ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/eor

Invited Review in Celebration of the 50th Anniversary of EURO

Fifty years of maintenance optimization: Reflections and perspectives

Joachim Arts [a], Robert N. Boute [b,c,d,*], Stijn Loeys [b], Heletjé E. van Staden [e]

[a] Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg, Luxembourg
[b] Faculty of Economics and Business, KU Leuven, Belgium
[c] Technology and Operations Management, Vlerick Business School, Belgium
[d] Flanders Make@KU Leuven, Belgium
[e] Michael Smurfit Graduate Business School, University College Dublin, Ireland

ARTICLE INFO

Keywords:
Maintenance
Reliability
Anniversary

ABSTRACT

On the occasion of the 50th anniversary of the Association of European Operational Research Societies (EURO), we share our perspectives and reflections on maintenance research. We review the main methods and techniques for optimizing when and what to maintain, providing concrete examples as illustrations. We also discuss the optimization of the logistics support system surrounding the act of maintenance. In doing so, we highlight the multidisciplinary nature of maintenance research and its interface with other domains, such as spare parts inventory management, production scheduling, and transportation planning. We support our reflections with basic text-mining analyses of the archive of the *European Journal of Operational Research*, the journal published in collaboration with EURO. With this paper, we introduce interested researchers to maintenance optimization and share opportunities to close the gaps between the current state of research and real-world needs.

1. Introduction

On the occasion of the 50th anniversary of EURO, the Association of European Operational Research Societies, in 2025, we share our reflections on maintenance research. Maintenance is an essential part of industrial applications. Especially for expensive assets, it is more cost-effective to maintain and repair than to purchase new assets when they fail or break down. Examples of such "capital goods" are heavy machinery, material handling equipment, windmills, or vehicles such as ships, trains, or airplanes. Maintenance planning aims to maximize asset availability (reduce downtime) while controlling or minimizing maintenance expenses. An unforeseen breakdown due to a failure can result in expensive downtime. Downtime can be prevented with timely preventive maintenance. These preventive maintenance interventions also come at a cost for the technician/repair crew and the spare parts used. The total costs of maintenance and unavailability of a capital asset over its lifetime (typically one to several decades) can elevate to a multiple of the acquisition price. The question is thus: when and what to maintain preventively.

Maintenance optimization is studied within operational research as well as in several engineering disciplines. Operational research focuses on optimizing processes to prevent and deal with failure, while engineering focuses on understanding the physics of failure to predict its occurrence. Each community has its respective strengths. The engineering community understands the physics of failure, and the operational research community understands the stochastic processes induced by different maintenance policies. These two perspectives are complementary: One must understand failure mechanisms to prevent and deal with failures. As data on failures, by nature, is scarce in maintenance environments (an abundance of failure data indicates overly poor decision-making in the past), knowledge of physical failure mechanisms is crucial in selecting and calibrating proper degradation models and time to failure distributions. Strong contributions in the field often include collaboration across these communities, e.g., Elwany et al. (2011).

The European perspective on maintenance research differs slightly from the American one. European "Operational" Research is traditionally more oriented towards decision support to real-life problems (Bertrand et al., 2023). This approach intends to include all relevant aspects to explain processes' behavior and actual performance. In the *European Journal of Operational Research (EJOR)*, for instance, Poppe et al. (2018) proposes a hybrid policy that combines corrective, periodic, and condition-based maintenance, offering a smooth transition towards implementing condition-based maintenance in practice. Keizer et al. (2016) studies condition-based maintenance for complex multi-unit systems with both redundancy and economic dependencies. In

* Corresponding author at: Faculty of Economics and Business, KU Leuven, Belgium.
E-mail address: robert.boute@vlerick.com (R.N. Boute).

https://doi.org/10.1016/j.ejor.2024.07.002
Received 13 October 2023; Accepted 3 July 2024
Available online 5 July 2024

Please cite this article as: Joachim Arts et al., *European Journal of Operational Research*, https://doi.org/10.1016/j.ejor.2024.07.002

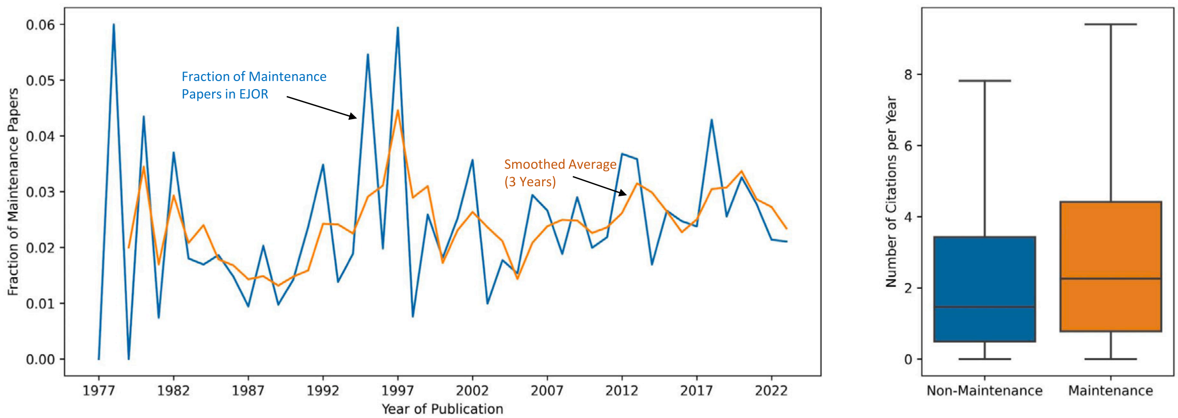

Fig. 1. Left panel: Fraction of EJOR papers devoted to maintenance per year. Right panel: Distribution of the yearly number of citations for (non-)maintenance papers published in EJOR (p-value: 0.011768).

contrast, American "Operations" Research tends to work on more stylized problems and uses these to build scientific knowledge and insights useful for knowledge transfer. Only those aspects of the problems that are assumed relevant from the perspective of the method and technique dealt with are included so that essential trade-offs become very explicit. For instance, the classical paper of Eckles (1968), published in *Operations Research*, uses a stylized partially observable Markov decision process to find the optimal maintenance policy where the system's state is not exactly known. Another example is the study of Maillart (2006) in *IIE Transactions*, which derives structural properties of optimal maintenance policies for systems with perfect information, which are used to motivate heuristic policies when information is imperfect.

Both the European and American perspectives on maintenance research have merit. Stylized models may serve as a stepping stone to more practical-oriented applications. As a European journal, EJOR has the "applicability" of EU research at its heart. American OR journals have more tradition in deriving structural results of a stylized model that may not necessarily have immediate applicability to decision support. However, this is changing, with several American journal editors stressing the importance of applied research. This change leads to more analytical papers that are empirically grounded or applied to company data.

We browsed through the archive of EJOR, the journal published in collaboration with EURO, and performed basic text-mining inspired by Song et al. (2019).[1] The fraction of EJOR publications devoted to maintenance is around 2.4%. There is no clear trend over the years, although there is a slight peak around the end of the nineties (see Fig. 1, left panel). Interestingly, maintenance-related publications are significantly (with a p-value of 0.011768) more frequently cited (on average, 3.76 citations per year since publication) than other non-maintenance publications in EJOR (on average, 3.03 citations per year since publication); see Fig. 1, right panel. We provide the top 15 most cited maintenance publications in Appendix B.

We do not intend to provide an exhaustive review or classification of papers related to maintenance and reliability. We refer the interested reader to existing reviews, such as, for instance, Olde Keizer et al. (2017) and De Jonge and Scarf (2020). Instead, we highlight the main methods and techniques for maintenance optimization and illustrate them with concrete examples. We then discuss the optimization of the maintenance logistics support system, such as spare parts inventory management or resource planning required to perform maintenance. We conclude by reflecting on the gaps between the current state of research and real-world needs.

[1] We explain our text-mining analysis in Appendix A, as we believe it may also serve future reviews. We also share our sources and source code on https://github.com/LoeysS/50yEJORMaintenance

Table 1
Maintenance strategies organized by timing and type uncertainty (Arts, 2019; Stoneham, 1998).

Type \ Timing	Known	Unknown
Known	Periodic, Age/Usage-based, modificative maintenance	Condition-based with real-time condition monitoring
Unknown	Condition-based with periodic inspections	Breakdown corrective maintenance

2. Maintenance optimization: Data and methods

Maintenance operations consist mainly of upgrading or replacing parts of an asset. The optimization thereof focuses on two decisions: *When* and *what* to maintain. These two decisions are subject to two major sources of uncertainty. The first uncertainty concerns the timing of the failure, which the act of maintenance is intended to prevent. This uncertainty depends on the lifetime of the asset or component, which is stochastic by nature. Preventive maintenance at known, predetermined moments can prevent the high cost of unplanned corrective maintenance associated with breakdown. The second uncertainty concerns what will be maintained. Sometimes, this uncertainty may only become known during the asset inspection. This latter uncertainty is often overlooked, but it distinguishes maintenance environments from production: When a production job starts, the materials and processes required to produce the product are known in advance. Frequently, neither is known when a maintenance job starts.

Table 1 classifies maintenance strategies based on these two uncertainty dimensions. Preventive maintenance replaces worn components before they fail to preserve and restore system reliability. When such preventive maintenance is scheduled at fixed maintenance intervals based on, for instance, time (age) or production volume (usage), the timing and content of the maintenance intervention are both known. When preventive maintenance is scheduled based on tracking a system's condition, the exact timing depends on the stochastic degradation behavior. When the condition is periodically monitored upon inspection, the timing is known, but not the content.

These maintenance strategies require accurate failure estimations to determine when to optimally intervene. Several approaches exist,

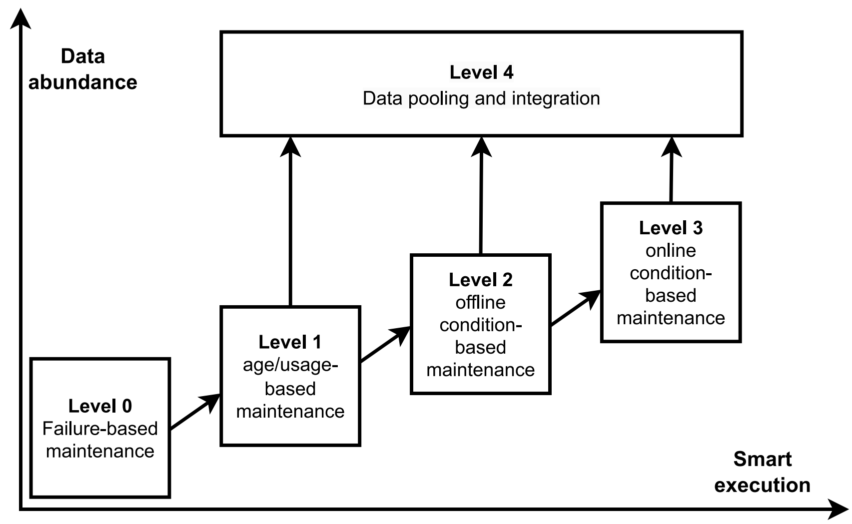

Fig. 2. We position the evolution of maintenance in terms of data availability and the type of methods used to process the data to inform decision-making.

depending on the data and the domain expertise available. We propose a maintenance maturity framework in Fig. 2 that captures the relationship between the methods applied in maintenance decision-making and the (abundance of) data used. We define maturity in terms of the volume of data that is available and the maintenance interpretation thereof using customized feedback controls. Here, we refer to the use of customized feedback controls as smart execution (Boute & Van Mieghem, 2021).

Level 0 in our framework refers to failure-based maintenance, where no data is collected, and corrective maintenance is performed when a breakdown is observed. In Level 1, age- or usage-based maintenance, historical failure data are used periodically to optimize the decision-making parameters, such as maintenance intervals. Levels 2 and 3 refer to condition-based maintenance (CBM), where information on the component's condition is collected to estimate the component's degradation level. Monitoring a system's condition is becoming increasingly accessible due to decreasing sensor and related data monitoring costs. Research, therefore, increasingly incorporates such data into the methods used to estimate the failure process of a deteriorating system. Such estimates can determine when to optimally intervene based on degradation thresholds and up-to-date information measurements. Online learning methods circumvent the sequential approach of predicting and optimizing by jointly estimating the failure process and optimizing maintenance decisions. Accordingly, the main difference between Levels 2 (offline CBM) and 3 (online CBM) is that the thresholds used at Level 3 are updated using the current condition information, whereas, in Level 2, the thresholds rely on historical data only. Level 4 refers to using data from various sources, including multiple machines and external environmental factors, to learn and predict system failures. When a company finds itself in the lower right triangle, it possesses sufficient smart(er) maintenance expertise but lacks the amount and/or quality of data to improve its maintenance decision-making. Similarly, when a company finds itself in the upper left triangle, it has the infrastructure to collect high-quality data but fails to use them to improve its maintenance decision-making.

Advances in data collection and algorithmic capabilities (moving towards the top right corner of Fig. 2), as enabled by digital technologies such as the Internet of Things (IoT), cloud computing and data analytics, collectively termed Industry 4.0, have resulted in renewed potential to optimize maintenance decision-making. For example, vast amounts of asset condition data can be captured by remotely monitoring complex assets via sensors and the IoT. Digital asset AI profiles, often called digital twins, can be developed using this data to model asset conditions and expected deterioration behavior. Such a digital duplicate of the physical asset can be used to simulate and evaluate the impact of maintenance decisions in a risk-free environment.

However, data from sensor observations used as input to maintenance optimization models may suffer from inherent sensor quality limitations or outside interference. In such cases, human input may complement sensor observations for a more complete and reliable digital profile of the monitored asset. In fact, humans are not only the end users of the IoT systems and services but are becoming active elements of the Internet through mobile devices, a term coined as the Internet of People (IoP). Specifically, IoP can (1) confirm or reject IoT data, upon which digital profiles may learn from the IoP action, (2) augment IoT data by providing an additional view with regards to the asset being monitored and (3) provide data where IoT data is unavailable.

Integration of data collection, maintenance decisions, and human input requires a central and holistic overview of the monitored assets to successfully coordinate operations. Digital control towers make this feasible. A digital control tower is a central information platform that collects real-time data from relevant entities to provide full visibility of the asset network. More extensive data overviews in the control tower translate into more efficient and effective operations coordination.

We believe that maintenance optimization models should be informed by understanding the physical process governing the systems under consideration. Such an approach requires a combination of engineering and operations expertise. Accordingly, we review the main techniques to optimize the maintenance strategies based on the available data in this section. We distinguish between estimating the failure process of a deteriorating system (Section 2.1) from the methods that use these estimates to determine when to intervene optimally (Section 2.2), concluding with online failure estimation and maintenance optimization (Section 2.3). The distinctions allow us to discuss the merits and demerits of treating failure estimation and maintenance optimization separately. We illustrate the different approaches through two examples, focusing on age- and condition-based replacement. We build upon each example in the subsequent sections to illustrate the increasing complexity of failure estimation and maintenance decision-making.[2]

[2] Programmed solutions per example are provided on https://github.com/LoeysS/50yEJORMaintenance

Table 2
Replacement age data of filaments, measured in weeks since their installation.

time to replacement	15	25	25	25	25	25	22	20	22	14	20	25	25	20	25	25	25
censored	0	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1

Table 3
Data of the filaments' age (in weeks since their installation) when their impedance has increased by 1Ω.

Filament 1	age	0.1	0.5	3.7	8.5	13.9	18.8	19.3	20.5	22.3	24.8
	excess impedance	1	2	3	4	5	6	7	8	9	10
Filament 2	age	0.3	0.9	6.4	10.3	10.4	11.5	14.8	22.1	24.7	27.2
	excess impedance	1	2	3	4	5	6	7	8	9	10
Filament 3	age	4.8	14.7	18.9	23.2	30.6	31.2	33.0	39.0	39.9	44.2
	excess impedance	1	2	3	4	5	6	7	8	9	10

Example 1 (*Age-based Replacement*). Interventional X-ray (IXR) machines make images of patients during medical procedures. A critical component is the filament that generates the X-rays. The filament is preventively replaced every 25 weeks unless it fails before that. The hospital has data on the age of different filaments at the time of their replacement (see Table 2).

The price to replace a filament preventively is $C_p = €500$. The replacement cost upon failure includes rescheduling medical procedures and other adverse events, estimated at $C_c = €6000$. The hospital would like to know whether the policy of replacing preventively at 25 weeks can be improved by lowering or raising the replacement age threshold.

Example 2 (*Condition-based Replacement*). The condition of a filament is given by its impedance, measured every time a current is run through the filament. The filament fails when its impedance is 10 Ω higher relative to its impedance when it is new. The hospital has data of moments when impedance increases by a whole Ohm for three filaments that were used until failure (see Table 3). The cost of a preventive and corrective replacement is identical to Example 1. The hospital would like to know after which impedance increase they should replace a filament.

2.1. Failure process estimation

Several approaches exist to predict when failures will occur if maintenance is not performed. The time until failure can be predicted based on historical failure data, or one can understand how assets degrade based on their condition. We discuss estimating time-to-failure models in Section 2.1.1 and degradation models in Section 2.1.2. These models correspond to Levels 1 and 2 of our maturity framework in Fig. 2.

2.1.1. Time-to-failure models

Predictions of the time to failure are usually framed as the probability that a system is still operational after a time t. This probability is called the reliability at time t and denoted $R(t)$. The field of reliability engineering studies the estimation and computation of reliability functions for complex engineering systems and is at the intersection of probability theory and engineering. An important concept in reliability engineering is the instantaneous failure probability of a system at age t, also called the hazard rate. If the random variable T with density $f(\cdot)$ and distribution $F(\cdot)$ denotes the lifetime of a component, then the hazard rate $h(t)$ is given by:

$$h(t) = \lim_{\epsilon \to 0} \mathbb{P}(T \le t + \epsilon \mid T \ge t)/\epsilon = f(t)/R(t).$$

Rather than estimating $R(t)$ directly, reliability engineers often estimate $h(t)$, given that the reliability is easily recovered from the hazard rate via simple integration as $R(t) = \exp\left(-\int_0^t h(x)dx\right)$. The hazard rate can be estimated through non-parametric and parametric approaches; see, e.g., Lewis (1996) and Ebeling (2001).

An important concern in estimating time-to-failure distributions is that relatively few data are available, and such data are often censored as a result of failure avoidance from preventive maintenance interventions. Parametric estimations can deal with censored data, and the distributional family choice can be informed by engineering knowledge. For example, the family of Weibull distributions is often used to model time-to-failure because it arises naturally as the limiting distribution of the minimum of a set of random variables. Such a minimum is important as the lifetime of an engineering system is given by the minimum lifetime of all its constituent components, the so-called weakest link.

The data to estimate the time-to-failure distribution is often given in the form of Example 1. That is, only data of the time to replacement of n components, denoted $x_1, \ldots, x_n$, and replacement data for each component i, be it preventive (censored, $c_i = 1$) or corrective replacement (uncensored $c_i = 0$), is available. Maximum Likelihood Estimation (MLE) can estimate the parameter θ of a given family of distributions as follows. Let $f(\cdot \mid \theta)$, $F(\cdot \mid \theta)$ and $R(\cdot \mid \theta) = 1 - F(\cdot \mid \theta)$ respectively be the density, distribution, and reliability function for the family under consideration. The log-likelihood $\mathcal{L}(\theta \mid \mathbf{x}, \mathbf{c})$ of a given sample $\mathbf{x} = (x_1, \ldots, x_n)$ $\quad \mathbf{c} = (c_1, \ldots, c_n)$ is then

$$\mathcal{L}(\theta \mid \mathbf{x}, \mathbf{c}) = \sum_{i=1}^n \ln\left(\mathbf{1}_{c_i=1} R(x_i \mid \theta) + \mathbf{1}_{c_i=0} f(x_i \mid \theta)\right), \tag{1}$$

where $\mathbf{1}_x$ is the indicator function which is 1 if x is true and 0 otherwise. The maximum likelihood estimator is then $\hat{\theta} := \arg\max_\theta \mathcal{L}(\theta \mid \mathbf{x}, \mathbf{c})$. Good parametric choices should generally be made based on comprehension of the physics of failure. Tinga (2013) can provide relevant guidance.

Example 3 (*Age-based Replacement Cont. from Example 1*). We let the random variable T denote the lifetime of an IXR filament. Suppose we assume that T follows a gamma distribution, i.e., $\mathbb{P}(T > t) = R(t \mid \alpha, \beta) = \int_t^\infty f(x \mid \alpha, \beta)dx$, where $f(x \mid \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1}\exp(-\beta x)}{\Gamma(\alpha)}$ is the probability density function of T and $\Gamma(z) = \int_0^\infty x^{z-1}\exp(-x)dx$ is the gamma function. The log-likelihood in Eq. (1) now becomes $\mathcal{L}(\alpha, \beta \mid \mathbf{x}, \mathbf{c}) = \sum_{i=1}^n \ln\left(\mathbf{1}_{c_i=1} R(x_i \mid \alpha, \beta) + \mathbf{1}_{c_i=0} f(x_i \mid \alpha, \beta)\right)$. A standard non-linear programming solver will show that $\mathcal{L}(\alpha, \beta \mid \mathbf{x}, \mathbf{c})$ is maximized in $(\hat{\alpha}, \hat{\beta}) = (9.5331, 0.3471)$ where the log-likelihood is $\mathcal{L}(\hat{\alpha}, \hat{\beta} \mid \mathbf{x}, \mathbf{c}) = -12.8814$. Thus $(\hat{\alpha}, \hat{\beta}) = (9.5331, 0.3471)$ are the maximum likelihood estimates of the parameters of a gamma distribution to model the censored lifetime data in Table 2. This time-to-failure distribution will be used to determine the optimal maintenance interval in Section 2.2.

In practical settings, it is often useful to update the distribution of the Remaining Useful Life (RUL) based on recent data, including multi-dimensional sensor readings. Depending on the nature of the equipment, there is much engineering literature on these topics; see Lei et al. (2018) for a recent overview. Methods from machine learning and AI such as Neural Networks (of many different architectures) are

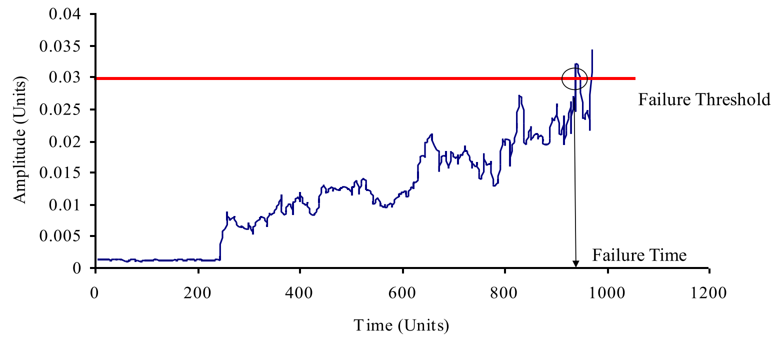

Fig. 3. A sample degradation path of a ball bearing.

recently gaining traction; see e.g. De Pater and Mitici (2023) and Liu and Gryllias (2020). Such methods are used to transform multidimensional sensor data into a one- or low-dimensional health indicator of RUL prediction. A major challenge is that there is much data on healthy equipment, but still very little data on equipment that has run to failure. These methods generate online RUL predictions, but the literature on how to incorporate changing estimates in decision-making is sparse.

2.1.2. Degradation models

Rather than only observing the asset/component's age, one can also improve the failure prediction by understanding how assets/components degrade based on their condition. What and how you measure the component's condition or degradation level depends heavily on the asset technology. The condition of a brake pad, for example, a component of disc brakes used in automotive and other applications, is its thickness. The thickness of brake pads can be measured periodically when a vehicle enters the maintenance shop. With modern sensing technology, however, it is also possible to continuously monitor such wear over time. In either case, we need to model the way the thickness of the brake pad evolves over time and use data to fit degradation models. With the installation of sensors in most modern high-tech equipment, it is possible to monitor changes in vibration amplitude, temperature, light intensity, concentration of contaminants in lubrication fluids, deformation, and position of parts relative to each other, generating numerous data points.

A degradation process is a stochastic process $X(t)$. We will assume for convenience and without loss of generality that $X(0) = 0$. There is a threshold L such that a component fails at time $T = \inf\{t \mid X(t) \geq L\}$; see Fig. 3. We assume that $\lim_{t\to\infty} X(t) \geq L$ so that a component fails almost surely during its usage. A common model of degradation occurs when $X(t)$ has stationary independent increments. That is when, for any $a > 0$ and $b \geq 0$, $X(a+b) - X(a)$ has the same distribution independent of $X(t)$ for all $t \leq a$. The stationary independent increments assumption also implies that degradation grows linearly in the sense that $\mathbb{E}[X(t)] = at$ for some $a > 0$. Many real degradation processes may not grow linearly but grow linearly after an appropriate transformation. An example of such a transformation is the logarithmic transformation for degradation that grows exponentially (e.g. Elwany et al., 2011). It is convenient to assume that a degradation increment comes from a distributional family in the class of linear exponential distributions. As such, $X(t)$ will be in this distributional family for all t. Examples include the normal/Gaussian (Elwany et al., 2011), (compound) Poisson (Drent et al., 2023), inverse Gaussian (Chen et al., 2015; Ye & Chen, 2014), gamma (Bautista et al., 2022; Van Noortwijk, 2009), and other distributions. Some of these references also discuss the physical interpretation of processes (e.g., Ye and Chen (2014) and Drent et al. (2023)).

Estimation of a degradation process using maximum likelihood estimation generally proceeds as follows. The parameters of $X(t)$ in a distribution family from the linear exponential class can generally be written as θt, where θ may be a vector. Data is often given as a degradation level at certain ages, as in Example 2. For our estimation, it is convenient to transform this data into degradation increments and the time between them. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be those increments and $\mathbf{t} = (t_1, \ldots, t_n)$ be the time associated with each increment. That is, x_i is the degradation accumulated during t_i time. Then, by hypothesis, x_i is a draw from a distribution family with parameter θt_i. Let us denote this distribution's probability mass or density function by $f(\cdot \mid \theta)$. Then, the log-likelihood $\mathcal{L}(\theta \mid \mathbf{x}, \mathbf{t})$ of the sample $(\mathbf{x}, \mathbf{t})$ is given by

$$\mathcal{L}(\theta \mid \mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} \ln\left(f(x_i \mid \theta t_i)\right). \tag{2}$$

The maximum likelihood estimate of the parameter θ is now given by $\hat{\theta} = \arg\max_\theta \mathcal{L}(\theta \mid \mathbf{x}, \mathbf{t})$. It can usually be obtained using either a standard non-linear programming solver or by numerically solving the Karush Kuhn Tucker conditions associated with the optimization problem $\max_\theta \mathcal{L}(\theta \mid \mathbf{x}, \mathbf{t})$. An application of this estimation procedure is given in the example below.

Example 4 (*Condition-based Replacement Cont. from Example 2*)**.** First, we transform the data in Table 3, which is given as total degradation and ages. The same data in terms of increments (we have $x_i = 1$ for all i) and time of increments $\mathbf{t}$, as shown in Table 4 for the condition data of filament 1.

If we assume degradation is a Poisson process with intensity λ, then the likelihood of sample i is given by $f(x_i, t_i \mid \lambda) = \frac{(\lambda t_i)^{x_i}}{x_i!}\exp(-\lambda t_i) = \lambda t_i \exp(-\lambda t_i)$ since $x_i = 1$ for all i. The log-likelihood function now becomes $\mathcal{L}(\lambda \mid \mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} \ln(f(x_i, t_i \mid \lambda))$ which is maximized in $\hat{\lambda} = 0.31185$ Ω per time unit with maximized log-likelihood $\mathcal{L}(\hat{\lambda} \mid \mathbf{x}, \mathbf{t}) = -44.9996$. This degradation process can then be used to determine when (and what) to maintain, which is discussed in the following section.

2.2. Maintenance optimization

Given failure process estimates, models used to optimize the timing of maintenance actions can be separated into two main paradigms: renewal reward theory and Markov decision processes. The renewal reward approach identifies repeating cycles in a stochastic system for which a decision rule has been established. In Markov decision processes (MDPs), there is, a priori, neither an established decision rule nor a repeating cycle, only possible system states and decision options. MDPs can be used to prove that a certain type of decision rule is optimal, whereas renewal reward processes can be used to find the optimal parameters of a given decision rule. We discuss each approach separately. Note that these approaches correspond to Levels 1 and 2 of Fig. 2.

Table 4
Increment times data (in weeks) for the first 21 measurements of filament 1.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	…
t_i	0.1	0.4	3.2	4.8	5.4	4.9	0.5	1.2	1.8	2.5	0.3	0.6	5.5	3.9	0.1	1.1	3.3	7.3	2.6	2.5	4.8	…

2.2.1. Renewal reward theory

A renewal process is a counting process with two i.i.d. sequences of random variables, namely a sequence of rewards or costs (e.g., the cost of a maintenance intervention), W_i, and the time between the occurrence of each successive reward (e.g., the time between maintenance visits), X_i, with common distribution $\mathbb{P}(X_i \leq x) = F(x)$. Each sequence is i.i.d., but X_i and W_i may be correlated. It is usual to interpret $\mathbb{E}[X_i]$ as the *expected cycle length* (ECL) and $\mathbb{E}[W_i]$ as the *expected cycle costs* (ECC). Let $S_i = \sum_{k=1}^{i} X_k$ denote the time that the ith renewal occurs with $S_0 = 0$ by convention. The number of cycles (or renewals) up to time t is given by $N(t) = \max\{k \in \mathbb{N} \mid S_k \leq t\}$. Then the total reward (cost) up to time $t > 0$ is denoted by $Y(t)$ and satisfies

$$Y(t) = \sum_{i=1}^{N(t)} W_i.$$

Here, $Y(t)$ is a renewal reward process. The renewal reward theorem (e.g., Ross, 1996) states that

$$\lim_{t\to\infty} \frac{Y(t)}{t} = \lim_{t\to\infty} \frac{\mathbb{E}[Y(t)]}{t} = \frac{\mathbb{E}[W_i]}{\mathbb{E}[X_i]} = \frac{ECC}{ECL}. \tag{3}$$

Given that the conditions for a renewal process hold, the optimal replacement age and expected maintenance cost per time unit for both failure- and age-based replacement policies can be determined using such a renewal reward process. In a failure-based policy, X_i represents the time to failure of a component. The maintenance cycle is then the time between two successive failures and corresponding corrective maintenance actions. The maintenance cycle has associated with it the expected cycle cost (ECC) and the expected cycle length (ECL), expressed in time units.

An age-replacement policy replaces a component when it reaches some age τ. Suppose that the lifetimes of each component i are T_i and $\{T_i\}_{i=1}^{\infty}$ is i.i.d. We can conceive of the cost under such a policy as a renewal reward process. The time between two successive maintenance interventions is now denoted as $X_i = \min(\tau, T_i)$ and $\{X_i\}_{i=1}^{\infty}$ is an i.i.d. sequence. If the component fails before it is preventively replaced at age τ, a cost $C_c > C_p$ is incurred for unplanned corrective maintenance while C_p is incurred for planned replacement at age τ. With probability $1 - F_T(\tau) = \mathbb{P}(T_i > \tau)$, the cost in a given maintenance cycle is C_p, whereas the cost is C_c with probability $F_T(\tau) = \mathbb{P}(T_i \leq \tau)$:

$$W_i = \begin{cases} C_p & \text{with probability } 1 - F_T(\tau), \\ C_c & \text{with probability } F_T(\tau). \end{cases}$$

The average cost per time unit for a given τ is then given by $g(\tau) = ECC/ECL$, where

$$ECC = \mathbb{E}[W_i] = F_T(\tau)C_c + (1 - F_T(\tau))C_p \quad \text{and} \quad ECL = \mathbb{E}[X_i] = \mathbb{E}[\min(T, \tau)].$$

The value for τ can be optimized by setting $dg(\tau)/d\tau = 0$ and solving for τ or numerically minimizing $g(\tau)$ directly.

Examples in literature include the seminal work by Barlow and Hunter (1960) on optimal age-based policies. Recent advances include extending the use of the renewal reward process in age-based maintenance to multi-components (Arts & Basten, 2018), unpunctual maintenance (Sanoubar et al., 2021), and population heterogeneity (Dursun et al., 2022).

The time between renewals must be i.i.d. for renewal theory to apply. In the age-replacement policy, renewals correspond to replacements, but this will not always be the case. For example, replacements for block policies are not renewal points; see Barlow and Proschan (1996) Section 3.3. It is common to define cycles that are not i.i.d. and use a renewal reward approximation based on the false assumption that cycles are independent. Relevant examples include evaluating opportunistic maintenance given age-based and CBM policies for single components in the presence of other components (e.g. Poppe et al., 2018; Zhu et al., 2017) and for joint maintenance of multi-components (e.g. Peng & Zhu, 2017; Zhu et al., 2015).

Renewal reward processes are also applied to condition-based maintenance policies (CBM), as in redundant multi-component systems (Zhang et al., 2020) and partially observable systems (Kim & Makis, 2013; Van Staden & Boute, 2021). (Poppe et al., 2018) use renewal theory to determine a multi-component CBM policy in a real case of a compressor manufacturer that offers after-sales services, and Jardine and Tsang (2005) offer many small real case studies.

Example 5 (*Age-Based Replacement Cont. From Examples 1 and 3*). For an age replacement policy for an IXR filament with threshold τ we have $ECC = 6000 \cdot F(\tau \mid \alpha, \beta) + 500(1 - F(\tau \mid \alpha, \beta))$, with $\alpha = 9.5331$ and $\beta = 2.8808$ obtained in Example 3. The expected cycle length is

$$\begin{aligned} ECL &= \int_0^{\tau} x f(x \mid \alpha, \beta) dx + \tau(1 - F(\tau \mid \alpha, \beta)) \\ &= \int_0^{\tau} \alpha\beta f(x \mid \alpha + 1, \beta) dx + \tau(1 - F(\tau \mid \alpha, \beta)) \\ &= \alpha\beta F(\tau \mid \alpha + 1, \beta) + \tau(1 - F(\tau \mid \alpha, \beta)). \end{aligned}$$

The weekly cost for a given replacement threshold τ is $g(\tau) = ECC/ECL$. Numerically minimizing for τ gives the optimal $\tau^* = 12.2$ weeks and $g(\tau^*) = €49.29$ per week. The current policy of replacing every 25 weeks has cost $g(25) = €126.30$ per week. Optimization of the age replacement parameter saves $(g(25) - g(\tau^*))/g(25) = 61\%$ relative to the current solution with $\tau = 25$ weeks.

2.2.2. Markov decision processes

Markov decision processes (MDPs) deal with decision-making over time under uncertainty that is sequentially revealed to a decision-maker. The elements of an MDP are the state space in which the decision maker finds herself, such as, e.g., the machine's age or condition, the action space of possible actions the decision maker can take, such as performing maintenance or not, and a stochastic mechanism by which the decision-maker finds herself in a new state after one time period, dependent on the current state and action taken. The decision-maker incurs a cost based on the state and action taken in each period. These elements can be arranged in a tuple $(\mathcal{S}, \mathcal{A}, P, c)$ where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, P is the transition law, and $c : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the cost function that determines the cost incurred in state $x \in \mathcal{S}$ when decision $a \in \mathcal{A}$ is taken.

The decision-maker may have different objectives and/or time horizons. For instance, she may want to minimize maintenance costs or maximize machine uptime. Here, we will focus on a decision-maker who seeks to minimize the average cost incurred per period over an infinite horizon. The decision maker then seeks a policy $\pi : \mathcal{S} \to \mathcal{A}$, informing the decision maker of what to do in each possible state to achieve the objective. Such a policy will usually satisfy a set of equations known as the Bellman optimality equations.

We illustrate this with a condition-based maintenance example. Suppose the condition of a component at age $t \in \mathbb{R}_+$ is given by $X(t)$ and $X(t)$ is a process with independent stationary increments as described in Section 2.1.2. $X(t)$ takes values in the state-space $\mathcal{S}$. By convention, we will let 0 denote the good as new state and $L > 0$ denote the failed state. Suppose further that the component can only be replaced on weekends. Then, the decision-maker observes $X(t)$ each weekend and must decide whether to replace the component. Let X_t denote the degradation level of the component at the end of week $t \in \mathbb{N}$. Then if the decision-maker decides not to replace, she will find the component the week after in state $X_{t+1} = \min(X_t + Z_t, L)$, where $\{Z_t\}_{t=1}^{\infty}$ is an i.i.d. sequence of non-negative random variables. If the component is replaced, the decision-maker will find the component next week in a state $X_{t+1} = \min(0 + Z_1, L) = \min(Z_1, L)$. Thus, the action space is given by $\mathcal{A} = \{1, 0\}$ where 1 denotes the decision to replace and 0 denotes the decision not to replace. We assume that the decision-maker has to replace the component when it is in the failed state L, i.e., $\mathcal{A}_L = \{1\}$ where $\mathcal{A}_x$ denote the decision possible in state $x \in \mathcal{S}$.

For illustration purposes, we will assume that Z_i is distributed on the integers such that $\mathcal{S} = \{0, 1, 2, \dots, L\}$. Let A_t denote the decision taken in weekend t. Then,

$$p_{ij}^a = \mathbb{P}(X_{t+1} = j \mid A_t = a, X_t = i) = \begin{cases} \mathbb{P}(Z = j - i) & \text{if } a = 0, \quad j < L \\ \mathbb{P}(Z \geq j - i) & \text{if } a = 0, \quad j = L \\ \mathbb{P}(Z = j) & \text{if } a = 1, \quad j < L \\ \mathbb{P}(Z \geq j) & \text{if } a = 1, \quad j = L. \end{cases} \tag{4}$$

P denotes the collection of all these transition probabilities and is called the transition law. The cost function associated with taking decision $a \in \mathcal{A}$ in state $x \in \mathcal{S}$ is given by

$$c(x, a) = \begin{cases} 0 & \text{if } a = 0 \quad x < L \\ C_p & \text{if } a = 1 \quad x < L \\ C_c & \text{if } a = 0 \quad x = L, \end{cases} \tag{5}$$

where C_p and $C_c > C_p$ denote the cost of preventive and corrective replacement, respectively.

The average cost over an infinite horizon of a given policy π is defined as

$$g^\pi = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=1}^{T} c(X_t, A_t) \right],$$

where $\mathbb{E}_\pi$ denotes the expectation taken with respect to the two discrete-time stochastic processes, X_t and A_t, induced by the policy π. The decision-maker seeks the optimal policy $\pi^* = \arg\min_\pi g^\pi$. The optimal cost rate, $g^* = g^{\pi^*}$ with policy π^* satisfy the Bellman optimality equations

$$\begin{aligned} V(x) + g^* &= c(x, \pi^*(x)) + \mathbb{E}[V(X_{t+1}) \mid X_t = x, \quad A_t = \pi^*(x)] \\ &= \min_{a \in \mathcal{A}_x} \left\{ c(x, a) + \mathbb{E}[V(X_{t+1}) \mid X_t = x, \quad A_t = a] \right\} \\ &= \min_{a \in \mathcal{A}_x} \left\{ c(x, a) + \sum_{j \in \mathcal{S}} p_{x,j}^a V(j) \right\} \end{aligned} \tag{6}$$

which hold for all $x \in \mathcal{S}$. The function $V(x)$ is called the (relative) value function and can be determined algorithmically through value iteration, policy iteration, or linear programming (Puterman, 1994). The value iteration solves the Bellman equations iteratively by computing the successive approximations $V_1, V_2, \dots$ through the recursion

$$V_n(x) = \min_{a \in \mathcal{A}_x} \left\{ c(x, a) + \sum_{j \in \mathcal{S}} p_{x,j}^a V_{n-1}(j) \right\}, \tag{7}$$

with the initial condition $V_0(x) \equiv 0$ for all $x \in \mathcal{S}$. Let the policy π_n be defined by $\pi_n(x) = \arg\min_{a \in \mathcal{A}_x} \{c(x, a) + \sum_{j \in \mathcal{S}} p_{x,j}^a V_{n-1}(j), \sum_{j \in \mathcal{S}} p_{x,j}^a V_{n-1}(j)\}$ for $x \in \mathcal{S}$. Then the policy π_n has a cost-rate within $\epsilon > 0$ of g^* if $\max_{x \in \mathcal{S}}(V_n(x) - V_{n-1}(x)) - \min_{x \in \mathcal{S}}(V_n(x) - V_{n-1}(x)) < \epsilon$ and g^* equals $(\max_{x \in \mathcal{S}}(V_n(x) - V_{n-1}(x)) + \min_{x \in \mathcal{S}}(V_n(x) - V_{n-1}(x)))/2$ within a precision of ϵ. Thus for a given $\epsilon > 0$ one can iterate Eq. (7) until the criterion above is satisfied. At that point, the Bellman Eqs. (6) are also satisfied within ϵ tolerance.

The MDP approach has also been used to model imperfect repairs (Kurt & Kharoufeh, 2010) and multiple component systems with redundancy (Andersen et al., 2022; Keizer et al., 2016). Semi-MDPs can be used to model the maintenance decision in continuous-time, as in Drent et al. (2019), Huang and Guo (2011). MDPs can also be applied to periodic age-based maintenance policies. Bayesian learning through MDPs can be used to advance the age-based maintenance decision given unplanned corrective maintenance actions (Van Staden et al., 2022).

Example 6 (*Condition-based Maintenance Cont. from Examples 2 and 4*)**.** Suppose that any replacement of the IXR should happen on the weekend when there are no scheduled patient procedures. The condition is measured every weekend and the filament fails at $L = 10$. The distribution of a degradation increment Z is given by $\mathbb{P}(Z = x) = \exp(-\lambda)\lambda^x / x!$, with the estimated degradation rate $\lambda = 0.31185\ \Omega$ per week obtained in Example 4. The transition law can be computed with Eq. (4) and the cost function using Eq. (5). Appendix C provides the transition probability matrices of the transition law. The optimal replacement policy can be computed with value iteration using Eq. (7) with optimality tolerance $\epsilon = 10^{-6}$. This yields $g^* =$ €22.04 per week, and the IXR unit should be replaced when the degradation is 7 Ω or higher, i.e., $\pi(x) = 1$ if $x \geq 7$ and $\pi(x) = 0$ otherwise. This is also called a threshold policy with threshold 7. The lifetime and condition data in Examples 1 and 2 come from the same process. Thus, we find that condition-based maintenance saves (49.29–22.04)/49.29 = 55% relative to the filament's optimal age-based maintenance policy derived in Example 5.

2.3. Online failure estimation and maintenance optimization

The aforementioned approaches work well, provided sufficient data is available for failure estimation. To increase the data pool, an implicit assumption often used is that the degradation process parameters of all components are identical. An alternative approach recently gaining momentum is learning and updating each individual component's degradation process parameters over time through cumulative signal observations obtained via condition monitoring of a heterogeneous pool of components. The most current understanding of the degradation process is then used to customize the maintenance planning for each individual component. We explain how to update one's understanding of the degradation process of an individual component in a Bayesian manner in Section 2.3.1. We use this updating procedure to formulate a partially observable Markov decision process (POMDP) that integrates failure estimation and maintenance optimization in Section 2.3.2. Such a learning and decision model approach corresponds to Levels 3 and 4 of Fig. 2.

2.3.1. Bayesian degradation process learning

Consider a heterogeneous population of components that degrades with independent increments from the linear exponential family of distributions. The degradation process parameters differ for each individual component but, across the population, they can be fitted to some exogenously given distribution. (We will illustrate later how to estimate such a distribution.) Specifically, assume that degradation follows a Poisson process and the degradation rates of the individual components can be fitted to a gamma distribution with shape α_0 and inverse scale β_0. The choice of these distributions is not arbitrary. To obtain a tractable model, the distribution of the degradation increments must come from the linear exponential family (Morris, 1982). Only then does a tractable conjugate distribution exist to model the population heterogeneity. A Poisson process where the rates follow a gamma distribution is just one such choice that we will use to illustrate the approach.

Let X_t denote the component's degradation level at age t. This monitored condition can be observed, but the Poisson rate Λ of this component's degradation process is unknown. All we initially assume is that Λ follows a gamma distribution with shape α_0 and inverse scale β_0 (notation $\Lambda \sim \Gamma(\alpha_0, \beta_0)$), so that $\mathbb{P}(\Lambda \leq z) = \int_0^z \frac{\beta_0^{\alpha_0} y^{\alpha_0 - 1} \exp(-\beta_0 y)}{\Gamma(\alpha_0)} dy$. By observing a component's degradation over time, $\boldsymbol{X}_t = (X_0, X_1, \dots, X_t)$,

Table 5
Condition data of two additional filaments.

Filament 4	age	0.5	2.1	3.8	3.9	4.8	9.2	9.4	11.9	13	13.5
	impedance excess	1	2	3	4	5	6	7	8	9	10
Filament 5	age	0.3	2.4	19.3	29.8	31.9	33.7	35.8	46.0	57.8	58.0
	impedance excess	1	2	3	4	5	6	7	8	9	10

we want to increase our knowledge of Λ as expressed by its distribution $P(\Lambda \leq z|\boldsymbol{X}_t)$

Suppose we observe the degradation level every week. After the first week, we know the realization x_1 of X_1. Now we can update the belief of the distribution of the Poisson parameter Λ, knowing that x_1 is the realization of X_1: $\mathbb{P}(\Lambda \leq z \mid X_1 = x_1)$. Applying Bayes' theorem, we find:

$$\mathbb{P}(\Lambda \leq z \mid X_1 = x_1) = \int_0^z \frac{(\beta_0 + x_1)^{\alpha_0+1} y^{\alpha_0-1+1} \exp(-(\beta_0 + x_1)y)}{\Gamma(\alpha_0 + 1)} dy,$$

that is $\{\Lambda|X_1 = x_1\} \sim \Gamma(\alpha_0 + 1, \beta_0 + x_1)$. After observing $X_t = x_t$, the same argument can be repeated to find $\{\Lambda|X_t = x_t\} \sim \Gamma(\alpha_0 + t, \beta_0 + x_t)$. Alternatively, $\{\Lambda|X_t = x_t\} \sim \Gamma(\alpha_t, \beta_t)$ with $\alpha_t = \alpha_0 + t$ and $\beta_t = \beta_0 + x_t$. The coefficient of variation of $\{\Lambda|X_t = x_t\}$ decreases over time and is given by $\frac{1}{\sqrt{\alpha_0+t}}$, i.e., an observer becomes more certain about the actual degradation parameter Λ of an individual component by observing its degradation process over time.

In Bayesian statistics, $\Gamma(\alpha_0, \beta_0)$ is called the prior distribution of Λ, and $\Gamma(\alpha_t, \beta_t)$ its posterior distribution. The initial parameters α_0 and β_0 are hyperparameters and model how the degradation rate varies from one component to another. We illustrate a possible way of estimating α_0 and β_0 in Example 7 below. By contrast, α_t and β_t encode our knowledge of the degradation rate of an individual component after observing its degradation for t time units.

The distribution of a degradation increment of an individual component given the past degradation of this component, i.e., $\mathbb{P}(X_{t+1} - X_t \leq x \mid X_t = x_t)$, is called the posterior predictive distribution. It can be found using the law of total probability. The posterior distribution for the case of Poisson degradation with a gamma prior is known to be a negative binomial distribution with shape parameter $r_t = \alpha_t$ and success probability $p_t = 1/(\beta_t + 1)$.

Example 7 (*Estimation of Hyperparameters Cont. from Example 2*)**.** Consider the case where the failure data of two additional filaments is collected during operation (see Table 5).

Suppose we assume that the degradation process of the five filaments (whose monitored degradation paths are shown in Tables 3 and 5) follows five Poisson processes whose random rate Λ across the components follows a $\Gamma(\alpha_0, \beta_0)$ distribution. Let $t_{\max,i}$ and $x_{\max,i}$ denote the maximum age and degradation level observed for filament i. The number of times the impedance excess of component i increased in $t_{\max,i}$ time units has a Poisson distribution with mean $\Lambda t_{\max,i}$. By the scaling property of the gamma distribution, $x_{\max,i}$ will be a draw from a Poisson distribution with a gamma prior on the rate given by $\Gamma(\alpha_0, \beta_0/t_{\max,i})$, which as indicated above, has a negative binomial distribution with shape $r = \alpha_0$ and success probability $p = \frac{t_{\max,i}}{\beta_0+t_{\max,i}}$. Thus the log-likelihood can be expressed as

$$\mathcal{L}(\mathbf{x}_{\max}, \mathbf{t}_{\max} \mid \alpha_0, \beta_0) = \sum_{i=1}^{5} \ln \left[\binom{x_{\max,i} + \alpha_0 - 1}{x_{\max,i}} \left(\frac{t_{\max,i}}{\beta_0 + t_{\max,i}} \right)^{x_{\max,i}} \left(\frac{\beta_0}{\beta_0 + t_{\max,i}} \right)^{\alpha_0} \right],$$

which is maximized in the MLE estimates $\hat{\alpha}_0 = 7.4097$ and $\hat{\beta}_0 = 21.6919$. These are the initial estimates of the prior distribution Λ. By observing the component's degradation over time, these parameters are updated by applying Bayes' theorem.

2.3.2. Partially observable Markov decision processes

The most recent Bayesian degradation estimates can be integrated into maintenance optimization in real-time. Partially Observable Markov Decision Processes (POMDP) provide a framework for this integration. In contrast to the conventional MDPs described in Section 2.2.2, a POMDP has parts of the state space that cannot be directly observed. The component's degradation process parameters are part of the current state but are not directly observable by the decision maker. The only two parts of the state space that the decision-maker can observe are the degradation level and age of the current component. Knowledge about the unobservable degradation parameter is encoded in the most recent belief distribution, i.e., in $\mathbb{P}(\Lambda \leq z \mid X_t = x_t)$.

Consider the same problem setting described in Section 2.2.2, except that the Poisson process degradation rate of each component is unknown and can be fitted to a gamma distribution across all components with shape α_0 and inverse scale β_0. It is convenient to define the random variable $Z(x,t)$ as the distribution of a degradation increment over a week, conditional on the current component having degraded to level x in t weeks, i.e.

$$\mathbb{P}(Z(x,t) \leq z) := \mathbb{P}(X_{t+1} - X_t \leq z \mid X_t = x).$$

We saw in Section 2.3.1 that $Z(x,t)$ has a negative binomial distribution with shape $\alpha_t = \alpha_0 + t$ and success probability $p_t = 1/(\beta_t + 1) = 1/(\beta_0 + x + 1)$. The same approach in Section 2.2.2 can now be followed to find that the Bellman equations are given by

$$V(x,t) + g^* = \begin{cases} \min \begin{array}{l} \{\ \mathbb{E}[V(\max(L, x + Z(x,t), t+1))], C_p \\ +\mathbb{E}[V(\max(L, Z(0,0)), 1)]\ \} \end{array}, & \text{if } x < L \\ C_c + \mathbb{E}[V(\max(L, Z(0,0)), 1)], & \text{if } x = L. \end{cases}$$

As before, the Bellman optimality equations can be solved by value or policy iteration or linear programming after truncating the age dimension of the state space to a sufficiently high value.

Example 8 (*Solution of POMDP, Cont. from Examples 2, 4, and 6*)**.** From Example 7, we know that the hyperparameters are estimated to be $\alpha_0 = 7.4097$ and $\beta_0 = 21.6919$. Solving the POMDP now gives $g^* =$ €20.49 per week, and the optimal replacement policy is a threshold policy given by $\pi(x,t) = 1$ if $x \geq T(t)$ and $\pi(x,t) = 0$ if $x < T(t)$, where the age-dependent thresholds are given by $T(t) = 7$ for $t \in \{1, \ldots, 25\}$, $T(t) = 8$ for $t \in \{26, \ldots, 146\}$ and $T(t) = 9$ for $t > 146$. Compared to Example 6, the weekly cost is reduced by (22.04−20.49)/20.49 = 7.55%. However, some care should be taken in comparing these numbers as they are computed under different assumptions about reality because (i) more data is available in this example compared to Example 6, and (ii) this example assumes that components are statistically distinguishable whereas Example 6 assumes components are statistically indistinguishable.

The approach in Section 2.3.2 works when there are linear sufficient statistics that summarize the entire degradation path (x and t in our example). Several more sophisticated degradation processes have been studied this way, e.g., Elwany et al. (2011) for Brownian motion, Chen et al. (2015) for the Inverse Gaussian processes, and Drent et al. (2023) for compound Poisson processes. All these authors consider continuous-time deterioration signal updating and determine optimal control limit policies for component replacement. Van Oosterom, Peng, and Van

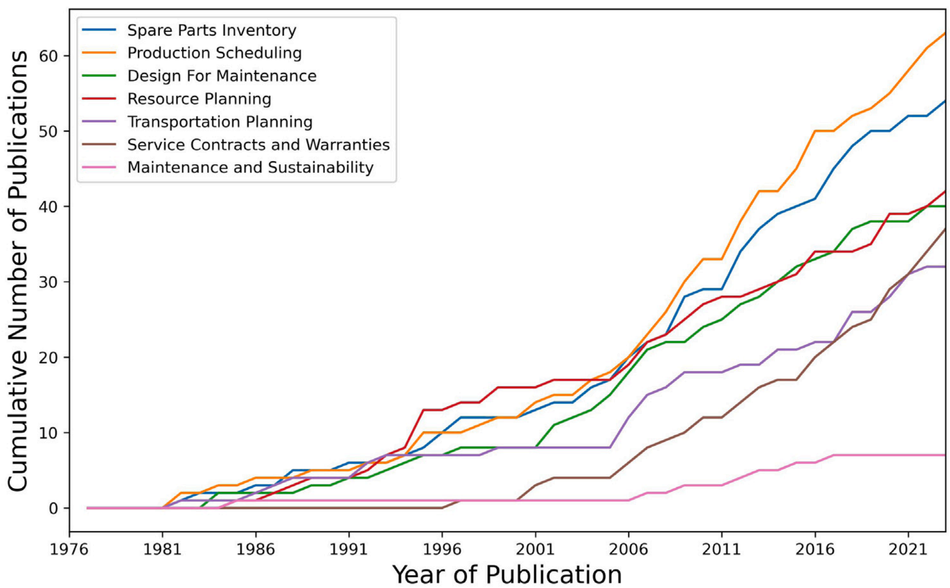

Fig. 4. Cumulative number of EJOR publications on maintenance and its interface with other domains.

Houtum (2017) makes more generic assumptions on the degradation process, but population heterogeneity is constrained to lie in a finite set. Other papers that consider situations in which the system state is only partially observable are Gamiz et al. (2023), Kim and Makis (2013), Maillart (2006), Van Oosterom, Maillart, and Kharoufeh (2017), Van Staden and Boute (2021), Zhang and Zhang (2023), and Deep et al. (2023). Drent et al. (2023) provide and extensive case study of an interventional X-ray machine from Philips electronics as well as a real-life degradation data set. Elwany et al. (2011) provide a real case of ball bearing degradation.

Bayesian updating and POMPDs are also used in age-based maintenance. Dursun et al. (2022) learn the probability that a component is drawn from either a weak or strong pool. Drent, Kapodistria, and Boxma (2020) use failure and censored observations to update the time-to-failure distribution of a component for an age-based replacement policy.

3. Interfaces with other domains

Maintenance optimization interfaces with several other domains. Besides optimizing the maintenance timing and content, an entire logistics support system is required to perform maintenance, such as spare parts inventory or resource planning. Maintenance is also increasingly offered as part of service contracts or warranties, and there is a link between maintenance and sustainability. These interfaces highlight the multidisciplinary nature of maintenance research. Fig. 4 shows the cumulative number of publications in EJOR devoted to such interfaces, indicating increased attention in the last fifteen years. In what follows, we discuss each of these interfaces.

3.1. Spare parts inventory management

Maintenance planning is closely linked to spare parts inventory management. Indeed, maintenance and component replacement can only be performed when the spare parts necessary to perform maintenance — also referred to as the “service parts” — are available. The unavailability of a machine part can jeopardize its maintenance, inducing time delays and, in turn, higher costs. As a result, the highly stochastic nature of machine breakdowns forces the need for inventory buffers. However, it is common for companies of moderate size to carry thousands of different spare parts in inventory. That means considerable capital is tied up if only one extra part is held for each item. This capital investment requirement has led to specialized spare parts inventory models that focus on improving the part availability whilst limiting the inventory investment. We acknowledge that many companies implement vendor-managed inventory and may have negotiating clout that requires the spare part to be delivered within a specified time window. In those cases, spare parts inventory management remains critical from the vendor’s perspective. Stein (2010) provides an industry perspective case study on how ASML, a vendor in the semiconductor industry, does this.

Maintenance parts have specific properties that render their inventory management different from many other products. The most prominent characteristic is that demand is intermittent. Specifically, demand is often zero for several consecutive periods, and it is only positive occasionally when the corresponding part is replaced. The maintenance policy and breakdowns thus dictate the spare part consumption and inventory management. Spare part inventory management is the first successful application of multi-echelon inventory management; the METRIC model is short for Multi-Echelon Technique for Recoverable Item Control (Sherbrooke, 1968). An extensive literature on spare parts inventory models has been consolidated in the books by Sherbrooke (2006), Muckstadt (2004), and Van Houtum and Kranenburg (2015).

Our text-mining analysis reveals 54 EJOR papers with both “`maintenance`” and either “`spare parts`”, “`inventory`”, or “`stocking`” in their title, abstract, or keywords. Most papers consider the benefits of joint optimization of maintenance and inventory control, e.g., Wang (2012), Kabir and Al-Olayan (1996), and Keizer et al. (2017). They show how significant savings can be obtained by optimizing the maintenance planning and the timing of ordering spare components.

We found some additional EJOR papers on the maintenance-inventory interface that are not included in our aforementioned 54 papers. Thomas and Osaki (1978), for instance, optimize the inventory policy for a preventive maintenance policy, but they do not explicitly reference `maintenance` in their title, abstract, or keywords. We refer to Hu et al. (2018) for a recent review of spare parts inventory control models.

A promising research avenue in this field uses maintenance policy information to forecast spare part demand and improve inventory control (see, e.g., Van der Auweraer et al., 2019, for a review). Romeijnders et al. (2012), for instance, study a real case at Fokker Services (aerospace spare parts). They take the additional repair information into account to forecast demand for a spare part. The rationale is that the ability to recognize what causes a change in the demand for spare parts, contrary to existing methods, should lead to better demand forecasts. With the growing data collection and processing opportunities, future research may be devoted to using these data to 'predict the unpredictable.'

3.2. Production scheduling

There is also an interface between production and maintenance planning. Machines are unavailable for production when faced with maintenance. Maintenance can also improve the production rate or speed (and postponing maintenance may have an adverse effect). The production level also influences the deterioration rate, so maintenance planning and production decisions must be jointly optimized. We found 63 papers in EJOR on the interface between `production` and `maintenance`. Our text-mining analysis initially revealed 107 maintenance papers with the stem `product`, but most merely discussed maintenance in a production facility or an environment selling products.

The literature on this interface addresses how production should be modified to incorporate maintenance. This question can be answered from a scheduling or profit maximization perspective.

Production scheduling aims to complete a set of jobs such that their tardiness is minimized. A typical assumption is the constant availability of machinery. However, one should integrate additional maintenance constraints in a scheduling model to incorporate machine downtime due to maintenance. Geurtsen et al. (2023) reviews the literature on the integration of maintenance with resource and production scheduling. For instance, the production schedule can be modified by adding a deterioration factor that is reset by performing maintenance (see, e.g., Gara-Ali et al., 2016; Lalla-Ruiz & Voß, 2016; Wang et al., 2018). Other papers focus on maintenance scheduling that needs to be performed in a flexible time window or before the machine age passes a certain threshold within a given maintenance interval (see, e.g. Topal & Ramazan, 2010), or by optimizing the maintenance interval as well (see, e.g., Xia et al., 2012).

In environments where the revenue from production is time-dependent, such as wind turbines or cloud computing, the production rate needs to be determined dynamically over time. The production rate additionally impacts the machine deterioration. Uit het Broek et al. (2020) dynamically adjusts the production rate based on the machine condition and the revenue potential. Drent et al. (2024) extend their model by optimizing the preventive maintenance interval for a heterogeneous machine population using Bayesian learning.

3.3. Design for maintenance

The ability to maintain an asset largely depends on its design. Smets et al. (2012) introduce a "Design for Availability" framework to cost-effectively optimize the availability of capital goods throughout their entire lifetime and illustrate their framework at a global manufacturer of capital goods in the food processing industry. We found 9 EJOR papers that study the design of the equipment for maintenance. Papers studying system design usually focus on system reliability, e.g., Bei et al. (2019). Redundancy design, which increases system reliability, is, therefore, a related problem. Our text mining revealed 31 EJOR papers with `redundancy` and `reliability` in the title, abstract, or keywords. It appears that design for maintenance is a young field within Operational Research without standardized terminology.

An important question in this domain is the design of so-called line replaceable units (LRUs). An LRU is a collection of connected parts in a system that is replaced when any part of the LRU fails. Companies use LRUs as a mechanism to reduce system downtime after a failure. The design of LRUs determines how fast a replacement is performed, such that a smart design reduces replacement and downtime costs. A firm must purchase/repair an LRU upon failure, and large LRUs are more expensive to purchase/repair. Hence, a firm seeks to design LRUs to minimize the average costs per time unit. Examples of research in this domain include Parada Puig and Basten (2015), Lambert (2007), and Driessen et al. (2024). Van Geel (2018) and Van Deursen (2020) provide extensive case studies such as Thales Radar Systems and Canon Printing, respectively.

The EU Green Deal established the 'Right to Repair' to make the European economy circular and resource-efficient. To prolong product lifecycles, consumers should have the right to repair their (electronic) devices instead of discarding them. In the past, there has been little incentive to create repairable products. The marginal cost of repair was often too high, and the marginal benefit was low. EU's Right to Repair may inflate product prices. It will force manufacturers to design their products for repairability. And for each design, they must forecast how many repairs to expect. Incorporating their cost into the product price can be seen as a warranty or service contract, where the right to repair is guaranteed during a specific period. One of the consequences is that a higher-priced product while making repairs cheaper, will indeed incentivize the number of repairs.

Modular design for maintenance and repair is also gaining broader traction. "Fairphone" is an interesting example of a manufacturer of modular smartphones that are designed for reuse and recycling. Fairphone's mission is to design longer-lasting products that are easier to repair.[3] Users can repair the phone by replacing slot-in modules using a standard screwdriver, and they offer a recycling program that facilitates and stimulates returns of old devices and their modules to reduce electronic waste. In line with the servitization trend discussed below in Section 3.6, they also offer subscription contracts that include an all-inclusive repair & swap service and a lifetime warranty. Compared to similar smartphones, however, there is a price premium.

3.4. Resource planning

Maintenance requires resources such as technicians, tools, and spare parts. Such resources may be scarce relative to the maintenance jobs needed in a given setting, especially for a pool of assets, each consisting of multiple components. In a simple setting with failure-based maintenance, the question is how many resources are needed to guarantee a requisite level of asset availability. This problem is known as the machine repairmen problem and has an extensive literature (see, e.g., Haque & Armstrong, 2007, for an overview). Text mining indicates 9 EJOR papers with `machine repair` in title, abstract, or keywords, while there are 33 with `capacity` and `maintenance`.

Under condition-based maintenance, a maintenance manager must decide how to allocate resources to different maintenance priorities (Olde Keizer et al., 2017). This setting gives rise to a restless multi-armed bandit problem that is challenging to solve without identifying and exploiting additional problem-specific structures (see, e.g., Demirci et al., 2024; Glazebrook et al., 2005; Larrnaaga et al., 2016).

Resource constraints can also pose problems when no stochasticity is involved in the condition of assets. This happens, for example, in the maintenance of infrastructures where shutting an asset down may affect whether other assets can operate. An excellent example of resource-constrained maintenance planning is Urbani et al. (2023).

[3] https://www.fairphone.com/en/story/

3.5. Transportation planning

The resources needed to perform maintenance are often in different locations from the assets that require maintenance. Either the asset or the resources must be transported for maintenance. This creates interesting challenges at the interface of transportation and maintenance planning. Their intersection seems to be a trending field in the least fifteen years, as shown in Fig. 4.

In some cases, the assets travel and the maintenance resources are stationary. Examples include aircraft, rolling stock, and naval vessels maintained in a hangar, maintenance track, or dry dock. The challenge here is to jointly design a transportation schedule and maintenance locations such that assets are in the right location when they need maintenance; see, e.g., Tönissen and Arts (2020), Tönissen et al. (2019), Feo and Bard (1989), and Gopalan (2014).

In most other cases, the assets are stationary (e.g., production equipment), and the maintenance resources (e.g., technicians with tools and spare parts) travel to different sites to perform maintenance (e.g. Pham & Kiesmüller, 2024). The condition of the asset should inform the transportation plan of the resource. The natural formulation for such a problem is a Markov Decision Process that suffers from the curse of dimensionality when solved to optimality using dynamic programming. Yet, there are some examples where small-scale instances are solved and used to develop heuristics (e.g., Drent, Keizer, & Van Houtum, 2020; Lagos et al., 2020; Sanoubar et al., 2023) or where approximate techniques such as reinforcement learning are used to find good policies (e.g., Costa et al., 2023).

3.6. Service contracts and warranties

As the field of maintenance optimization matured, companies discovered the potential of offering maintenance as a service. This shifted original equipment manufacturers (OEMs) towards a "servitization" strategy to provide after-sales maintenance during the lifetime of the equipment. In its extreme form, users do not acquire the equipment. Rather, they only buy the use of it. This phenomenon is known as 'power by the hour', where the OEM controls the uptime and maintenance associated with the equipment's usage. Such after-sales services generate stable revenues, enhance customer relations, and establish higher barriers to competition. They can be provided on-demand or through the implementation of service contracts. Service contracts and warranties provide some repair or maintenance element for a specified period. Where warranties are usually included in the purchase price, service contracts cost extra.

Contracts and warranties have been discussed in EJOR since the turn of the century, with an increase around 2006 (see Fig. 4). To date, 51 EJOR publications (around 10% of the maintenance publications in EJOR) are devoted to "`warranties`" (18), "`contracts`" (30), or both (3). Based on a review covering 44 journal publications in 2001–2011, Shafiee and Chukova (2013) identifies EJOR as the journal with the second-highest share of papers on warranty and maintenance.

Service contracts and warranties are characterized by an inverted business cycle. The revenues generated from the contracts are collected upfront, while the costs attached to the services are incurred during the contract. To ensure profitability, the key questions relate to (1) reducing the maintenance and servicing costs during the contract and (2) appropriately pricing the contract.

To minimize the expected total warranty cost for a pre-specified period, Yeh and Lo (2001) derive the optimal policy, defined by the number of preventive maintenance actions, corresponding maintenance degrees, and the maintenance schedule. Another example is given by Van Staden et al. (2022), who use a sequential 'predict, then optimize approach' to minimize the expected maintenance costs of a machine over its maintenance contract period, using readily available operational intervention data.

Different approaches have been adopted to optimize the price of a warranty or service contract. Jackson and Pascual (2008) develop a non-cooperative game model to determine the pricing structure in the contract and the number of customers to service that maximizes profits. Wang et al. (2020) studies the design and pricing of warranties with differentiated lengths and prices with a multinomial logit model to describe customer choice behaviors. They show that a cost-plus-margin pricing policy, with the same profit margins for all warranty options, optimizes the expected warranty profit. Huber and Spinler (2012, 2014) derive theoretical pricing insights using utility theory. Finally, Deprez et al. (2021) uses concepts from insurance pricing to predict the number of maintenance interventions and their cost to differentiate the price of full-service contracts.

3.7. Maintenance and sustainability

Sustainability is on the agenda of many European manufacturers to reduce greenhouse gas (GHG) emissions and combat climate change. The European Union (EU) aims to reduce the continent's net GHG emissions by at least 55% by 2030, compared to 1990 levels, and achieve net zero by 2050. Accordingly, the European Commission has adopted the EU Green Deal, a set of proposals to guide the EU's climate, energy, transport, and taxation policies towards their climate goals.

Considering climate goals, maintenance, by design, promotes sustainability. Maintenance intends to extend the lifetime of equipment and assets. Therefore, it alleviates the need to produce new equipment and reduces the emissions of their production. Also, periodically lubricating and servicing helps with energy conservation. Poorly lubricated machines require more energy due to the higher power usage. However, being sustainable also means that you expend fewer resources on replacements and repairs and release fewer environmental emissions. Maintenance requires materials and travel that harm the environment. Maintenance's Scope 3 emissions include the spare part production, their transportation, and the technician's travel to the maintenance site.

Research that explicitly includes sustainability objectives or GHG emissions is limited to date. We found only seven EJOR papers on maintenance with "`environmental`" in their title, abstract, or keywords, and no EJOR papers with "`circularity`" or "`recycling`". Wu et al. (2023) is one of the few papers that explicitly include GHG emissions in the optimization of maintenance policies. They optimize replacement policies acknowledging the GHG emissions produced during the initial manufacturing stage and the emissions associated with accumulated running hours during operations.

There are, however, several avenues to make maintenance more sustainable. Predictive maintenance using condition-monitored data aims to prevent unnecessary and *too early* preventive service interventions. Accurate prediction of what should be maintained and when can reduce the frequency of maintenance tasks without suffering any notable downturn in performance. Remotely monitoring the machines' operation via sensor technologies can also reduce the travel that comes with inspection or maintenance. Such a "digital control tower", in analogy to the airport control tower, provides (quasi) real-time visibility and can warn of upcoming failures before they happen (Boute & Van Mieghem, 2021). Through a *digital twin* of its physical operation, real-time analysis and optimization can prescribe decision-making where users decide based on what intelligent agents recommend. Moreover, when the same party monitors multiple factories at nearby sites, maintenance interventions can be combined in the same travel.

Promoting circular material flow and waste reduction can be achieved via changes in design, consumption (such as reuse, refurbishment or remanufacturing) or return (such as recycle and recover) of assets. While Section 3.3 addressed design for maintenance, a final directive for sustainable maintenance is recycling or remanufacturing of the service parts. Several authors and authorities, including the UN, believe offering product-service systems will launch considerable savings in material and energy consumption (Colen & Lambrecht, 2013). When OEMs assume control of after-sales activities, they become responsible for waste disposal, component replacement, and energy use. With the right contractual incentives, OEMs will incorporate after-sales resource use in their decision-making, effectively reducing the environmental impact during the entire equipment life cycle. Developing product-service systems will help companies comply with and surpass increasing environmental obligations. In practice, we observe the launch of energy-saving services, refurbishment and recycling activities, and efforts to increase equipment reliability (Colen & Lambrecht, 2013).

4. Closing the gap between research and practice

Although maintenance research has been well-studied in the past fifty years, many open questions remain to close the gaps between the current state of research and real-world needs. Companies have an abundance of data at their disposal. In the last few years, OEMs have increasingly been placing sensors on their machinery, allowing for operational data collection. Such data offers a potential competitive advantage in the form of, for example, the calibration of a condition-based maintenance policy. Data analytics and machine learning are also evolving fast to leverage these data for online learning. As a result, both the timing and the content of maintenance should no longer be restricted to a predefined plan and can become more flexible and adaptive to the monitored condition. This approach will increase useful lifetime and reduce costs.

However, those who have worked with actual maintenance data know this comes with ample challenges. First, the frustration of cleaning and preparing the data. Operational data are hardly ‘clean:’ they may be intermittent, noisy, or have outliers. Moreover, condition data is typically multi-dimensional. One machine has a plethora of sensors, generating multiple time series related to various machine conditions. Standard machine learning techniques may not find meaningful patterns. To leverage these techniques, it will be mandatory to transform the multi-dimensional sensor data into low-dimensional data, such as a single health indicator. Mapping or clustering these multi-dimensional data into useful indicators may require domain knowledge beyond operational research or data science techniques. As we have advocated earlier, we believe integrating engineering into operational research could handle that.

The main challenge with maintenance data, however, remains in the lack of sufficient ‘useful’ data that can be used to effectively predict and prevent failure. Machines are designed to last long and preventive maintenance intends to avoid failures. Therefore, the abundance of condition data contrasts with the scarcity of effective failure data, needed to learn when a failure will happen. Data from a few machines may not suffice to learn this failure behavior. The lack of sufficient data to accurately estimate failures is known as the ‘small data’ problem. One approach to cope with the small data problem includes *smart* data pooling across components or machines (e.g., Deprez et al., 2023; Dursun et al., 2022; Van Staden et al., 2022). This corresponds to Level 4 in our maturity framework in Fig. 2.

The collection of data from multiple machines and customers is facilitated through service contracts. These can be offered by the OEM. In those cases, data ownership should be carefully considered. The customer generates all the data and thus, in a way, owns them, but the OEM needs them to optimize their contracts. As a result, the after-sales business model is shifted from the customers’ side to the OEM that offers these service contracts. In doing so, OEMs largely monopolize the after-sales market and the operational data generated. This shift has adverse effects on the competitiveness of independent after-sales service providers, who do not have access to this type of conditional data.

An alternative possible solution to overcome the lack of relevant data is federated learning (Li et al., 2020). In federated learning, each entity only shares its local model parameters, instead of its entire dataset, with the developed machine-learning model. This approach can consolidate parameters from the datasets of various independent service providers, eliminating the economies of scale effect on the data frontier as enjoyed by the larger OEMs.

We may draw inspiration from insurance companies with more comprehensive experience — and, therefore, maybe more maturity — in data collection than manufacturing. After all, a maintenance contract can be likened to an insurance policy covering the maintenance costs during a specific period. Insurance companies also offer products that are under-discovered in maintenance, such as a bonus-malus system that adjusts the price paid by a customer according to their individual claim history. Its application to maintenance contracts and related maintenance policies is an untapped research area.

We are optimistic and excited about the next fifty years of maintenance research. The multidisciplinary nature and the numerous opportunities to leverage data and algorithms into smart decision-making are expected to make impactful contributions to this important field and the broad maintenance community.

CRediT authorship contribution statement

Joachim Arts: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Robert N. Boute:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Stijn Loeys:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Heletjé E. van Staden:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Appendix A. Basic text-mining to browse through the EJOR archive

We started with a list of DOI numbers of all EJOR publications from 1977–2023 received from the publisher (19,995 records). This also included information on the title, authors, year of publication, and number of Scopus citations (per 7 March 2023). We retrieved the abstracts and author-provided keywords from Scopus using the DOI number via the Scopus API.[4] The title, abstract, and keywords were then concatenated in a separate column with basic text cleaning, i.e., lower-casing all words and removing punctuation marks. Subsequently, all words were stemmed using the NLTK-stemming library. For example, the stem of “`maintenances`” is “`mainten`”.

To extract the relevant papers on maintenance, we searched for the entries with the stemmed version of the word “`maintenance`” (`mainten`). This resulted in 506 publications. We considered including reliability-related publications with the stemmed version of “`reliable`” (`reliabl`), but this resulted in papers relying on, for instance, a *reliable* method.

Appendix B. Fifteen most cited EJOR papers on maintenance

See Table 6.

Appendix C. Transition matrices for Example 6

The collection of transition probabilities, P, for Example 6 is calculated using Eq. (4) and the data provided in the example. The resulting transition probability p_{ij}^{a} can be organized into two matrices as given in Box I

[4] https://dev.elsevier.com/

Table 6
The fifteen most cited EJOR publications on maintenance.

Rank	Title	Authors	Year	Citations
1	Remaining useful life estimation - A review on the statistical data-driven approaches	Si, X.-S., Wang, W., Hu, C.-H., Zhou, D.-H.	2011	1449
2	A survey of maintenance policies of deteriorating systems	Wang, H.	2002	1353
3	Imperfect maintenance	Pham, H., Wang, H.	1996	826
4	A survey of maintenance models for multi-unit systems	Cho, D.I., Parlar, M.	1991	558
5	Problem structuring methods in action	Mingers, J., Rosenhead, J.	2004	444
6	Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods	Zhang, Z., Si, X., Hu, C., Lei, Y.	2018	301
7	On the application of mathematical models in maintenance	Scarf, P.A.	1997	243
8	Maintenance of continuously monitored degrading systems	Liao, H., Elsayed, E.A., Chan, L.-Y.	2006	243
9	Sequential condition-based maintenance scheduling for a deteriorating system	Dieulle, L., Bérenguer, C., Grall, A., Roussignol, M.	2003	230
10	Condition-based maintenance policies for systems with multiple dependent components: A review	Olde Keizer, M.C.A., Flapper, S.D.P., Teunter, R.H.	2017	220
11	Best practice analysis of bank branches: An application of DEA in a large Canadian bank	Schaffnit, C., Rosen, D., Paradi, J.C.	1997	211
12	A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution	Si, X.-S., Wang, W., Chen, M.-Y., Hu, C.-H., Zhou, D.-H.	2013	199
13	An integrated production and preventive maintenance planning model	Aghezzaf, E.H., Jamali, M.A., Ait-Kadi, D.	2007	184
14	Maintenance management decision making	Pintelon, L.M., Gelders, L.F.	1992	181
15	Maintenance models in warranty: A literature review	Shafiee, M., Chukova, S.	2013	177

$$
P^0 = \begin{pmatrix} p^0_{00} & p^0_{01} & \cdots & p^0_{0L} \\ p^0_{10} & p^0_{11} & \cdots & p^0_{1L} \\ \vdots & \vdots & \ddots & p^0_{2L} \\ p^0_{L0} & p^0_{L1} & \cdots & p^0_{LL} \end{pmatrix}
$$
$$
= \begin{pmatrix}
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0356 & 0.0040 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2283 & 0.0396 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7321 & 0.2679 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000
\end{pmatrix}
$$

and

$$
P^1 = \begin{pmatrix} p^1_{00} & p^1_{01} & \cdots & p^1_{0L} \\ p^1_{10} & p^1_{11} & \cdots & p^1_{1L} \\ \vdots & \vdots & \ddots & \vdots \\ p^1_{L0} & p^1_{L1} & \cdots & p^1_{LL} \end{pmatrix}
$$
$$
= \begin{pmatrix}
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.7321 & 0.2283 & 0.0356 & 0.0037 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000
\end{pmatrix}.
$$

Box I.

References

Andersen, J. F., Andersen, A. R., Kulahci, M., & Nielsen, B. F. (2022). A numerical study of Markov decision process algorithms for multi-component replacement problems. *European Journal of Operational Research*, *299*(3), 898–909.

Arts, J. (2019). *Lecture notes: Maintenance modeling and optimization*: *Technical report*, Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg.

Arts, J., & Basten, B. (2018). Design of multi-component periodic maintenance programs with single-component models. *IISE Transactions*, *50*(7), 606–615.

Barlow, R. E., & Hunter, L. C. (1960). Optimum preventive maintenance policies. *Operations Research*, *8*, 90–100.

Barlow, R. E., & Proschan, F. (1996). *Mathematical theory of reliability*. SIAM.

Bautista, L., Castro, I. T., & Landesa, L. (2022). Condition-based maintenance for a system subject to multiple degradation processes with stochastic arrival intensity. *European Journal of Operational Research*, *302*(2), 560–574.

Bei, X., Zhu, X., & Coit, D. W. (2019). A risk-averse stochastic program for integrated system design and preventive maintenance planning. *European Journal of Operational Research*, *276*(2), 536–548.

Bertrand, J. W. M., Fransoo, J. C., & Udenio, M. (2023). Model-based research. In C. Karlsson (Ed.), *Research methods for operations and supply chain management* (3rd ed.). London: Routledge.

Boute, R. N., & Van Mieghem, J. A. (2021). Digital operations: Autonomous automation and the smart execution of work. *Management and Business Review*, *1*(1).

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

Colen, P. J., & Lambrecht, M. (2013). Product service systems: exploring operational practices. *The Service Industries Journal*, *33*, 501–515.

Da Costa, P., Verleijsdonk, P., Voorberg, S., Akcay, A., Kapodistria, S., Van Jaarsveld, W., & Zhang, Y. (2023). Policies for the dynamic traveling maintainer problem with alerts. *European Journal of Operational Research*, *305*, 1141–1152.

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

De Pater, I., & Mitici, M. (2023). Developing health indicators and RUL prognostics for systems with few failure instances and varying operating conditions using a LSTM autoencoder. *Engineering Applications of Artificial Intelligence*, *117*, Article 105582.

Deep, A., Zhou, S., Veeramani, D., & Chen, Y. (2023). Partially observable Markov decision process-based optimal maintenance planning with time-dependent observations. *European Journal of Operational Research*, *311*(2), 533–544.

Demirci, E. Z., Arts, J., & Houtum, G.-J. v. (2024). A restless bandit approach for capacitated condition based maintenance scheduling. *Flexible Services and Manufacturing Journal*, 1–29.

Deprez, L., Antonio, K., Arts, J., & Boute, R. N. (2023). Data-driven preventive maintenance for a heterogeneous machine portfolio. *Operations Research Letters*, *51*(2), 163–170.

Deprez, L., Antonio, K., & Boute, R. (2021). Pricing service maintenance contracts using predictive analytics. *European Journal of Operational Research*, *290*(2), 530–545.

Drent, C., Drent, M., & Arts, J. (2024). Condition-based production for stochastically deteriorating systems: optimal policies and learning. *Manufacturing & Service Operations Management*, *articles in advance*.

Drent, C., Drent, M., Arts, J., & Kapodistria, S. (2023). Real-time integrated learning and decision making for cumulative shock degradation. *Manufacturing and Service Operations Management*, *25*(1), 235–253.

Drent, C., Kapodistria, S., & Boxma, O. (2020). Censored lifetime learning: Optimal Bayesian age-replacement policies. *Operations Research Letters*, *48*(6), 827–834.

Drent, C., Kapodistria, S., & Resing, J. (2019). Condition-based maintenance policies under imperfect maintenance at scheduled and unscheduled opportunities. *Queueing Systems*, *93*, 269–308.

Drent, C., Keizer, M. O., & Van Houtum, G.-J. (2020). Dynamic dispatching and repositioning policies for fast-response service networks. *European Journal of Operational Research*, *285*(2), 583–598.

Driessen, J., de Kruijff, J., Arts, J., & Van Houtum, G.-J. (2024). Optimal design of line replaceable units. *Naval Research Logistics*, *71*, 189–209.

Dursun, I., Akcay, A., & Van Houtum, G.-J. (2022). Age-based maintenance under population heterogeneity: Optimal exploration and exploitation. *European Journal of Operational Research*, *301*, 1007–1020.

Ebeling, C. (2001). *Introduction to reliability and maintainablity engineering* (2nd ed.). McGraw-Hill.

Eckles, J. E. (1968). Optimum maintenance with incomplete information. *Operations Research*, *16*(5), 1058–1067.

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

Feo, T. A., & Bard, J. F. (1989). Flight scheduling and maintenance base planning. *Management Science*, *35*(12), 1415–1432.

Gamiz, M. L., Limnios, N., & Del Carman Segovia-Garcia, M. (2023). Hidden markov models in reliability and maintenance. *European Journal of Operational Research*, *304*, 1242–1255.

Gara-Ali, A., Finke, G., & Espinouse, M.-L. (2016). Parallel-machine scheduling with maintenance: Praising the assignment problem. *European Journal of Operational Research*, *252*(1), 90–97.

Geurtsen, M., Didden, J. B., Adan, J., Atan, Z., & Adan, I. (2023). Production, maintenance and resource scheduling: A review. *European Journal of Operational Research*, *305*(2), 501–529.

Glazebrook, K. D., Mitchell, H., & Ansell, P. (2005). Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research*, *165*(1), 267–284.

Gopalan, R. (2014). The aircraft maintenance base location problem. *European Journal of Operational Research*, *236*(2), 634–642.

Haque, L., & Armstrong, M. J. (2007). A survey of the machine interference problem. *European Journal of Operational Research*, *179*(2), 469–482.

Hu, Q., Boylan, J. E., Chen, H., & Labib, A. (2018). OR in spare parts management: A review. *European Journal of Operational Research*, *266*(2), 395–414.

Huang, Y., & Guo, X. (2011). Finite horizon semi-Markov decision processes with application to maintenance systems. *European Journal of Operational Research*, *212*(1), 131–140.

Huber, S., & Spinler, S. (2012). Pricing of full-service repair contracts. *European Journal of Operational Research*, *222*(1), 113–121.

Huber, S., & Spinler, S. (2014). Pricing of full-service repair contracts with learning, optimized maintenance, and information asymmetry. *Decision Sciences*, *45*(4), 791–815.

Jackson, C., & Pascual, R. (2008). Optimal maintenance service contract negotiation with aging equipment. *European Journal of Operational Research*, *189*(2), 387–398.

Jardine, A. K., & Tsang, A. H. (2005). *Maintenance, replacement, and reliability: theory and applications*. CRC Press.

Kabir, A. Z., & Al-Olayan, A. S. (1996). A stocking policy for spare part provisioning under age based preventive replacement. *European Journal of Operational Research*, *90*(1), 171–181.

Keizer, M. C. O., Teunter, R. H., & Veldman, J. (2016). Clustering condition-based maintenance for systems with redundancy and economic dependencies. *European Journal of Operational Research*, *251*(2), 531–540.

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

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

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

Lagos, C., Delgado, F., & Klapp, M. A. (2020). Dynamic optimization for airline maintenance operations. *Transportation Science*, *54*(4), 998–1015.

Lalla-Ruiz, E., & Voß, S. (2016). Modeling the parallel machine scheduling problem with step deteriorating jobs. *European Journal of Operational Research*, *255*(1), 21–33.

Lambert, A. J. D. (2007). Optimizing disassembly processes subjected to sequence–dependent cost. *Computers & Operations Research*, *34*(2), 536–551.

Larrnaaga, M., Ayesta, U., & Verloop, I. M. (2016). Dynamic control of birth-and-death restless bandits: application to resource-allocation problems. *IEEE/ACM Transactions on Networking*, *24*(6), 3812–3825.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mechanical Systems and Signal Processing*, *104*, 799–834.

Lewis, E. (1996). *Introduction to reliability engineering*. John Wiley & Sons.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50–60.

Liu, C., & Gryllias, K. (2020). Unsupervised domain adaptation based remaining useful life prediction of rolling element bearings. In *PHM society European conference, vol. 5, no. 1* (pp. 10–10).

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

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

Muckstadt, J. A. (2004). *Analysis and algorithms for service parts supply chains*. Springer Science & Business Media.

Olde Keizer, M. C. A., Flapper, S. D. P., & Teunter, R. H. (2017). Condition-based maintenance policies for systems with multiple dependent components: A review. *European Journal of Operational Research*, *261*, 405–420.

Parada Puig, J. E., & Basten, R. J. I. (2015). Defining line replaceable units. *European Journal of Operational Research*, *247*(1), 310–320.

Peng, H., & Zhu, Q. (2017). Approximate evaluation of average downtime under an integrated approach of opportunistic maintenance for multi-component systems. *Computers & Industrial Engineering*, *109*, 335–346.

Pham, D., & Kiesmüller, G. (2024). *Transportation Science*, *58*(2), 499–519.

Poppe, J., Boute, R. N., & Lambrecht, M. R. (2018). A hybrid condition-based maintenance policy for continuously monitored components with two degradation thresholds. *European Journal of Operational Research*, *268*, 515–532.

Puterman, M. L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. NY: Wiley.

Romeijnders, W., Teunter, R., & Van Jaarsveld, W. (2012). A two-step method for forecasting spare parts demand using information on component repairs. *European Journal of Operational Research*, *220*(2), 386–393.

Ross, S. M. (1996). *Stochastic processes* (2nd ed.). John Wiley & Sons.

Sanoubar, S., He, K., Maillart, L. M., & Prokopyev, O. A. (2021). Optimal age-replacement in anticipation of time-dependent, unpunctual policy implementation. *IEEE Transactions on Reliability*, *70*(3), 1177–1192.

Sanoubar, S., de Jonge, B., Maillart, L. M., & Prokopyev, O. A. (2023). Optimal condition-based maintenance via a mobile maintenance resource. *Transportation Science*, *57*(6), 1646–1670.

Shafiee, M., & Chukova, S. (2013). Maintenance models in warranty: A literature review. *European Journal of Operational Research*, *229*(3), 561–572.

Sherbrooke, C. C. (1968). METRIC: A multi-echelon technique for recoverable item control. *Operations Research*, *16*(1), 122–141.

Sherbrooke, C. C. (2006). *Optimal inventory modeling of systems: Multi-echelon techniques*: *vol. 72*, Springer Science & Business Media.

Smets, L. P., Van Houtum, G.-J., & Langerak, F. (2012). Design for availability: A holistic approach to create value for manufacturers and customers of capital goods. *Journal of Systems Science and Systems Engineering*, *21*, 403–421.

Song, J.-S., Van Houtum, G.-J., & Van Mieghem, J. A. (2019). Capacity and inventory management: Review, trends, and projections. *Manufacturing & Service Operations Management*, *22*(1), 36–46.

Stein, W. (2010). Spare parts planning at ASML. *eSCF Operations Practices*.

Stoneham, D. (1998). *Maintenance management and technology handbook*. Elsevier Advanced Technology.

Thomas, L., & Osaki, S. (1978). An optimal ordering policy for a spare unit with lead time. *European Journal of Operational Research*, *2*(6), 409–419.

Tinga, T. (2013). *Principles of loads and failure mechanisms: Applications in maintenance, reliability and design*. Springer.

Tönissen, D., & Arts, J. (2020). The stochastic maintenance location routing allocation problem for rolling stock. *International Journal of Production Economics*, *230*, Article 107826.

Tönissen, D. D., Arts, J. J., & Shen, Z.-J. (2019). Maintenance location routing for rolling stock under line and fleet planning uncertainty. *Transportation Science*, *53*(5), 1252–1270.

Topal, E., & Ramazan, S. (2010). A new MIP model for mine equipment scheduling by minimizing maintenance cost. *European Journal of Operational Research*, *207*(2), 1065–1071.

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

Urbani, M., Brunelli, M., & Punkka, A. (2023). An approach for bi-objective maintenance scheduling on a networked system with limited resources. *European Journal of Operational Research*, *305*(1), 101–113.

Van der Auweraer, S., Boute, R. N., & Syntetos, A. A. (2019). Forecasting spare part demand with installed base information: A review. *International Journal of Forecasting*, *35*(1), 181–196.

Van Deursen, R. W. H. (2020). *A tool to define line replaceable units to improve maintenance* (Master's thesis), Eindhoven University of Technology, https://pure.tue.nl/ws/portalfiles/portal/152640976/Master_Thesis_Rik_van_Deursen.pdf.

Van Geel, G. C. J. (2018). *Improving the LRU definition process* (Master's thesis), Eindhoven University of Technology, https://pure.tue.nl/ws/portalfiles/portal/107222615/Master_Thesis_Glenn_van_Geel.pdf.

Van Houtum, G.-J., & Kranenburg, B. (2015). *Spare parts inventory control under system availability constraints*: *vol. 227*, Springer.

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., Maillart, L. M., & Kharoufeh, J. P. (2017). Optimal maintenance policies for a safety-critical system and its deteriorating sensor. *Naval Research Logistics*, *64*, 399–417.

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

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

Van Staden, H. E., Deprez, L., & Boute, R. N. (2022). A dynamic "predict, then optimize" preventive maintenance approach using operational intervention data. *European Journal of Operational Research*, *302*(3), 1079–1096.

Wang, W. (2012). A stochastic model for joint spare parts inventory and planned maintenance optimisation. *European Journal of Operational Research*, *216*(1), 127–139.

Wang, T., Baldacci, R., Lim, A., & Hu, Q. (2018). A branch-and-price algorithm for scheduling of deteriorating jobs and flexible periodic maintenance on a single machine. *European Journal of Operational Research*, *271*(3), 826–838.

Wang, X., Zhao, X., & Liu, B. (2020). Design and pricing of extended warranty menus based on the multinomial logit choice model. *European Journal of Operational Research*, *287*(1), 237–250.

Wu, S., Wu, D., & Peng, R. (2023). Considering greenhouse gas emissions in maintenance optimisation. *European Journal of Operational Research*, *307*, 1135–1145.

Xia, T., Xi, L., Zhou, X., & Lee, J. (2012). Dynamic maintenance decision-making for series–parallel manufacturing system based on MAM–MTW methodology. *European Journal of Operational Research*, *221*(1), 231–240.

Ye, Z.-S., & Chen, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics*, *56*(3), 302–311.

Yeh, R. H., & Lo, H.-C. (2001). Optimal preventive-maintenance warranty policy for repairable products. *European Journal of Operational Research*, *134*(1), 59–69.

Zhang, N., Fouladirad, M., Barros, A., & Zhang, J. (2020). Condition-based maintenance for a K-out-of-N deteriorating system under periodic inspection with failure dependence. *European Journal of Operational Research*, *302*(2), 560–574.

Zhang, H., & Zhang, W. (2023). Analytical solution to a partially observable machine maintenance problem with obvious failures. *Management Science*, *69*(7), 3993–4015.

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

Zhu, Q., Peng, H., & Van Houtum, G.-J. (2015). A condition-based maintenance policy for multi-component systems with a high maintenance setup cost. *OR Spectrum*, *109*, 1007–1035.