Real-Time Integrated Learning and Decision Making for Cumulative Shock Degradation

Collin Drent
Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands, PO BOX 513, 5600 MB, c.drent@tue.nl

Melvin Drent
Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, the Netherlands, PO BOX 513, 5600 MB, m.drent@tue.nl

Joachim Arts
Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg, L-1359 Luxembourg-City, Luxembourg, joachim.arts@uni.lu

Stella Kapodistria
Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands, PO BOX 513, 5600 MB, s.kapodistria@tue.nl

Problem Definition: Unexpected failures of equipment can have severe consequences and costs. Such unexpected failures can be prevented by performing preventive replacement based on real-time degradation data. We study a component that degrades according to a compound Poisson process and fails when the degradation exceeds the failure threshold. An online sensor measures the degradation in real-time, but interventions are only possible during planned downtime. **Academic / Practical Relevance:** We characterize the optimal replacement policy that integrates real-time learning from the online sensor. We demonstrate the effectiveness in practice with a case study on interventional X-ray machines. The data set of this case study is made available with this article. As such, it can serve as a benchmark data set for future studies on stochastically deteriorating systems. **Methodology:** The degradation parameters vary from one component to the next but cannot be observed directly; the component population is heterogeneous. These parameters must, therefore, be inferred by observing the real-time degradation signal. We model this situation as a partially observable Markov decision process (POMDP) so that decision making and learning are integrated. We collapse the information state space of this POMDP to three dimensions so that optimal policies can be analyzed and computed tractably. **Results:** The optimal policy is a state dependent control limit. The control limit increases with age but may decrease as a result of other information in the degradation signal. Numerical case study analyses reveal that integration of learning and decision making leads to cost reductions of 10.50% relative to approaches that do not learn from the real-time signal and 4.28% relative to approaches that separate learning and decision making. **Managerial Implications:** Real-time sensor information can reduce the cost of maintenance and unplanned downtime by a considerable amount. The integration of learning and decision making is tractably possible for industrial systems with our state space collapse. Finally, the benefit of our model increases with the amount of data available for initial model calibration while additional data is much less valuable for approaches that ignore population heterogeneity.

Key words: partially observable Markov decision process; optimal policies; Bayesian learning; maintenance, data-driven operations

History: October 10, 2020; Revisions: April 25, 2022; June 22, 2022

1. Introduction

Advanced technical systems are critical for the smooth operation of public services such as public transport (e.g., aircraft, rolling stocks), utilities (e.g., power plants), and health care (e.g., MRI-scanners, X-ray machines) as well as for the primary processes of companies (e.g., data centers, lithography machines). Unavailability and failure of these systems – especially when unplanned – have severe consequences and can even lead to immediate safety hazard. In 2017, for example, the failure of a deteriorating propeller blade led to the crash of a KC-130T aircraft with 16 casualties (Insinna and Ziezuliwicz 2018). A year before that, the nuclear power plant in Doel, Belgium, had to be shut down unexpectedly twice. This caused concerns for citizens in Belgium, the Netherlands and Germany, as well as expensive power imports by the operator Electrabel (Van Soest 2016). When unplanned downtime does not lead to immediate safety hazard, the consequences can still be severe financially. Indeed, recent studies indicate that unplanned downtime costs industrial manufacturers alone $50 billion annually (Wall Street Journal Custom Studios 2017, Coleman et al. 2017). Almost half of that unplanned downtime is caused by equipment failures. Minimizing equipment failures is thus pivotal in reducing both downtime costs and safety hazards. This also seems to be well-understood by executives in asset management. In a recent survey, they perceive unplanned equipment failures as the most important risk to manage (Pacquin 2014).

Condition-based maintenance (CBM) is the most advanced maintenance strategy that asset managers can implement to reduce unplanned equipment failures. Based on real-time condition monitoring, CBM initiates maintenance actions only if equipment failure is imminent, thereby trading off costly premature interventions with costly tardy replacements. Conventional CBM approaches separate estimation and optimization (De Jonge and Scarf 2020). Asset managers first estimate a statistical model of the degradation behavior of an asset based on historical degradation data, and they subsequently optimize the decision of when to perform maintenance based on this model. This separation necessitates the implicit assumption that one asset is statistically identical to any other asset of the same type. Maintenance actions for each asset of the same type are consequently equivalent. Through our collaboration with Philips Healthcare, a major medical equipment manufacturer and service provider, we find however that this assumption is violated in practice. Figure 1 below shows two historical degradation paths of a component that is critical for the operation of an interventional X-ray (IXR) system. While these two degradation paths are from the same type of component, they clearly differ from one another and their maintenance decisions should be tailored accordingly. This illustrates that the traditional approach that hinges on the assumption of a homogeneous asset population is no longer appropriate in practice, increases failure risk, and reduces the useful lifetime of assets. Our case study with Philips Healthcare in Section 7 substantiates these claims.

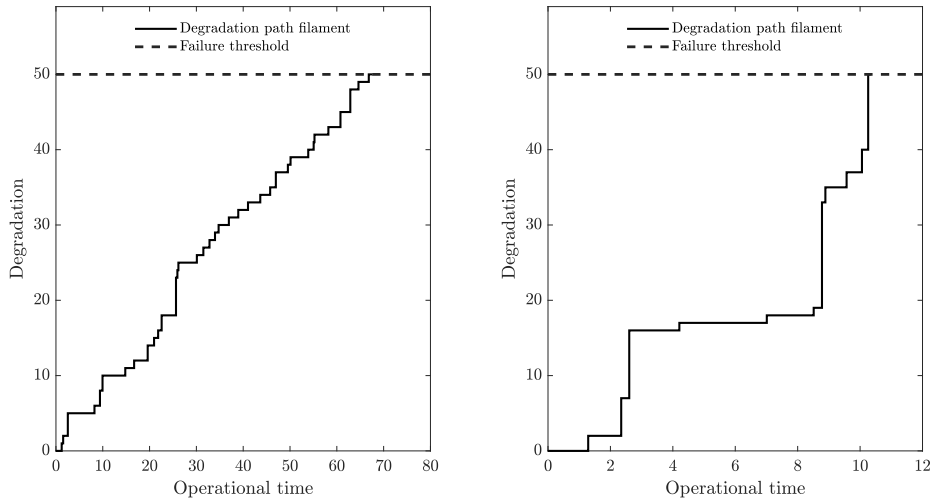


Figure 1 Two historical IXR filament degradation paths.

The separation between estimation and optimization was necessitated by the fact that degradation data of a critical asset could only be obtained by performing expensive measurements during (planned) asset downtime. Recently built high-tech systems, including the IXR systems of our case study, are now increasingly equipped with integrated sensor technology that allows degradation data of a critical component to be gathered at hardly any additional cost (Olsen and Tomlin 2020). These sensors measure the degradation of components in real-time and are integrated into the Internet-of-Things (IoT) (Price Waterhouse Coopers 2014). Data that is generated by sensors and relayed in real-time through the IoT renders the assumption that assets are statistically identical unnecessary. Indeed, real-time degradation data allows us to learn degradation behavior on the individual component level and tailor our decision making accordingly. This assumption has been partially relaxed before, but only in settings where measurements are possible at planned downtimes and not in real-time; see Elwany et al. (2011), Kim and Makis (2013), Chen et al. (2015), and Van Oosterom et al. (2017). When the condition of a system can only be measured at planned downtimes, the amount of information that can be learned from the degradation level is limited compared to the situation with real-time data. Accordingly, attention in the few papers mentioned above is restricted to population heterogeneity within a finite set of possibilities or heterogeneity in degradation drift only, with drift defined as the expected degradation increment per unit of time. By contrast, this paper uses the *entire* degradation path of each component. This allows us to infer higher-order properties of the degradation behavior of the individual component, in particular the volatility, defined as the variance of a degradation increment per unit time (cf. Figure 1). We show numerically that this additional information has substantial value. Grounded on the

widespread adoption of sensor technology and its integration in the IoT, this paper integrates estimation and decision making in real-time to tailor maintenance intervention decisions to high-tech systems individually.

We consider a single component that is critical for operating the system in which it is installed. This component is subjected to stochastic shock deterioration (Esary and Marshall 1973, Sobczyk 1987). Such shock models assume that components accumulate random amounts of damage due to shocks that occur randomly over time. These models are appropriate for, e.g., certain metal and ceramic components in trains, aircraft, and medical equipment (including the IXR systems of our case study) that only deteriorate at events at which they are subjected to shocks. Stochastic wear models, as opposed to shock models, assume that components deteriorate continuously as long as they are being used in operation (Kharoufeh and Cox 2005). Such models are more appropriate for, e.g., rotating machinery and bearings (e.g., Elwany et al. 2011). In this paper, we model random shock deterioration through a general jump process (Van Noortwijk 2009). More specifically, we assume that the sequence of random shocks arrives as a Poisson process with a randomly varying shock size, that is, degradation accumulation is modeled as a compound Poisson process. The compounding distribution is quite general; the only restriction we impose is that this distribution belongs to the natural exponential family with non-negative support. This family includes many well-known distributions such as the geometric distribution and the Poisson distribution.

Components are subject to compound Poisson degradation, but the parameters of the Poisson process as well as the compounding distribution vary from one component to the next. That is, the population of components is heterogeneous. These parameters cannot be observed directly and they therefore need to be learned by observing the degradation signal that is relayed in real-time through sensors on the component. Although we observe the degradation level of a component continuously through condition monitoring, we can only interfere with the system at equally spaced discrete decision epochs (i.e. at planned downtimes). The costs to replace a component after failure are much higher than before because they include the costs of unplanned downtime. This decision problem can be modeled by a partially observable Markov decision process (POMDP). The entire past degradation path of a component is relevant state information in this setting. Dealing with the entire degradation path can lead to tractability issues. We circumvent these issues by using conjugate prior pairs to model the heterogeneity of the component population. We further collapse the state space by identifying structure in the prior to posterior updating procedure. This enables us to tractably compute optimal policies as well as prove structural results about optimal policies.

This paper makes the following contributions:

1. We tractably model the situation where components are heterogeneous in their degradation process by using conjugate prior pairs. We collapse the high dimensional state space to only

3 dimensions while retaining all relevant information. This collapse gives insight into how all relevant information in a real-time degradation signal can be parsimoniously represented. Furthermore, this collapse makes the model both tractable numerically and amenable to structural analysis.

2. We characterize the optimal replacement policy as a threshold replacement policy where the threshold is increasing in the age of a component and furthermore depends on the volatility of the observed degradation signal.
3. In a first simulation study, we study (i) the benefits of explicitly modeling heterogeneity, (ii) the benefits of integrating learning with decision making, and (iii) the impact of the amount of available historical degradation data for estimation of the population heterogeneity on their performance. The results of this simulation study indicate that the integration of learning and decision making leads to excellent results with gaps of only 0.60% on average relative to an oracle that knows the true population heterogeneity. By contrast, ignoring heterogeneity altogether leads to average gaps of 15.02% relative to an oracle that knows the true population heterogeneity. Failing to integrate learning with decision making leads to average gaps of 7.08% relative to that same oracle. Furthermore, we show that models that ignore population heterogeneity do *not* perform appreciably better when the amount of historical degradation data for model calibration increases.
4. In a second simulation study, we assess the performance of the optimal policy (under real-time, perfect data) when applied in a setting where (i) the degradation signal is not perfect, and (ii) the degradation signal is not relayed in real-time. The results indicate that having access to data in real-time is valuable, while at the same time, this real-time data need not be perfect to achieve excellent performance.
5. We demonstrate the efficacy of integrated learning and decision making on a real data set of X-ray tube degradation in an IXR machine. This large normalized data set is made available together with the paper to benchmark future approaches to perform maintenance on degrading systems. This is the first openly available data set containing real degradation data of components from a heterogeneous population. We find that integrated learning can save around 10.50% compared to approaches without learning and around 4.28% compared to an approach where learning is separated from decision making.

The remainder of the paper is organized as follows. We start with a brief literature review in Section 2, and subsequently present the model formulation in Section 3. We characterize the optimal replacement policy in Section 4 and we report on the results of a comprehensive simulation study 5. We discuss the application of our approach to alternate settings in which the degradation signal is imperfect or not relayed in real-time in Section 6. We establish the practical value of our

model in Section 7, where we discuss a real life case study on an IXR machine of Philips Healthcae. We provide concluding remarks in Section 8.

2. Literature review

Maintenance and replacement models have been studied extensively in both management and engineering literature and have been reviewed over the decades (Pierskalla and Voelker 1976, Valdez-Flores and Feldman 1989, Scarf 1997, Wang 2002). The question of when to intervene to maintain or replace a component based on its condition or degradation has been studied already by Derman (1963) and Kolesar (1966). Both establish that, under specific conditions on the degradation process, there exists a threshold such that it is optimal to replace a component if and only if the degradation level of the component is found to have exceeded that threshold. Since these early results, many results about the optimality of threshold policies for the replacement of deteriorating components under many model variations have appeared; see e.g., Ross (1969), Kao (1973), Rosenfield (1976), Benyamini and Yechiali (1999), Makis and Jiang (2003), Maillart (2006) and Kurt and Kharoufeh (2010). Within this literature, the degradation process is assumed to be known to the decision maker and the focus is on proving structural properties of optimal replacement policies.

Elwany et al. (2011) started to study models in which the parameters of the degradation process are only partially known to the decision maker, and differ from component to component. They consider a stochastic wear model and model the degradation process as a Brownian motion in which the drift parameter is initially unknown and comes from a known prior distribution upon replacement of a component. In the same spirit as our model, this prior distribution models the heterogeneity of the component population. During each planned downtime, the degradation is measured and the belief state regarding the drift parameter of the current component is updated. In this situation, the optimal policy is shown to be a threshold policy where the threshold is an increasing function of the age of the component. The intuition is that as a component ages without failing, it is more likely that the drift is not very high. Since then, this result has been extended, yet only to other stochastic wear degradation processes including the inverse Gaussian process (Chen et al. 2015), the gamma process (Zhang et al. 2016), and Markovian processes with a prior distribution on a finite number of possible transition matrices (Van Oosterom et al. 2017). By contrast, in this paper, we are the first to study a stochastic shock model. In particular, our paper studies the situation where all parameters of the stochastic shock degradation process are initially unknown and come from a prior distribution. Furthermore, our model requires that the decision maker can access the entire degradation signal since the last decision epoch, whereas Elwany et al. (2011), Chen et al. (2015), Zhang et al. (2016), Van Oosterom et al. (2017) only require that the decision maker can access the current degradation level. That is, our model exploits the entire

degradation signal, whereas existing models only use measurements that are made just prior to decision epochs. This allows us to infer not only the drift of the degradation process but also higher order properties, in particular the volatility, all of which are unknown to decision makers in real life. Thus, we add to the existing literature described above, by considering a stochastic shock model where more relevant information can be used to learn the parameters, but also more parameters need to be learned (i.e., of both the arrival process and damage distribution).

All the papers in the previous paragraph use, as do we, a partially observable Markov decision process (POMDP) to study their models. This approach has also been used to study when expensive condition measurements should be made (e.g., Kim and Makis 2013, Van Staden and Boute 2021, Khaleghei and Kim 2021) and when spare part inventories should be replenished in anticipation of failures (e.g., Li and Ryan 2011). A more related and recent paper that also involves a POMDP is Kim (2016). He studies a condition based maintenance setting akin to ours but in which the initial priors of the POMDP are mis-specified. The decision maker then seeks a policy that is robust against such mis-specifications. Furthermore, POMDPs have been frequently used in the literature on inventory management where learning is somehow involved. In particular to learn unknown demand distributions from (censored) observations in a single location (e.g., Azoury 1985, Chen and Plambeck 2008, Chen 2010) and across multiple locations (e.g., Chen et al. 2017), as well as to learn inventory levels when records are inaccurate (e.g., DeHoratius et al. 2008, Mersereau 2013) or when there is unobserved inventory shrinkage (e.g., Chen 2021, Li et al. 2022).

The usage of condition monitoring to improve decision making is not limited to the condition based maintenance settings that we study. In fact, there exists a broad stream of literature that studies how condition monitoring can improve general decision making; our paper contributes to this literature stream as well. For example, Uit het Broek et al. (2020) study how the production rate of a production system, which affects the degradation rate of such systems, should be dynamically adjusted based on condition monitoring such that production profits are maximized. Other examples of applications of condition monitoring include spare parts provision based on degradation of components (Olde Keizer et al. 2017), managing rental cars based on their condition (Slaugh et al. 2016), and simultaneously optimizing maintenance and production schedules for multiple products based on machine condition data (Batun and Maillart 2012).

3. Model formulation

In this section we describe the degradation model as well as the integrated learning problem of learning the degradation behavior of a component and deciding when to replace it.

3.1. Compound Poisson degradation

We consider a component that degrades as random shocks arrive. Shocks arrive as a Poisson process and the damage that accumulates during a shock is random variable, i.e., degradation is a compound Poisson process. The Poisson intensity of shock arrivals is denoted by λ. The random amount of damage incurred by a shock follows the law of a member of the one-parameter (denoted by ϕ) exponential family, supported on $\mathbb{R}_+$, where $\mathbb{R}_+$ denotes the non-negative real line. Hence, the probability density or mass function of this random amount can be expressed in the form

$$f(x|\phi) = h(x)e^{\phi T(x) - A(\phi)}, \tag{1}$$

where $T(x)$ is the sufficient statistic, and $h(x)$ and $A(\phi)$ are known functions. We assume that $T(x) \coloneqq x$, which enables a state space collapse in our optimization problem (see Section 3.3). In the literature, this family of distributions is often referred to as the linear (due to the linear sufficient statistic) exponential family or natural exponential family and was first introduced by Morris (1982). This class encompasses many well-known distributions used in maintenance such as the geometric distribution, the Poisson distribution, the gamma distribution with known shape parameter, the negative binomial distribution with known number of failures, and the binomial distribution with known number of trials (see Morris (1982) for a complete overview). The following two examples illustrate how the geometric distribution with unknown success probability $p \in (0,1)$ and the Poisson distribution with unknown mean $\mu > 0$, respectively, can be expressed in the canonical form of the natural exponential family.

Example 1 (Geometric distribution). Let the damages be geometrically distributed with (unknown) success probability $p \in (0,1)$, with support $\mathbb{N}_0 \coloneqq \{0,1,\ldots\}$. The probability mass function of the random amount of damage, denoted by $f(x|p)$, then takes the form

$$f(x|p) = (1-p)^x p = e^{\ln(1-p)x - \ln(1/p)}. \tag{2}$$

Comparing Equation (2) with Equation (1), we find that $h(x) = 1$, $T(x) = x$, $\phi = \ln(1-p)$, and $A(\phi) = \ln(1/p) = \ln\left(1/(1-e^{\phi})\right)$ for the geometric distribution. ♦

Example 2 (Poisson distribution). Let the damages be Poisson random variables with (unknown) mean $\mu > 0$. The probability mass function of the random amount of damage, denoted by $f(x|\mu)$, then takes the form

$$f(x|\mu) = \frac{\mu^x e^{-\mu}}{x!} = \frac{1}{x!} e^{\ln(\mu)x - \mu}. \tag{3}$$

Comparing Equation (3) with Equation (1), we find that $h(x) = \frac{1}{x!}$, $T(x) = x$, $\phi = \ln(\mu)$, and $A(\phi) = \mu = e^{\phi}$ for the Poisson distribution. ♦

For simplicity and due to its practical appeal (see Section 7), throughout this paper, we use the geometric distribution with support $\mathbb{N}_0$ to illustrate further results – building further on Example 1 – but we emphasise that all structural results hold for *any* compounding distribution whose probability density or mass function can be expressed in the form displayed in (1) with $T(x) = x$.

The degradation level is observed continuously, but it is only possible to interfere with the system at equally spaced decision epochs. These decision epochs correspond to planned downtimes. For convenience, we rescale time such that the time between two decision epochs equals 1. Furthermore, there exists a threshold $\xi \in \mathbb{N}_+$, where $\mathbb{N}_+ \coloneqq \{1, 2, \ldots\}$, such that a component has failed if its degradation is equal to or exceeds ξ.

Let $N_{(0,t]} \equiv N_t$ denote the total number of shocks received by a component from the start of its life (i.e., from the installation of the component) up to its *operational* age t. We write here explicitly "operational" to emphasize that this age is the cumulative time that the current component is in use. Components of technical systems such as aircraft and rolling stock, and also the IXR system in our case study that we discuss in Section 7, only sustain shocks when these systems are in use. For aircraft and rolling stock, this is the case when they are in operation, while for the IXR system in our case study, this is during a working day. For components that exhibit this natural on-off behavior in practice, we can – and we will in our case study – interpret the age t as the cumulative operational age. Observe that in light of the memory-less property of the Poisson process, this is also a justifiable interpretation. The number of shocks that arrive between age $t-1 \in \mathbb{N}_0$ and $t \in \mathbb{N}_0$ is denoted by $K_{(t-1,t]} \coloneqq N_t - N_{t-1}$. Observe that integer ages of components coincide with decision epochs. Moreover, let Y_i denote the damage incurred at the $i-$th shock since the installation of the component. The compound Poisson process at component age $t \in \mathbb{N}_0$ satisfies

$$X_t = \sum_{i=1}^{N_t} Y_i = X_{t-1} + \sum_{i=N_{t-1}+1}^{N_t} Y_i, \quad t \in \mathbb{N}_0, \tag{4}$$

where $X_0 = N_0 = 0$, and by definition $\sum_{i=1}^{0} \cdot \equiv 0$. Furthermore, let $\boldsymbol{Y}_t \coloneqq \left(Y_{N_{t-1}+1}\ Y_{N_{t-1}+2}\ \ldots\ Y_{N_t}\right)$ and let $Z_{(t-1,t]} \coloneqq \sum_{i=N_{t-1}+1}^{N_t} Y_i$.

3.2. Learning the degradation model

We assume that each component stems from a heterogeneous population of components in which each component has different degradation parameters λ and ϕ, which are unknown to the decision maker. Hence, the degradation parameters differ from one component to the next. This reflects the fact that the degradation process may be affected by the individual component's endogenous conditions. We treat the parameters λ and ϕ as random variables, denoted by Λ and Φ, which can be inferred with increasing accuracy by observing the degradation signal of the component in a Bayesian manner.

Λ has a Gamma distribution with shape α_0 and scale β_0 (notation: $\Lambda \sim \text{Gamma}(\alpha_0, \beta_0)$) and Φ is distributed according to the general prior for a member of the exponential family (parameterized by a_0 and b_0), with $\alpha_0, \beta_0, a_0, b_0 > 0$, with prior density distribution denoted by $f_\Lambda(\lambda|\alpha_0, \beta_0)$ and $f_\Phi(\phi|a_0, b_0)$, respectively. We refer to α_0, β_0, a_0, and b_0 as the hyperparameters. Upon the installation of a new component, the parameters of the compound Poisson degradation process, λ and ϕ, are drawn from these distributions. Let k_t denote the observed number of shocks a components has sustained between ages $t-1$ and t, i.e., k_t is the realization of $K_{(t-1,t]}$. Furthermore let $\boldsymbol{y}_t \coloneqq \begin{pmatrix} y_t^1 & y_t^2 & \dots & y_t^{k_t} \end{pmatrix}$ be the array of the observed amounts of damage of the shocks sustained, i.e., $\boldsymbol{y}_t$ is the realization of $\boldsymbol{Y}_t$. Finally, let $z_t \coloneqq \sum_{i=1}^{k_t} y_t^i$, be the sustained damage between ages $t-1$ and t, i.e., z_t is the realization of $Z_{(t-1,t]}$. The tuple $\boldsymbol{\theta}_t \coloneqq (k_t, \boldsymbol{y}_t)$ is then the observed degradation signal of a component between ages $t-1$ and t.

The sequential Bayesian updating procedure (also referred to as prior-to-posterior updating, (Ghosh et al. 2007)) works as follows. When a new component is installed, there is no observed degradation signal accumulated yet, and hence Λ and Φ follow independent prior distributions, respectively. This joint prior density distribution, denoted by $f^0_{\Lambda,\Phi}(\lambda, \phi) \coloneqq f_\Lambda(\lambda|\alpha_0, \beta_0) \cdot f_\Phi(\phi|a_0, b_0)$, may be obtained from historical or testing data. (In Section C of the Appendix we discuss an appropriate estimation procedure.) At component age t, we use the observed degradation signal $\boldsymbol{\theta}_t$ and the joint posterior density distribution of Λ and Φ updated at component age $t-1$, say $f^{t-1}_{\Lambda,\Phi}(\lambda, \phi) \coloneqq f_{\Lambda,\Phi}(\lambda, \phi|\boldsymbol{\theta}_0, \dots, \boldsymbol{\theta}_{t-1})$, to derive the newly updated joint posterior distribution of Λ and Φ, denoted by $f^t_{\Lambda,\Phi}(\lambda, \phi) \coloneqq f_{\Lambda,\Phi}(\lambda, \phi|\boldsymbol{\theta}_0, \dots, \boldsymbol{\theta}_t)$.

For tractability purposes, so-called conjugate pairs, which have the appealing computational property that the posterior is in the same family as the prior, are often of interest in prior-to-posterior updating. It is well-known that the gamma distribution is a conjugate prior distribution for the Poisson distribution and that a member of the exponential family has a conjugate prior whose density can be expressed in the form (cf. Ghosh et al. 2007)

$$f_\Phi(\phi|a_t, b_t) = H(a_t, b_t) e^{a_t \phi - b_t A(\phi)}. \tag{5}$$

However, since we infer the joint distribution of Λ and Φ using the same observed degradation signal, it is not evident which form the joint posterior distribution of Λ and Φ takes. Proposition 1 shows that this joint posterior distribution at component age t can be decomposed into two independent distributions of the same form with updated parameters that only depend on the information obtained in the last period ($\boldsymbol{\theta}_t$). All the proofs are given in Appendix A.

PROPOSITION 1. *Given the last observed degradation signal at component age t, $\boldsymbol{\theta}_t = (k_t, \boldsymbol{y}_t)$, and the joint prior distribution $f^{t-1}_{\Lambda,\Phi}(\lambda, \phi) = f_\Lambda(\lambda|\alpha_{t-1}, \beta_{t-1}) \cdot f_\Phi(\phi|a_{t-1}, b_{t-1})$, the joint posterior distribution, $f^t_{\Lambda,\Phi}(\lambda, \phi)$, is equal to $f_\Lambda(\lambda|\alpha_{t-1} + k_t, \beta_{t-1} + 1) \cdot f_\Phi(\phi|a_{t-1} + z_t, b_{t-1} + k_t)$.*

Proposition 1 induces a simple scheme to infer the true parameters of the degradation process of a component with increasing accuracy. The following example illustrates how the parameter of the compounding distribution can be inferred in the case of geometrically distributed damages.

Example 3 (Geometric distribution, continuation of Example 1). We endow a prior on the canonical parameter ϕ with density

$$f_\Phi(\phi|a_t,b_t) = H(a_t,b_t)e^{b_t\phi - a_t A(\phi)} = H(a_t,b_t)e^{\phi b_t}(1-e^\phi)^{a_t},$$

or equivalently, parameterized in terms of p using $p = 1-e^\phi$,

$$f_P(p|a_t,b_t) = H(a_t,b_t)p^{a_t}(1-p)^{b_t},$$

in which we recognize, after normalization, the beta distribution with shape parameter $a_t - 1$ and scale parameter $b_t - 1$. Note also that $a_t = a_0 + \sum_{i=1}^t z_i$ and $b_t = b_0 + \sum_{i=1}^t k_i$. ♦

We now determine the posterior predictive distribution at component age t of the random variable $Z_{(t,t+1]}$ given the learned information contained in $\alpha_t,\ \beta_t,\ a_t$, and b_t.

Lemma 1. *The posterior predictive distribution at component age t of the random variable $Z_{(t,t+1]}$ given the joint posterior distribution of Λ and Φ, $f_\Lambda(\lambda|\alpha_t,\beta_t)\cdot f_\Phi(\phi|a_t,b_t)$, is equal to:*

$$\begin{aligned}&\mathbb{P}\Big[Z_{(t,t+1]} = z|\alpha_t,\beta_t,a_t,b_t\Big]\\ &\qquad = \sum_{k=0}^{\infty}\int_{-\infty}^{+\infty} f^{(k)}(z|\,\Phi=\phi)f_\Phi(\phi|a_t,b_t)d\phi\binom{k+\alpha_t-1}{k}\left(\frac{1}{\beta_t+1}\right)^k\left(\frac{\beta_t}{\beta_t+1}\right)^{\alpha_t}, \quad (6)\end{aligned}$$

where $f^{(k)}(z|\,\Phi=\phi)$ denotes the k-fold convolution of the probability density (or mass) function of the random variable $\{Y|\,\Phi=\phi\}$.

Lemma 1 can be used to construct an updated posterior predictive distribution at each component's age t of the next observed damage increment in real-time based on the observed degradation signal. Hence, the posterior distribution of the degradation parameters of the system is a Markov process whose evolution is induced by the degradation trajectory of the current component.

At first sight, the posterior predictive distribution in Lemma 1 seems rather intractable due to the convolution term involved. Fortunately, members of the natural exponential family with a linear sufficient statistic are closed under convolution with itself and hence possess a tractable form (Morris 1982). Upon insertion of the expression for this convolution term and the corresponding conjugate prior, the posterior predictive distribution reduces to a closed-form expression that can be used for computational purposes. This is illustrated in the example below.

EXAMPLE 4 (GEOMETRIC DISTRIBUTION, CONTINUATION OF EXAMPLE 3). In this example we use the parameterization in terms of unknown success probability p, which we treat as a random variable denoted by P. Due to the discrete nature of the geometric distribution and as $p \in (0,1)$, we have by Lemma 1 that

$$\begin{aligned}\mathbb{P}\Big[Z_{(t,t+1]} = z \Big| K_{(t,t+1]} = k, \alpha_t, \beta_t, a_t, b_t\Big] &= \int_0^1 \mathbb{P}\Big[\sum\nolimits_{i=1}^{k} Y_i = z \Big| P = p\Big] f_P(p|a_t, b_t) dp \\ &= \frac{1}{\mathrm{B}(a_t, b_t)} \binom{z+k-1}{z} \int_0^1 p^k (1-p)^z p^{a_t - 1} (1-p)^{b_t - 1} dp \\ &= \frac{\mathrm{B}(k + a_t, z + b_t)}{\mathrm{B}(a_t, b_t)} \binom{z+k-1}{z}, \end{aligned} \tag{7}$$

where $\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$ is the beta function. Note that the distribution of $K_{(t,t+1]}$ is a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate follows a Gamma(α_t, β_t) distribution, which is known to be the negative binomial distribution with success probability $q = \frac{1}{\beta_t + 1}$ and $r = \alpha_t$ number of required successes. Hence, we have

$$\mathbb{P}\Big[K_{(t,t+1]} = k \Big| \alpha_t, \beta_t\Big] = \binom{k + \alpha_t - 1}{k} \left(\frac{1}{\beta_t + 1}\right)^k \left(\frac{\beta_t}{\beta_t + 1}\right)^{\alpha_t}. \tag{8}$$

Unconditioning Equation (7) using Equation (8) yields the closed form expression of the posterior predictive distribution:

$$\begin{aligned}&\mathbb{P}\Big[Z_{(t,t+1]} = z | \alpha_t, \beta_t, a_t, b_t\Big] \\ &\qquad = \sum_{k=0}^{\infty} \frac{\mathrm{B}(k + a_t, z + b_t)}{\mathrm{B}(a_t, b_t)} \binom{z+k-1}{z} \binom{k + \alpha_t - 1}{k} \left(\frac{1}{\beta_t + 1}\right)^k \left(\frac{\beta_t}{\beta_t + 1}\right)^{\alpha_t}. \end{aligned}$$

♦

3.3. Markov decision process formulation

Each component will incur a cost due to either corrective or preventive replacement. If the degradation level at a decision epoch is greater than or equal to the failure threshold ξ, then the failed component is replaced correctively at cost c_u. If the degradation level at a decision epoch does not exceed ξ, then we can either perform preventive replacement at cost c_p, or continue to the next decision epoch at no cost. We assume that replacements take negligible time and that $0 < c_p < c_u < \infty$ to avoid trivial cases.

Recall that each component stems from a heterogeneous population that includes components with different degradation parameters λ and ϕ. Note that these parameters cannot be observed; only the degradation signal $\boldsymbol{\theta}_i$ for $i = 1, \cdots, t$ is observable at component age t. We will therefore use a POMDP to model the integrated problem of learning the degradation parameters of a component and deciding when to replace. First we observe that due to the results in the previous section, the

information state of a component at age t can be represented by $(\alpha_t, \beta_t, a_t, b_t)$. Furthermore, the decision maker knows the current degradation level x. The state at decision epoch τ is therefore given by $(x, \alpha, \beta, a, b, \tau)$ where x denotes the degradation level of the component that is in service and α, β, a, and b encode the most current degradation information of the component that is in service. This six dimensional state representation can be collapsed into an equivalent four dimensional state representation (x, n, t, τ) where n denotes the number of shocks that the current component has sustained and t denotes its age. Indeed observe that by Proposition 1 we have $\alpha = \alpha_0 + n$, $\beta = \beta_0 + t$, $a = a_0 + x$, and $b = b_0 + n$. This representation is insightful: All the information in the degradation signal is encoded in the total degradation level, the number of shocks sustained, and the age of the component. The crucial assumption for this collapse is that the sufficient statistic $T(x)$ for the damage distribution is equal to x, see Equation (1). If the sufficient statistic would not be linear in the damage, then the state space would need to include all individual damage arrivals.

We are interested in finding the optimal replacement policy $\pi^* = \{\pi_\tau\}_{\tau \in \mathbb{N}_0}$ that minimizes the total expected discounted cost of corrective and preventive replacements over an infinite horizon, where the costs are discounted by a factor $\gamma \in (0, 1)$. This policy is a sequence of decision rules that prescribe whether or not to perform preventive maintenance if $x < \xi$. By Proposition 1.2.2. of Bertsekas (2007), there exists an optimal Markov (deterministic) policy depending only on state (x, n, t) independent of the decision epoch $\tau \in \mathbb{N}_0$. Let

$$V(\boldsymbol{s}) = \inf_{\pi \in \Pi} \lim_{T \to \infty} \mathbb{E}_{\boldsymbol{s}} \left[\sum_{\tau=1}^{T} \gamma^\tau C\left(\boldsymbol{S}_\tau, \pi(\boldsymbol{S}_\tau)\right) \right] \tag{9}$$

denote the expected total discounted cost given that the process starts in state $\boldsymbol{s} = (x, n, t)$ where $\boldsymbol{S}_\tau$ denotes the state of the component operating at decision epoch τ, Π denotes the set of Markov (deterministic) policies, $\mathbb{E}_{\boldsymbol{s}}$ is the conditional expectation given that the process starts in state $\boldsymbol{s} = (x, n, t)$, and $C(\boldsymbol{s}, \pi(\boldsymbol{s}))$ denotes the cost function defined as

$$C(\boldsymbol{s}, \pi(\boldsymbol{s})) = \begin{cases} c_p & \text{if } x < \xi \text{ and } \pi(\boldsymbol{s}) = \text{replace}, \\ 0 & \text{if } \ x < \xi \text{ and } \pi(\boldsymbol{s}) = \text{continue}, \\ c_u & \text{if } x \geq \xi. \end{cases}$$

Due to this state space collapse, the posterior distribution of $Z_{(t,t+1]}$ is function of the state $\boldsymbol{s} = (x, n, t)$. Therefore we will use the shorthand notation $Z(\boldsymbol{s})$ to indicate the random variable $Z_{(t,t+1]}$ given the state $\boldsymbol{s} = (x, n, t)$. Similarly, we use the shorthand notation $K(\boldsymbol{s})$ and $Y(\boldsymbol{s})_i$ for the random variable $K_{(t,t+1]}$ and the random variables $\{Y_i\}_{i \in \mathbb{N}}$, given state $\boldsymbol{s} = (x, n, t)$. Note that $Z(\boldsymbol{s})$ and $K(\boldsymbol{s})$ are dependent random variables. This is intuitively clear as damage can only accumulate when the component sustains shocks. To keep our notation concise, it is convenient to define the multivariate random variable $A(\boldsymbol{s}) \coloneqq (Z(\boldsymbol{s}), K(\boldsymbol{s}), 1)$. The conditional distributions (on $\boldsymbol{s}$) of the

first two entries of this multivariate random variable can be determined with Proposition 1 and Lemma 1, while the use of the third entry shortly becomes apparent. The optimal replacement policy π^* satisfies the Bellman optimality equations:

$$V(\boldsymbol{s}) = \begin{cases} c_u + \gamma \mathbb{E}\left[V(\boldsymbol{s_0} + A(\boldsymbol{s}_0))\right], & \text{if } x \geq \xi, \\ \min\Big\{ c_p + \gamma \mathbb{E}\left[V(\boldsymbol{s}_0 + A(\boldsymbol{s}_0))\right]; \gamma \mathbb{E}_{\boldsymbol{s}}\left[V(\boldsymbol{s} + A(\boldsymbol{s}))\right] \Big\}, & \text{if } x < \xi, \end{cases} \tag{10}$$

where $\boldsymbol{s}_0 := (0, 0, 0)$.

The first case in Equation (10) follows because failed components must be replaced correctively at cost c_u. If the component's degradation level is less than ξ, we can either perform preventive replacement at cost c_p, or leave the component in operation until the next decision epoch at no cost. Upon preventive or corrective replacement, the parameters of the new component are unknown and need to be learned anew as the replacement component ages. The aging of the component is captured in the third entry of the random variable $A(\boldsymbol{s})$, where $\boldsymbol{s} = (\cdot, \cdot, t)$ always moves to $(\cdot, \cdot, t+1)$ at the next decision epoch, regardless of the decision. In the remainder of this paper, we refer to $V(\boldsymbol{s})$ as the value function.

4. Optimal replacement policy

In this section, we present structural results of the optimal replacement policy. We first prove that the value function behaves monotonically in its state variables, which we then use to prove that the optimal replacement policy is a control-limit policy with certain monotonic properties. We highlight these theoretical results with numerical examples.

In Lemma 1, we establish how the decision maker can update the posterior predictive distribution of the upcoming degradation increment by utilizing the observed degradation signal $\boldsymbol{s}$. The next result, Proposition 2, presents two properties that provide insight into what the decision maker knows about the distribution of future degradation increments conditionally on the observed degradation signal.

PROPOSITION 2. *The random variables $Z(\boldsymbol{s})$ satisfy the following stochastic orders:*

(i) $Z(\boldsymbol{s})$ is stochastically decreasing in t in the usual stochastic order;

(ii) $Z(\boldsymbol{s})$ is stochastically increasing in x in the usual stochastic order.

Part (i) of Proposition 2 shows that older components will accumulate – in expectation – less damage than younger components. As a component ages without failing, the decision maker infers that large increments are unlikely to happen for this component. The intuition behind Part (ii) of Proposition 2 is that when more damage has accumulated already, then we should expect more damage to accumulate in the future.

In contrast to x and t, there is no monotone stochastic ordering of $Z(\boldsymbol{s})$ in n in general, as is shown in Example 5. In this example we show that when the damages are geometrically distributed, the first two moments have opposing monotonic behaviors. This contrasts the prerequisite for stochastic ordering that these are aligned (see, e.g., Shaked and Shanthikumar 2007, Theorem 1.A.3).

EXAMPLE 5 (GEOMETRIC DISTRIBUTION, CONTINUATION OF EXAMPLE 4). Consider again the geometric compounding distribution with a beta prior. Let $b_0 - \alpha_0 - 1 > 0$, then the expectation of $Z(\boldsymbol{s})$.

$$\mathbb{E}[Z(\boldsymbol{s})] = \frac{a(\alpha_0 + n)}{\beta(b_0 + n - 1)}$$

increases in n. However, the second moment

$$\mathbb{E}[Z(\boldsymbol{s})^2] = \frac{a(\alpha_0 + n)}{\beta^2(b_0 + n - 1)^2} \left(\beta^3 \left(\frac{2(a+1)}{b_0 + n - 2} + \beta(b_0 + n) + 2b_0 + a - \beta + 2n \right) + a(\alpha_0 + n) \right)$$

is not monotonically increasing. (The derivation of these two moments is provided in Appendix B). For example, choosing $b_0 = 2.62$, $\alpha_0 = \beta = 1$ and $a = 10$ yields that the second moment is increasing for all $n \leq 15$ and decreasing for all $n \geq 16$. We can therefore conclude that there is no monotone stochastic ordering of the random variables $Z(\boldsymbol{s})$ in n. ♦

The following two properties of the value function are essential to establish the structure of the optimal replacement policy. As Proposition 2 is pivotal to establish these two properties, there is again no monotonic behavior of the value function in n in general.

LEMMA 2. *The value function $V(\boldsymbol{s})$ is*

(i) non-increasing in t;

(ii) non-decreasing in x.

We are now in the position to establish the main result of this section: the optimality of a control limit policy, as well as the monotonic structure of this control limit.

THEOREM 1. *At each component age $t \in \mathbb{N}_0$, for a given number of shock arrivals $n \in \mathbb{N}_0$, there exists a control limit $\delta^{(n,t)} \leq \xi$, such that the optimal action is to carry out a preventive replacement if and only if $x \geq \delta^{(n,t)}$. The control limit $\delta^{(n,t)}$ is monotonically non-decreasing in t, for all n.*

The optimal control limit is non-decreasing in t because as the component ages without failing, the decision maker is increasingly assured that the current component degrades slowly relative to the general population of components.

Figure 2 illustrates the control limit policy and its monotonic behavior, while Figure 3 displays an example in which this optimal control limit is non-monotonic in the number of shocks sustained. In

these examples, we again use the geometric distribution as choice for the compounding distribution, the solid black line depicts the observed degradation path as the component ages, and a star denotes a shock arrival. The dashed line depicts the optimal control limit, where at each decision epoch, the optimal action is to carry out a preventive replacement if the degradation level is at or above the optimal control limit. The failure level ξ is depicted with the dot-dashed line, so that a corrective replacement has to be performed if the solid black line is at or above this dot-dashed line.

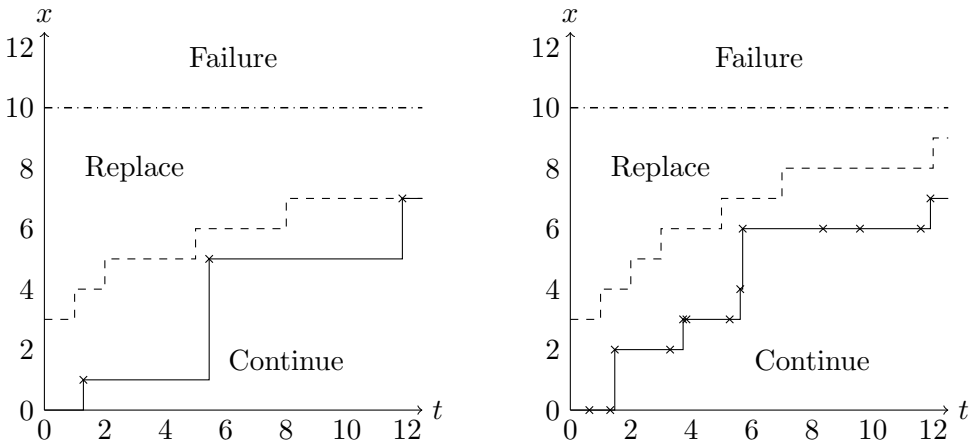


Figure 2 **Two sample degradation paths and the optimal replacement policy, with $\alpha_0 = \beta_0 = a_0 = b_0 = 1$, $c_u = 4$, $c_p = 1$ and $\xi = 10$. For the compound Poisson processes that generate the sample paths, we have $\lambda = 0.25$ and $\phi = 0.2$ (left), and $\lambda = 1.5$ and $\phi = 0.8$ (right).**

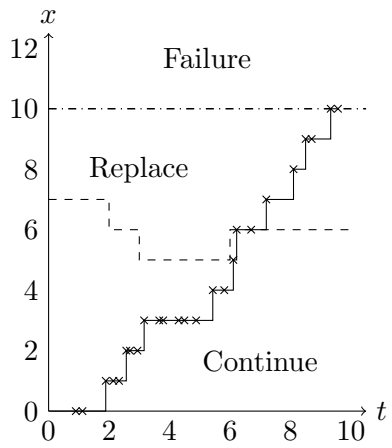


Figure 3 **Sample degradation path and the optimal replacement policy, with $\alpha_0 = \beta_0 = 1$ and $a_0 = 60$ and $b_0 = 10$, $c_u = 4$, $c_p = 1$ and $\xi = 10$. For the compound Poisson process that generates the sample path, we have $\lambda = 1.75$ and $\phi = 0.9$.**

The structural results are not only intuitive and convenient for the implementation of an optimal policy in practice. They can also be exploited to decrease the computational burden of finding the optimal policy by employing existing algorithms that rely on these monotonicity properties such as the monotone policy iteration algorithm (see Puterman 2005, Section 6.11.2).

5. Simulation study

This section reports the results of a comprehensive simulation study. Although the established structural results hold for any one-parameter member of the exponential family with positive support, we assume in this section, motivated by practice, that the damages are geometrically distributed. This simulation study starts from the premise that the true hyperparameters of the degradation behavior are unknown to the decision maker; they only have access to historical degradation data for model calibration. This premise differs from previous contributions that use Bayesian techniques to model real-time learning, where hyperparameters for prior distributions are generally assumed to be given (e.g., Chen 2010). This latter approach is, however, arguably not the case in practice. Indeed, decision makers only have access to historical degradation data that they should leverage in order to estimate hyperparameters. The main objective of this simulation study is, therefore, twofold:

1. To examine the value of integrating learning and decision making, which takes into account explicitly that degradation of components is heterogeneous (value of integration).
2. To assess how the amount of available historical degradation data that is used for model calibration affects the performance (value of data).

To assess the value of integration and data, we define three heuristic approaches, all of which start from the same historical degradation data, but differ in (i) the model calibration, (ii) the learning, and (iii) the integration of learning and decision making. We shall compare the performance of each approach with the performance of an oracle who does know the true hyperparameters of the degradation behavior of different components. The oracle thus follows the replacement policy that solves the Bellman optimality equations in (10), calibrated with the true hyperparameters. We denote the oracle policy by $\pi_{\mathcal{O}}$.

For each instance of the simulation study, which we describe in detail later, the true hyperparameters of the gamma and beta prior distribution of Λ and Φ that model the population heterogeneity are denoted by $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{a}$, and $\tilde{b}$. These are known only to the oracle. The historical degradation data that serves as starting point for the heuristic approaches of that same instance is obtained by simulating degradation paths of components whose degradation parameters λ and ϕ are drawn from a gamma distribution with the true hyperparameters $\tilde{\alpha}$ and $\tilde{\beta}$, and a beta distribution with $\tilde{a}$ and $\tilde{b}$, respectively. We now proceed with defining the three heuristic approaches, after which we describe the simulation set-up and discuss the results.

5.1. Offline approach

The first heuristic approach ignores both the population heterogeneity and the real-time degradation signal, which is the current state-of-the-art. This approach assumes that the degradation of each component upon installation follows a compound Poisson process with the same parameters λ and ϕ. Under this assumption, the decision maker faces the classical replacement problem for which we know that the optimal replacement policy is given by a stationary control limit (e.g., Kolesar 1966). The optimal control limit can be readily found by solving the following one dimensional Bellman optimality equations:

$$V(x) = \begin{cases} c_u + \gamma \mathbb{E}\left[V(\tilde{Z}(\lambda,\phi))\right], & \text{if } x \geq \xi, \\ \min\left\{c_p + \gamma \mathbb{E}\left[V(\tilde{Z}(\lambda,\phi))\right]; \gamma \mathbb{E}\left[V(x+\tilde{Z}(\lambda,\phi))\right]\right\}, & \text{if } x < \xi, \end{cases} \tag{11}$$

where $\tilde{Z}(\lambda,\phi) := \sum_{i=1}^{K(\lambda)} Y_i(\phi)$ denotes the degradation increment in between two consecutive decision epochs, $K(\lambda)$ is a Poisson distributed random variable with parameter λ and $Y_i(\phi)$ are geometrically distributed with parameter ϕ.

Hence, under the first approach, the decision maker approximates the degradation parameters by the corresponding point estimates $\bar{\lambda}$ and $\bar{\phi}$, which are obtained using Maximum Likelihood Estimation (MLE) based on the available historical degradation data, and then solves the optimality equations (11) with those estimates to obtain a single control limit, that is used for all components. In the remainder of this section, we refer to this approach as the offline approach because it ignores both the population heterogeneity and the real-time degradation signal. We denote this approach by $\pi_{\mathcal{N}}$ since it is the most naive heuristic approach of all three approaches, yet it is common practice.

5.2. Open-loop feedback approach

The second heuristic approach does utilize the real-time degradation signal to update the point estimates of the degradation parameters of an individual component. As such, this approach takes into account that the population of components is heterogeneous. This approach is calibrated as follows. Given the historical degradation paths, we estimate the initial hyperparameters α_0, β_0, a_0, and b_0, by maximizing the likelihood of those degradation paths being induced by components stemming from a population whose heterogeneity is modeled through a gamma and beta distribution with these hyperparameters. Thus, based on the available historical degradation data, we estimate the initial hyperparameters α_0, β_0, a_0, and b_0, using MLE (further details regarding this MLE procedure are relegated to Appendix C).

After calibration, this heuristic approach works as follows: At component age t, the decision maker updates the information state encoded in $f_\Lambda(\lambda|\alpha_t,\beta_t)$ and $f_\Phi(\phi|a_t,b_t)$ in the same way as in

the original Bayes model (cf. Proposition 1). The decision maker then updates the point estimates of the degradation increment based on the minimum mean square error (MMSE) estimator. In a Bayesian setting, MMSE estimates correspond to posterior means. Hence, at component age t, the MMSE estimates for the degradation parameters are given by $\bar{\lambda}_t := \int_0^\infty u f_\Lambda(u|\alpha_t, \beta_t)du = \alpha_t/\beta_t$ and $\bar{\phi}_t := \int_0^1 u f_\Phi(u|a_t, b_t)du = a_t/(a_t + b_t)$. The decision maker then computes a control limit by solving the optimality equations (11), where the parameters of the Poisson and of the geometric distribution of the degradation increment $\tilde{Z}(\lambda, \phi)$ are now given by $\bar{\lambda}_t$ and $\bar{\phi}_t$, respectively. Although the second approach partly captures the Bayesian learning benefits, it does not integrate learning with optimization but rather solves a myopic optimization problem repeatedly with the latest point estimates of the degradation parameters. Such a myopic policy is known as an open-loop feedback policy in the stochastic control literature, see, e.g., Bertsekas (2007). In the remainder of this section, we therefore refer to this approach as the open-loop feedback policy, denoted by $\pi_{\mathcal{F}}$.

5.3. Integrated Bayes approach

The third heuristic approach is the most sophisticated approach. It is similar to $\pi_{\mathcal{O}}$ in that it follows the replacement policy that solves the Bellman optimality equations in (10). However, as we have argued before, the true hyperparameters that model the population heterogeneity are unknown to the decision maker in practice. As such, this approach differs from $\pi_{\mathcal{O}}$ in that we calibrate this approach using MLE, in the same way as how we calibrate $\pi_{\mathcal{F}}$. That is, we estimate the initial hyperparameters α_0, β_0, a_0, and b_0, using MLE based on the available historical degradation data. These estimated hyperparameters are then used as input for finding the optimal replacement policy through solving the Bellman optimality equations in (10). We henceforth refer to this approach as the integrated Bayes approach, denoted by $\pi_{\mathcal{I}}$. It is important to note that this approach is precisely how the results of this paper should be applied in practice, where the hyperparameters describing the heterogeneity of the component population should be estimated from historical degradation data.

5.4. Results

The main performance metric in this simulation study is the gap between the long run average cost rate induced by the oracle policy and the offline approach, the open-loop feedback approach, and the integrated Bayes approach, where the latter three models are calibrated based on MLE. More formally, we are interested in $\%GAP_\pi = 100 \cdot (C_\pi - C_{\pi_{\mathcal{O}}})/C_{\pi_{\mathcal{O}}}$, where C_π is the long run average cost rate of approach $\pi \in \{\pi_{\mathcal{I}}, \pi_{\mathcal{F}}, \pi_{\mathcal{N}}\}$.

To achieve the two main objectives stated in the beginning of this section, we set up a large test-bed consisting of instances obtained through all combinations of the parameter values in Table 1, with $\xi = 20$ and $c_p = 1$. To vary the heterogeneity in the population of components, we naturally

vary the coefficient of variation of the gamma and beta prior distributions of Λ and Φ, respectively, while keeping their respective means fixed at 1 and 0.5 for all instances.

Table 1 Input parameter values for simulation study.

	Input parameter	No. of choices	Values
1	Coefficient of variation of prior gamma distribution, cv_Λ	2	0.3, 0.6
2	Coefficient of variation of prior beta distribution, cv_Φ	2	0.01, 0.02
3	Corrective maintenance cost, c_u	2	5, 10
4	Number of simulated degradation paths, ν	2	10, 50

For each instance of the test-bed, we determine the true hyperparameters $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{a}$, and $\tilde{b}$, based on the coefficient of variation of the gamma and beta distribution of that instance (and their fixed means).

We use these hyperparameters to compute $\pi_{\mathcal{O}}$ and we also simulate a number of degradation paths corresponding to that instance. We subsequently use these simulated degradation paths to calibrate approaches $\pi_{\mathcal{I}}$, $\pi_{\mathcal{F}}$ and $\pi_{\mathcal{N}}$. That is, we use MLE to estimate α_0, β_0, a_0, and b_0, in case of $\pi_{\mathcal{I}}$ and $\pi_{\mathcal{F}}$, and point estimates $\bar{\lambda}$ and $\bar{\phi}$, in case of $\pi_{\mathcal{N}}$. We then simulate $15 \cdot 10^3$ components, where upon installation of a new component, its degradation parameters are drawn from a gamma and beta distribution with the true hyperparameters $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{a}$ and $\tilde{b}$. For each simulated component, we keep track of the relevant cost rates and subsequently calculate $\%GAP_\pi$ for each approach $\pi \in \{\pi_{\mathcal{I}}, \pi_{\mathcal{F}}, \pi_{\mathcal{N}}\}$. We repeat this procedure 30 times for each instance of the test-bed to ensure that the confidence intervals of the cost rates are sufficiently small. Throughout this simulation study, we use a discount factor of 0.99 in solving the corresponding optimality equations of each approach, and we truncate both state variables n and t at a sufficiently large value (i.e., 40) in computing the optimal policy under both $\pi_{\mathcal{I}}$ and $\pi_{\mathcal{O}}$.

The results of the simulation study are summarized in Table 2. In this table, we present the minimum, average, and maximum $\%GAP_\pi$ for each approach $\pi \in \{\pi_{\mathcal{I}}, \pi_{\mathcal{F}}, \pi_{\mathcal{N}}\}$. We first distinguish between subsets of instances with the same value for a specific input parameter of Table 1 and then present the results for all instances.

The following main observations can be drawn from Table 2: First, $\pi_{\mathcal{I}}$ yields excellent results with a gap of only 0.60% on average relative to the oracle. Both heuristic approaches $\pi_{\mathcal{F}}$ and $\pi_{\mathcal{N}}$ perform poorly, with gaps of 7.08% and 15.02% on average relative to the oracle. Ignoring both the degradation signal and the heterogeneity in the population can be quite detrimental, as gaps of up to 24.04% relative to the oracle do occur under $\pi_{\mathcal{N}}$. Although learning the degradation signal is beneficial, failing to integrate this with decision making directly can still lead to gaps with the oracle of up to 11.43%. All three approaches $\pi_{\mathcal{I}}$, $\pi_{\mathcal{F}}$, and $\pi_{\mathcal{N}}$ seem to perform worse when the

Table 2 Results of simulation study

		$\%GAP_\pi$								
		$\pi_{\mathcal{I}}$			$\pi_{\mathcal{F}}$			$\pi_{\mathcal{N}}$		
Input	Value	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
cv_Λ	0.3	0.05	0.58	1.26	3.52	6.14	10.05	7.91	13.97	22.93
	0.6	0.15	0.63	1.34	5.24	8.02	11.43	10.16	16.08	24.04
cv_Φ	0.01	0.05	0.46	1.02	3.52	5.54	8.11	7.91	11.31	15.25
	0.02	0.25	0.75	1.34	5.83	8.62	11.43	13.85	18.73	24.04
c_u	5	0.05	0.48	1.16	3.52	5.66	7.74	7.91	12.01	15.73
	10	0.24	0.72	1.34	5.21	8.50	11.43	10.83	18.03	24.04
ν	10	0.42	0.95	1.34	3.97	7.23	11.43	8.63	15.46	24.04
	50	0.05	0.26	0.49	3.52	6.93	11.36	7.91	14.58	23.00
Total		0.05	0.60	1.34	3.52	7.08	11.43	7.91	15.02	24.04

heterogeneity in the population increases, and when the cost of performing corrective maintenance becomes higher.

Second, it is generally believed that increasing the amount of data available for model calibration leads to better decisions. However, when the underlying assumptions of the model are wrong, then this may not be true. Indeed, the performance of both heuristic approaches $\pi_{\mathcal{F}}$ and $\pi_{\mathcal{N}}$ does not increase considerably in the number of simulated degradation paths that serve as input to these models. The performance of $\pi_{\mathcal{I}}$ increases however significantly when this number increases. Furthermore, even when the amount of available data for estimating the population heterogeneity is limited, the integrated Bayes approach, $\pi_{\mathcal{I}}$, still yields excellent results.

6. Alternate settings

We have so far assumed that the degradation signal is relayed in real-time and that it provides a perfect observation of the actual degradation level. These assumptions are in line with what we have observed at our industrial partner who instigated this research (see also the case study in the next section). We note that there may be practical settings where such assumptions are not justified. It is however intractable to relax these assumptions and consequently compute optimal policies within our current modelling framework. In this section, we therefore study the performance of our integrated Bayes approach when it is applied to settings where the degradation signal is imperfect or relayed periodically. We do so in a simulation study that follows the same procedure as in the previous section.

6.1. Imperfect degradation signal

In line with most of the research on imperfect condition monitoring (e.g., Maillart 2006, Kim and Makis 2013), we model an imperfect degradation signal by constructing a state-observation matrix $\mathbf{Q}$ that captures the stochastic relationship between the actual degradation level X_t and

the (imperfect) observation, denoted with $\bar{X}_t$, that the decision maker observes. More formally, let $\mathbf{Q} \coloneqq \left(q_{ij}\right)_{\xi\times\xi}$, whose entry $q_{ij} \coloneqq \mathbb{P}\left[\bar{X}_t = j \mid X_t = i\right]$ is equal to

$$\frac{\varphi_i(j|\sigma)}{\sum_{j=0}^{\xi-1}\varphi_i(j|\sigma)}, \tag{12}$$

where $\varphi_x(y|\sigma)$ is the density function of a normal random variable with mean x and standard deviation σ evaluated at y. Recall that in our simulation study, we assume that the damages are integer valued, so that we have ξ non-failed states, i.e. $0, 1, \ldots, \xi-1$, that are not observable, and one failed state that is observable. The matrix $\mathbf{Q}$ applies to the ξ non-failed states. Constructing this matrix using the parameterization in (12) is an approach often used in literature (see, e.g., Maillart 2006, Kim and Makis 2013, Liu et al. 2021), where the standard deviation σ is a measure of the noise (or imperfectness) of the observations when the system has not failed yet. By varying the value of σ we can thus investigate the robustness of our approach to the level of imperfectness of the observations. Note that by setting σ equal to zero, we are in the situation that we have a perfect observation of the actual degradation level.

6.2. Intermittent degradation signal

When the degradation signal is relayed periodically at decision epochs, the decision maker has only access to the degradation level of the current component x_t (and its age t). However, to apply the integrated Bayes approach we require the number of sustained loading epochs n_t too. To resolve this, we rely on the following recursive formula to obtain a proxy for n_t:

$$\bar{n}_t = \frac{x_t(a_0 + x_{t-1} - 1)}{b_0 + \bar{n}_{t-1}}, \qquad t > 0, \tag{13}$$

and $\bar{n}_0 = 0$. We round to the nearest integer in case $\bar{n}_t$ is not integral. This procedure has intuitive appeal. To obtain a proxy for n_1 at $t = 1$, we divide the degradation level x_1 by the expected damage per loading epoch given the initial hyperparameters a_0 and b_0 that are obtained through the MLE procedure for model calibration, i.e. $\frac{b_0}{(a_0-1)}$. (Observe that this is the expectation of a geometric random variable whose parameter is beta distributed with parameters a_0 and b_0, see Appendix B for the derivation.) We then apply the updating rules of Proposition 1 using this proxy, that is $a_1 = a_0 + x_1$ and $b_1 = b_0 + \bar{n}_1$, and consequently apply the same logic at $t = 2$ to obtain a proxy for n_2. This procedure repeats until the component is replaced.

6.3. Results

The long run average cost rate of the integrated Bayes approach applied to the imperfect degradation signal setting and the intermittent degradation signal setting are denoted $C^{\text{imp}}_{\pi_{\mathcal{I}}}$ and $C^{\text{int}}_{\pi_{\mathcal{I}}}$, respectively. We are interested in how the integrated Bayes approach performs in these alternate

settings. Therefore we compare said costs rates with the cost rate of the integrated Bayes approach that does have access to the true degradation signal in real time. That is, we compute $\%VAL_{\text{int}} = 100 \cdot (C^{\text{int}}_{\pi_{\mathcal{I}}} - C_{\pi_{\mathcal{I}}})/C_{\pi_{\mathcal{I}}}$ and $\%VAL_{\text{imp}} = 100 \cdot (C^{\text{imp}}_{\pi_{\mathcal{I}}} - C_{\pi_{\mathcal{I}}})/C_{\pi_{\mathcal{I}}}$ for each instance of the test bed. The instances of our test bed are identical to the test bed of the previous section, see Table 1. In the test bed of the imperfect degradation signal, we additionally vary σ over 5 different levels, i.e. $\sigma \in \{0.25, 0.5, 0.75, 1, 1.25\}$. Note that in these settings an optimal policy cannot use information contained in the real-time signal, whereas the integrated Bayes approach does use this information. Therefore $C_{\pi_{\mathcal{I}}}$ is a practical lower bound on the cost rate of the *optimal* policy for both alternate settings.

The average $\%VAL_{\text{int}}$ and the average $\%VAL_{\text{imp}}$ are presented in Table 3 and Table 4, respectively. As before, we first distinguish between subsets of instances with the same value for a specific input parameter and then present the results for all instances.

Table 3 Results for intermittent degradation signals

Input parameter	Value	$\%VAL_{\text{int}}$
Coefficient of variation of prior gamma distribution, cv_Λ	0.3	9.20
	0.6	20.11
Coefficient of variation of prior beta distribution, cv_Φ	0.01	17.78
	0.02	11.53
Corrective maintenance cost, c_u	5	8.95
	10	20.36
Number of simulated paths, ν	10	15.82
	50	13.49
Total		14.66

Table 4 Results for imperfect degradation signals

Input parameter	Value	$\%VAL_{\text{imp}}$
Coefficient of variation of prior gamma distribution, cv_Λ	0.3	4.09
	0.6	8.58
Coefficient of variation of prior beta distribution, cv_Φ	0.01	6.30
	0.02	6.38
Corrective maintenance cost, c_u	5	4.62
	10	8.05
Number of simulated paths, ν	10	6.43
	50	6.24
Standard deviation of the noise, σ	0.25	0.01
	0.5	2.16
	0.75	4.93
	1	9.25
	1.25	15.33
Total		6.34

From Table 3 we see that when the degradation signal is not relayed in real-time, our integrated Bayes approach performs relatively poorly with a $\%VAL_{\text{int}}$ of almost 15 percent on average. This

also implies that relaying a degradation signal in real-time rather than only periodically has considerable value. Indeed, it allows the decision maker to not only learn the drift of a degradation signal (encoded in the degradation level and the age) but also the volatility of the degradation signal; the latter can only be inferred if one has access to the individual arrivals of loading epochs and their corresponding damages in between decision epochs.

By contrast, our integrated Bayes approach performs quite well when the degradation signal is imperfect. Indeed, from Table 4 we see that $\%VAL_{\text{imp}}$ is around 6 percent on average, and even remains below 5 percent for moderate levels of noise.

7. Case study

IXR systems are used by physicians for minimally-invasive image-guided procedures to diagnose and treat diseases in nearly every organ system. X-ray tubes (denoted by the rectangle in Figure 4) are the most expensive replacement components of an IXR system and therefore of major concern. Philips Healthcare produces the IXR system and does maintenance and service for many hospitals that use the IXR system.

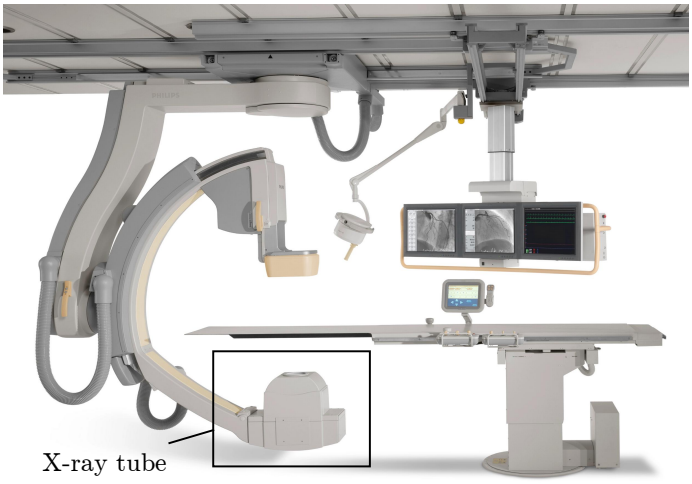


Figure 4 Example of IXR system with X-ray tube denoted by rectangle. (Philips 2020)

Unexpected downtime incidents have a major impact, especially for the patients whose medical procedure is cut short or postponed. The cost of premature maintenance is also substantial. Medical imaging equipment have list-prices on the order of one million US dollars and the annual maintenance expenses of such equipment are around 10% of the list-price (ECRI 2013). Since these medical imaging systems generally last up to 10 years, roughly half of the total cost of ownership of such a system (excluding downtime costs) consists of maintenance costs.

Philips healthcare faces the challenge of replacing the expensive X-ray tube before failures occur but also to maximize their useful lifetime. Below we describe a case study on the most critical component of an IXR system: the X-ray tube. The full data set for this case-study is included in the online companion to this article.

This section is organized as follows: We describe the dominant failure mechanism of the X-ray tube (Section 7.1), give a description of the data set (Section 7.2), illustrate the operation of the optimal policy with real data (Section 7.3) and compare the three approaches outlined in Section 5 on real data with a bootstrapping study (Section 7.4).

7.1. IXR Filaments

A failure analysis performed by Philips Healthcare indicates that X-ray tube failures are predominantly caused by worn out filaments (Albano et al. 2019, Section 5.3.3.4). These tungsten filaments are heated to a high temperature by a voltage differential such that they emit electrons. These electrons are then accelerated by a high voltage potential differential towards the target so that they emit X-rays when they hit the target. The X-rays are then used to produce the desired image during image guided medical procedures. This process is depicted in Figure 5.

The tungsten evaporates slowly when the filament heats up. This filament usually develops a "hot-spot" at the thinnest location. The evaporation causes the hotspot to become thinner with every image taken. This continues until the tungsten melts at the hot-spot and the filament fails; see Covington (1973). The degradation state of a filament can be inferred from the resistance of

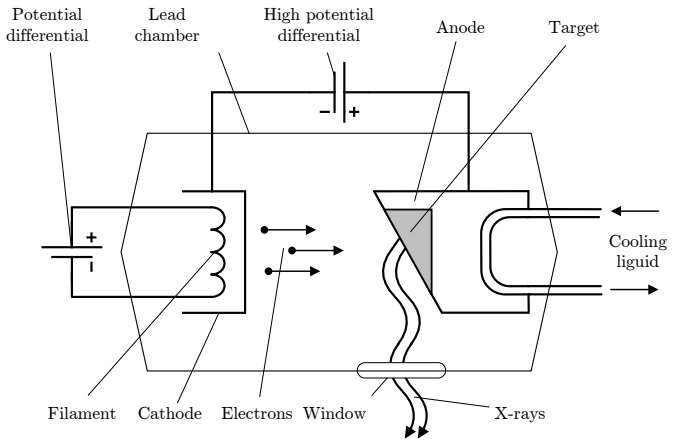


Figure 5 Simplified X-ray tube schematic.

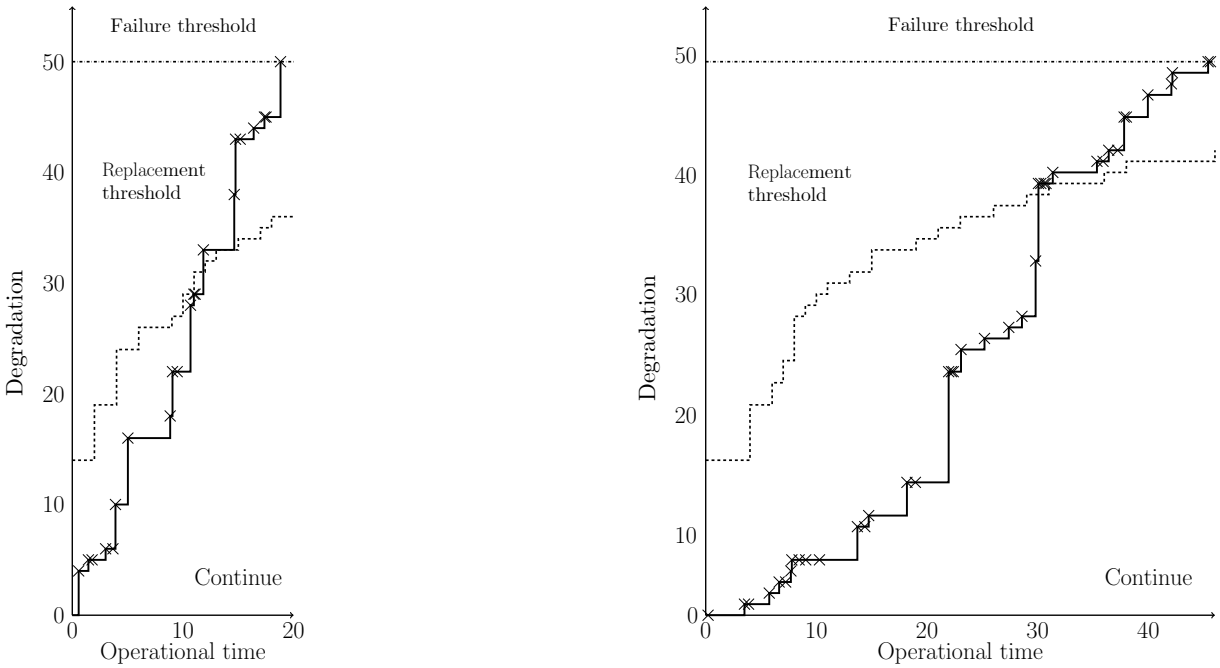
the filament. Philips Healthcare performed an engineering analysis to derive a single-dimensional health indicator that perfectly measures the degradation state over time of a filament (Albano et al. 2019, Section 7.9.7.12). This single-dimensional health indicator, the degradation, is recorded in a database each time that an IXR system is used for an image-guided procedure.

7.2. Degradation data of IXR X-ray tubes

The data set of X-ray tube degradation consists of 52 time-series of degradation levels. Let $\mathcal{I}$ be the set of all X-ray tubes for which there is available data; $|\mathcal{I}| = 52$. The time series of a single X-ray tube $i \in \mathcal{I}$ is denoted by $\mathcal{J}_i$. Each datum $j \in \mathcal{J}_i$ in such a time-series is a tuple $(t_j, x_j)_i$ of the age of the X-ray tube t_j and the degradation level x_j at that age. Each tuple $(t_j, x_j)_i$ of X-ray tube i is generated when an IXR system is used and each time-series consists of 20.000-300.000 data points originating from a time period of 2-5 years. Due to confidentiality reasons we have left-truncated the data and normalized the data. That is, all time-series start with $x_0 = 0$ for $t_0 = 0$, and end at $x_{|\mathcal{J}_i|} = 50$ (i.e., $\xi = 50$) for all $i \in \mathcal{I}$. For each time-series, we computed the inter-arrival times (i.e., $t_j - t_{j-1}$) between succeeding data points and damage increments per data point (i.e., $x_j - x_{j-1}$). By removing outliers in the inter-arrival times due to either weekends, nights, or other prolonged non-operational periods, we transformed the original time-series into time-series based on the operational age of each X-ray tube (see also the discussion in Section 3.1 on operational age). We then removed data points for which the image-guided procedure was considered too short to wear out the X-ray tube (i.e., shock arrival is regarded as non-critical). For the resulting time-series, the assumption that shocks arrive as a Poisson process is not rejected based on a Kolmogorov-Smirnov test. Additionally, we normalized the time such that one unit of time corresponds to roughly the operational time that is considered as minimally achievable for performing maintenance practices from a practical perspective (e.g., sending a service engineer to the location of the hospital). Furthermore, based on the Akaike information criterion, the damage size distribution is within the class of the applicable probability distributions best represented by the geometric distribution. Finally, pair-wise Kolmogorov-Smirnov tests in our data set show that the parameters of the distributions for the inter-arrival time and damage size, differed from one component to another (i.e., there is heterogeneity).

7.3. Illustration of optimal replacement policy

Figure 6 shows two examples of the optimal replacement policy, applied a-posteriori to two time-series of the filament data set. In these examples, and also throughout the rest of this section, the ratio $\frac{c_u}{c_p}$ is set to 5 based on discussions with Philips Healthcare. Furthermore, the prior values of the hyperparameters are estimated using the MLE method described in Appendix C applied on the remaining time-series, which resulted in $\alpha_0 = 44.88$, $\beta_0 = 32.43$, $a_0 = 4.29$, and $b_0 = 4.76$ for this example. Figure 6 visually confirms that there is heterogeneity in both the shock arrival process and damage size distribution. In this example, the optimal replacement policy prescribes to preventively replace the X-ray tube at $x = 29$ and $n = 14$ at $t = 12$ (left) and $x = 39$ and $n = 28$ at $t = 31$ (right), respectively, such that the useful lifetime of the X-ray tube is best utilized. The examples illustrate how integration of learning and decision making allows for the X-ray tube to be replaced early when necessary and late when possible.



(a) Replacement when $(x, n, t) = (29, 14, 12)$ with $\delta^{(14,12)} = 29$.

(b) Replacement when $(x, n, t) = (39, 28, 31)$ with $\delta^{(28,31)} = 39$.

Figure 6 Examples of optimal replacement policy applied to IXR filament degradation paths.

7.4. Bootstrapping study

The goal of this section is to illustrate the optimal replacement policy, and to assess the performance of the integrated Bayes approach compared to both the offline and open-loop feedback approach described in Section 5 on real life data. The main performance metric is therefore the relative cost savings that can be attained by using the integrated Bayes approach instead of the two heuristic approaches commonly used in practice and described in Section 5. More formally, we are interested in $\%SAV_\pi = 100 \cdot (C_\pi - C_{\pi_\mathcal{I}})/C_\pi$, where C_π is the average cost rate of approach $\pi \in \{\pi_\mathcal{F}, \pi_\mathcal{N}\}$. In addition, we are interested in the impact of the amount of available historical degradation data on the $\%SAV_\pi$ of the two heuristic approaches.

We evaluate the performance metric $\%SAV_\pi$ retroactively on the data set described in Section 7.2. To evaluate the impact of the amount of historical data available we bootstrap the amount of available data. That is, we first sample with replacement $s \in \{5, 10, 15\}$ time-series from the data set $\mathcal{I}$. We then use these time-series to estimate the required parameters for the integrated Bayes approach as well as for the heuristic approaches using MLE (see Appendix C for further details regarding the MLE procedure). We then implement and evaluate both the integrated Bayes approach and the heuristic approaches on the remaining time-series that were not used for the estimation. This procedure is repeated 150 times per choice of s, such that the confidence intervals on the average cost rates are sufficiently small, resulting in 450 bootstrap instances. The resulting

average values for $\%SAV_\pi$ are reported in Table 5. Observe that this bootstrapping study indicates the induced savings when the integrated Bayes approach is implemented instead of the offline approach or the open-loop feedback approach in a real life setting.

Table 5 Results of bootstrapping study

	$\%SAV_\pi$	
Number of time series	$\pi_{\mathcal{F}}$	$\pi_{\mathcal{N}}$
5	4.68	10.46
10	4.08	10.35
15	4.07	10.70
Total	4.28	10.50

Table 5 shows that the integrated Bayes approach reduces the average cost rate with 10.50% on average compared to the state-of-the-art approach. Moreover, cost savings of 4.28% can be attained by integrating the learning with the decision making instead of using a data-driven approach that does not integrate the two. The results furthermore indicates that the attainable savings are not significantly influenced by the amount of available historical degradation data. Hence, the integrated Bayes approach does not need a large amount of historical data to perform well. These results highlight the value of the proposed method to integrate learning and decision making in a real life setting.

8. Conclusion

In this paper, we have considered the condition based maintenance of components that are subject to compound Poisson degradation, where the compounding distribution is a member of the one-parameter non-negative exponential family. The cost to replace a component after failure is much higher than before failure, because it includes the costs of unplanned downtime. Motivated by practice, we have assumed that the population of components is heterogeneous. That is, the degradation parameters of components vary from one component to the other.

Since the degradation parameters of an individual component cannot be observed directly and need to be learned by observing its degradation signal, we have modeled this replacement problem as a partially observable Markov decision process. The entire past degradation path of a component is relevant state information in this setting, which can lead to tractability issues. We have shown that we can circumvent these issues by using conjugate prior pairs to model the heterogeneity of the component population. This allowed us to collapse a high dimensional state space to a 3 dimensional state space while retaining all relevant information. This collapse enabled us to tractably compute optimal policies as well as to characterize the optimal replacement policy. We have characterized the optimal replacement policy as a threshold replacement policy, where the

threshold is increasing in the age of a component. Furthermore, we have shown that the threshold also depends on the volatility of the observed degradation signal of a component. This volatility can only be captured by observing the *entire* degradation signal in real-time. Although we have assumed geometric compounding throughout this paper, we show that these structural properties hold for a wide variety of compounding distributions.

We performed two comprehensive simulation studies. In the first study, we assessed (i) the benefits of explicitly modeling heterogeneity, (ii) the value of integrating learning with decision making, and (iii) the impact of the amount of available historical degradation data for model calibration on their performance. Our results indicate that the integration of learning and decision making leads to excellent results with optimality gaps of only 0.60% on average. By contrast, ignoring heterogeneity leads to average optimality gaps of 15.02% while failing to integrate learning with decision making leads to average optimality gaps of 7.08%. Furthermore, we have shown that models that ignore population heterogeneity do *not* perform appreciably better when the amount of historical degradation data for model calibration increases. In the second simulation study, we assessed the performance of the optimal policy (under real- time, perfect data) when applied to settings where the degradation signal (i) is not perfect or (ii) not relayed in real-time. Our results indicate that real-time access to degradation data is valuable but this real-time degradation data need not be perfect to achieve excellent performance.

Finally, we have established the practical value of integrated learning and decision making based on a real life data set of filament deterioration in an IXR machine. We find that integrated learning can save up to 10.50% compared to approaches without learning and up to 4.28% compared to an approach where learning is separated from decision making. This normalized data set with is made available to benchmark future approaches to perform maintenance on stochastically deteriorating systems.

Appendix A: Proofs

Proof of Proposition 1. The joint posterior distribution of Λ and Φ at component age t is proportional to the product of the joint likelihood function and the joint prior distributions on Λ and Φ at component age $t-1$. The joint likelihood of observing $\boldsymbol{\theta}_t$ given (λ, ϕ), denoted by $\mathcal{L}(\boldsymbol{\theta}_t|\lambda,\phi)$, is equal to

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}_t|\lambda,\phi) &:= \mathbb{P}\Big[K_{(t-1,t]} = k_t, \boldsymbol{Y}_t = \boldsymbol{y}_t | \Lambda = \lambda, \Phi = \phi\Big] \\ &= \frac{\lambda^{k_t} e^{-\lambda}}{k_t!} \prod_{i=1}^{k_t} \Big[h(y_t^i) e^{\phi y_t^i - A(\phi)}\Big] \\ &= \frac{\lambda^{k_t} e^{-\lambda}}{k_t!} e^{\phi \sum_{i=1}^{k_t} y_t^i - k_t \cdot A(\phi)} \prod_{i=1}^{k_t} h(y_t^i).\end{aligned}$$

where $\prod_{i=1}^{0} \cdot \equiv 1$. This yields

$$\begin{aligned}&f_{\Lambda,\Phi}(\lambda,\phi|\boldsymbol{\theta}_0,\ldots,\boldsymbol{\theta}_t) \\ &\propto \mathcal{L}(\boldsymbol{\theta}_t|\lambda,\phi) \cdot f_\Lambda(\lambda|\alpha_{t-1},\beta_{t-1}) \cdot f_\Phi(\phi|a_{t-1},b_{t-1}) \\ &= \frac{\lambda^{k_t} e^{-\lambda}}{k_t!} e^{\phi \sum_{i=1}^{k_t} y_t^i - k_t \cdot A(\phi)} \prod_{i=1}^{k_t} h(y_t^i) \cdot \frac{\beta_{t-1}^{\alpha_{t-1}} \lambda^{\alpha_{t-1}-1} e^{-\beta_{t-1}\lambda}}{\Gamma(\alpha_{t-1})} \cdot H(a_{t-1},b_{t-1}) e^{a_{t-1}\phi - b_{t-1}A(\phi)} \\ &\propto \lambda^{\alpha_{t-1}+k_t-1} e^{-(\beta_{t-1}+1)\lambda} H(a_{t-1},b_{t-1}) e^{\left(a_{t-1}+\sum_{i=1}^{k_t} y_t^i\right)\phi - \left(b_{t-1}+k_t\right)A(\phi)},\end{aligned}$$

which is after normalization (over hyperparameters) equal to

$$\frac{(\beta_{t-1}+1)^{(\alpha_{t-1}+k_t)} \lambda^{\alpha_{t-1}+k_t-1} e^{-(\beta_{t-1}+1)\lambda}}{\Gamma(\alpha_{t-1}+k_t)} \cdot H(a_{t-1} + \sum_{i=1}^{k_t} y_t^i, b_{t-1}+k_t) e^{\left(a_{t-1}+\sum_{i=1}^{k_t} y_t^i\right)\phi - \left(b_{t-1}+k_t\right)A(\phi)}, \quad (14)$$

where $\Gamma(\cdot)$ denotes the gamma function and $H(a_{t-1} + \sum_{i=1}^{k_t} y_t^i, b_{t-1}+k_t)$ is the new normalization factor with the updated hyperparameters. Observe that the joint posterior distribution of Λ and Φ in Equation (14) is equal to the product of a Gamma$(\alpha_{t-1}+k_t, \beta_{t-1}+1)$ distribution and the general prior with updated hyperparameters $(a_{t-1} + \sum_{i=1}^{k_t} y_t^i, b_{t-1}+k_t)$ for a member of the exponential family, which completes the proof. □

Proof of Lemma 1. We first consider the posterior predictive distribution conditioned on $K_{(t,t+1]}$,

$$\mathbb{P}\Big[Z_{(t,t+1]} = z \Big| K_{(t,t+1]} = k, \alpha_t, \beta_t, a_t, b_t\Big] = \int_0^\infty f^{(k)}(z|\,\Phi=\phi) f_\Phi(\phi|a_t,b_t) d\phi, \quad (15)$$

where $f^{(k)}(z|\,\Phi=\phi)$ denotes the k-fold convolution of probability density (or mass) function of the random variable $\{Y|\,\Phi=\phi\}$. The distribution of $K_{(t,t+1]}$ is a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate follows a Gamma(α_t,β_t) distribution, which is known to be the negative binomial distribution with $p = \frac{1}{\beta_t+1}$ and $r = \alpha_t$. Hence, we have

$$\mathbb{P}\Big[K_{(t,t+1]} = k \Big| \alpha_t, \beta_t\Big] = \binom{k+\alpha_t-1}{k} \left(\frac{1}{\beta_t+1}\right)^k \left(\frac{\beta_t}{\beta_t+1}\right)^{\alpha_t}. \quad (16)$$

Unconditioning Equation (15) using Equation (16) yields the desired result. □

Proof of Proposition 2. In this proof, we use the following two definitions quantifying the concept of one random variable being "bigger" than another random variable.

Definition 1 (1.A.1 Definition, Shaked and Shanthikumar (2007)). A random variable X is stochastically larger than a random variable Y in the usual stochastic order, denoted by $X \geq_{\text{st}} Y$, if and only if

$$\mathbb{P}[X \geq x] \geq \mathbb{P}[Y \geq x], \quad \text{for all } x \in \mathbb{R}. \tag{17}$$

Definition 1 implies that random variable Y is less likely than random variable X to take on large values, where large means any value greater than x, and that this holds for all $x \in \mathbb{R}$.

Definition 2 (1.C.1 Definition, Shaked and Shanthikumar (2007)). Let X and Y be continuous (discrete) random variables with probability densities f and g, respectively, such that

$$\frac{g(t)}{f(t)} \text{ increases in } t \text{ over the union of the supports of } X \text{ and } Y$$

(here $a/0$ is taken to be equal to ∞ whenever $a > 0$), or, equivalently,

$$f(x)g(y) \geq f(y)g(x), \quad \text{for all } x \leq y.$$

Then X is said to be smaller than Y in the likelihood ratio order (denoted by $X \leq_{\text{lr}} Y$). Note that the likelihood ratio order is stronger than the usual stochastic order, as such, if $X \leq_{\text{lr}} Y$ then $X \leq_{\text{st}} Y$, cf. (Shaked and Shanthikumar 2007, Theorem 1.C.1).

We now proceed with the proofs of parts (i) and (ii) of Proposition 2:

(i) We consider two different component ages, t^+ and t^- $(t^+ > t^-)$, with the same degradation level x and the same total number of shocks, n, received by the component. By Proposition 1, we have $\beta_{t^+} > \beta_{t^-}$, whereas all other hyperparameters (i.e., α_t, a_t and b_t) remain fixed when only t changes. Note that

$$\mathbb{P}[K_{(t,t+1]} = u \,|\, \Lambda = \lambda] = e^{-\lambda}\frac{\lambda^u}{u!} \quad \text{and} \quad \mathbb{P}[\Lambda = \lambda \,|\, \alpha, \beta] = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}.$$

We can then show that the random variables $\Lambda(\alpha, \beta) = \{\Lambda \,|\, \alpha, \beta\}$ are stochastically increasing in α and stochastically decreasing in β. This easily follows from the appropriate likelihood ratio.

We now show that the random variables $\mathcal{K}(\lambda) = \{K_{(t,t+1]} \,|\, \Lambda = \lambda\}$ are stochastically increasing in λ, i.e., $\lambda \leq \lambda'$ implies $\mathcal{K}(\lambda) \leq_{\text{st}} \mathcal{K}(\lambda')$. Consider the likelihood ratio

$$\frac{\mathbb{P}[\mathcal{K}(\lambda') = u]}{\mathbb{P}[\mathcal{K}(\lambda) = u]} = \frac{\mathbb{P}[K_{(t,t+1]} = u \,|\, \Lambda = \lambda']}{\mathbb{P}[K_{(t,t+1]} = u \,|\, \Lambda = \lambda]} = e^{-(\lambda'-\lambda)}\left(\frac{\lambda'}{\lambda}\right)^u.$$

It is immediately evident that for $\lambda \leq \lambda'$, the ratio is increasing in u. This yields the likelihood ratio order and as a consequence the usual stochastic order.

All in all, for $t^+ > t^-$, $\beta_{t^+} > \beta_{t^-}$, thus $\Lambda(\alpha, \beta_{t^+}) \leq_{\text{st}} \Lambda(\alpha, \beta_{t^-})$. From the stochastic monotonicity of $\mathcal{K}(\Lambda)$, the above yields $\mathcal{K}^{(t^+)} := \mathcal{K}(\Lambda(\alpha, \beta_{t^+})) \leq_{\text{st}} \mathcal{K}(\Lambda(\alpha, \beta_{t^-})) := \mathcal{K}^{(t^-)}$, cf. (Shaked and Shanthikumar 2007, Theorem 1.A.2.). Finally, as the parameters associated with the random variables Y_i remain fixed, it is evident that

$$Z_{(t^+,t^++1]} = \sum_{i=1}^{\mathcal{K}^{(t^+)}} Y_i \leq_{\text{st}} \sum_{i=1}^{\mathcal{K}^{(t^-)}} Y_i = Z_{(t^-,t^-+1]},$$

cf. (Shaked and Shanthikumar 2007, Theorem 1.A.4.). In conclusion, for $\boldsymbol{s} = (x, n, t)$, it follows that $Z(\boldsymbol{s})$ is stochastically decreasing in t.

(*ii*) We consider two degradation levels x^+ and x^- ($x^+ > x^-$) at the same component age t when we have observed the same total number of shocks n. By Proposition 1, we have $a_{x^+} > a_{x^-}$, whereas all other hyperparameters are equivalent as they remain fixed when only x changes.

It is well-known that a member of the one-parameter exponential family is likelihood-ratio ordered (and thus stochastically ordered in the usual stochastic order) according to its parameter (see e.g., Karlin and Rubin 1956, Bapat and Kochar 1994). Hence, we know that i) the random variables $\mathcal{Y}(\phi) = \{Y \mid \Phi = \phi\}$ are stochastically non-decreasing in ϕ, i.e., $\phi \leq \phi'$ implies $\mathcal{Y}(\phi) \leq_{\text{st}} \mathcal{Y}(\phi')$, and ii) the random variables $\Phi(a,b) = \{\Phi \mid a,b\}$ are stochastically non-decreasing in a, i.e., $a \leq a'$ implies $\Phi(a,b) \leq_{\text{st}} \Phi(a',b)$. Then, by Theorem 6 of Huang and Mi (2020), which relates the stochastic order of a posterior predictive distribution with the stochastic order of the posterior and the corresponding conditional distribution, we can conclude that $\mathcal{Y}^{(x^+)} := \mathcal{Y}(\Phi(a_{x^+}, b)) \geq_{\text{st}} \mathcal{Y}(\Phi(a_{x^-}, b)) =: \mathcal{Y}^{(x^-)}$. Finally, as the parameters associated with the random variables $K_{(t,t+1]}$ remain fixed, it is evident that

$$Z^{(x^+)}_{(t,t+1]} := \sum_{i=1}^{K_{(t,t+1]}} \mathcal{Y}_i^{(x^+)} \geq_{\text{st}} \sum_{i=1}^{K_{(t,t+1]}} \mathcal{Y}_i^{(x^-)} =: Z^{(x^+)}_{(t,t+1]}.$$

In conclusion, for $\boldsymbol{s} = (x,n,t)$, it follows that $Z(\boldsymbol{s})$ is stochastically increasing in x.

□

Proof of Lemma 2. We prove part (*i*) and omit the proof of part (*ii*) as its proof structure follows verbatim.

For $\boldsymbol{s} = (x,n,t) \in \mathbb{N}_0^3$, let $V^m(\boldsymbol{s})$ denote the value function at the m-th iteration of the value iteration algorithm, so that the value iteration algorithm produces the sequence $\{V^m(\boldsymbol{s})\}_{m \in \mathbb{N}_0}$. We use induction on the steps of the value iteration algorithm as a proof technique. Since our state space is countable, $\gamma \in (0,1)$, and costs are bounded from above, the value function is guaranteed to converge point-wise to the optimal value function that satisfies Equation (10) (i.e., $V^m(\boldsymbol{s}) \to V(\boldsymbol{s})$ for all $\boldsymbol{s} \in \mathbb{N}_0^3$ as $m \to \infty$) from any arbitrary starting position through the value iteration algorithm (Bertsekas 2007, Proposition 1.2.1).

For $\boldsymbol{s} = (x,n,t) \in \mathbb{N}_0^3$, we set $V^0(\boldsymbol{s}) = 0$. Note that $V^0(\boldsymbol{s})$ is non-increasing in t. We assume that the theorem holds for the m-th iteration, i.e., $V^m(\boldsymbol{s})$ is non-increasing in t. Then according to Equation (10), we have

$$V^{m+1}(\boldsymbol{s}) = \begin{cases} c_u + \gamma \mathbb{E}\left[V^m(\boldsymbol{s}_0 + A(\boldsymbol{s}_0))\right], & \text{if } x \geq \xi, \\ \min\Big\{c_p + \gamma \mathbb{E}\left[V^m(\boldsymbol{s}_0 + A(\boldsymbol{s}_0))\right]; \gamma \mathbb{E}_{\boldsymbol{s}}\left[V^m(\boldsymbol{s} + A(\boldsymbol{s}))\right]\Big\}, & \text{if } x < \xi. \end{cases} \tag{18}$$

Since $V^m(\boldsymbol{s})$ is non-increasing in t by the induction hypothesis and the random variable $A(\boldsymbol{s})$ is stochastically decreasing in t in the usual stochastic order (cf. Proposition 2 and the proof therein), the expectation $\mathbb{E}_{\boldsymbol{s}}\left[V^m(\boldsymbol{s} + A(\boldsymbol{s}))\right]$ holds this property as well (cf. Shaked and Shanthikumar 2007, Theorem 1.A.3). Because the terms of the right-hand side of Equation (18) are non-increasing in t, $V^{m+1}(\boldsymbol{s})$ is also non-increasing in t. Due to the point-wise convergence in which the structure of $V^m(\boldsymbol{s})$ is preserved, we may conclude that $V(\boldsymbol{s})$ is also non-increasing in t (cf. Puterman 2005, Theorem 6.3.1.). □

Proof of Theorem 1. Preventive replacement is optimal when the following equation holds

$$c_p + \gamma \mathbb{E}\left[V(A(\boldsymbol{s}_0))\right] \leq \gamma \mathbb{E}_{\boldsymbol{s}}\left[V(\boldsymbol{s} + A(\boldsymbol{s}))\right]. \tag{19}$$

Since the expectation $\mathbb{E}\left[V(A(\boldsymbol{s}_0))\right]$ is constant with respect to x, the left-hand side of Inequality (19) is constant with respect to x. Based on part (ii) of Lemma 2, the right-hand side of Inequality (19) is non-decreasing in x. Hence, if the optimal decision is to carry out preventive replacement in state $(\delta^{(n,t)}, n, t)$, then the same decision is optimal for any state (x, n, t) with $x \geq \delta^{(n,t)}$, which implies the control limit policy.

Similarly, the right-hand side of Inequality (19) is non-increasing in t by part (i) of Lemma 2. Hence, if it optimal to carry out a preventive replacement in state $(\delta^{(n,t)}, n, t)$, then it is optimal to carry out a preventive replacement in any state $(\delta^{(n,t)}, n, t')$ with $t' < t$, which implies that $\delta^{(n,t)}$ is monotonically non-decreasing in t. □

Appendix B: Derivation of $\mathbb{E}[Z(\boldsymbol{s})]$ and $\mathbb{E}[Z(\boldsymbol{s})^2]$ for the compound Poisson process

We first repeat some notation. Let $\boldsymbol{s} = (x, n, t) \in \mathbb{N}_0^3$ and let $Z(\boldsymbol{s}) = \sum_{i=1}^{K(\boldsymbol{s})} Y_i(\boldsymbol{s})$, where $K(\boldsymbol{s})$ is a Poisson random variable unknown rate λ, and $\{Y_i(\boldsymbol{s})\}_{i\in\mathbb{N}}$ is a sequence of independent and (identically) geometrically distributed random variables with unknown success probability p, independent from $K(\boldsymbol{s})$. We endow p and λ with prior beta and gamma distributions, denoted by P and Λ, respectively. We can then find the first and second moment of $Z(\boldsymbol{s})$ by repeatedly using the tower property of expectations and the law of total variance.

We have

$$\mathbb{E}[Y_i(\boldsymbol{s})] = \mathbb{E}\big[\mathbb{E}[Y_i(\boldsymbol{s}) \mid P]\big] = \mathbb{E}\left[\frac{P}{1-P}\right] = \frac{\int_0^1 \frac{p}{1-p} p^{a_t-1}(1-p)^{b_t-1} dp}{\mathrm{B}(a_t, b_t)} = \frac{\mathrm{B}(a_t+1, b_t-1)}{\mathrm{B}(a_t, b_t)} = \frac{a_t}{b_t - 1}, \tag{20}$$

where the fourth equality follows from manipulating the term inside the integral to the density of another beta distribution.

Proceeding with $\mathbb{E}[K(\boldsymbol{s})]$, we have

$$\mathbb{E}[K(\boldsymbol{s})] = \mathbb{E}\big[\mathbb{E}[K(\boldsymbol{s}) \mid \Lambda] = \mathbb{E}[\Lambda] = \frac{\alpha_t}{\beta_t}. \tag{21}$$

Combining (20) and (21) and using the updating rules of Proposition 1 for $\boldsymbol{s} = (x, n, t)$, we have

$$\mathbb{E}[Z(\boldsymbol{s})] = \mathbb{E}\left[\sum\nolimits_{i=1}^{K(\boldsymbol{s})} Y_i(\boldsymbol{s})\right] = \mathbb{E}[K(\boldsymbol{s})]\mathbb{E}[Y_i(\boldsymbol{s})] = \frac{a_0 + x}{b_0 + n - 1}\frac{\alpha_0 + n}{\beta_0 + t}.$$

For the second moment, we first compute the variance. We have

$$\begin{aligned}
\mathrm{Var}[Z(\boldsymbol{s})] = \mathrm{Var}\left[\sum\nolimits_{i=1}^{K(\boldsymbol{s})} Y_i(\boldsymbol{s})\right] &= \mathbb{E}\left[\mathrm{Var}\left[\sum\nolimits_{i=1}^{K(\boldsymbol{s})} Y_i(\boldsymbol{s}) \mid K(\boldsymbol{s})\right]\right] + \mathrm{Var}\left[\mathbb{E}\left[\sum\nolimits_{i=1}^{K(\boldsymbol{s})} Y_i(\boldsymbol{s}) \mid K(\boldsymbol{s})\right]\right] \\
&= \mathbb{E}\big[K(\boldsymbol{s})\big]\mathrm{Var}\big[Y_i(\boldsymbol{s})\big] + \mathrm{Var}\big[K(\boldsymbol{s})\big]\mathbb{E}\big[Y_i(\boldsymbol{s})\big] \\
&= \frac{\alpha_t}{\beta_t}\mathrm{Var}\big[Y_i(\boldsymbol{s})\big] + \frac{a_t}{b_t - 1}\mathrm{Var}\big[K(\boldsymbol{s})\big],
\end{aligned} \tag{22}$$

where $\mathbb{E}\left[Y_i(\boldsymbol{s})\right]$ and $\mathbb{E}\big[K(\boldsymbol{s})\big]$ are computed in (20) and (21), respectively. We now proceed with the two variances in (22). For $\mathrm{Var}\big[Y_i(\boldsymbol{s})\big]$ we have

$$\begin{aligned}
\mathrm{Var}\big[Y_i(\boldsymbol{s})\big] &= \mathbb{E}\Big[\mathrm{Var}\big[Y_i(\boldsymbol{s}) \mid P\big]\Big] + \mathrm{Var}\Big[\mathbb{E}\big[Y_i(\boldsymbol{s}) \mid P\big]\Big] \\
&= \mathbb{E}\left[\frac{P}{(1-P)^2}\right] + \mathrm{Var}\left[\frac{P}{1-P}\right]
\end{aligned} \tag{23}$$

For $\mathbb{E}\left[\frac{P}{(1-P)^2}\right]$, we have

$$\mathbb{E}\left[\frac{P}{(1-P)^2}\right] = \frac{\int_0^1 \frac{p}{(1-p)^2} p^{a_t-1}(1-p)^{b_t-1}dp}{\mathrm{B}(a_t,b_t)} = \frac{\mathrm{B}(a_t+1,b_t-2)}{\mathrm{B}(a_t,b_t)} = \frac{a_t(a_t+b_t-1)}{(b_t-2)(b_t-1)}, \tag{24}$$

where the second equality follows from manipulating the term inside the integral to the density of another beta distribution. Similarly, we can write

$$\begin{aligned}\mathrm{Var}\left[\frac{P}{1-P}\right] &= \mathbb{E}\left[\frac{P^2}{(1-P)^2}\right] - \mathbb{E}^2\left[\frac{P}{(1-P)}\right] \\ &= \frac{\mathrm{B}(a_t+2,b_t-2)}{\mathrm{B}(a_t,b_t)} - \left(\frac{a_t}{b_t-1}\right)^2 \\ &= \frac{a_t(a_t+1)}{(b_t-2)(b_t-1)} - \left(\frac{a_t}{b_t-1}\right)^2,\end{aligned} \tag{25}$$

which completes all the ingredients for $\mathrm{Var}\big[Y_i(\boldsymbol{s})\big]$. Using again the law of total variance, we can compute $\mathrm{Var}\big[K(\boldsymbol{s})\big]$ as follows

$$\mathrm{Var}\big[K(\boldsymbol{s})\big] = \mathbb{E}\Big[\mathrm{Var}\big[K(\boldsymbol{s}) \mid \Lambda\big]\Big] + \mathrm{Var}\Big[\mathbb{E}\big[K(\boldsymbol{s}) \mid \Lambda\big]\Big] = \mathbb{E}\left[\Lambda\right] + \mathrm{Var}\left[\Lambda\right] = \frac{\alpha_t}{\beta_t} + \frac{\alpha_t}{\beta_t^2}. \tag{26}$$

Plugging (24) and (25) in (23), and combining this with (26) in (22) gives us $\mathrm{Var}[Z(\boldsymbol{s})]$. We can then use the variance expansion $\mathrm{Var}[Z(\boldsymbol{s})] = \mathbb{E}[Z(\boldsymbol{s})^2] - \mathbb{E}^2[Z(\boldsymbol{s})]$ to obtain an explicit expression for the second moment in terms of the hyperparameters associated with state $\boldsymbol{s} = (x,n,t)$.

Appendix C: Maximum likelihood estimation

Let $\mathcal{I}$ denote the set containing all simulated sample paths (without maintenance). For each sample path $i \in \mathcal{I}$, we let x_i, n_i, and t_i, denote, respectively, the total damage accumulated by the component, the number of critical loading epochs sustained by the component, and the number of decision epochs that the component was operated until its failure. Thus the tuple $\boldsymbol{s}_i = (x_i, n_i, t_i)$ contains all relevant information for sample path i. Because the prior distributions of Φ and Λ are independent, we next look at their respective likelihoods given the set of simulated degradation paths $\mathcal{I}$ separately.

From Equation (15) we know that the likelihood of the degradation paths in $\mathcal{I}$ being induced by components stemming from a population whose degradation parameter Φ follows a beta prior distribution with parameters a_0 and b_0 is given by

$$\mathcal{L}_\Phi(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_{|\mathcal{I}|} \mid a_0,b_0) = \prod_{i=1}^{|\mathcal{I}|}\left[\frac{B(n_i+a_0,x_i+b_0)}{B(a_0,b_0)}\binom{x_i+n_i-1}{x_i}\right].$$

Observe that the number of critical loading epochs sustained by the i-th component of age t_i has a Poisson distribution with parameter $t_i\Lambda$, where $\Lambda \sim \Gamma(\alpha_0,\beta_0)$. By the scaling property of the gamma distribution, this number is therefore a continuous mixture of Poisson distributions where the mixing distribution of the Poisson rate follows a $\Gamma(\alpha_0,\beta_0/t_i)$ distribution, which is known to be a negative binomial (or Pascal) distribution with parameters $r=\alpha_0$ and $p=\frac{t_i}{\beta_0+t_i}$. Hence, the likelihood of the degradation paths in $\mathcal{I}$ being induced by components stemming from a population whose degradation parameter Λ follows a gamma prior distribution with parameters α_0 and β_0 is given by

$$\mathcal{L}_\Lambda(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_{|\mathcal{I}|} \mid \alpha_0,\beta_0) = \prod_{i=1}^{|\mathcal{I}|}\left[\binom{n_i+\alpha_0-1}{n_i}\left(\frac{t_i}{\beta_0+t_i}\right)^{n_i}\left(\frac{\beta_0}{\beta_0+t_i}\right)^{\alpha_0}\right].$$

We are interested in maximizing both likelihoods. For convenience, we maximize the log likelihoods. That is,

$$\arg\max_{\alpha_0,\beta_0,a_0,b_0} \ln \mathcal{L}_\Lambda(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_{|\mathcal{I}|} \,|\, \alpha_0,\beta_0) + \ln \mathcal{L}_\phi(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_{|\mathcal{I}|} \,|\, a_0,b_0),$$

which is a nonlinear multidimensional maximization problem that can be solved using standard numerical methods (e.g., the Nelder-Mead method).

References

Albano, M., E. Jantunen, G. Papa, U. Zurutuza. 2019. *The MANTIS Book: Cyber Physical System Based Proactive Collaborative Maintenance*. River Publishers Series in Automation, Control and Robotics.

Azoury, K.S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science* **31**(9) 1150–1160.

Bapat, R.B., S.C. Kochar. 1994. On likelihood-ratio ordering of order statistics. *Linear Algebra and Its Applications* **199** 281–291.

Batun, S., L.M. Maillart. 2012. Reassessing tradeoffs inherent to simultaneous maintenance and production planning. *Production and Operations Management* **21**(2) 396–403.

Benyamini, Z., U. Yechiali. 1999. Optimality of control limit maintenance policies under nonstationary deterioration. *Probability in the Engineering and Informational Sciences* **13**(1) 55–70.

Bertsekas, D.P. 2007. *Dynamic Programming and Optimal Control, Vol. II*. Athena Scientific.

Chen, L. 2010. Bounds and heuristics for optimal Bayesian inventory control with unobserved lost sales. *Operations Research* **58**(2) 396–413.

Chen, L. 2021. Fixing phantom stockouts: Optimal data-driven shelf inspection policies. *Production and Operations Management* **30**(3) 689–702.

Chen, L., A. Mersereau, Z. Wang. 2017. Optimal merchandise testing with limited inventory. *Operations Research* **65**(4) 968–991.

Chen, L., E.L. Plambeck. 2008. Dynamic inventory management with learning about the demand distribution and substitution probability. *Manufacturing & Service Operations Management* **10**(2) 236–256.

Chen, N., Z.S. Ye, Y. Xiang, L. Zhang. 2015. Condition-based maintenance using the inverse Gaussian degradation model. *European Journal of Operational Research* **243**(1) 190–199.

Coleman, C., S. Damodaran, M. Chandramouli, E. Deuel. 2017. Making maintenance smarter: Predictive maintenance and the digital supply network. Tech. rep., Deloitte.

Covington, E.J. 1973. Hot spot burnout of tungsten filaments. *Journal of the Illuminating Engineering Society* **2**(4) 372–380.

De Jonge, B., P.A. Scarf. 2020. A review on maintenance optimization. *European Journal of Operational Research* **285**(3) 805–824.

DeHoratius, N., A. Mersereau, L. Schrage. 2008. Retail inventory management when records are inaccurate. *Manufacturing & Service Operations Management* **10**(2) 257–277.

Derman, C. 1963. Optimal replacement and maintenance under Markovian deterioration with probability bounds on failure. *Management Science* **9**(3) 478–481.

ECRI. 2013. Healthcare product comparison system. Tech. rep., ECRI Institute, Plymouth Meeting, PA.

Elwany, A., N. Gebraeel, L.M. Maillart. 2011. Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research* **59**(3) 684–695.

Esary, J.D., A.W. Marshall. 1973. Shock models and wear processes. *The Annals of Probability* **1**(4) 627–649.

Ghosh, J.K., M. Delampady, T. Samanta. 2007. *An Introduction to Bayesian Analysis: Theory and Methods*. Springer Science & Business Media.

Huang, K., J. Mi. 2020. Applications of likelihood ratio order in Bayesian inferences. *Probability in the Engineering and Informational Sciences* **34**(1) 1–13.

Insinna, V., G. Ziezuliwicz. 2018. Investigation blames air force and navy for systemic failures in fatal marine corps c-130 crash that killed 16.

Kao, E. 1973. Optimal replacement rules when changes of state are semi-Markovian. *Operations Research* **21**(6) 1231–1249.

Karlin, S., H. Rubin. 1956. Distributions possessing a monotone likelihood ratio. *Journal of the American Statistical Association* **51**(276) 637–643.

Khaleghei, A., M. J. Kim. 2021. Optimal control of partially observable semi-Markovian failing systems: An analysis using a phase methodology. *Operations Research* **69**(4) 1282–1304.

Kharoufeh, J.P., S.M. Cox. 2005. Stochastic models for degradation-based reliability. *IIE Transactions* **37**(6) 533–542.

Kim, M.J. 2016. Robust control of partially observable failing systems. *Operations Research* **64**(4) 999–1014.

Kim, M.J., V. Makis. 2013. Joint optimization of sampling and control of partially observable failing systems. *Operations Research* **61**(3) 777–790.

Kolesar, P. 1966. Minimum cost replacement under Markovian deterioration. *Management Science* **12**(9) 694–706.

Kurt, M., J.P. Kharoufeh. 2010. Optimally maintaining a Markovian deteriorating system with limited imperfect repairs. *European Journal of Operational Research* **205**(2) 368–380.

Li, R., J.K. Ryan. 2011. A Bayesian inventory model using real-time condition monitoring information. *Production and Operations Management* **20**(5) 754–771.

Li, R., J. Song, S. Sun, X. Zheng. 2022. Fight inventory shrinkage: Simultaneous learning of inventory level and shrinkage rate. *Production and Operations Management* forthcoming.

Liu, X., Q. Sun, Z.S. Ye, M. Yildirim. 2021. Optimal multi-type inspection policy for systems with imperfect online monitoring. *Reliability Engineering & System Safety* **207** 107335.

Maillart, L.M. 2006. Maintenance policies for systems with condition monitoring and obvious failures. *IIE Transactions* **38** 463–475.

Makis, V., X. Jiang. 2003. Optimal replacement under partial observations. *Mathematics of Operations Research* **28**(2) 382–394.

Mersereau, A. 2013. Information-sensitive replenishment when inventory records are inaccurate. *Production and Operations Management* **22**(4) 792–810.

Morris, C.N. 1982. Natural exponential families with quadratic variance functions. *The Annals of Statistics* **10**(1) 65–80.

Olde Keizer, M.C.A., R.H. Teunter, J. Veldman. 2017. Joint condition-based maintenance and inventory optimization for systems with multiple components. *European Journal of Operational Research* **257**(1) 209 – 222.

Olsen, T.L., B. Tomlin. 2020. Industry 4.0: Opportunities and challenges for operations management. *Manufacturing & Service Operations Management* **22**(1) 113–122.

Pacquin, R. 2014. Asset management: The changing landscape of predictive maintenance. AberdeenGroup.

Philips. 2020. Allura Xper FD 10 - DS Interventional X-ray system. URL `https://www.usa.philips.com/healthcare/product/HC889021/diamondselectalluraxperfd10refurbishedxraysytem`.

Pierskalla, W.P., J.A. Voelker. 1976. A survey of maintenance models: the control and surveillance of deteriorating systems. *Naval Research Logistics* **23**(3) 353–388.

Price Waterhouse Coopers. 2014. Sensing the future of the internet of things. Digital IQ snapshot.

Puterman, M.L. 2005. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Rosenfield, D. 1976. Markovian deterioration with uncertain information. *Operations Research* **24**(1) 141–155.

Ross, S.M. 1969. A Markovian replacement model with a generalization to include stocking. *Management Science* **15**(11) 573–739.

Scarf, P.A. 1997. On the application of mathematical models in maintenance. *European Journal of Operational Research* **99**(3) 493–506.

Shaked, M., G.J. Shanthikumar. 2007. *Stochastic Orders*. Springer Science & Business Media.

Slaugh, V.W., B. Biller, S.R. Tayur. 2016. Managing rentals with usage-based loss. *Manufacturing & Service Operations Management* **18**(3) 429–444.

Sobczyk, K. 1987. Stochastic models for fatigue damage of materials. *Advances in Applied Probability* **19**(3) 652–673.

Uit het Broek, M.A.J., R.H. Teunter, B. De Jonge, J. Veldman, N.D. Van Foreest. 2020. Condition-based production planning: Adjusting production rates to balance output and failure risk. *Manufacturing & Service Operations Management* **22**(4) 792–811.

Valdez-Flores, C., R.M. Feldman. 1989. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics* **36**(4) 419–446.

Van Noortwijk, J.M. 2009. A survey of the application of Gamma processes in maintenance. *Reliability Engineering & System Safety* **94**(1) 2–21.

Van Oosterom, C., H. Peng, G.J. Van Houtum. 2017. Maintenance optimization for a Markovian deteriorating system with population heterogeneity. *IISE Transactions* **49**(1) 96–109.

Van Soest, H. 2016. Wanneer gaat het echt fout in Doel? Algeneen Dagblad (Dutch Newspaper), January 4.

Van Staden, H.E., R.N. Boute. 2021. The effect of multi-sensor data on condition-based maintenance policies. *European Journal of Operational Research* **290**(2) 585–600.

Wall Street Journal Custom Studios. 2017. How manufacturers achieve top quartile performance.

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

Zhang, L., Y. Lei, H. Shen. 2016. How heterogeneity influences condition-based maintenance for Gamma degradation process. *International Journal of Production Research* **54**(19) 5829–5841.