Lambrecht[1]

[1] *Research Center for Operations Mgt, Faculty of Economics and Business, KU Leuven*
[2] *Technology & Operations Management Area, Vlerick Business School, Belgium*

19th January 2018

**Abstract**

Condition-based maintenance (CBM) makes use of the actual condition of the component to decide when to maintain and/or replace the component, thereby maximising the lifetime of the machine, while minimising the number of service interventions. In this paper we combine CBM on one (monitored) component, with periodic preventive maintenance (PM) and corrective maintenance (CM) on the other components of the same machine/system. We implement two thresholds on the degradation level to decide when to service the monitored component: when the degradation level of the monitored component surpasses a first 'opportunistic' threshold, the monitored component will be serviced together with other components, for instance with a (planned) PM intervention, or upon breakdown of another component, requiring CM. In case none of these opportunities have taken place, and the degradation level surpasses a second 'intervention' threshold, an additional maintenance intervention is planned for the monitored component in order to prevent a failure. Both thresholds are optimised to minimise the total expected maintenance costs of the monitored component, or to minimise the downtime of the machine due to maintenance on the monitored component. We perform an extensive numerical experiment to demonstrate the potential gains of this hybrid policy with two thresholds compared to using a traditional PM policy, and we identify its key drivers of performance. We also benchmark our results when only one threshold is implemented. Our model is validated and applied at an OEM in the compressed air and generator industry.

**Keywords:** maintenance, condition-based maintenance, servitisation

---

## 1 Introduction and Literature Review

After-sales maintenance services are an effective way to keep manufacturing systems in good condition. The maintenance policies that are most frequently encountered, both in academia and in practice, are corrective maintenance, periodic maintenance, and condition-based maintenance; the latter two being preventive maintenance policies. For an overview of maintenance policies, we refer to for instance, Goossens and Basten [6], Tinga [16, 17], Van Horenbeek et al. [19], Wang [21].

*To appear in *European Journal of Operational Research*

Under a corrective maintenance (CM) policy, a part or component is replaced when it does not fulfil its intended purpose (properly) anymore, defined as a *failure*, which may lead to a reduced output or a breakdown of the machine. We define periodic maintenance (PM) as a type of usage-based preventive maintenance, where the total usage of a part is measured and maintenance is conducted after a certain amount of usage (e.g., running hours). Since the usage of equipment is usually scheduled, the moment that maintenance is performed can also be scheduled. If there is a large set-up cost associated with maintenance, it can be beneficial to interchange several parts simultaneously, which is known as block replacement. The period between two PM interventions is known as the maintenance interval (defined by a number of running hours of the machine). During this maintenance interval, the component can fail (due to its stochastic failure behaviour), at which an emergency CM intervention is performed. The optimisation of the maintenance interval is a trade-off between the cost of planned (preventive) and unplanned (corrective) maintenance interventions. Optimisation of a PM policy has first been treated by Barlow and Hunter [3]; nowadays, any standard textbook on reliability engineering or maintenance optimisation treats this topic [e.g., 10, 15].

Over the past years, many industrial companies have adopted CM and PM in their maintenance operations. Whereas these maintenance policies have shown to perform well, they actually do not (or at least not explicitly) take the actual state (or level of deterioration) of the components into account. Recently, monitoring degradation levels has become cheaper, and big data analytics are progressing rapidly, which led to condition-based maintenance (CBM) gaining interest to both practitioners as well as researchers. CBM is a maintenance strategy that performs maintenance based on the actual condition (degradation level) of the component. As such, CBM may improve the efficiency of the maintenance strategy, implementing a *dynamic* preventive maintenance strategy that is able to reduce redundant (too early) maintenance interventions, while at the same time reducing the number of unexpected failures (too late). This could lead to lower maintenance costs and higher availability of the machines. We refer to Jardine et al. [9], Peng et al. [14], and Ahmad and Kamaruddin [1] for reviews on CBM and its potential benefits.

In the literature, there are various ways to monitor the degradation process, and various ways to use these monitored data in the design of the maintenance policy. The most basic interpretation of condition monitoring is to assume that failures can be foreseen in advance [e.g., as in 11, 12]. In most cases, however, degradation is defined by a discrete number of qualitative (functional) degradation states [5, 8, 18], a set of quantitative degradation states [13, 22, 23], or the exact (continuous) degradation value [2, 7, 24, 25]. These monitored degradation data are then used to decide when the component should be replaced, typically by introducing thresholds on the degradation level. The most commonly used threshold is a 'control limit', to be interpreted as an *intervention* threshold: when the degradation of the component surpasses this level, a maintenance intervention is scheduled to prevent breakdown of the machine. Marseguerra et al. [13] and Barata et al. [2] apply Monte Carlo simulation to optimise the intervention threshold, where both the monitored degradation states and deterioration values have discrete values.

Initial research on CBM has primarily focused on single-component systems. Whereas this has proven to be useful to understand the potential benefits of CBM, most practical settings involve multiple components. In a multi-component setting, additional benefits can be reaped by bundling maintenance interventions, for instance by combining a PM or CM intervention with a replacement of a monitored component that is close to failure. This is known as *opportunistic* maintenance (OM). OM may reduce the setup and/or travel costs that are needed to perform maintenance. This can be enabled by defining an opportunistic threshold on the degradation level of the monitored component, which indicates that the component is eligible for maintenance when there is an opportunity to combine it with maintenance on another component.

Koochaki et al. [11, 12] implement an opportunistic threshold on three identical components, which is defined as a time window before the exact moment of failure, assuming failures can be foreseen in advance. de Jonge et al. [5] subdivide identical units into three functional degradation states (normal, alert and alarm): if the alarm state is reached, maintenance will be performed after a predefined time period, and all other components in the alert or alarm state will be maintained simultaneously. Van Der Duyn Schouten and Vanneste [18] categorise the functional degradation states of a system of identical components (good, doubtful, PM is due, failed); although they do not explicitly implement a threshold, they determine the number of components that should be in the 'doubtful' state to replace all components, in case one of them breaks down. Gürler and Kaya [8] extend their model with more degradation states. Zhu et al. [26] consider a multi-component system, in which all (non-identical) components are monitored. They propose a joint periodic maintenance interval, in which all components that have passed the opportunistic threshold are replaced. They consider soft failures and focus on reducing the high setup costs of maintenance. Zhu et al. [25] monitor the degradation level of one component and present a hybrid opportunistic maintenance policy that implements CBM on one critical component, while implementing a CM and PM policy on the other components of the machine. The CM and PM interventions on the non-monitored components generate opportunities to maintain the monitored component. Through renewal theory, the opportunistic threshold is optimised to minimise total maintenance costs.

The model that we present in this article is similar to and inspired by the model presented in Zhu et al. [25]. However, there are two key differences. First, whereas Zhu et al. [25] assume that the block replacement policy is rescheduled every time a maintenance intervention is performed on the monitored component, in our paper we assume that the block replacement interventions remain fixed over time. The rationale behind this assumption is that these PM interventions are scheduled to maintain the other (non-monitored) components and are thus independent of the interventions of the monitored component. As a result of this assumption, we can no longer rely on the same methodology as Zhu et al. [25]. Second, whereas Zhu et al. [25] consider one (opportunistic) threshold, we introduce a second (intervention) threshold on top of the opportunistic threshold. The second intervention threshold triggers a maintenance intervention on the monitored component when it becomes too risky to wait for an opportunity to bundle its maintenance, as the degradation level announces an imminent failure. The introduction of the intervention threshold trades off the cost of an extra maintenance intervention with the cost upon failure of the component. As we will show later in this paper, especially in environments where the breakdown cost is high, the benefit of having this second intervention threshold can be significant.

Our paper is not the first to introduce multiple thresholds. Wang et al. [22] also define both an opportunistic and intervention threshold. They model the deterioration process using continuous-time Markov chains, and optimise the threshold values through Monte Carlo simulation. Wijnmalen and Hontelez [23] define three thresholds: the first threshold allows joint maintenance with a component of the same type (saving on both the major and minor set-up cost), the second with a component of a different type (saving on the major set-up cost) and the last forces mandatory repair. They decompose the multi-component problem into single-component Markov decision problems to make it tractable and determine a stationary coordinated repair strategy for each component. Our paper differs from Wang et al. [22] and Wijnmalen and Hontelez [23] as we consider a hybrid policy, combining CBM on one monitored component with PM and CM on the other components of the same system. Furthermore, whereas these papers assume a discrete number of degradation states (allowing the use of Markov decision processes), we assume continuous degradation, which is appropriate in many settings.

We analyse a multi-component system (e.g., a compressor) and propose a hybrid opportunistic maintenance policy that combines condition-based maintenance on one monitored component with periodic preventive and corrective maintenance on the other components of the same machine. A key feature of such a hybrid policy is that allows a gradual implementation of CBM on one or a few monitored components, while keeping the current PM strategy on the rest of the machine. We treat the service interventions on the non-monitored components as exogenous to our policy, but we use them as opportunities to service the monitored component. In other words, we model a multi-component system, but our goal is to minimise the maintenance costs of the monitored component. A more detailed description of our model assumptions and objective function is given in Section 2.

Our maintenance policy works with an opportunistic and an intervention threshold on the degradation levels of the monitored component. Whereas the opportunistic threshold allows the monitored component to be jointly maintained with the non-monitored components, the intervention threshold triggers a maintenance intervention on the monitored component when no bundling opportunity could be seized and a failure is imminent. We provide analytical expressions to evaluate the expected maintenance costs of the monitored component and the machine downtime due to maintenance (or breakdown) of the monitored component, and determine the optimal threshold values. We also study the impact of each of the thresholds on the maintenance costs under CBM.

The model that we propose in this article has been validated and applied at an OEM (Original Equipment Manufacturer) in the compressed air, generator and pump industry that seeks to leverage this type of opportunistic maintenance. The OEM deploys a PM strategy on the majority of its components and implements CBM on a selected number of components.

The remainder of this paper is organised as follows. In the next section, we describe our model assumptions and introduce notations. Section 3 presents the expressions needed to evaluate the performance of our proposed hybrid policy with two thresholds. In Section 4 we evaluate these expressions when only one of the two thresholds is used. This not only serves as a benchmark for our policy with two thresholds, it also allows to identify the economic impact of each threshold. Section 5 presents the results of a numerical experiment to determine the gains of the proposed policy, and identify its key drivers. Finally, Section 6 discusses an illustrative example how the thresholds can be determined when multiple objectives are considered. Section 7 concludes.

# 2 Model description

Our model assumptions are inspired by the OEM with whom we collaborate. Table 1 summarises the notations used throughout the paper. We consider a system with a large number of components. The OEM pursues a periodic, usage-based, preventive maintenance policy on its installed base. Such usage-based periodic maintenance based on effective running hours is common for instance for rotating equipment, which degrades based on its effective use. Obviously, many components are not preventively maintained at all, but only upon failure. The components that are periodically preventively maintained, each have their own maintenance interval, dependent on its degradation rate and the costs related to a periodic or corrective maintenance intervention. The block-replacement policy operates with a 'base' periodic maintenance interval, which we denote by $\tau$, which is the minimum interval at which some components are always replaced (e.g., air filter), and each component's maintenance interval is a multiple of this base maintenance interval. The reason why each component's maintenance interval is a multiple of the base interval is that this allows combining a component-specific maintenance interval (based on its degradation rate), with the simultaneous replacement of some components due to the set-up cost associated with maintenance (block-replacement). That way, fast deteriorating consumables, such as filters, may have a higher frequency of periodic maintenance than slower deteriorating components that are only replaced at a multiple of this maintenance interval. The analogy with car maintenance can be made: some components are replaced every 15,000 km, whereas others only every 30,000 or 90,000 km. The minimum 'base' interval is determined based on the fast deteriorating consumables. In this paper we will treat the base PM interval $\tau$ as an exogenous constant parameter and assume that the PM schedule remains fixed over time, regardless of any maintenance interventions on other components.

Table 1: Nomenclature

| Notation | Description |
|---|---|
| $\tau$ | Base periodic maintenance interval |
| $\xi$ | Opportunistic threshold |
| $\varepsilon$ | Intervention threshold |
| $\lambda^{-1}$ | Average time between (random) machine failure, given a certain PM policy |
| $t_{\mathrm{p}}$ | Preparation time to schedule an extra maintenance intervention |
| $L_{\mathrm{c}}$ | Maintenance cycle length |
| $g(u)$ | Pdf of time $u$ until the next PM visit |
| $h(v)$ | Pdf of time $v$ until the next random machine failure |
| $\eta(t)$ | Degradation function of the monitored component |
| $f(l)$ | Pdf of lifetime $l$ of the monitored component |
| $f^1(l_\xi)$ | Pdf of degradation time between replacement and degradation level $\xi$ |
| $f^2(l_{\xi,\varepsilon})$ | Pdf of degradation time between degradation level $\xi$ and $\varepsilon$ |
| $f^3(l_{\varepsilon,\mathrm{F}})$ | Pdf of degradation time between degradation level $\varepsilon$ and failure |

When a component is maintained under CBM, that component is no longer included in the preventive block replacement policy. Instead, we continuously monitor the degradation level of that component and decide upon its maintenance based on the monitored degradation levels. The benefit of this hybrid policy is that only a limited number of components need to be continuously monitored, while the conventional block replacement policy can still be applied for the other components, and it allows a gradual implementation of CBM in the existing maintenance operations. Inspired by Zhu et al. [25], we leverage this hybrid strategy by using opportunities to bundle maintenance interventions on the monitored component with the PM and CM interventions on the other, non-monitored components. This allows reducing the number of interventions and its corresponding costs.

We denote the stochastic degradation behaviour of the monitored component by $\eta(t)$, and assume it obeys physical laws that are known – e.g. due to fatigue and corrosion – and thus can be remotely monitored [2]. For some type of environments – e.g., if the degradation level depends on the combination of many deterioration processes – this assumption may not hold and physical inspections are necessary to know the condition of the equipment. We assume that the stochastic degradation behaviour is independent of how the monitored component was maintained, as well as independent of the maintenance interventions performed on the other non-monitored components. We express $\eta(t)$ as a percentage of the critical degradation level at which the component fails: upon replacement, a component is considered as-good-as-new ($\eta(0) = 0$), and a failure at the end of its lifetime $l$ is defined as $\eta(l) = 1$.

To maintain and replace the monitored component, we can bundle its service with a service intervention of the other components, when that leads to a reduction of the setup costs. We distinguish two types of opportunities. First, when a PM intervention is scheduled (for other, non-monitored components), the technician may as well replace the monitored component. We denote this opportunity by $\mathrm{O}_{\mathrm{PM}}$. As the monitored component is no longer included in the PM policy, the schedule of the PM interventions is not based on the monitored component, but it is set for the other (non-monitored) components. Given that the degradation of the monitored component is independent of how it was maintained (i.e., by a PM intervention or any other type of intervention), this means that the degradation of the monitored component is largely independent of the PM interventions on the other, non-monitored components in the system. Hence, at the moment that the degradation level reaches the opportunistic threshold value (which is a random observation moment due to the stochastic degradation behaviour), the time until the next planned PM intervention is *approximately* uniformly distributed between then and $\tau$ time later. In Section 5.2 we provide a simulation experiment to validate this assumption that the time until the next planned PM intervention is approximately uniformly distributed. When the probability density function (pdf) of the time $u$ until the next planned PM intervention can be approximated by a uniform distribution, we have:

$$g(u) = \begin{cases} \frac{1}{\tau} & \text{if } 0 \leq u \leq \tau, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

A second opportunity to bundle service of the monitored component occurs when the machine fails (due to failures of other components), and requires CM. This opportunity is denoted by $\mathrm{O}_{\mathrm{CM}}$. When observing the entire multi-component system, many components are not maintained periodically and those that are periodically maintained, do not all have the same maintenance interval. In addition, there could be external (unpredictable) factors that could lead to failure of the component. According to the Palm-Khintchine theorem [4], the failure rate of a machine with a large number of components can then be modelled by a memoryless exponential function. Indeed, when we observe one component in isolation, it is inappropriate to model the time until the next system failure by an exponential distribution. However, the combination of a number of non-Poisson renewal processes will still have Poisson properties. We refer to Section 5.2 for a validation of this assumption. The pdf of the time $v$ until the next random machine failure can then be defined as:

$$h(v) = \lambda e^{-\lambda v}, \quad (2)$$

with $\lambda^{-1}$ the average time between two failures, which also depends on the executed PM policy.

To decide when to use these opportunities, we introduce an *opportunistic* threshold on the degradation level, which we denote by $\xi$, expressed as a percentage of the critical degradation level at which the component fails. Once the condition (level of degradation) $\eta(t)$ of the monitored component surpasses this threshold value $\xi$, any maintenance visit for other components (i.e., a CM or scheduled PM) will be used as an opportunity to replace the monitored component (respectively $\mathrm{O}_{\mathrm{CM}}$ or $\mathrm{O}_{\mathrm{PM}}$). If one of these opportunities is used to replace the monitored component, the total number of replacements of the component is higher compared to when only replacing as late as possible. We thus have a trade-off between a reduction of the component's lifetime and the reduction of a setup cost, and potential failure. There is no additional planning time required for the maintenance on the monitored component, as the service technician is scheduled to go on site anyhow for the CM or PM intervention.

We also introduce a second *intervention* threshold, denoted by $\varepsilon$ (also expressed as a percentage of the critical degradation level). In case there has not been any opportunity to bundle maintenance, but the monitored component is close to failure, this intervention threshold enables to schedule an *extra* maintenance intervention for this component (which we denote by E) in order to avoid its failure. At this point, the risk of waiting for an opportunistic maintenance is too high as the degradation level announces an imminent failure. In this case, we incur a preparation time $t_{\mathrm{p}}$ to schedule this extra maintenance intervention in a cost-effective way. When an opportunity to bundle the intervention with a PM or CM would occur during this planning period $t_{\mathrm{p}}$, the extra maintenance visit will be cancelled. Due to the random degradation behaviour, the monitored component can still fail during this planning period. We denote the corrective maintenance intervention upon failure of the monitored component by F. Recall that the preparation period $t_{\mathrm{p}}$ is not incurred in case of opportunistic maintenance; in that case the technician is on site and when it performs some extra maintenance tasks for the monitored component, no additional planning time is incurred. In Figure 1, we summarise the hybrid maintenance policy.

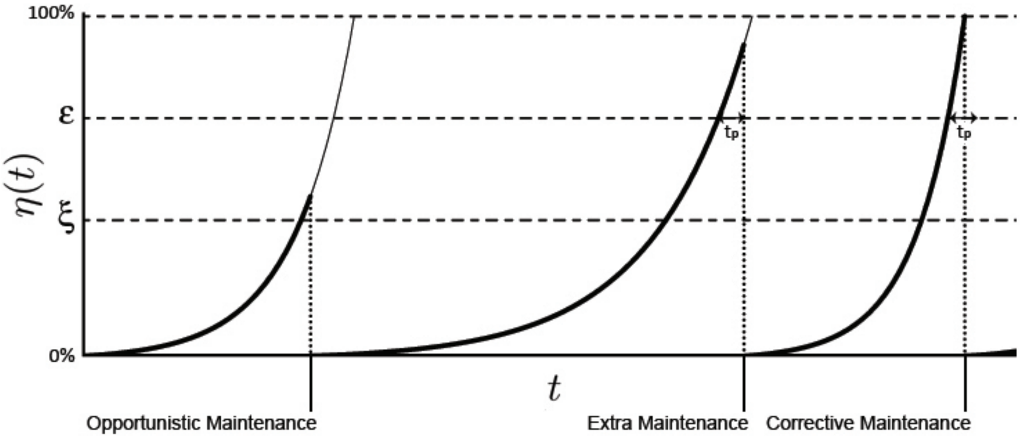


Figure 1: A hybrid condition-based maintenance policy with an opportunistic and intervention threshold

The value of the pair of thresholds $(\xi, \varepsilon)$ determines the likelihood that the monitored component will be maintained by one of the four types of service maintenance: $\mathrm{O_{PM}}$, $\mathrm{O_{CM}}$, E, and F. Let $\mathbb{P}(\mathrm{O_{PM}})$, $\mathbb{P}(\mathrm{O_{CM}})$, $\mathbb{P}(\mathrm{E})$, and $\mathbb{P}(\mathrm{F})$ denote the respective probabilities that the monitored component will be maintained by that type of maintenance intervention in a maintenance cycle. Low values of $(\xi, \varepsilon)$ generally induce many (potentially too early) replacements on the monitored component (i.e., higher values of $\mathbb{P}(\mathrm{O_{PM}})$, $\mathbb{P}(\mathrm{O_{CM}})$ and $\mathbb{P}(\mathrm{E})$), whereas high values of $(\xi, \varepsilon)$ tend to increase the risk of failure as the replacement may be too late (i.e., higher values of $\mathbb{P}(\mathrm{F})$). Also, a lower value of the opportunistic threshold $\xi$ allows for a higher intervention threshold $\varepsilon$ to prevent the same number of failures. In the next section we will quantify these probabilities.

Let $C_i$ with $i \in \{\mathrm{O}, \mathrm{E}, \mathrm{F}\}$ denote the cost to service the monitored component upon an opportunistic maintenance (O), an extra maintenance (E) or upon failure (F). The opportunistic maintenance cost consists of the cost to replace the monitored component (i.e., the cost of the component itself and the wrenching time). This cost is independent whether the replacement occurs together with a PM or CM intervention on another component. The cost of an extra intervention consists of an additional travel or setup cost, on top of the cost to replace the monitored component, as in this case an additional intervention is scheduled for the monitored component. We define this cost as $C_{\mathrm{E}} = (1 + \rho_{\mathrm{E}})C_{\mathrm{O}}$, with $\rho_{\mathrm{E}}$ the relative cost increase of an extra maintenance compared to an opportunistic maintenance that is due to the additional travel or setup. Finally, the maintenance cost upon failure consists of the replacement cost and the travel cost, and additionally the costs related to the breakdown of the component, e.g., reduced output or consequential damage. We define $C_\text{F} = (1 + \rho_\text{E} + \rho_\text{F})C_\text{O}$, with $\rho_\text{F}$ the percentage cost increase that is related to the costs due to the breakdown. In Section 5 we report how the optimal threshold values and the maintenance costs are driven by the values of $\rho_\text{E}$ and $\rho_\text{F}$.

Observe that we could also let $C_i$ denote the *downtime* to service the monitored component upon each of the four possible maintenance interventions, rather than its cost. The analysis for cost or downtime is generally the same, but it allows us to analyse both separately. In Section 6 we minimise both cost and downtime in a bi-objective function.

We determine the optimal threshold values that minimise the long run average maintenance costs per running hour for the monitored component, which is found by dividing the expected costs per maintenance cycle over the expected length of a maintenance cycle, $\mathbb{E}\left\{L_\text{c}\right\}$, expressed in running hours. A maintenance cycle is defined as the time between two replacements of the monitored component; it depends on the degradation process as well as the type of maintenance on the monitored component in each cycle. We can split the length of a maintenance cycle into two parts: first, the time from replacement until the degradation level reaches the opportunistic threshold $\xi$, denoted by the stochastic variable $L_\xi$, which is independent of the type of maintenance; and second, the degradation time from the opportunistic threshold $\xi$ to performing one of the four types of maintenance, which we denote by $L_{\xi,(\cdot)}$, which clearly does depend on the maintenance type. This results in the following objective function:

$$\begin{aligned} &\underset{\xi,\varepsilon}{\text{minimise}} && \frac{\mathbb{P}(\text{O}_\text{PM})\cdot C_\text{O} + \mathbb{P}(\text{O}_\text{CM})\cdot C_\text{O} + \mathbb{P}(\text{E})\cdot C_\text{E} + \mathbb{P}(\text{F})\cdot C_\text{F}}{\mathbb{E}\left\{L_\text{c}\right\}}, \\ &\text{subject to} && 0 \le \xi \le \varepsilon \le 1, \end{aligned} \tag{3}$$

with

$$\begin{aligned} \mathbb{E}\left\{L_\text{c}\right\} = {} & \mathbb{E}\left\{L_\xi\right\} + \mathbb{P}(\text{O}_\text{PM})\cdot\mathbb{E}\left\{L_{\xi,\text{O}_\text{PM}}\right\} + \mathbb{P}(\text{O}_\text{CM})\cdot\mathbb{E}\left\{L_{\xi,\text{O}_\text{CM}}\right\} + \mathbb{P}(\text{E})\cdot\mathbb{E}\left\{L_{\xi,\text{E}}\right\} \\ & + \mathbb{P}(\text{F})\cdot\mathbb{E}\left\{L_{\xi,\text{F}}\right\}. \end{aligned}$$

The pair of threshold values $(\xi, \varepsilon)$ determines the probabilities $\mathbb{P}(\cdot)$ how the monitored component is maintained, as well as its expected maintenance length $\mathbb{E}\left\{L_\text{c}\right\}$. In the next section we derive expressions for these probabilities.

# 3 Probabilities of each type of maintenance in a cycle

We determine the probabilities of each type of maintenance for the monitored component in a maintenance cycle (i.e., an opportunistic maintenance with a PM or CM intervention on one of the non-monitored components, an extra maintenance, or upon failure) as follows. Define the lifetime $l$ of the monitored component by the time to degrade from $\eta(0) = 0$ (as-good-as-new) to $\eta(l) = 1$ (failure). In our case-study company, the degradation levels $\eta(t)$ are obtained through sensors installed in the equipment. The degradation process can be any random process, as long as it is monotonic, e.g., a Gamma process, compound Poisson process, inverse Gaussian process, etc. As a result of this degradation process, the lifetime $L$ is a random variable; we denote its probability density function by $f(l)$. The time to failure can be defined by any distribution function. Assuming the degradation function to be monotonic, we can divide the lifetime $L$ of the monitored component into three parts,

$$L = L_{\xi} + L_{\xi,\varepsilon} + L_{\varepsilon,\mathrm{F}}, \tag{4}$$

where $L_{\xi}$ defines the time to degrade from replacement to the opportunistic threshold $\xi$; $L_{\xi,\varepsilon}$ the degradation time from the opportunistic threshold $\xi$ to the intervention threshold $\varepsilon$ (in this time window, opportunities can be seized to bundle maintenance); and $L_{\varepsilon,\mathrm{F}}$ the degradation time from the intervention threshold $\varepsilon$ to failure of the monitored component (in this time window, an extra maintenance intervention can prevent failure if $l_{\varepsilon,\mathrm{F}} \geq t_{\mathrm{p}}$, while opportunities can still be seized to bundle maintenance if they occur before the extra maintenance intervention is executed). Denote $f^1(l_{\xi})$, $f^2(l_{\xi,\varepsilon})$ and $f^3(l_{\varepsilon,\mathrm{F}})$ their respective pdf.

The probabilities $f^1(l_{\xi})$, $f^2(l_{\xi,\varepsilon})$, $f^3(l_{\varepsilon,\mathrm{F}})$ and $f(l)$ can be derived as follows. Either we assume a completely random degradation process, where the probabilities $f^1(l_{\xi})$, $f^2(l_{\xi,\varepsilon})$, and $f^3(l_{\varepsilon,\mathrm{F}})$ are independent of each other. Based on the stochastic deterioration process the distribution of the time until failure can be determined. Likewise, the distribution of the degradation time in each of the three phases of the lifetime can be found for any given set of thresholds. These distributions change when the thresholds are adjusted. For some specific degradation processes like for instance the Gamma process, the degradation time can be analytically characterized [20].

An alternative approach is to assume a pre-specified distribution of the time to failure and a tractable degradation process. This indicates that the degradation is not completely random (i.e., there are several root causes that may lead to degradation) and implies a dependency between $f^1(l_{\xi})$, $f^2(l_{\xi,\varepsilon})$ and $f^3(l_{\varepsilon,\mathrm{F}})$. Indeed, suppose that the time to failure follows a pre-specified distribution, then the pdf of the time between the second threshold until failure is dependent on the time to reach the second threshold (and likewise the time to reach the first threshold) in order for the failure times to follow that pre-specified distribution. This is captured in the dependence of the second probability distribution with respect to the outcome of the first. Note however, that this does not preclude that a component first degrades slowly and then accelerates to failure. As the time until failure is random, this changing degradation behaviour remains possible in our model. We denote $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})$ the conditional pdf of $L_{\varepsilon,\mathrm{F}}$ on $l_{\xi,\varepsilon}$. As we only need the expected value $\mathbb{E}\{L_{\xi}\}$ in our analysis (rather than $f^1(l_{\xi})$), there is no explicit need to denote the conditional pdf of $L_{\xi,\varepsilon}$ and $L_{\varepsilon,\mathrm{F}}$ on $l_{\xi}$.

First, we define the probability that the monitored component is maintained through an opportunistic maintenance with a PM visit that is scheduled for other components ($\mathrm{O}_{\mathrm{PM}}$). The probability that this happens, corresponds to the situation that the degradation level surpassed the first opportunistic threshold value $\xi$ (which implies that a timespan longer than $l_{\xi}$ has passed since replacement), but the monitored component has not yet failed, nor it is replaced by an extra maintenance intervention, nor there has been an opportunity to bundle with a CM intervention in case of machine failure (whichever comes first). In other words, this opportunity happens when at $l_{\xi}$, the next PM intervention (scheduled $u$ time later) occurs before either the machine fails (which happens after $v$ time), or an extra maintenance on the monitored component is executed (after $l_{\xi,\varepsilon} + t_{\mathrm{p}}$ time), or the monitored component itself fails (after $l_{\xi,\varepsilon} + l_{\varepsilon,\mathrm{F}}$ time). With $g(u)$ the pdf of the time until the next PM intervention, and $h(v)$ the pdf of the time until the next machine failure, the probability that the monitored component is maintained together with a scheduled PM intervention is defined as:

$$\mathbb{P}(\mathrm{O_{PM}}) = \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}\int\limits_{u=0}^{l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}\left(\int\limits_{v=u}^{+\infty} h(v)dv\right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_{\mathrm{p}}\\ \int\limits_{u=0}^{l_{\xi,\varepsilon}+t_{\mathrm{p}}}\left(\int\limits_{v=u}^{+\infty} h(v)dv\right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \quad (5)$$

Likewise, we quantify the probability that the monitored component is maintained through an opportunistic maintenance with a CM intervention on any of the other components of the machine ($\mathrm{O_{CM}}$). This happens when at $l_\xi$, the next machine failure (which is $v$ time later) occurs before either executing the next PM intervention (which happens after $u$ time), or an extra maintenance on the monitored component is executed (after $l_{\xi,\varepsilon} + t_{\mathrm{p}}$ time), or the monitored component itself fails (after $l_{\xi,\varepsilon} + l_{\varepsilon,\mathrm{F}}$ time). Analogous to Eq. (5), we then define the probability of an opportunistic maintenance with a CM intervention on any of the other components of the machine, as:

$$\mathbb{P}(\mathrm{O_{CM}}) = \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}\int\limits_{v=0}^{l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}\left(\int\limits_{u=v}^{+\infty} g(u)du\right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_{\mathrm{p}}\\ \int\limits_{v=0}^{l_{\xi,\varepsilon}+t_{\mathrm{p}}}\left(\int\limits_{u=v}^{+\infty} g(u)du\right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \quad (6)$$

In case none of the above opportunities has taken place in the time window $l_{\xi,\varepsilon}$, an extra planned maintenance intervention is scheduled as soon as the degradation level surpasses the intervention threshold $\varepsilon$, for execution at time $l_{\xi,\varepsilon} + t_{\mathrm{p}}$. This extra maintenance intervention will effectively take place if the component does not fail during the planning period $t_{\mathrm{p}}$ and if no opportunities come along for an OM during this planning period. In other words, this happens when at $l_\xi$, the execution of the extra maintenance intervention (scheduled $l_{\xi,\varepsilon} + t_{\mathrm{p}}$ time later) occurs before either the next PM intervention (which happens after $u$ time), another component of the machine fails (after $v$ time), or the monitored component itself fails (after $l_{\xi,\varepsilon} + l_{\varepsilon,\mathrm{F}}$ time). Thus, the probability that the monitored component is maintained by means of an extra planned maintenance intervention, is given by:

$$\mathbb{P}(\mathrm{E}) = \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}0, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_{\mathrm{p}}\\ \int\limits_{u=l_{\xi,\varepsilon}+t_{\mathrm{p}}}^{+\infty} g(u)du \int\limits_{v=l_{\xi,\varepsilon}+t_{\mathrm{p}}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \quad (7)$$

Finally, we determine the probability that the monitored component fails, which means that the component is maintained upon its failure. This happens after the degradation surpassed the intervention threshold $\varepsilon$, and the component fails before the extra maintenance intervention is executed (i.e., when $l_{\varepsilon,\mathrm{F}} \le t_{\mathrm{p}}$), while no opportunity for OM occurred in the period $l_{\xi,\varepsilon} + l_{\varepsilon,\mathrm{F}}$. Hence we obtain the probability of a failure of the monitored component as:

$$\mathbb{P}(\mathrm{F}) = \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}\int\limits_{u=l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}^{+\infty} g(u)du \int\limits_{v=l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_{\mathrm{p}}\\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \quad (8)$$

Using the probabilities defined by Eqs.(5-8), we can derive the average maintenance costs per cycle, as defined in Eq. (3).

To determine the average cycle time, $\mathbb{E}\{L_{\mathrm{c}}\}$, we need to multiply the probability of each type of maintenance intervention with the cycle time from replacement until this maintenance intervention, which is $\mathbb{E}\{L_\xi\}$ plus, respectively, $u$ time for $\mathrm{O_{PM}}$, $v$ for $\mathrm{O_{CM}}$, $l_{\xi,\varepsilon} + t_{\mathrm{p}}$ for an extra maintenance, and

$l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}$ for a failure. This yields:

$$\begin{aligned}
\mathbb{E}\{L_{\mathrm{c}}\} &= \mathbb{E}\left\{L_{\xi}\right\} \\
&+ \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}\int\limits_{u=0}^{l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}} u\left(\int\limits_{v=u}^{+\infty} h(v)dv\right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}}\leq t_{\mathrm{P}} \\ \int\limits_{u=0}^{l_{\xi,\varepsilon}+t_{\mathrm{P}}} u\left(\int\limits_{v=u}^{+\infty} h(v)dv\right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}}> t_{\mathrm{P}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon} \\
&+ \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}\int\limits_{v=0}^{l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}} v\left(\int\limits_{u=v}^{+\infty} g(u)du\right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}}\leq t_{\mathrm{P}} \\ \int\limits_{v=0}^{l_{\xi,\varepsilon}+t_{\mathrm{P}}} v\left(\int\limits_{u=v}^{+\infty} g(u)du\right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}}> t_{\mathrm{P}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon} \\
&+ \int\limits_{l_{\xi,\varepsilon}=0}^{+\infty}\int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\begin{cases}(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\int\limits_{u=l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}^{+\infty} g(u)du\int\limits_{v=l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}}\leq t_{\mathrm{P}} \\ (l_{\xi,\varepsilon}+t_{\mathrm{P}})\int\limits_{u=l_{\xi,\varepsilon}+t_{\mathrm{P}}}^{+\infty} g(u)du\int\limits_{v=l_{\xi,\varepsilon}+t_{\mathrm{P}}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}}> t_{\mathrm{P}}\end{cases}\right) f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}.
\end{aligned} \tag{9}$$

We can further elaborate Eqs. (5-9) by substituting $g(u)$ and $h(v)$, as defined in (1) and (2). We relegate this derivation to Appendix A.

We are now in a position to evaluate and optimise the expected maintenance cost per running hour. As Eqs. (5-9) are analytically intractable, we rely on a numerical evaluation to determine the optimal pair of threshold values $(\xi^*, \varepsilon^*)$ that minimises the expected maintenance costs as follows. We first determine the probabilities of each of the four possible types of service interventions on the monitored component, as well as the expected maintenance cycle length, for each combination of threshold values $(\xi, \varepsilon)$, using Eqs. (5-9). We use Wolfram Mathematica 9.0.1.0 for the numerical calculation, although any other program that can numerically solve integrals can be used. Next, we use these probabilities to evaluate the expected cost performance of each combination of thresholds $(\xi, \varepsilon)$. The optimal pair of thresholds $(\xi^*, \varepsilon^*)$ is then the combination that leads to the lowest expected costs. This second step can be easily performed in any spreadsheet program. Although full enumeration is a time-consuming optimisation procedure, it is noteworthy to mention that the expressions in Eqs. (5-9) remain robust for a given failure behaviour of the monitored component, independent on the assumed cost parameters $C_i$. This makes it convenient to perform sensitivity analyses on the cost parameters (which may depend, for instance, on the operating circumstances) in a later phase.

Consider for instance a monitored component with a Weibull distributed failure behaviour, characterised by a scale factor $\alpha = 2{,}000$ and shape parameter $\beta = 1.5$ and a linear degradation behaviour. For each set of threshold values, Eqs. (5-9) define the probabilities and average time to each type of intervention. We use a full enumeration approach over the entire range of integer percentage values, which yields 5151 combinations of threshold values for which $\xi \leq \varepsilon$. For instance, for $\xi = 0.6$ (60% degradation) and $\varepsilon = 0.8$ (80% degradation), we have $\mathbb{P}(\mathrm{O_{PM}}) = 42\%$, $\mathbb{P}(\mathrm{O_{CM}}) = 15\%$, $\mathbb{P}(\mathrm{E}) = 27\%$, $\mathbb{P}(\mathrm{F}) = 16\%$, and the average time between two maintenance interventions with these thresholds, $\mathbb{E}\{L_{\mathrm{c}}\}$ $= 1{,}388$ hours.

Suppose the cost of the monitored component is \$200, the travel cost of the technician is \$100, and the additional cost in case of failure (e.g., due to consequential damage) is \$500. Then, for $\xi = 0.6$ and $\varepsilon = 0.8$, the long run average maintenance costs per running hour for the monitored component is given by (Eq. (3)):

$$\frac{42\%(\$200) + 15\%(\$200) + 27\%(\$200 + \$100) + 16\%(\$200 + \$100 + \$500)}{1,388} = 0.23\$. \quad (10)$$

Clearly, this number will change under modified cost values. For instance, if the component cost is only \$160, we find that an average maintenance costs per running hour reduces to:

$$\frac{42\%(\$160) + 15\%(\$160) + 27\%(\$160 + \$100) + 16\%(\$160 + \$100 + \$500)}{1,388} = 0.20\$. \quad (11)$$

We calculate these costs for each combination of threshold values. The optimal set of threshold values is then the combination that minimises these costs (which can be easily found in any spreadsheet program). As Eqs. (5-9) are calculated independent of the cost values, we can simply use the outcome of this analysis and complement it with the specific cost environment to obtain the corresponding maintenance costs. This approach has proven to be useful for our case study company, to whom we provided the outcome of Eqs. (5-9) for a range of failure behaviours that are commonly observed for their components. It allows them to evaluate the cost savings potential of CBM for any given component in a variety of cost environments.

In our example, the optimal combination is $(\xi^*, \varepsilon^*) = (65\%, 77\%)$ when the cost of the component is \$200, and $(\xi^*, \varepsilon^*) = (63\%, 76\%)$ when the cost of the component is \$160 (note that these results can also be derived from Figures 4 and 5 in our numerical section with $\rho_{\mathrm{E}} = 0.5$ and $\rho_{\mathrm{F}} = 2.5$ for a component cost of \$200, and $\rho_{\mathrm{E}} = 0.625$ and $\rho_{\mathrm{F}} = 3.125$ for a component cost of \$160).

As is evident from this example, one of the benefits of a full enumeration is that the outcome of Eqs. (5-9) complemented with the specific cost environment, allows a swift computation of the optimal set of threshold values. Instead of a full enumeration approach, we could as well use more sophisticated search algorithms or meta-heuristics, such as for instance (as an example of an easy search procedure) the steepest descent method. This method guarantees an optimal solution if each local optimum is a global optimum. Although our numerical results suggest that the cost function is convex in the threshold values, we cannot prove it and therefore a local search heuristic cannot guarantee the global optimal solution.

# 4 CBM with only one degradation threshold

The rationale behind using two degradation thresholds in the hybrid CBM policy is that it aims to reduce both the setup costs for each maintenance intervention (by using opportunities to bundle maintenance of the monitored component with maintenance on other components), as well as downtime costs (by implementing an intervention threshold that prevents an imminent failure if no opportunities have come across). In this section we evaluate the performance when only one of the two thresholds is implemented. In the next section we show how this analysis helps us to identify the value of each threshold, and understand in which setting each threshold adds most value.

## 4.1 Condition-based maintenance with intervention threshold only

The hybrid CBM policy with only an intervention threshold corresponds to the case where $\xi = \varepsilon$, which means there is no time window specified in which the monitored component can be maintained together with PM and CM interventions on the other components before the intervention threshold is reached: $l_{\xi,\varepsilon} = 0$. As a result, we can simplify Eqs. (5-9) by setting $f^2(l_{\xi,\varepsilon}) = 0$, and use $f^3(l_{\varepsilon,\mathrm{F}})$ instead of its conditional probability. We do, however, include the possibility for OM during the planning period when an opportunity occurs. The probabilities that the monitored component will be maintained by each of the four types of maintenance interventions then reduce to:

$$
\begin{aligned}
\mathbb{P}(\mathrm{O_{PM}}) &= \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \int_{u=0}^{l_{\varepsilon,\mathrm{F}}} \left( \int_{v=u}^{+\infty} h(v)dv \right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_\mathrm{p} \\ \int_{u=0}^{t_\mathrm{p}} \left( \int_{v=u}^{+\infty} h(v)dv \right) g(u)du, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}})dl_{\varepsilon,\mathrm{F}}, \\
\mathbb{P}(\mathrm{O_{CM}}) &= \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \int_{v=0}^{l_{\varepsilon,\mathrm{F}}} \left( \int_{u=v}^{+\infty} g(u)du \right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_\mathrm{p} \\ \int_{v=0}^{t_\mathrm{p}} \left( \int_{u=v}^{+\infty} g(u)du \right) h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}})dl_{\varepsilon,\mathrm{F}}, \\
\mathbb{P}(\mathrm{E}) &= \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} 0, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_\mathrm{p} \\ \int_{u=t_\mathrm{p}}^{+\infty} g(u)du \int_{v=t_\mathrm{p}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}})dl_{\varepsilon,\mathrm{F}}, \\
\mathbb{P}(\mathrm{F}) &= \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \int_{u=l_{\varepsilon,\mathrm{F}}}^{+\infty} g(u)du \int_{v=l_{\varepsilon,\mathrm{F}}}^{+\infty} h(v)dv, & \text{if } l_{\varepsilon,\mathrm{F}} \le t_\mathrm{p} \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}})dl_{\varepsilon,\mathrm{F}}. \qquad (12)
\end{aligned}
$$

Likewise the expression for $\mathbb{E}\{L_\mathrm{c}\}$ in (9) can be simplified in a similar way.

### 4.2 Condition-based maintenance with opportunistic threshold only

The hybrid CBM policy with only an opportunistic threshold corresponds to the case where $\varepsilon = 1$, which means that $l_{\varepsilon,\mathrm{F}} = 0$, and the monitored component fails if there is no opportunity for OM during the period $l_{\xi,\varepsilon} = l_{\xi,\mathrm{F}}$. As such, $\mathbb{P}(\mathrm{E}) = 0$ and by setting $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon}) = 0$ in Eqs. (5-9), we obtain:

$$
\begin{aligned}
\mathbb{P}(\mathrm{O_{PM}}) &= \int_{l_{\xi,\varepsilon}=0}^{+\infty} \left( \int_{u=0}^{l_{\xi,\varepsilon}} \left( \int_{v=u}^{+\infty} h(v)dv \right) g(u)du \right) f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}, \\
\mathbb{P}(\mathrm{O_{CM}}) &= \int_{l_{\xi,\varepsilon}=0}^{+\infty} \left( \int_{v=0}^{l_{\xi,\varepsilon}} \left( \int_{u=v}^{+\infty} g(u)du \right) h(v)dv \right) f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}, \\
\mathbb{P}(\mathrm{F}) &= \int_{l_{\xi,\varepsilon}=0}^{+\infty} \left( \int_{u=l_{\xi,\varepsilon}}^{+\infty} g(u)du \int_{v=l_{\xi,\varepsilon}}^{+\infty} h(v)dv \right) f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \qquad (13)
\end{aligned}
$$

It is noteworthy to highlight that this policy in fact coincides with a CBM policy without remotely monitoring the degradation level, but where the technician physically inspects the component when doing the maintenance on other components and then decides if it needs replacement or not.

# 5 Numerical analysis of the cost performance under CBM

## 5.1 Experimental setup

We analyse the performance of our hybrid CBM policy using reference data (of the component's failure behaviour and cost environment) of an OEM active in the compressed air industry, which are representative for their aftermarket business. We set up an extensive numerical experiment to generate insight in the potential savings that can be realised by implementing our hybrid CBM policy. This is done by analysing the maintenance cost performance when the monitored component is maintained under CBM, and benchmark it against the policy where the monitored component is included in the usage-based block replacement PM policy. In this section we report our findings.

In our numerical experiment we use the Weibull distribution to model the failure behaviour, which is commonly used to model real life data due to its flexibility in shape. The pdf of the lifetime $l$ (i.e., the time to failure) is then defined by:

$$f(l) = \frac{\beta}{\alpha}\left(\frac{l}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{l}{\alpha}\right)^{\beta}\right], \quad l \geq 0, \tag{14}$$

with $\alpha$ the characteristic life or scale factor and $\beta$ the shape parameter. A high $\alpha$-value represents a slow deteriorating component, and a high $\beta$-value means its lifetime can be relatively well estimated (low variance). Examples of fast deteriorating components are consumables such as air and oil filters. Slow deteriorating components are durable goods, such as the engine. The parameters of the Weibull distribution are usually found using failure data of the components.

Whereas the failure behaviour determines when the component fails, it does not determine how it degrades (from replacement until failure). In our experiment we assume a linear degradation, but it could be extended to other degradation processes. The benefit of a tractable degradation process is that it allows an analytical derivation of the distribution of the degradation time. In Appendix B we provide a derivation of $\mathbb{E}\{L_\xi\}$, $f^2(l_{\xi,\varepsilon})$, $f^3(l_{\varepsilon,\mathrm{F}})$, and $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})$ under a linear degradation process and a Weibull distributed time to failure. We analysed an exponential degradation behaviour as well; this yielded similar managerial insights, with the only notable difference in the optimal threshold values.

In our numerical experiment, we let the scale factor have values $\alpha \in \{1,000; 2,000; ...; 20,000\}$ RH (running hours), and the shape parameter $\beta \in \{1.5; 3\}$. We assume the base PM interval $\tau = 1,000$ RH and the average time between (random) machine failures $\lambda^{-1} = 2,000$ RH. The preparation time for an extra maintenance intervention $t_\mathrm{p} = 150$ RH.

The cost factors $\rho_\mathrm{E}$ and $\rho_\mathrm{F}$, respectively denoting the percentage cost increase of an extra maintenance compare to an opportunistic maintenance, and the percentage cost increase of a failure compared to an extra maintenance, can have a wide range of values, depending on the operating conditions of the company and the component. At the OEM in the compressed air industry, we consider three types of components that can be considered for CBM, each with its own range of $\rho_\mathrm{E}$ and $\rho_\mathrm{F}$ values: (1) Key components that are usually expensive and vital to operating the machine (i.e., durable goods). These components generally have a long lifetime. The cost of travel to maintain these components is usually negligible compared to the cost of a failure of the component, and the cost to replace them (i.e., the cost of the spare part and the wrenching cost). The value of $\rho_\mathrm{E}$ for these components is close to zero, and the value of $\rho_\mathrm{F}$ is usually between 1 and 5. (2) Consumables, which are generally relatively cheap, but crucial for a good functioning of the machine (e.g., filters). These components typically have a shorter lifetime. For those components, the cost of travel is significant compared to the replacement cost of the component, with $\rho_\mathrm{E}$ between 0.5 and 4. Although the component is crucial for the functioning of the machine, its failure will not result in a lot of consequential damage to other components and the value of $\rho_\mathrm{F}$ will be similar to that of the key component, between 1 and 5. These crucial consumables will usually perform relatively well under a block replacement policy, and the benefits of maintaining these components under a CBM policy are often lower. (3) Supporting consumables that are not directly required to operate the machine, but they support the well-functioning of it (e.g., a condensate drain). They are relatively cheap but when they fail, the cost of consequential damage is very high. Whereas the cost of travel $\rho_{\mathrm{E}}$ will range between 1 and 5, the cost of failure $\rho_{\mathrm{F}}$ can reasonably range between 5 and 30. Despite their lower unit cost, these components are good candidates for CBM to avoid expensive failures.

We benchmark the cost performance when the monitored component is maintained under our hybrid CBM policy against its cost performance when it is maintained under a usage-based PM policy. In the latter case, the (optimised) maintenance interval for the monitored component is $\tau^*$, which is a multiple of the base maintenance interval $\tau$, indicating that the monitored component is replaced during PM interventions every time interval $\tau^*$ (and it has a CM upon failure). When the monitored component would be maintained under a periodic PM policy, an intervention cost is incurred upon every maintenance interval $\tau^*$, which includes the unit component cost and the wrenching time to replace the monitored component. As the PM intervention is already scheduled for the other components, its setup (e.g. travel) cost is not included in the maintenance costs for the monitored component. Upon failure of the monitored component, a CM intervention is required, which includes an additional breakdown cost and travel cost to replace the monitored component. The optimisation of the maintenance interval $\tau^*$ that minimises the corresponding maintenance costs when the monitored component is maintained under PM is then defined by:

$$\begin{aligned} \underset{\tau^*}{\text{minimise}} \quad & \frac{C_{\mathrm{O}} + C_{\mathrm{F}} \cdot \mathbb{E}\left\{\mathcal{Y}\right\}_{\tau^*}}{\tau^*} \\ \text{subject to} \quad & \tau^* = k\tau \qquad k \in \mathbb{Z}^+, \end{aligned} \tag{15}$$

with $\mathbb{E}\left\{\mathcal{Y}\right\}_{\tau^*}$ the expected number of failures between two subsequent PM interventions in the interval $\tau^*$. Observe that when $\tau^* = \infty$, it is optimal to do only corrective maintenance for the monitored component (replacement upon failure). The value of $\mathbb{E}\left\{\mathcal{Y}\right\}_{\tau^*}$ can be calculated as:

$$\begin{aligned} \mathbb{E}\left\{\mathcal{Y}\right\}_{\tau^*} = & \int_{t_1=0}^{\tau^*} f(t_1)\mathrm{d}t_1 + \int_{t_1=0}^{\tau^*} \int_{t_2=0}^{(\tau^*-t_1)} f(t_2)\mathrm{d}t_2 f(t_1)\mathrm{d}t_1 \\ & + \sum_{i=3}^{\infty} \int_{t_1=0}^{\tau^*} \int_{t_2=0}^{(\tau^*-t_1)} \cdots \int_{t_i=0}^{(\tau^*-t_1 \ldots -t_{i-1})} f(t_i)\mathrm{d}t_i \ldots f(t_2)\mathrm{d}t_2 f(t_1)\mathrm{d}t_1, \end{aligned} \tag{16}$$

where $t_1$ is the time between the PM intervention and the first failure (and thus CM intervention), $t_2$ is the time between the first and second failure, etc. The intuition behind Eq. (16) is the following: its first term defines the probability that at least one failure will occur during $\tau^*$ time, the second term defines the probability that at least two failures occur during $\tau^*$ time, etc. For instance, the probability of at least two failures during $\tau^*$ time, corresponds to the situation when the first failure occurs after $t_1 < \tau^*$ time and the second failure occurs after $t_2 < \tau^* - t_1$ time, so that the second failure occurs before $\tau^*$ time have passed, i.e., before the next PM intervention.

Due to the convexity of the cost function of a PM policy defined in Eq. (15), the optimal value of $\tau^*$ can be found using a simple search algorithm.

In what follows we first validate our solution method and our model assumptions (Section 5.2) . We then provide the cost reductions that can be maintained when our hybrid policy with two thresholds is implemented instead of maintaining the component using a traditional periodic maintenance policy

(Section 5.3). Finally we show the value of having two thresholds instead of one (Section 5.4).

## 5.2 Validation of the solution method

We first validate our assumption that the time until the next planned PM intervention is approximately uniformly distributed. We have simulated the environment without the assumption that the time until the next planned PM intervention is uniformly distributed (denote this the 'real' environment), and optimized its cost performance. We then evaluated how the optimised threshold values under the assumptions of our model (i.e. assuming time until the next planned PM intervention is uniformly distributed) performs in this 'real' environment. Our simulation experiment considers 10,000,000 iterations, which is equivalent to 10,000,000 replacements of the monitored component. We perform a full enumeration approach over the entire range of integer percentage values, which yields 5,151 combinations of threshold values for which $\xi \leq \varepsilon$. The simulation experiment was done in Arena 14.5 (Rockwell Automation).

Table 2 reports that the average cost performance using these threshold values is less than 1% higher, compared to the optimized performance without that assumption. This result holds under a range of parameter values (the table reports the setting that is mostly assumed throughout our paper: the failure behaviour defined by $\alpha = 2,000$ and $\beta = 1.5$, with different values for the relative travel and breakdown cost; this setting matches with the Figures in the remainder of this section). This somehow provides a justification of our assumptions to approximate the time until the next planned PM intervention to be uniformly distributed.

Table 2: When optimizing the threshold values under our model assumptions, the costs are less then 1% higher compared to the optimized performance without these assumptions (results are reported under a failure behaviour defined by $\alpha$ = 2,000 and $\beta$ = 1.5).

| | | $\rho_F$ | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 5 | 10 |
| $\rho_E$ | 0 | 0.00% | 0.07% | 0.01% | 0.04% | 0.03% |
| | 1 | 0.03% | 0.76% | 0.38% | 0.32% | 0.29% |
| | 2 | 0.56% | 1.54% | 0.99% | 0.72% | 0.37% |
| | 5 | 0.50% | 0.77% | 0.61% | 1.03% | 0.79% |
| | 10 | 0.49% | 0.35% | 0.24% | 0.79% | 0.58% |

We have also set up a numerical experiment to validate how many non-preventively maintained components in a machine are required to assume that the time between failures of two arbitrary components in the machine is exponentially distributed (as the non-preventively maintained components will constitute the majority of the machine failures). We restricted our experiment to components with identical failure behaviour, although we are aware as mentioned earlier in our paper, that there are a number of additional elements that strengthen the effect of randomness of the failures, such as different operating conditions or additional preventive maintenance interventions. In our numerical experiment we assume that each component has the same Weibull distributed failure behaviour. We set up a Monte Carlo simulation experiment to record the time between failures when one or more identical components fail independently of each other. We generated results for a range of 1 to 10 identical components with different values of the shape parameter $\beta$ of the Weibull distribution (i.e., $\beta \in \{1.5; 3\}$), which are also the values encountered at our case study company. The scale parameter $\alpha$ of the Weibull distribution has no influence on the convergence. We considered a warm-up of 1,000,000 failures, before recording 1000 failures as a reference. We similarly generated 1000 random realisations of the exponential distribution. When $\beta = 1.5$, we found that as of 3 components we cannot statistically differentiate between both samples. In case $\beta = 3$, which indicates a lower variance of the time to failure, we found this conclusion to be true as of 4 components. This means that it suffices to have around 4 or more components that are non-preventively maintained, to justify the use of the exponential distribution in this regard.

Finally, we have performed a simulation analysis to verify how well our model assumptions perform in the 'real' environment, compared to the model assumptions made in Zhu et al. [25] when applied in the real environment. Zhu et al. [25] make the assumption that the block replacement policy is rescheduled after each failure. This allows to approximate the optimal threshold values using renewal theory. In order to have a fair comparison of both models, we extended the model of Zhu et al. [25] with the extra threshold, as Zhu et al. [25] propose only an opportunistic threshold. As can be observed in Table 3 (for an example with failure behaviour defined by $\alpha = 2,000$ and $\beta = 1.5$), in some cost environments our model performs better (reported by the negative values – in our example up to 6%), whereas in other environments their model outperforms slightly (up to 1,5% in our example).

Table 3: Cost comparison of our heuristic solution compared to the assumptions made by Zhu et al. [25], when two thresholds are applied (failure behaviour defined by $\alpha$ = 2,000 and $\beta$ = 1.5).

| | | $\rho_F$ | | | | |
|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 5 | 10 |
| $\rho_E$ | 0 | 0.00% | 0.07% | 0.01% | 0.04% | 0.03% |
| | 1 | -0.03% | 0.76% | -4.34% | 0.32% | 0.26% |
| | 2 | 0.56% | 1.54% | -0.04% | 0.60% | 0.10% |
| | 5 | -0.27% | 0.15% | -0.17% | -6.17% | -0.54% |
| | 10 | 0.45% | 0.24% | 0.07% | -1.93% | 0.36% |

Whereas this provides a comparison on the heuristic solution, the benefit of our model compared to Zhu et al. [25] is the addition of the extra intervention threshold. Our model operates with both an opportunistic and an intervention threshold, whereas Zhu et al. [25] only considers the opportunistic threshold. This is also acknowledged in Zhu et al. [25], who refers to our model: "They show that the average cost can be reduced significantly if this extra type of maintenance action is much cheaper than the normal corrective maintenance action. Their contributions are based on a different approximate evaluation procedure and a different degradation process". Our two-threshold approach makes our paper different from Zhu et al. [25], but more importantly it can result in lower costs. The extra intervention threshold is meant to avoid expensive downtime costs in case no opportunities have come across (initiated by the opportunistic threshold). In Figure 8 we will report the value of having this extra intervention threshold: especially in environments where the breakdown cost is high, the benefit of having this extra intervention threshold can be significant.

## 5.3 Cost performance of CBM policy with two thresholds

When CBM is implemented with two thresholds, the expected cost performance of the monitored component per running hour is defined by Eq. (3). The comparable cost performance of the monitored component under a PM policy is defined by Eq. (15).

Figure 2 reports the percentage reduction of the maintenance costs of the monitored component when the hybrid CBM policy is implemented compared to the situation where the monitored component is part of the PM interventions for one specific failure behaviour (defined by $\alpha = 2,000$ and $\beta = 1.5$). The cost reductions are expressed as a function of the travel cost $\rho_{\mathrm{E}}$ and the breakdown cost $\rho_{\mathrm{F}}$. It illustrates how the cost saving potential of CBM is strongly dependent on these cost ratios. These cost reductions can be explained by the dynamics of the thresholds in response to these cost ratios:

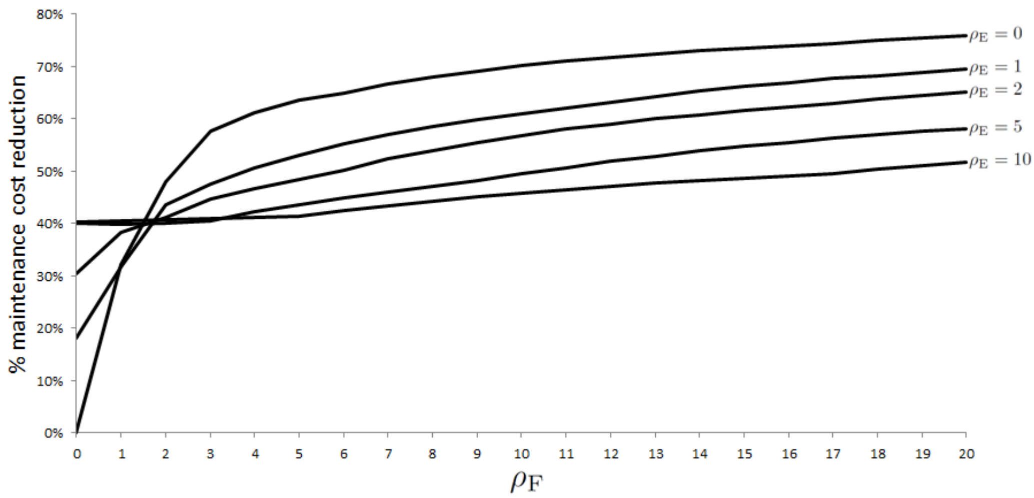


Figure 2: Percentage maintenance cost reduction of CBM vs. PM as a function of the breakdown cost $\rho_{\mathrm{F}}$ for different values of travel cost $\rho_{\mathrm{E}}$ (failure behaviour defined by $\alpha = 2{,}000$ and $\beta = 1.5$)

- In the extreme (but non-realistic) case where $\rho_{\mathrm{E}} = \rho_{\mathrm{F}} = 0$, the optimal policy is the CM-only policy as there are no benefits from bundling travel or avoiding breakdown. This corresponds to $\tau^* = \infty$ when the component is maintained in PM, and $\xi^* = \varepsilon^* = 1$ when the component is maintained under CBM (and thus CBM nor PM provide benefits over only CM interventions).
- When the breakdown cost $\rho_{\mathrm{F}}$ increases, CBM reduces the number of failures by either reducing the value of the optimal intervention threshold $\varepsilon^*$, which comes at the expense of an additional travel cost; or by reducing the value of the opportunistic threshold $\xi^*$, where the benefits of opportunistic maintenance (less travel costs) are traded off against the shorter used lifetime of the component. As both options reduce the total number of breakdowns, the benefits of CBM will always increase with the breakdown cost $\rho_{\mathrm{F}}$ (irrespective of the travel cost $\rho_{\mathrm{E}}$). This is illustrated in Figure 2.
- When the travel cost $\rho_{\mathrm{E}}$ increases, CBM can reduce the number of travels by extending the opportunistic time window to seize more opportunities for OM. This is done by either having a lower value of the optimal opportunistic threshold $\xi^*$, or by a higher value of the intervention threshold $\varepsilon^*$. A lower value of $\xi^*$ shortens the lifetime of the components, which leads to more unit component and wrenching costs per time unit. A higher value of $\varepsilon^*$ increases the risk of failures. In case of a low breakdown cost, the value of the intervention threshold $\varepsilon^*$ is typically high and thus by setting a lower value of the opportunistic threshold $\xi^*$, CBM reduces the number of travels. Under these circumstances (i.e., low $\rho_{\mathrm{F}}$ values), higher travel costs will lead to more benefits of CBM. In contrast, in case of high breakdown cost (high $\rho_{\mathrm{F}}$ values), both threshold values are low. A higher value of the intervention threshold is in that case not desired, and further reducing the opportunistic threshold shortens the lifetime of the components even more. This reduces the relative benefits of CBM and explains why the benefits of CBM may even go down with increasing $\rho_{\mathrm{E}}$ values when the value of $\rho_{\mathrm{F}}$ is high. Figure 3 provides the same information as Figure 2, but illustrates the latter effect better.

Figures 4 and 5 illustrate the above mentioned dynamics on the optimal threshold values for the settings described in Figures 2 and 3.

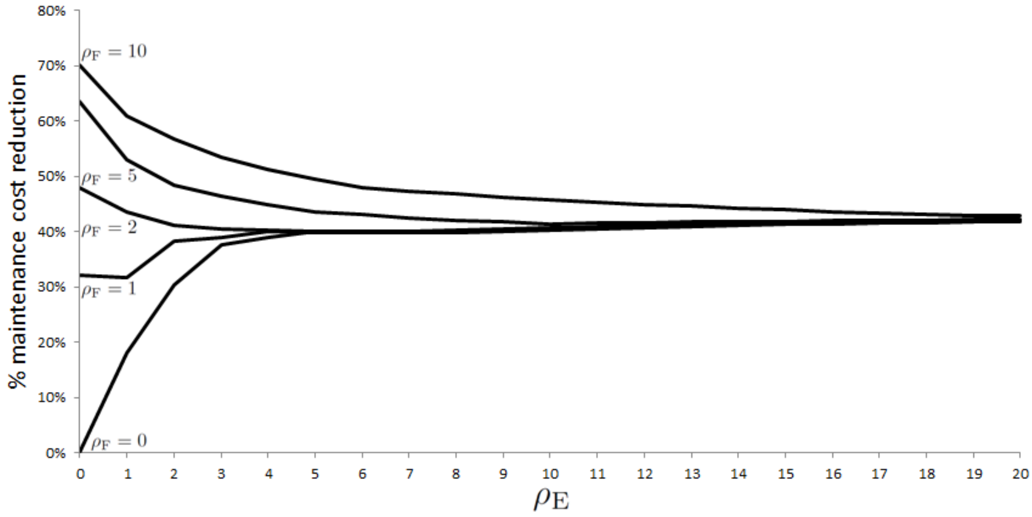


Figure 3: Percentage maintenance cost reduction of CBM vs. PM as a function of the travel cost $\rho_E$ for different values of breakdown cost $\rho_F$ (failure behaviour defined by $\alpha = 2{,}000$ and $\beta = 1.5$)

In Figure 6 we illustrate how the failure behaviour of the monitored component impacts the benefits of CBM. We observe higher gains for slowly degrading components (i.e., higher $\alpha$-values of the Weibull distributed failure behaviour in (14)). For slowly degrading components, there are more opportunities for OM, and there is more time to prevent failure as soon as the intervention threshold is reached. The gains of CBM also increase with a higher variance on the failure time (i.e., a lower $\beta$-value of the Weibull distributed failure behaviour). When the uncertainty of failure times is high, PM is less cost-efficient as it is hard to determine the optimal maintenance interval: maintenance can be much too early, or much too late. In that case, the value of monitoring is higher as it helps determining the right timing of the maintenance, which enables a longer lifetime of the components without incurring too many failures.

Although the above results are case-specific for our company case study, they do provide an order of magnitude of the cost savings by implementing our hybrid CBM policy. As our solution procedure is versatile to quantify the savings for any given cost environment, the strength of our methodology is that a variety of companies can use it to calculate the potential benefit of introducing CBM in their specific situation.

## 5.4 Cost performance of CBM with one threshold

In the previous section we highlighted the potential benefits of implementing a hybrid CBM policy with two thresholds. In this section we determine how much of the gains of CBM are lost when we omit one of the two thresholds. This information is valuable to understand which threshold adds most value in which environments, and what the incremental value is of having two thresholds instead of one (either the opportunistic or the intervention threshold).

In general, the opportunistic threshold has more economic value in environments where the travel cost $\rho_E$ is high, because the opportunistic threshold aims to bundle maintenance interventions and reduce total travel costs. The intervention threshold has more value when the breakdown cost $\rho_F$ is high, as the intervention threshold aims to prevent a breakdown. However, we observe some interaction effects.

Figure 7 reports the percentage cost increase in the monitored component's maintenance costs when

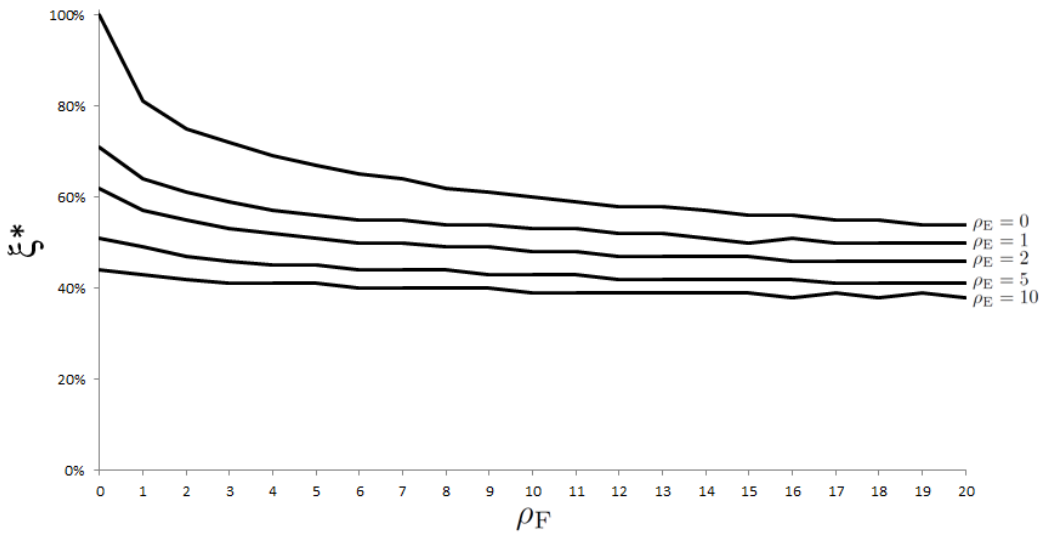


Figure 4: Optimal opportunistic threshold value $\xi^*$ as a function of the breakdown cost $\rho_{\mathrm{F}}$ for different values of travel cost $\rho_{\mathrm{E}}$ (failure behaviour defined by $\alpha$ = 2,000 and $\beta$ = 1.5)

we implement CBM with only the intervention threshold, compared to the two-threshold policy (represented as *iso-percentage-cost-increase* curves). This means that we drop the opportunistic threshold and we only seize opportunities of OM during the preparation period of the extra maintenance. As expected, dropping the opportunistic threshold increases the component's maintenance costs when the travel cost $\rho_{\mathrm{E}}$ is high. Especially in an environment where $\rho_{\mathrm{F}}$ is low, the maintenance costs rapidly increase with the value of $\rho_{\mathrm{E}}$. We also observe that the breakdown cost $\rho_{\mathrm{F}}$ only marginally impacts the economic value of the opportunistic threshold, especially when the travel cost $\rho_{\mathrm{E}}$ is low. Although not shown, we also found in our results that the opportunistic threshold is more valuable when the average failure time is high. In that case, there are more opportunities for OM in the presence of an opportunistic threshold.

Figure 8 reports the percentage cost increase in the monitored component's maintenance costs when we implement CBM with only the opportunistic threshold, compared to the costs in the two-threshold policy. As expected, when we omit the intervention threshold, the increase in maintenance costs is higher for high values of $\rho_{\mathrm{F}}$, as the intervention threshold prevents a breakdown. We also observe that for a given value of $\rho_{\mathrm{F}}$, the economic value of the intervention threshold goes down when the travel cost $\rho_{\mathrm{E}}$ increases. Indeed, for components with a high travel cost $\rho_{\mathrm{E}}$, the opportunistic threshold $\xi^*$ is low. This at the same time prevents a large number of failures and hence, dropping the intervention threshold has less impact on maintenance costs. Finally, we also found in our results that the intervention threshold has more value when the average time to failure is short (i.e., a low $\alpha$-value). In that case CBM benefits less from OM, and most gains are reaped by implementing an intervention threshold.

By combining Figures 7-8 with Figures 2-3, it can be quantified whether CBM with only one threshold still performs better than PM. Whereas Figures 2-3 quantify the percentage cost reductions of CBM with two thresholds compared to the PM policy, Figures 7-8 quantify how much these maintenance costs increase again when we drop one of the two thresholds.

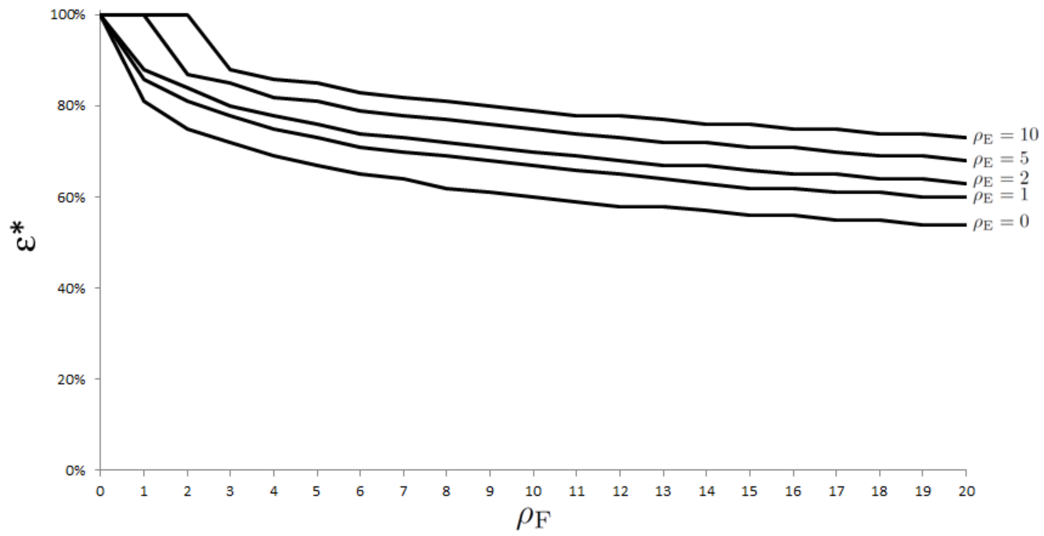


Figure 5: Optimal intervention threshold value $\varepsilon^*$ as a function of the breakdown cost $\rho_\mathrm{F}$ for different values of travel cost $\rho_\mathrm{E}$ (failure behaviour defined by $\alpha = 2{,}000$ and $\beta = 1.5$)

# 6 Setting CBM thresholds with multiple objectives

In the previous sections we focused on maintenance costs incurred by the OEM. Customers however are mainly interested in the impact of a maintenance policy on the downtime of their machines. Unfortunately, both points of view may be in conflict. Consider the following numerical example.

Assume the labour cost of a technician is \$100 per hour and the unit cost of the monitored component is \$50. To replace the component (preventively or upon failure), it takes the technician on average two hours to travel on site, and 30 minutes on average for the replacement itself (i.e., the wrenching time). Upon failure, the additional downtime for the customer is on average 6 hours (which means the machine is down for a total of 8 hours before the technician is on site), and there is an additional cost of \$400 due to consequential damage on other components. Assume the failure behaviour of the monitored component is Weibull distributed with $\alpha = 2{,}000$ and $\beta = 1.5$ and degradation is linear. We use the solution method explained in Section 3 to determine the optimal threshold values.

When the OEM has a full service contract, he pays for the costs of the technician and the components, but not for costs incurred due to the downtime of the machine. That means that $\rho_\mathrm{E} = \frac{2\times\$100}{\$50+{}^1\!/{}_2\times\$100} = 2$ and $\rho_\mathrm{F} = \frac{\$400}{\$50+{}^1\!/{}_2\times\$100} = 4$. Setting the optimal thresholds based on these parameters, results in $\xi^* = 0.52$ and $\varepsilon^* = 0.78$. This results in a cost reduction of 47% on the maintenance costs of the monitored component compared to the maintenance cost in a PM policy (see also Figure 9).

The user of the equipment also benefits from implementing this CBM policy. Because the number of breakdowns can be reduced, he will face less downtime. More specifically, the downtime costs will reduce by 60% when the monitored component is maintained using the above CBM policy compared to a PM policy. However, from the perspective of the user, he would prefer to reduce the downtime even more by setting different values of the thresholds. Indeed, from his perspective, $\rho_\mathrm{E} = 0$ (there is no additional downtime on top of the wrenching time in case of an extra maintenance) and $\rho_\mathrm{F} = \frac{8\text{hours}}{{}^1\!/{}_2\text{hour}} = 16$. That means that from the perspective of the equipment user, the optimal thresholds should be set at $\xi^* = 0.56$ and $\varepsilon^* = 0.56$ (interestingly, from the perspective of the equipment user, there is no benefit of having an opportunistic threshold). This leads to a 74% reduction in downtime compared to the PM policy

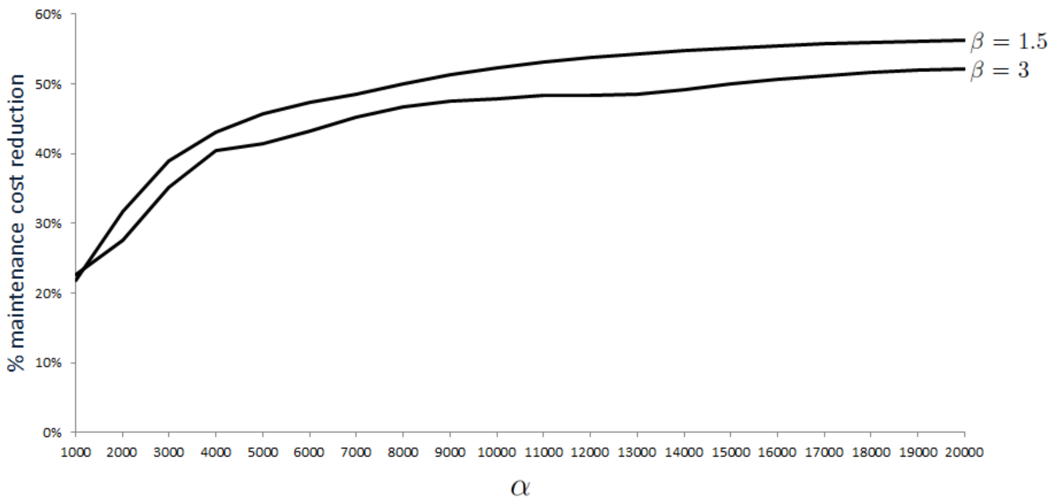


Figure 6: Percentage maintenance cost reduction of CBM vs. PM as a function of $\alpha$ (determining the avg time to failure) for different values of $\beta$ (related to the variance of the time to failure), with $\rho_{\mathrm{E}}=\rho_{\mathrm{F}}=1$

(see Figure 9), while the reduction of the maintenance costs is only 29%. This illustrates the potential conflict.

In search of a good servitisation strategy, these multiple objectives play a role. CBM helps to reduce the maintenance costs for the OEM, as well the downtime for the equipment user. Our analysis can be used for both. As such the results presented in the previous section can be interpreted from the perspective of maintenance cost reduction (which is of interest to the OEM providing the service), or from the perspective of downtime reduction (affecting the equipment user). However, as both are differently impacted by an extra maintenance or failure (leading to different values of $\rho_{\mathrm{E}}$ and $\rho_{\mathrm{F}}$), each stakeholder will opt for a different optimal set of threshold values. Figure 10 illustrates the impact of different thresholds on both performance objectives. It also shows the Pareto Efficient Frontier, which contains the optimal trade-offs between both objectives. The exact positioning on the frontier depends on negotiations between the OEM and the equipment user.

# 7 Conclusions

In this paper we introduced a hybrid multi-component opportunistic maintenance policy, combining condition-based maintenance (CBM) on one monitored component, with periodic preventive maintenance (PM) and corrective maintenance (CM) on the other components. Our policy consists of two thresholds on the level of degradation: when the degradation surpasses a first ‘opportunistic’ threshold, the component is maintained as soon as a PM or CM occurs on the other components; once it surpasses a second ‘intervention’ threshold, an additional maintenance intervention is planned to replace the component in order to avoid failure. We provide analytical expressions to quantify the expected maintenance costs for each set of thresholds. Using an enumeration procedure, we then determine the optimal set of thresholds that minimises the expected maintenance costs per running hour.

This hybrid CBM policy has the following advantages over the PM policy. First, monitoring the degradation helps to cope better with the uncertainty in the component’s failure behaviour, enabling a longer used lifetime, while limiting the planned and/or unplanned maintenance interventions and the

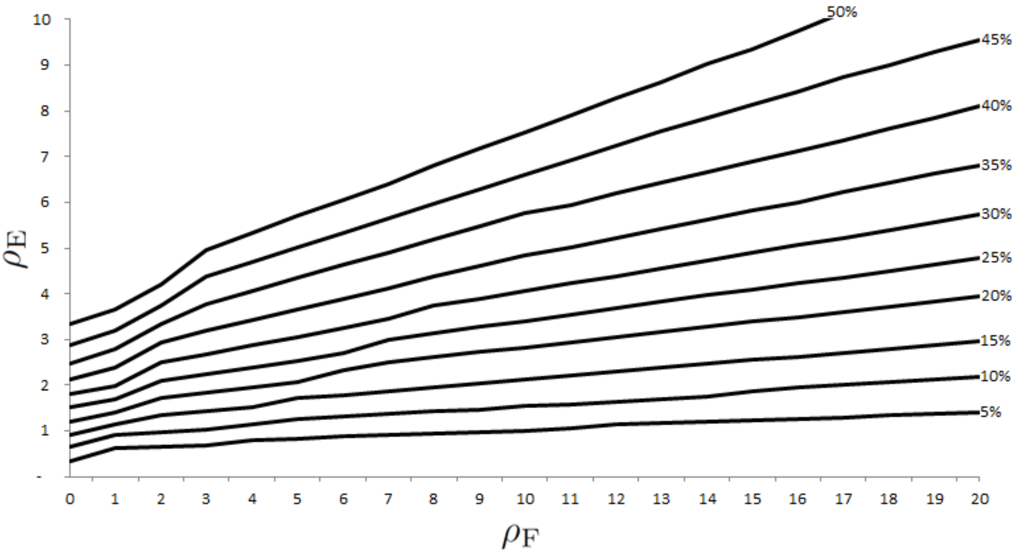


Figure 7: Sensitivity analysis: percentage cost increase if the opportunistic threshold is dropped (failure behaviour defined by $\alpha$=2,000; $\beta$=1.5)

number of breakdowns. Second, by bundling the maintenance of the monitored component with other maintenance interventions, setup costs, such as travel to the machine, can be reduced. The key drivers of the cost reductions of our hybrid CBM policy with two thresholds are thus the cost of travel and the cost of a breakdown, compared to the replacement cost of the component.

We find that our hybrid opportunistic maintenance has the potential to offer significant cost savings, while offering a smooth transition to implementing CBM in the current maintenance operations. Slowly degrading components with highly volatile failure times and high breakdown costs provide the biggest gains of such a CBM implementation. The value of the opportunistic threshold is high in environments where the travel costs are high. The intervention threshold is valuable in environments where the breakdown costs are high, especially when the travel cost is low. Whereas the opportunistic threshold has more economic value for components with a longer lifetime, the intervention threshold has more value when the lifetime of the component is short.

Our model can be extended along the following lines. First, it could be investigated whether upon an extra maintenance intervention the subsequent PM intervention should be advanced and combined with this extra maintenance intervention in order to save travel costs. This begs the question whether only the subsequent PM intervention should be rescheduled, or all of the future PM interventions. An optimal policy could probably be obtained through for instance a dynamic programming approach. Second, whereas our model assumes that the degradation state can be perfectly monitored, it is worthwhile to investigate how the model would perform under misspecification errors; e.g., the monitored degradation value may not coincide with the real degradation value (noise), or the component fails for a reason that is not captured by the monitored degradation data. Finally, although our CBM policy with thresholds is practical to implement, machine-learning techniques that predict upcoming failures might increase the potential of continuous monitoring.

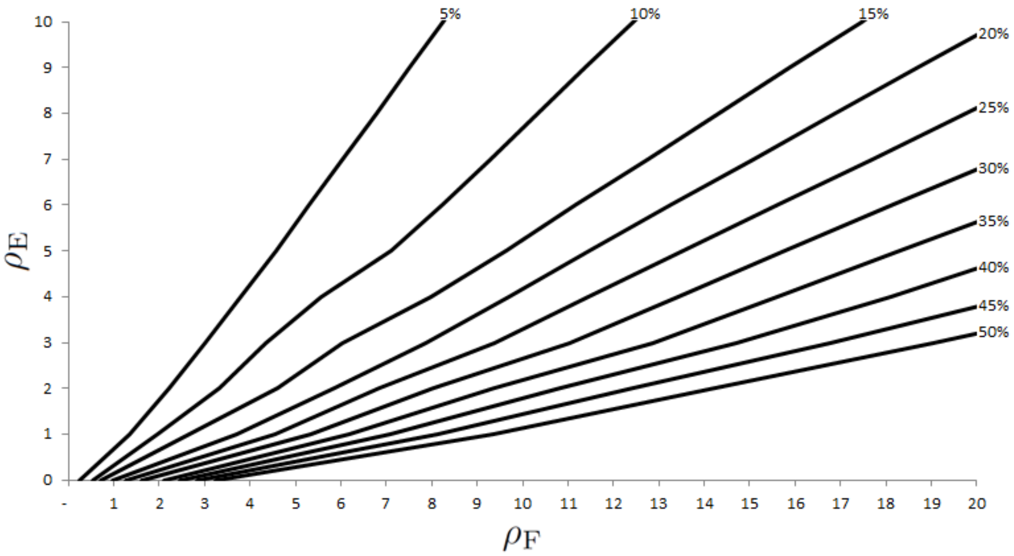


Figure 8: Sensitivity analysis: percentage cost increase if the intervention threshold is dropped (failure behaviour defined by $\alpha$=2,000; $\beta$=1.5)

# References

[1] Ahmad, R. and Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1):135–149.

[2] Barata, J., Soares, C. G., Marseguerra, M., and Zio, E. (2002). Simulation modelling of repairable multi-component deteriorating systems for 'on condition' maintenance optimisation. *Reliability Engineering & System Safety*, 76(3):255–264.

[3] Barlow, R. and Hunter, L. (1960). Optimum preventive maintenance policies. *Operations Research*, 8(1):90–100.

[4] Cox, D. R. and Miller, H. D. (1977). *The theory of stochastic processes*, volume 134. CRC Press.

[5] de Jonge, B., Klingenberg, W., Teunter, R., and Tinga, T. (2016). Reducing costs by clustering maintenance activities for multiple critical units. *Reliability engineering & system safety*, 145:93–103.

[6] Goossens, A. and Basten, R. (2015). Exploring maintenance policy selection using the analytic hierarchy process; an application for naval ships. *Reliability Engineering & System Safety*, 142:31–41.

[7] Gupta, A. and Lawsirirat, C. (2006). Strategically optimum maintenance of monitoring-enabled multi-component systems using continuous-time jump deterioration models. *Journal of Quality in Maintenance Engineering*, 12(3):306–329.

[8] Gürler, Ü. and Kaya, A. (2002). A maintenance policy for a system with multi-state components: an approximate solution. *Reliability Engineering & System Safety*, 76(2):117–127.

[9] Jardine, A. K., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510.

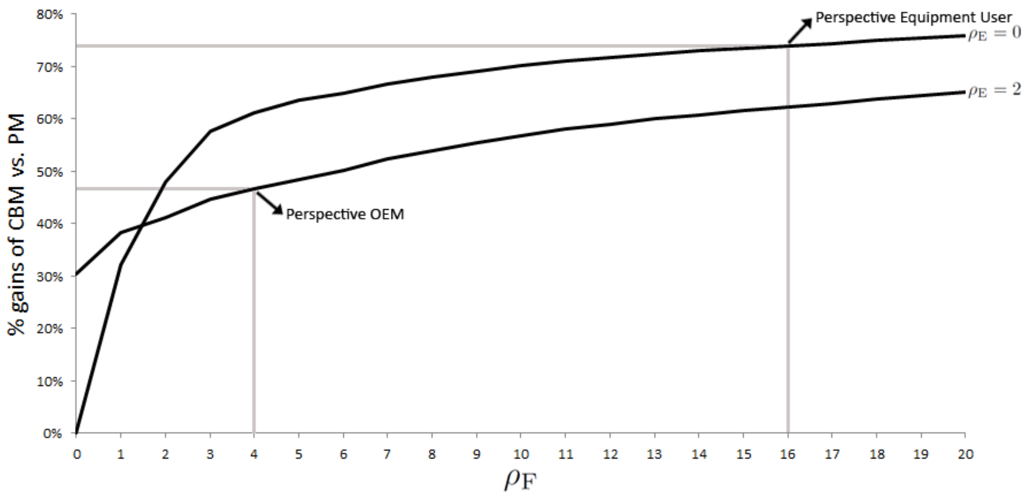


Figure 9: Percentage gains of CBM vs. PM for $\rho_E = \{0, 2\}$ and different values of $\rho_F$ (failure behaviour defined by $\alpha$ = 2,000 and $\beta$ = 1.5)

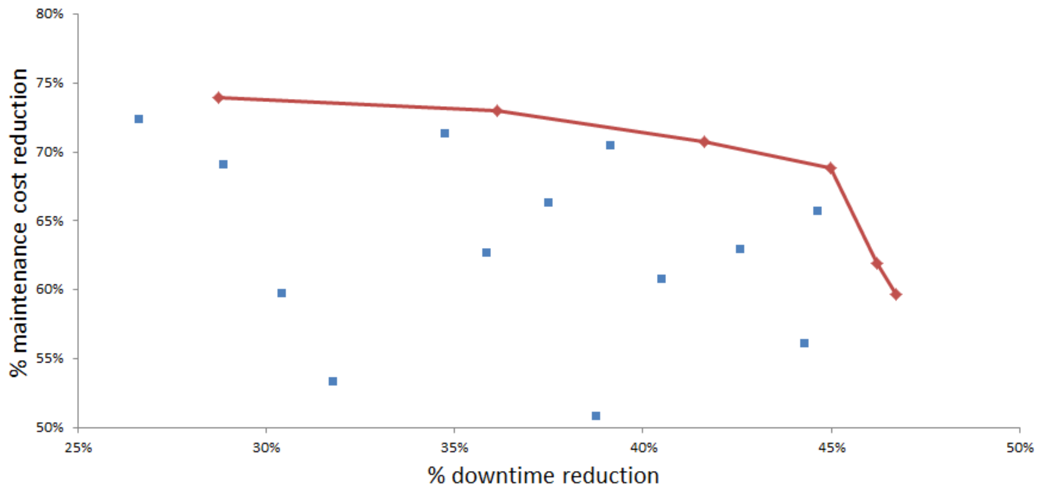


Figure 10: Percentage cost reduction vs. downtime reduction for different combinations of $\xi$ and $\varepsilon$ (failure behaviour defined by $\alpha$ = 2,000 and $\beta$ = 1.5)


[10] Jardine, A. K. S. and Tsang, A. H. C. (2006). *Maintenance, Replacement, and Reliability. Theory and Applications*. Dekker Mechanical Engineering. CRC Press, Boca Raton (FL).

[11] Koochaki, J., Bokhorst, J. A., Wortmann, H., and Klingenberg, W. (2012). Condition based maintenance in the context of opportunistic maintenance. *International Journal of Production Research*, 50(23):6918–6929.

[12] Koochaki, J., Bokhorst, J. A., Wortmann, H., and Klingenberg, W. (2013). The influence of condition-based maintenance on workforce planning and maintenance scheduling. *International Journal of Production Research*, 51(8):2339–2351.

[13] Marseguerra, M., Zio, E., and Podofillini, L. (2002). Condition-based maintenance optimization by means of genetic algorithms and monte carlo simulation. *Reliability Engineering & System Safety*, 77(2):151–165.

[14] Peng, Y., Dong, M., and Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50(1-4):297–313.

[15] Pintelon, L. and Van Puyvelde, F. (2006). *Maintenance Decision Making*. Acco, Leuven (Belgium).

[16] Tinga, T. (2010). Application of physical failure models to enable usage and load based maintenance. *Reliability Engineering and System Safety*, 95:1061–1075.

[17] Tinga, T. (2013). *Principles of Loads and Failure Mechanisms: Applications in Maintenance, Reliability and Design*. Springer Science & Business Media.

[18] Van Der Duyn Schouten, F. A. and Vanneste, S. G. (1993). Two simple control policies for a multicomponent maintenance system. *Operations research*, 41(6):1125–1136.

[19] Van Horenbeek, A., Buré, J., Cattrysse, D., Pintelon, L., and Vansteenwegen, P. (2013). Joint maintenance and inventory optimization systems: A review. *International Journal of Production Economics*, 143(2):499–508.

[20] van Noortwijk, J. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety*, 94(1):2 – 21.

[21] Wang, H. (2002). A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3):469–489.

[22] Wang, L., Zheng, E., Li, Y., Wang, B., and Wu, J. (2009). Maintenance optimization of generating equipment based on a condition-based maintenance policy for multi-unit systems. In *Control and Decision Conference, 2009. CCDC'09. Chinese*, pages 2440–2445. IEEE.

[23] Wijnmalen, D. J. and Hontelez, J. A. (1997). Coordinated condition-based repair strategies for components of a multi-component maintenance system with discounts. *European Journal of Operational Research*, 98(1):52–63.

[24] Zhou, X., Xi, L., and Lee, J. (2006). A dynamic opportunistic maintenance policy for continuously monitored systems. *Journal of Quality in Maintenance Engineering*, 12(3):294–305.

[25] Zhu, Q., Peng, H., Timmermans, B., and van Houtum, G.-J. (2017). A condition-based maintenance model for a single component in a system with scheduled and unscheduled downs. *International Journal of Production Economics*, 193:365–380.

[26] Zhu, Q., Peng, H., and van Houtum, G.-J. (2015). A condition-based maintenance policy for multi-component systems with a high maintenance setup cost. *OR Spectrum*, 37:1007–1035.

# Appendix

## A Detailed derivation of the probability of maintenance actions

In Eqs. (5-8) we define the probabilities of each type of maintenance intervention, which is needed to determine the average cost performance per maintenance cycle and the average length of a maintenance cycle. In this appendix we further elaborate these formulas, by replacing $g(u)$ and $h(v)$ using Eqs. (1-2). The benefit of this derivation is that it simplifies the numerical calculations considerably.

$$\mathbb{P}(\mathrm{O}_{\mathrm{PM}}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \frac{1}{\tau\lambda} \begin{cases} \begin{cases} e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(-1+e^{(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_{\mathrm{p}} \\ \begin{cases} e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(-1+e^{(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon}) dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon}) dl_{\xi,\varepsilon},$$

$$\mathbb{P}(\mathrm{O}_{\mathrm{CM}}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left(1 + \frac{1}{\tau\lambda} q(l_{\xi,\varepsilon}, l_{\varepsilon,\mathrm{F}}, \lambda, \tau)\right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon}) dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon}) dl_{\xi,\varepsilon},$$

$$q(l_{\xi,\varepsilon}, l_{\varepsilon,\mathrm{F}}, \lambda, \tau) = \begin{cases} \begin{cases} e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(1-e^{(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}+(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}-\tau)\lambda), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} \leq \tau \\ -1+e^{-\tau\lambda}, & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_{\mathrm{p}} \\ \begin{cases} e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(1-e^{(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}+(l_{\xi,\varepsilon}+t_{\mathrm{p}}-\tau)\lambda), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} \leq \tau \\ -1+e^{-\tau\lambda}, & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}} \end{cases},$$

$$\mathbb{P}(\mathrm{E}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} 0, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_{\mathrm{p}} \\ \begin{cases} \frac{e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(-l_{\xi,\varepsilon}-t_{\mathrm{p}}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} \leq \tau \\ 0, & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon}) dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon}) dl_{\xi,\varepsilon},$$

$$\mathbb{P}(\mathrm{F}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \begin{cases} \frac{e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(-l_{\xi,\varepsilon}-l_{\varepsilon,\mathrm{F}}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} \leq \tau \\ 0, & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_{\mathrm{p}} \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > t_{\mathrm{p}} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon}) dl_{\varepsilon,\mathrm{F}} f^2(l_{\xi,\varepsilon}) dl_{\xi,\varepsilon},$$

$$\mathbb{E}\left\{L_{\mathrm{c}}\right\}=\mathbb{E}\left\{L_{\xi}\right\}$$

$$+\int_{l_{\xi,\varepsilon}=0}^{+\infty}\int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\frac{1}{\tau\lambda^2}q(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)\right)f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}$$

$$+\int_{l_{\xi,\varepsilon}=0}^{+\infty}\int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(\frac{1}{\lambda}+\frac{1}{\tau\lambda^2}r(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)\right)f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}$$

$$+\int_{l_{\xi,\varepsilon}=0}^{+\infty}\int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(s(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)\right)f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}$$

$$+\int_{l_{\xi,\varepsilon}=0}^{+\infty}\int_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty}\left(t(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)\right)f^3(l_{\varepsilon,\mathrm{F}}\mid l_{\xi,\varepsilon})dl_{\varepsilon,\mathrm{F}}f^2(l_{\xi,\varepsilon})dl_{\xi,\varepsilon},$$

with

$$q(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)=$$
$$\begin{cases}\begin{cases}e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(-1+e^{(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}\le\tau\\ e^{-\tau\lambda}(-1+e^{\tau\lambda}-\tau\lambda), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}\le t_{\mathrm{p}}\\ \begin{cases}e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(-1+e^{(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}\le\tau\\ e^{-\tau\lambda}(-1+e^{\tau\lambda}-\tau\lambda), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}>t_{\mathrm{p}}\end{cases},$$

$$r(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)=$$
$$\begin{cases}\begin{cases}e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(2-2e^{(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}+(2(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})-\tau)\lambda+(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}-\tau)(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda^2), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}\le\tau\\ e^{-\tau\lambda}(2-2e^{\tau\lambda}+\tau\lambda), & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}\le t_{\mathrm{p}}\\ \begin{cases}e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(2-2e^{(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}+(2(l_{\xi,\varepsilon}+t_{\mathrm{p}})-\tau)\lambda+(l_{\xi,\varepsilon}+t_{\mathrm{p}}-\tau)(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda^2), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}\le\tau\\ e^{-\tau\lambda}(2-2e^{\tau\lambda}+\tau\lambda), & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}>t_{\mathrm{p}}\end{cases},$$

$$s(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)=$$
$$\begin{cases}0, & \text{if } l_{\varepsilon,\mathrm{F}}\le t_{\mathrm{p}}\\ \begin{cases}\frac{(l_{\xi,\varepsilon}+t_{\mathrm{p}})e^{-(l_{\xi,\varepsilon}+t_{\mathrm{p}})\lambda}(-l_{\xi,\varepsilon}-t_{\mathrm{p}}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}\le\tau\\ 0, & \text{if } l_{\xi,\varepsilon}+t_{\mathrm{p}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}>t_{\mathrm{p}}\end{cases},$$

$$t(l_{\xi,\varepsilon},l_{\varepsilon,\mathrm{F}},\lambda,\tau)=$$
$$\begin{cases}\begin{cases}\frac{(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})e^{-(l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}})\lambda}(-l_{\xi,\varepsilon}-l_{\varepsilon,\mathrm{F}}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}\le\tau\\ 0, & \text{if } l_{\xi,\varepsilon}+l_{\varepsilon,\mathrm{F}}>\tau\end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}}\le t_{\mathrm{p}}\\ 0, & \text{if } l_{\varepsilon,\mathrm{F}}>t_{\mathrm{p}}\end{cases}.$$

Similarly, when only the intervention threshold is used, Eq. (12) can be further simplified by replacing $g(u)$ and $h(v)$ using Eqs. (1-2):

$$\mathbb{P}(\mathrm{O_{PM}}) = \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \frac{1}{\tau\lambda} \begin{cases} \begin{cases} e^{-l_{\varepsilon,\mathrm{F}}\lambda}(-1+e^{l_{\varepsilon,\mathrm{F}}\lambda}), & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}), & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} e^{-t_\mathrm{p}\lambda}(-1+e^{t_\mathrm{p}\lambda}), & \text{if } t_\mathrm{p} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}), & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}},$$

$$\mathbb{P}(\mathrm{O_{CM}}) = \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( 1 + \frac{1}{\tau\lambda} \begin{cases} \begin{cases} e^{-l_{\varepsilon,\mathrm{F}}\lambda}(1-e^{l_{\varepsilon,\mathrm{F}}\lambda}+(l_{\varepsilon,\mathrm{F}}-\tau)\lambda), & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ -1+e^{-\tau\lambda}, & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} e^{-t_\mathrm{p}\lambda}(1-e^{t_\mathrm{p}\lambda}+(t_\mathrm{p}-\tau)\lambda), & \text{if } t_\mathrm{p} \leq \tau \\ -1+e^{-\tau\lambda}, & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}},$$

$$\mathbb{P}(\mathrm{E}) = \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} 0, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} \frac{e^{-t_\mathrm{p}\lambda}(-t_\mathrm{p}+\tau)}{\tau}, & \text{if } t_\mathrm{p} \leq \tau \\ 0, & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}},$$

$$\mathbb{P}(\mathrm{F}) = \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \begin{cases} \frac{e^{-l_{\varepsilon,\mathrm{F}}\lambda}(-l_{\varepsilon,\mathrm{F}}+\tau)}{\tau}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}},$$

$$\begin{aligned}
\mathbb{E}\{L_\mathrm{c}\} &= \mathbb{E}\left\{L_\xi\right\} \\
&+ \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \frac{1}{\tau\lambda^2} \begin{cases} \begin{cases} e^{-l_{\varepsilon,\mathrm{F}}\lambda}(-1+e^{l_{\varepsilon,\mathrm{F}}\lambda}-l_{\varepsilon,\mathrm{F}}\lambda), & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}-\tau\lambda), & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} e^{-t_\mathrm{p}\lambda}(-1+e^{t_\mathrm{p}\lambda}-t_\mathrm{p}\lambda), & \text{if } t_\mathrm{p} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}-\tau\lambda), & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}} \\
&+ \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \frac{1}{\lambda} + \frac{1}{\tau\lambda^2} \begin{cases} \begin{cases} e^{-l_{\varepsilon,\mathrm{F}}\lambda}(2-2e^{l_{\varepsilon,\mathrm{F}}\lambda}+(2l_{\varepsilon,\mathrm{F}}-\tau)\lambda+(l_{\varepsilon,\mathrm{F}}-\tau)l_{\varepsilon,\mathrm{F}}\lambda^2), & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ e^{-\tau\lambda}(2-2e^{\tau\lambda}+\tau\lambda), & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} e^{-t_\mathrm{p}\lambda}(2-2e^{t_\mathrm{p}\lambda}+(2t_\mathrm{p}-\tau)\lambda+(t_\mathrm{p}-\tau)t_\mathrm{p}\lambda^2), & \text{if } t_\mathrm{p} \leq \tau \\ e^{-\tau\lambda}(2-2e^{\tau\lambda}+\tau\lambda), & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}} \\
&+ \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} 0, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ \begin{cases} \frac{t_\mathrm{p} e^{-t_\mathrm{p}\lambda}(-t_\mathrm{p}+\tau)}{\tau}, & \text{if } t_\mathrm{p} \leq \tau \\ 0, & \text{if } t_\mathrm{p} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}} \\
&+ \int\limits_{l_{\varepsilon,\mathrm{F}}=0}^{+\infty} \left( \begin{cases} \begin{cases} \frac{l_{\varepsilon,\mathrm{F}} e^{-l_{\varepsilon,\mathrm{F}}\lambda}(-l_{\varepsilon,\mathrm{F}}+\tau)}{\tau}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq \tau \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > \tau \end{cases}, & \text{if } l_{\varepsilon,\mathrm{F}} \leq t_\mathrm{p} \\ 0, & \text{if } l_{\varepsilon,\mathrm{F}} > t_\mathrm{p} \end{cases} \right) f^3(l_{\varepsilon,\mathrm{F}}) dl_{\varepsilon,\mathrm{F}}.
\end{aligned}$$

When only the opportunistic threshold is used, Eq. (13) can be elaborated by replacing $g(u)$ and $h(v)$ using Eqs. (1-2):

$$\mathbb{P}(\mathrm{O_{PM}}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \frac{1}{\tau\lambda} \begin{cases} e^{-l_{\xi,\varepsilon}\lambda}(-1+e^{l_{\xi,\varepsilon}\lambda}), & \text{if } l_{\xi,\varepsilon} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}), & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon},$$

$$\mathbb{P}(\mathrm{O_{CM}}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} 1 + \frac{1}{\tau\lambda} \begin{cases} e^{-l_{\xi,\varepsilon}\lambda}(1-e^{l_{\xi,\varepsilon}\lambda}+(l_{\xi,\varepsilon}-\tau)\lambda), & \text{if } l_{\xi,\varepsilon} \leq \tau \\ -1+e^{-\tau\lambda}, & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon},$$

$$\mathbb{P}(\mathrm{F}) = \int_{l_{\xi,\varepsilon}=0}^{+\infty} \begin{cases} \frac{e^{-l_{\xi,\varepsilon}\lambda}(-l_{\xi,\varepsilon}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon} \leq \tau \\ 0, & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon},$$

$$\begin{aligned} &\mathbb{E}\{L_\mathrm{c}\} = \mathbb{E}\{L_\xi\} \\ &+ \int_{l_{\xi,\varepsilon}=0}^{+\infty} \frac{1}{\tau\lambda^2} \begin{cases} e^{-l_{\xi,\varepsilon}\lambda}(-1+e^{l_{\xi,\varepsilon}\lambda}-l_{\xi,\varepsilon}\lambda), & \text{if } l_{\xi,\varepsilon} \leq \tau \\ e^{-\tau\lambda}(-1+e^{\tau\lambda}-\tau\lambda), & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}, \\ &+ \int_{l_{\xi,\varepsilon}=0}^{+\infty} \left( \frac{1}{\lambda} + \frac{1}{\tau\lambda^2} \begin{cases} e^{-l_{\xi,\varepsilon}\lambda}(2-2e^{l_{\xi,\varepsilon}\lambda}+(2l_{\xi,\varepsilon}-\tau)\lambda+(l_{\xi,\varepsilon}-\tau)l_{\xi,\varepsilon}\lambda^2), & \text{if } l_{\xi,\varepsilon} \leq \tau \\ e^{-\tau\lambda}(2-2e^{\tau\lambda}+\tau\lambda), & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} \right) f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}, \\ &+ \int_{l_{\xi,\varepsilon}=0}^{+\infty} l_{\xi,\varepsilon} \begin{cases} \frac{e^{-l_{\xi,\varepsilon}\lambda}(-l_{\xi,\varepsilon}+\tau)}{\tau}, & \text{if } l_{\xi,\varepsilon} \leq \tau \\ 0, & \text{if } l_{\xi,\varepsilon} > \tau \end{cases} f(l_{\xi,\varepsilon})dl_{\xi,\varepsilon}. \end{aligned}$$

# B Expressions under a linear degradation process and a Weibull distributed time to failure

In this appendix, we derive $\mathbb{E}\{L_\xi\}$, $f^2(l_{\xi,\varepsilon})$, $f^3(l_{\varepsilon,\mathrm{F}})$, and $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})$ as a function of the opportunistic threshold $\xi$ and intervention threshold $\varepsilon$ for a linear degradation behaviour $\eta(t)$. Recall that $\mathbb{E}\{L_\xi\}$ is defined as the expected degradation time until it reaches the opportunistic threshold $\xi$, $f^2(l_{\xi,\varepsilon})$ the pdf of the degradation time from the opportunistic threshold $\xi$ to the intervention threshold $\varepsilon$, $f^3(l_{\varepsilon,\mathrm{F}})$ the pdf of the degradation time from the intervention threshold $\varepsilon$ until failure, and $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})$ the latter pdf conditional on $l_{\xi,\varepsilon}$.

Without loss of generality, we have defined $\eta(0) = 0$ and $\eta(l) = 1$, which means that upon replacement a component is as-good-as-new and it fails when the degradation is 100%. We assume the pdf of the lifetime $l$ (i.e., the degradation time between replacement and failure) to be a two-parameter Weibull distribution $f(l) = \frac{\beta}{\alpha}\left(\frac{l}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{l}{\alpha}\right)^{\beta}\right]$, $l \geq 0$, with $\alpha$ the characteristic life (scale factor) and $\beta$ the shape parameter.

Under a linear degradation process, $l_{\xi,\varepsilon}$ and $l_{\varepsilon,\mathrm{F}}$ are proportional to the lifetime $l$. Hence, by multiplying the degree of degradation between two degradation levels with the scale parameter $\alpha$ (which is respectively ($\varepsilon - \xi$) for $f^2(l_{\xi,\varepsilon})$ and $(1-\varepsilon)$ for $f^3(l_{\varepsilon,\mathrm{F}})$), we can obtain the pdf of the degradation time between these two levels:

$$f^2(l_{\xi,\varepsilon}) = \frac{\beta}{(\varepsilon - \xi)\,\alpha}\left(\frac{l_{\xi,\varepsilon}}{(\varepsilon - \xi)\,\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{l_{\xi,\varepsilon}}{(\varepsilon - \xi)\,\alpha}\right)^{\beta}\right],\ \ l_{\xi,\varepsilon} \geq 0, \tag{17}$$

$$f^3(l_{\varepsilon,\mathrm{F}}) = \frac{\beta}{(1 - \varepsilon)\,\alpha}\left(\frac{l_{\varepsilon,\mathrm{F}}}{(1 - \varepsilon)\,\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{l_{\varepsilon,\mathrm{F}}}{(1 - \varepsilon)\,\alpha}\right)^{\beta}\right],\ \ l_{\varepsilon,\mathrm{F}} \geq 0. \tag{18}$$

The expected time to reach the opportunistic threshold under a linear degradation process is then:

$$\mathbb{E}\left\{L_\xi\right\} = \xi\mathbb{E}\left\{L\right\} = \xi\alpha\Gamma\left(1 + \frac{1}{\beta}\right), \tag{19}$$

with $\alpha\Gamma\left(1 + \frac{1}{\beta}\right)$ the average of a Weibull distribution.

To determine $f^3(l_{\varepsilon,\mathrm{F}} \mid l_{\xi,\varepsilon})$, we assume that $l_{\xi,\varepsilon}$ and $l_{\varepsilon,\mathrm{F}}$ are perfectly correlated (although our hybrid policy and solution methodology theoretically works without that assumption, we use this assumption to provide numerical results). Under this assumption, $l_{\varepsilon,\mathrm{F}} = \left(\frac{1-\varepsilon}{\varepsilon-\xi}\right) l_{\xi,\varepsilon}$, which defines the proportional relation between $l_{\xi,\varepsilon}$ and $l_{\varepsilon,\mathrm{F}}$.