UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Data-driven preventive maintenance for a heterogeneous machine portfolio

Deprez, L.; Antonio, K.; Arts, J.; Boute, R.

DOI
10.1016/j.orl.2023.01.006

Publication date
2023

Document Version
Final published version

Published in
Operations Research Letters

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

Citation for published version (APA):
Deprez, L., Antonio, K., Arts, J., & Boute, R. (2023). Data-driven preventive maintenance for a heterogeneous machine portfolio. *Operations Research Letters*, *51*(2), 163-170. https://doi.org/10.1016/j.orl.2023.01.006

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Nov 2026

Operations Research Letters 51 (2023) 163–170

Contents lists available at ScienceDirect

Operations Research Letters

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

Data-driven preventive maintenance for a heterogeneous machine portfolio

Laurens Deprez [a,*], Katrien Antonio [b,c], Joachim Arts [a], Robert Boute [b,d,e]

[a] Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg, Luxembourg
[b] Faculty of Economics and Business, KU Leuven, Belgium
[c] Faculty of Economics and Business, University of Amsterdam, the Netherlands
[d] Technology & Operations Management Area, Vlerick Business School, Belgium
[e] VCCM, Flanders Make, Belgium

ARTICLE INFO

Article history:
Received 24 August 2021
Received in revised form 29 November 2022
Accepted 4 January 2023
Available online 10 January 2023

Keywords:
Preventive maintenance
Data pooling
Proportional hazards
Small data

ABSTRACT

We describe a data-driven approach to optimize periodic maintenance policies for a heterogeneous portfolio with different machine profiles. When insufficient data are available per profile to assess failure intensities and costs accurately, we pool the data of all machine profiles and evaluate the effect of (observable) machine characteristics by calibrating appropriate statistical models. This reduces maintenance costs compared to a stratified approach that splits the data into subsets per profile and a uniform approach that treats all profiles the same.

© 2023 Elsevier B.V. All rights reserved.

1. Introduction

Despite advances in condition monitoring, time- or age-based periodic preventive maintenance is still common practice in many companies. Under this policy, preventive maintenance (PM) interventions are scheduled at a periodic interval, either measured in calendar time or running hours. This interval is determined by trading off the failure rate and corresponding failure costs against the cost of performing preventive maintenance. Most literature assumes that the failure behaviour, or failure intensity, is known, in line with [5]. In practice, however, the failure distribution may not be known and should be estimated based on historical data. This may introduce some challenges. A small machine population may not provide sufficient maintenance and failure data for an accurate estimation. The lack of sufficient data to accurately estimate regression parameters is known as the 'small data' problem. Even when the machine population is of adequate size, the machine population might be heterogeneous and as such may contain multiple different machine profiles. This heterogeneity can be caused by different running conditions, environmental factors or different manufacturing plants where the machines were assembled and might induce different failure intensities and costs. Estimation by stratifying, or splitting the data, per machine profile might again induce the small data problem. Aggregating these machine data without taking into account the heterogeneity in the machine population leads to a maintenance policy that is optimal for an average machine, but it is not tailored to a particular machine (profile). The novelty of our approach is that we introduce a multivariate time-to-failure and cost model such that the data of all machines (across the different machine profiles) can be pooled. This allows differentiating the PM policy over the different machine profiles, while at same time obtains accurate time-to-failure and maintenance cost estimates for a heterogeneous machine population. The latter is not possible when stratifying data per machine profile.

Service providers, who maintain the assets of their customers, can make use of such an approach. Their maintenance portfolio provides them potentially with ample historical maintenance and failure data, yet data per machine profile might be limited. Our objective is to optimize the periodic maintenance interval by explicitly considering the heterogeneity in the machine portfolio that is induced by observable machine characteristics. Our policy optimization applies to a finite horizon, motivated by the finite duration of service maintenance contracts or (extended) warranties [22,15]. We focus on observable machine characteristics that remain constant during the contract horizon, e.g. running conditions, country, operating environment and industry type. This categorizes each machine in a specific machine profile, specified by these machine characteristics. By tailoring the maintenance policy to the

* Corresponding author. 6 Rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Luxembourg.
E-mail address: laurens.deprez@kuleuven.be (L. Deprez).

https://doi.org/10.1016/j.orl.2023.01.006
0167-6377/© 2023 Elsevier B.V. All rights reserved.

machine profile, the resulting cost optimization may lead to higher profit margins or lower contract prices.

Population heterogeneity has been studied by Dursun et al. [15], Abdul-Malak et al. [1], and de Jonge et al. [10], among others. In [15], [1], and [10], parts (or machines) originate from multiple, mostly two, different sub-populations with different failure distributions. It is unknown to which sub-population a spare part belongs but the failure distribution of each sub-population itself is known. In their analysis, the sub-population of the part is inferred by observing its failure behaviour and adapting the maintenance policy accordingly. In contrast, we consider the case where it is known to which sub-population machines belong as machines are labelled by their machine profile, characterized by observable machine characteristics. The underlying failure distribution, however, is not known and we infer the failure behaviour and maintenance costs for each machine profile from data. Drent et al. [14] also deals with an unknown failure distribution but assumes population homogeneity. In their paper the failure distribution and maintenance policy are inferred by means of a Bayesian approach as information accumulates during the machine operation. The information accumulation process consists of both censored (i.e., preventive replacements) and uncensored (i.e., corrective replacements) observations of the underlying lifetime distribution. This leads to an inherent exploration-exploitation trade-off. Exploration (i.e., a longer periodic maintenance interval) increases the probability of corrective replacement, but at the same time leads to accumulating more, valuable, information. In our approach there is no exploration-exploitation trade-off as the data over all machine profiles has been collected prior to the moment of estimation.

Our methodology relies on a data set with historical failure and maintenance data of a heterogeneous machine portfolio. We *learn* the failure behaviour and costs and assess the effect of (observable) machine characteristics by calibrating appropriate statistical models. The calibrated statistical models enable the optimization of each machine profile's periodic maintenance interval. We study how the resulting maintenance policies are more cost-effective than (1) a *uniform* approach that disregards the machine profiles, and (2) a *stratified* approach that splits the data per machine profile to take into account each machine profile.

The next section introduces our reliability model, the optimality condition for the periodic maintenance interval and the predictive models to calibrate the failure behaviour and costs. Section 3 numerically evaluates how and when our data pooling approach reduces maintenance costs. Section 4 concludes.

2. Maintenance policies for a heterogeneous machine population

We first set out the details of our reliability model. We then specify the optimality condition for a periodic, preventive maintenance policy over a finite horizon for a heterogeneous machine population. Finally we establish the statistical models and their calibration to data to estimate the failure rate, and the costs of failure and preventive maintenance.

2.1. Reliability model

We consider the optimization of the number of preventive maintenance interventions, n, for a machine during a finite horizon $[0, \Delta t]$, e.g. the coverage period of a service agreement. In our analysis we will assume that failures require minimal, as-good-as-old, corrective maintenance (see e.g. [4,19,27,13,2]) and planned, preventive maintenance actions are perfect. The latter justifies a periodic maintenance policy, where preventive maintenance is executed with a fixed interval $\Delta t_{pm} = \frac{\Delta t}{n+1}$, at times $t_k = k\Delta t_{pm}$ $(k = 1, \ldots, n)$. Each machine is characterized by a set of machine characteristics or covariates $\boldsymbol{x}$, e.g. operating conditions, industry type, or country, that remain the same during the planning horizon.

Failure times We specify the machine-dependent failure intensity function $\lambda(t)$ in each maintenance interval under the Cox proportional hazards assumption [8] with the same baseline failure intensity function $\lambda_0(t)$ for each machine and $\boldsymbol{\beta}$ representing the impact of the machine characteristics $\boldsymbol{x}$. The proportional hazards model is a versatile model to analyze the effect of operating conditions (or covariates) on the lifetime of a system. Kumar and Westberg [17] have shown that models from the proportional hazards family appear to be the better ones for analyzing the effect of the covariates. It is also very convenient to add terms or cross products. The practical value of the proportional hazards model has been demonstrated, for instance, by Barabadi et al. [3] in a case study in mining equipment to identify the covariates that influence the reliability. We acknowledge that, like any multi-variate model, there is a risk of mis-specification in proportional hazard models. We show in Section 3.5 that the benefits of data pooling outweigh this risk.

The machine-dependent failure intensity function $\lambda(t)$ defines a non-homogeneous Poisson process [21] for the arrival of failures. With t the time since the last PM intervention, the machine-specific failure intensity function $\lambda(t)$ is then characterized by

$$\lambda(t) = \lambda_0(t) \exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) \text{ for } t \in [0, t_k - t_{k-1}) \ (\forall k), \tag{1}$$

where $t_0 = 0$ and $t_{n+1} = \Delta t$, respectively referring to the start and end of the planning horizon. Assuming perfect preventive maintenance, the failure intensity function $\lambda(t)$ is the same for each inter-PM interval.

The baseline failure intensity $\lambda_0(t)$ can take any parametric form. We define $\lambda_0(t)$ by a Weibull failure intensity function, which is regularly used in the reliability literature to model failure times [6,25,26], with scale parameter $\alpha \in \mathbb{R}_0^+$ and shape $\gamma \in \mathbb{R}^+$,

$$\lambda_0(t) = \gamma \alpha^\gamma t^{\gamma-1}.$$

Costs of maintenance The expected preventive maintenance cost, $c_p(\boldsymbol{x})$, and expected failure cost, $c_f(\boldsymbol{x})$ are modelled with a gamma generalized linear model (GLM) to account for the impact of the machine characteristics $\boldsymbol{x}$. [11] argues that gamma models are appropriate to model cost data given their positive support and right-skewed distribution. We denote the scale of the gamma distribution by θ_f for the failure costs. For the expected failure costs, the GLM with categorical explanatory variables $\boldsymbol{x}$ is specified by (see [23] for an overview)

$$c_f(\boldsymbol{x}) = \exp\left(\boldsymbol{\beta}_f' \cdot (1, \boldsymbol{x})\right) = \exp\left(\beta_{f,0} + \sum_{j=1}^{q} \beta_{f,j} x_j\right), \tag{2}$$

with $\boldsymbol{\beta}_f' \cdot (1, \boldsymbol{x})$ the linear predictor. The impact of the machine characteristics $\boldsymbol{x}$ on the cost of failure is captured by $\boldsymbol{\beta}_f$. We remark that the first component of the vector $\boldsymbol{\beta}_f$ acts as an intercept and consequently the length of the vector $\boldsymbol{\beta}_f$ is one plus the length of vector $\boldsymbol{x}$, i.e. $1 + q$. The expected preventive maintenance costs $c_p(\boldsymbol{x})$ are similarly specified. Their respective parameters are denoted with subscript p. We remark that we set the parameters for the cost of preventive maintenance $c_p(\boldsymbol{x})$ and the cost of failure $c_f(\boldsymbol{x})$ such that $c_p(\boldsymbol{x}) < c_f(\boldsymbol{x})$ for each combination of machine characteristics $\boldsymbol{x}$.

Total maintenance costs The expected total maintenance cost over the horizon $[0, \Delta t]$ for a policy with preventive maintenance at

(consecutive) times $\{t_1, t_2, ..., t_n\}$ ($\subset [0, \Delta t]$), on a machine with characteristics $\boldsymbol{x}$ is then defined by,

$$\begin{aligned} C_{\boldsymbol{x}}(\{t_1, t_2, ..., t_n\}) &= c_f(\boldsymbol{x}) \sum_{k=0}^{n} \int_0^{t_{k+1}-t_k} \lambda(t)dt + nc_p(\boldsymbol{x}) \\ &= c_f(\boldsymbol{x}) \sum_{k=0}^{n} \int_0^{t_{k+1}-t_k} \lambda_0(t)\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})dt + nc_p(\boldsymbol{x}) \\ &= c_f(\boldsymbol{x}) \exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) \sum_{k=0}^{n} \Lambda_0(t_{k+1} - t_k) + nc_p(\boldsymbol{x}), \end{aligned} \tag{3}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)du$. The integral $\int_0^{t_{k+1}-t_k} \lambda(u)du$ represents the expected number of failures between t_k and t_{k+1}. The latter is a property of non-homogeneous Poisson processes [18]. By summing over all preventive maintenance intervals we obtain all failures during the horizon $[0, \Delta t]$. Equation (3) can be simplified for periodic maintenance with fixed interval $\Delta t_{pm} = \frac{\Delta t}{n+1}$. In this case, the expected total maintenance costs $C_{\boldsymbol{x}}(n)$ is a function of the number of preventive maintenance interventions, n,

$$C_{\boldsymbol{x}}(n) = c_f(\boldsymbol{x}) \exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})(n+1)\Lambda_0(\Delta t_{pm}) + nc_p(\boldsymbol{x}). \tag{4}$$

Observe that the impact of the machine characteristics $\boldsymbol{x}$ on the failure rate, i.e. $\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})$ is equivalent to an increase in the expected failure costs with the same factor, $\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})$.

2.2. Optimality condition for a differentiated, periodic policy

We denote $n^\star$ the optimal number of PMs that minimizes the total maintenance costs in Eq. (4) during the planning horizon $[0, \Delta t]$. Although $n^\star$ depends on the machine characteristics $\boldsymbol{x}$, we will adopt $n^\star$ instead of $n^\star(\boldsymbol{x})$ to simplify notation.

Proposition 1. *If the baseline failure intensity $\lambda_0(t)$ is strictly increasing in t ($\in \mathbb{R}^+$), then the optimal number of PMs, $n^\star$, is the smallest $n \in \mathbb{N}_0$ that satisfies,*

$$(n+1)\Lambda_0\left(\frac{\Delta t}{n+1}\right) - (n+2)\Lambda_0\left(\frac{\Delta t}{n+2}\right) \leq \frac{c_p(\boldsymbol{x})}{c_f(\boldsymbol{x})\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})}. \tag{5}$$

Proof. The second order derivative of the total maintenance cost (for $n \in \mathbb{R}^+$) is

$$\frac{\mathrm{d}^2 C_{\boldsymbol{x}}(n)}{\mathrm{d}n^2} = c_f(\boldsymbol{x})\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})\frac{\Delta t^2}{(n+1)^3}\lambda_0'\left(\frac{\Delta t}{n+1}\right),$$

with $\lambda_0'(t)$ the first order derivative of $\lambda_0(t)$. If $\lambda_0(t)$ is strictly increasing, i.e. $\lambda_0'(t) > 0$ ($\forall t \in \mathbb{R}^+$), then $C_{\boldsymbol{x}}(n)$ is convex on $\mathbb{R}^+$. The optimal n^* satisfies the first order optimality condition, i.e. it is the smallest n ($\in \mathbb{N}_0$) for which $\Delta_1[C_{\boldsymbol{x}}](n) \geq 0$ where the forward difference of the cost function $\Delta_1[C_{\boldsymbol{x}}](n) = C_{\boldsymbol{x}}(n+1) - C_{\boldsymbol{x}}(n)$. □

2.3. Calibration of predictive models to estimate failure intensity and costs

The optimal periodic maintenance policy depends on the failure intensity and the costs of failure and maintenance of each machine profile (as defined by the characteristics $\boldsymbol{x}$). We now describe the statistical models and their calibration to estimate this failure intensity and the costs given a data set with failure and maintenance records.

Time-to-failure model From the failure intensity function for each inter-PM interval, $\lambda(t)$, (Eq. (1)) and the timings of the PM interventions, $\{t_1, t_2, ..., t_n\}$, we define the failure intensity function $\lambda_T(t)$ in absolute time, i.e. since the start of the observation horizon,

$$\lambda_T(t) = \begin{cases} \lambda_0(t)\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) & \text{if } 0 \leq t < t_1 \\ \lambda_0(t-t_1)\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) & \text{if } t_1 \leq t < t_2 \\ \lambda_0(t-t_2)\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) & \text{if } t_2 \leq t < t_3 \\ \vdots & \\ \lambda_0(t-t_n)\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x}) & \text{if } t_n \leq t \leq \Delta t. \end{cases}$$

The timings of the PM interventions, or the PM interval, should be known to set up the expression for $\lambda_T(t)$. This is not an issue, however, since the PM interventions are planned upfront. Denote $R(t|t_-)$ the reliability, or survival function, for the time-to-next-failure, with t_- the time of the previous failure,

$$R(t|t_-) = \exp\left(-\int_{t_-}^{t} \lambda_T(u)du\right).$$

To calibrate the time-to-failure model and consequently find estimates for the parameters of the baseline failure intensity function $\lambda_0(t)$, i.e. α and γ, and the impact of the covariates $\boldsymbol{\beta}$, we maximize the time-to-failure log-likelihood. The events of interest are the failures as well as the end of the observation horizon for each machine in the data. The latter acts as a censoring event. Each event is characterized by the vector $(t_f, t_{f,-}, \boldsymbol{x}, \delta)$, where t_f is the event time, $t_{f,-}$ the time of the previous event, $\boldsymbol{x}$ the machine characteristics and $\delta \in \{0, 1\}$ where $\delta = 0$ indicates a censored event, i.e. the end of the observation horizon, and $\delta = 1$ a failure. Summing over failure events j on machine i in the data provides the expression for the log-likelihood for the time-to-failure data,

$$\mathcal{L}(\alpha, \gamma, \boldsymbol{\beta}) = \sum_{i=1}^{N}\sum_{j=1}^{f_i} \delta_j \log\left(\lambda_T(t_{f,j})\right) + \log\left(R(t_{f,j}|t_{f,j,-})\right), \tag{6}$$

where N is the total number of machines and f_i is the number of failures (including the end of the observation horizon) on machine i. Consequently, the failures contribute with the logarithm of the probability density function and the end of the observation horizon with the logarithm of the reliability. Maximizing the log-likelihood $\mathcal{L}(\alpha, \gamma, \boldsymbol{\beta})$ for the time-to-failure model leads to estimates for the parameters α and γ of the baseline failure intensity function $\lambda_0(t)$ and the impact of the covariates $\boldsymbol{\beta}$.

Costs model We calibrate separate gamma generalized linear models (GLMs), as specified in Eq. (2), for the expected preventive maintenance costs $c_p(\boldsymbol{x})$ and the expected failure costs $c_f(\boldsymbol{x})$ taking into account the machine characteristics $\boldsymbol{x}$. This provides estimates for $\boldsymbol{\beta}_p$ and $\boldsymbol{\beta}_f$, the impact of the machine characteristics $\boldsymbol{x}$ on the preventive maintenance costs and on the failure costs respectively, and for θ_p and θ_f, the scale parameter of the gamma distribution of the preventive maintenance costs and of the failure costs respectively.

Benchmark approaches The approach described above makes use of all available failure and maintenance data by pooling the data across all machines. We therefore refer to this approach as the *pooling* approach. Yet, by specifying and calibrating the impact of the machine characteristics $\boldsymbol{x}$ on the failure intensity and costs, our approach can differentiate the optimal periodic policy per machine

profile. We specify two benchmark approaches. First, a *uniform* approach that aggregates the data, but disregards the machine profiles, calibrates the time-to-failure and costs models ignoring the machine characteristics $\boldsymbol{x}$. This is identical to setting $\boldsymbol{\beta} = 0$ when optimizing the log-likelihood in Eq. (6). Similarly, we also ignore the machine characteristics $\boldsymbol{x}$ when calibrating the cost models. This gives us only estimates for $\beta_{p,0}$, θ_p, $\beta_{f,0}$ and θ_p. This approach also makes use of all available data, but the models are calibrated as if all machines would have no machine-specific characteristics, resulting in a uniform PM policy with identical, optimized number of PM interventions $n^\star$ for each machine (profile). Second, we also benchmark against a *stratified* approach. For this approach, we split or stratify the data in subsets that only contain data on a single machine profile, i.e. combination of machine characteristics, and then calibrate the time-to-failure and costs models in similar fashion to the *uniform* approach, i.e. by ignoring the machine characteristics $\boldsymbol{x}$, for each subset. The models for each machine profile serve as input to optimize the number of PM interventions. Although these policies are capable of differentiating over the different machine profiles, they use less data to calibrate the models. The latter can lead to less accurate estimates and inferior cost performance of the resulting maintenance policies. In our numerical analysis we will also benchmark against the *oracle* approach. The oracle knows the distribution of the failure behaviour and costs, and their associated parameters and is hence equivalent to assuming an infinite number of observations. Here we use the true distributions (rather than estimates) to find the optimal number of preventive maintenance interventions for each machine profile. These oracle policies serve as a lower bound on the costs.

3. Results and insights

We set up a numerical experiment to assess the value of differentiating the PM policy per machine profile by pooling the data using the approach described in the previous section. To do so, we generate a data set with maintenance and failure records from a heterogeneous machine portfolio with different machine profiles. The simulation engine to generate this data set is described in Section 3.1. We calibrate the parameters of the time-to-failure and cost models making use of maximum likelihood estimation (described in Section 2.3), and apply the optimality condition (Proposition 1) to prescribe the optimal number of preventive maintenance interventions for each machine profile. We report the cost performance of this approach in Section 3.2 and illustrate how each approach performs with limited amounts of data in Section 3.3. In Section 3.4 we check whether it is actually worth differentiating the PM policy at all compared to adopting a uniform PM policy that is identical across all machine profiles. Finally, Section 3.5 studies model mis-specification where the fitted model has a different parametric form from the model from which the data were simulated.

3.1. Simulation engine

We generate a data set of failures and maintenance records for a heterogeneous machine portfolio, following the reliability model introduced in Section 2.1. We consider a portfolio with 240 machines, of which 90% is observed for $\Delta t_i = 5$ years and 10% has a shorter history of $\Delta t_i \overset{d}{\sim} U[1, 5]$ years. We characterize each machine by 4 features, $\boldsymbol{x} = (x_{1,1}, x_{1,2}, x_{1,3}, x_{1,4})^T \in \{0, 1\}^4$ which are randomly assigned following a uniform distribution. This leads to 16 different machine profiles, each occurring in the portfolio with equal probability.

The preventive maintenance interventions in the data set are executed periodically with an interval of $\Delta t_{\text{PM}} = 1$ year. Recurrent failure times in a PM interval are generated by inverse transform sampling from the failure distribution determined by the failure intensity function in Eq. (1) [20,7,16,24]. The costs of failures and preventive maintenance interventions are positive, follow a right-skewed distribution and are dependent of machine characteristics $\boldsymbol{x}$. To accommodate for these properties, they are sampled from a gamma GLM [12,9].

The simulation parameters that we have used to generated failure times and costs are summarized in Supplementary Table (Appendix A). Our parameters set the mean time-to-failure for the machine profile $\boldsymbol{x} = (0000)$ equal to 1.27 years, if the machine would not be maintained preventively. The expected costs $c_p(\boldsymbol{x})$ and $c_f(\boldsymbol{x})$ for machine profile $\boldsymbol{x} = (0000)$ are respectively 30 and 300. Without loss of generality, we let the costs of a PM be independent of the machine characteristics, as the optimality condition (Eq. (5)) only depends on the ratio of $c_p(\boldsymbol{x})$ and $c_f(\boldsymbol{x})$. In Supplementary Table (Appendix B), we display an extract of the failure and maintenance records in our simulated data set. In Sections 3.2-3.4, the data are generated from correctly specified models, i.e., of the same parametric form. Section 3.5 studies model mis-specification where the fitted model has a different parametric form from the model from which the data were simulated.

3.2. What is the value of data pooling?

Table 1 reports for a single simulated data set the optimal number of preventive maintenance interventions, $n^\star$, during a time horizon of $\Delta t = 5$ years for each of the 16 machine profiles, as determined by the different approaches. These different approaches all rely on the optimality condition in Proposition 1, yet with different estimations of the failure behaviour and the costs. We refer to Section 2.3 for a discussion of these approaches. Depending on the (simulated) data set, the estimations of the failure behaviour and the costs, and therefore also the prescribed number of PM interventions, may slightly differ. To report the cost performance of each of these approaches, we therefore generated 100 data sets with identical simulation parameters. Table 2 reports the average total maintenance costs for each approach (making use of Eq. (4)), as well as the empirical 95%-confidence interval over the 100 simulation runs. The costs are given with respect to the costs of the *oracle* approach by means of a multiplicative ratio. Furthermore, we include the cost performance for an `average` machine profile. The maintenance cost for the `average` machine profile is determined by taking the average of the costs of all machine profiles with equal weights. Since the *oracle* serves as a lower bound, all other approaches will have higher or at best equal costs. The *oracle* provides the optimal differentiated maintenance strategy over the different machine profiles. To obtain these policies, however, we would need infinite amount of data to exactly know the underlying failure behaviour and costs.

The maintenance costs obtained under the *stratified* approach are on average only 5% higher compared to the *oracle* for the `average` machine profile (see Table 2). However, for machine profile $\boldsymbol{x} = (1100)$ the cost performance resulting from a *stratified* approach is worse, i.e. 8.3% on average, compared to the *oracle*. For a specific data set, we also observe that there can be a large discrepancy between the prescribed number of PMs by the *oracle* and the *stratified* approach, e.g. the *oracle* and the *stratified* approach respectively prescribe 13 and 6 PMs for machine profile $\boldsymbol{x} = (0100)$ (see Table 1). In general, there is also a lot of spread in the cost performance under the *stratified* approach. The 97.5% quantiles for the maintenance costs are very high, up to 74.4% higher than the costs obtained by the *oracle* (see profile $\boldsymbol{x} = (0011)$ in Table 2). The lack of sufficient data per profile due to the stratification of the data set, may lead to poor and volatile performance.

The *uniform* approach pools the data across machine profiles, yet without tailoring the maintenance policies to each machine

Table 1
Prescribed number of PMs during a contract horizon of 5 years for a single data set.

machine profile	oracle	pooling	stratified	uniform
0 0 0 0	10	9	11	10
0 0 0 1	6	6	5	10
0 0 1 0	8	8	7	10
0 0 1 1	5	5	4	10
0 1 0 0	13	11	6	10
0 1 0 1	9	8	6	10
0 1 1 0	11	9	9	10
0 1 1 1	7	6	5	10
1 0 0 0	14	13	13	10
1 0 0 1	9	9	10	10
1 0 1 0	11	11	13	10
1 0 1 1	7	7	5	10
1 1 0 0	18	15	21	10
1 1 0 1	12	11	10	10
1 1 1 0	15	13	15	10
1 1 1 1	10	9	5	10

profile. This approach alleviates the lack of data, but also leads to an average loss of 5% compared to the *oracle*, in this case due to lack of differentiation in the PM policies. Although the average performance of the *uniform* approach and the *stratified* approach is (almost) the same, the spread on the costs under the *uniform* approach is much smaller. This is also observed from the smaller 97.5% quantiles for specific machine profiles compared to the *stratified* approach.

The *pooling* approach makes use of all the data, and also differentiates the maintenance policies over the different machine profiles. This leads to considerably better performance compared to the *stratified* and *uniform* approaches, only having a loss of 0.7% on average with respect to the *oracle*. The improvement is not only for the `average` machine, it also is the case for the different machine profiles. Furthermore, the spread in performance, as quantified by the 95%-confidence intervals, is considerably smaller compared to the *stratified* and *uniform* approach, both for an `average` machine and for each individual machine profile.

3.3. How does data pooling overcome the small data problem?

Although the *pooling* and *stratified* approaches have a similar goal, i.e. tailoring the maintenance policy to its machine profile, they use the data set in a different way. While the *pooling* approach uses all the data over the different profiles, relying on the assumption of proportional hazards for the failure behaviour and the assumption of a GLM for the costs, the *stratified* approach only considers the data per profile completely disjoint from the others. Splitting the data set per machine profile produces smaller subsets of data, potentially inducing a small data problem. The consequent underperformance of the *stratified* approach is due to the fact that insufficient data may be available *per machine profile* to estimate the failure behaviour and costs accurately from the data. Clearly, if sufficient data would be available for each machine profile, both the *stratified* and *pooling* approach converge to the *oracle*. Yet, the rate at which both approaches converge is different.

Fig. 1 demonstrates this convergence by displaying the relative costs of the *stratified* and the *pooling* approach with respect to the *oracle* when we gradually increase the number of machines in the data set, and thus the number of failure and maintenance records that are used to estimate the failure behaviour and costs. We start from a data set generated by 10 machines observed during 5 years and consider increases of 10 machines at a time. These increases correspond to an additional 50 machine years of historical data (recall that the mean time-to-failure for machine profile $\boldsymbol{x} = (0000)$ is equal to 1.27 years). We report the average cost performance over 40 simulated data sets with identical parameters and focus on the `average` machine and worst performing machine profile, i.e. the machine profile that has the highest relative costs with respect to the *oracle* at any given size of the data set. Fig. 1 shows how both the *stratified* and the *pooling* approach converge to the *oracle* costs. Yet, the rate of convergence for the *pooling* is much higher than the *stratified* approach. Also the 90%-confidence intervals for the *pooling* approach are smaller and shrink faster than the *stratified* approach. To get insight in the rate of convergence, we consider the relative costs, averaged over the 40 simulated data sets, of the `average` profile as a function of the number of machines, and we look for a such that,

$$\text{average relative costs} = \frac{a}{\text{number of machines}} + 1.$$

We find fitted values for a equal to 1.715 and 13.341 for the *pooling* and *stratified* approach respectively. The ratio of these values indicates that the *pooling* approach converges to the *oracle* over seven and half times faster than the *stratified* approach in the number of machines on average. It shows how the *pooling* approach requires much less data to obtain adequate performance.

Another downside of the *stratified* approach is that it cannot prescribe the number of PMs for a machine profile that is not available in the data set. For instance, if we consider a data set generated by only 10 machines, not all 16 profiles will be represented in the data. This also explains the sharp cost increase of the worst performing machine profile under the *stratified* approach when the machine portfolio is small (see Fig. 1, panel (b)). In that case, increasing the number of machines also increases the number of machine profiles in the data set, and with that also the likelihood of a worse performing profile. This is not the case under a *pooling* approach. Even when certain combinations of machine characteristics, i.e. machine profiles, may not be observed in the data, the *pooling* approach can still prescribe the preferred number of PM interventions. This property is essential for an OEM or a service provider when they expand their maintenance portfolio with machine profiles, for which no data is yet available.

3.4. When is it valuable to differentiate the PM policy?

Whereas our numerical experiment has shown how our *pooling* approach is capable to differentiate the PM policy per machine profile, it is worthwhile to check whether differentiating the maintenance policy is actually worth the effort compared to adopting a uniform PM policy that is the same for all machines. We do so by careful analysis of the optimality condition in Proposition 1. Specifically, we consider the equivalent condition under the assumption of a Weibull failure intensity function,

$$(\alpha \Delta t)^{\gamma} \frac{(n+2)^{\gamma-1} - (n+1)^{\gamma-1}}{((n+1)(n+2))^{\gamma-1}} \leq c_{ratio}(\boldsymbol{x}), \tag{7}$$

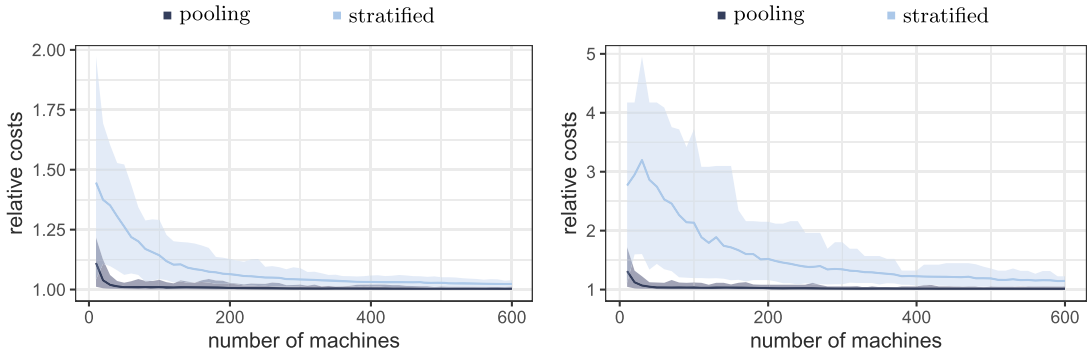
where

$$c_{ratio}(\boldsymbol{x}) = \frac{c_p(\boldsymbol{x})}{c_f(\boldsymbol{x}) \exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})}$$

denotes the machine profile specific cost ratio. This ratio provides an indication how many PM interventions are economic for a machine profile $\boldsymbol{x}$: a low value of $c_{ratio}(\boldsymbol{x})$ indicates that it is preferred to perform many PM interventions on this machine profile, as the cost of preventive maintenance is low compared to the cost of failure, and vice versa for a high value of $c_{ratio}(\boldsymbol{x})$. The left-hand-side of (7) depends on the baseline failure intensity, characterized by the Weibull scale α and shape γ, and the contract length Δt.

Table 2
Average costs (and empirical 95%-confidence interval) over 100 simulated data sets with identical parameters. Costs are determined exact using Eq. (4). We display the absolute costs for the *oracle* and the relative costs w.r.t. the *oracle* for the other approaches. The relative costs are determined by division with the costs realized under *oracle*. The `average` machine profile's costs are determined by taking the average over all machine profiles with equal weights.

machine profile	oracle	pooling (%)	stratified (%)	uniform (%)
0 0 0 0	634.09	100.5 (100,102.2)	104.2 (100,120.3)	100.3 (100,102.2)
0 0 0 1	415.9	100.2 (100,102.2)	105 (100,134.5)	109.5 (101.8,112.4)
0 0 1 0	513.71	100.3 (100,100.8)	104.6 (100,137.4)	102.8 (100,104.2)
0 0 1 1	334.48	100.3 (100,102.1)	107.5 (100,174.4)	121.6 (108.5,126.2)
0 1 0 0	822.79	100.9 (100,105)	105 (100,134.4)	103.2 (101.5,111)
0 1 0 1	542.25	100.3 (100,101.5)	105.6 (100,132.8)	101.6 (100,102.7)
0 1 1 0	668.46	100.6 (100,103.4)	103.7 (100,123.7)	100.3 (100,103.4)
0 1 1 1	438.12	100.3 (100,100.6)	104.9 (100,139.3)	107.5 (101.1,110)
1 0 0 0	866.42	100.9 (100,104.9)	103.7 (100,126.3)	104.6 (102.5,113.6)
1 0 0 1	570.88	100.3 (100,100.6)	103.8 (100,114.1)	101 (100,101.7)
1 0 1 0	704.05	100.6 (100,102.1)	103.9 (100,127.6)	100.7 (100,104.9)
1 0 1 1	462.11	100.3 (100,101.3)	104.2 (100,128.6)	105.6 (100.4,107.8)
1 1 0 0	1121.07	101.8 (100,111.5)	108.3 (100,136.6)	115.4 (111.5,130.8)
1 1 0 1	741.59	100.6 (100,103.3)	103.5 (100,119.8)	101.3 (100.2,106.7)
1 1 1 0	912.53	101.2 (100,106.6)	106.5 (100,135.6)	106.2 (103.7,116.4)
1 1 1 1	602.3	100.3 (100,101.2)	103.3 (100,116.9)	100.4 (100,101.2)
`average`	646.92	100.7 (100,103.5)	105 (101.1,111.7)	105 (104.6,108.5)



(a) `average` machine profile (b) worst machine profile

Fig. 1. Relative costs for the *pooling* approach and the *stratified* approach with respect to the *oracle* in function of the number of machines that generate data, averaged over 40 simulated data sets with identical parameters and with the empirical 90%-confidence interval shaded. Panel (a) Relative costs of the `average` machine profile; panel (b) Relative costs of the worst performing machine profile.

Both are independent of the machine profile $\boldsymbol{x}$. The right-hand-side $c_{ratio}(\boldsymbol{x})$ captures all the effects of the heterogeneity in the machine portfolio. It depends on the machine profile's costs of failure via $c_f(\boldsymbol{x})$ and preventive maintenance via $c_p(\boldsymbol{x})$ and the impact of the profile on the failure behaviour, $\exp(\boldsymbol{\beta}' \cdot \boldsymbol{x})$.

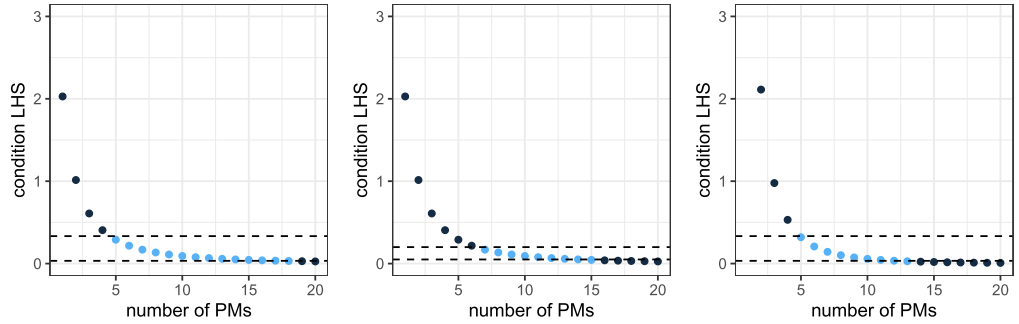
It is insightful to visualize this optimality condition. The dots in Fig. 2 visualize the left-hand-side of (7) for each value of n, the number of PM interventions performed during the contract horizon. The minimum and maximum value of $c_{ratio}(\boldsymbol{x})$ over the machine profiles, resp. $\min(c_{ratio}(\boldsymbol{x}))$ and $\max(c_{ratio}(\boldsymbol{x}))$, are indicated by dashed horizontal lines. These two horizontal lines define the region of the optimal number of PMs for the machine profiles, defined by $\boldsymbol{x}$. We highlight these PM policies in light blue. When there are many light blue dots, the optimal PM policies differ much across the different machine profiles. In the extreme case where there is only a single point in this region, then the same number of PMs is optimal for all machine profiles and differentiation of the PM policy is not required.

An extensive analysis of the optimality condition and the resulting number of differentiated PM policies, reveals that less heterogeneity in the machine portfolio, resulting in a smaller gap between $\min(c_{ratio}(\boldsymbol{x}))$ and $\max(c_{ratio}(\boldsymbol{x}))$ with the horizontal, dashed lines closer together, leads to less differentiation in the PM policies over the different machine profiles (see Fig. 2b). Also, a larger Weibull shape γ of the failure behaviour (while adapting scale α accordingly to maintain a constant mean-time-to-failure), makes the left-hand-side of the optimality condition steeper, i.e. the left-hand-side decreases faster (see Fig. 2c). This reduces the number of different optimal PM policies across the profiles and diminishes the effect of the heterogeneity of the machine population.

3.5. Model mis-specification vs data-pooling

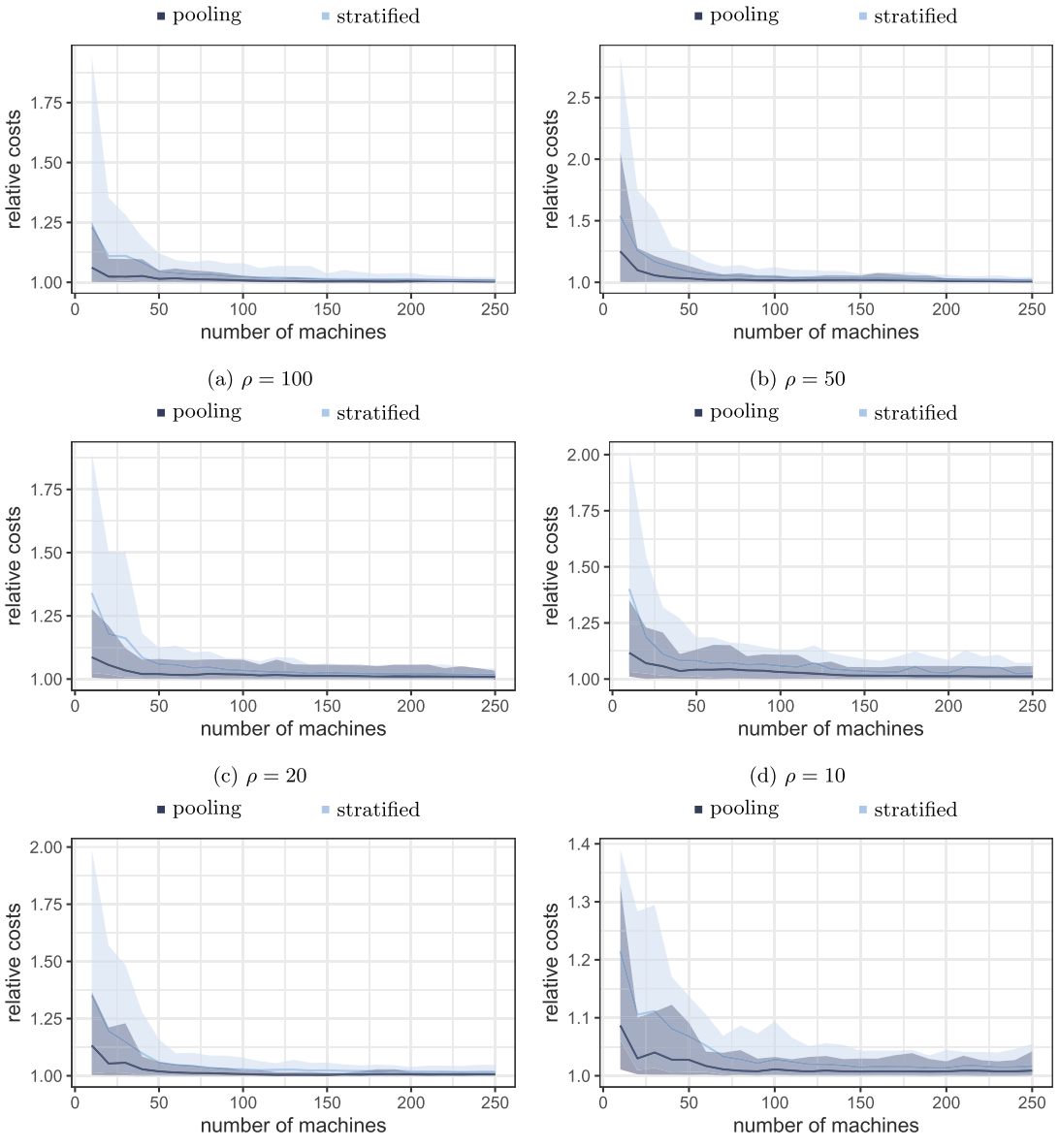
The data pooling approach requires the specification of relevant terms in (1). In certain contexts, a modeller may not correctly identify all the terms that should be present in a proportional hazard model. The stratified approach is immune to such mis-specifications, but requires large amounts of data to obtain accurate estimates. The data-pooling approach is susceptible to such mis-specifications but has the advantage of being able to pool the data. Below we consider a set-up to study which effect is dominant.

We consider a set-up with only two (binary) covariates, i.e. $x_1, x_2 \in \{0, 1\}$, impacting the time-to-failure. We simulate the failure times for each machine from following machine-specific failure intensity function $\tilde{\lambda}(t)$,



(a) Simulation parameters (b) Less heterogeneity (c) Larger Weibull shape γ

Fig. 2. Left-hand-side of the optimality condition (dots) and $\min(c_{ratio}(\boldsymbol{x}))$ and $\max(c_{ratio}(\boldsymbol{x}))$ (dashed horizontal lines). The available number of PMs determined by $\min(c_{ratio}(\boldsymbol{x}))$ and $\max(c_{ratio}(\boldsymbol{x}))$ are highlighted in light blue. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)



(a) $\rho = 100$ (b) $\rho = 50$

(c) $\rho = 20$ (d) $\rho = 10$

(e) $\rho = 5$ (f) $\rho = 2$

Fig. 3. Relative costs for the *pooling* approach and the *stratified* approach with respect to the *oracle* in function of the number of machines that generate data according to the set-up in Section 3.5, averaged over 20 simulated data sets with identical parameters and with the empirical 90%-confidence interval shaded.

$$\tilde{\lambda}(t) = \lambda_0(t) \exp\left(\beta x_1 + \beta x_2 + \frac{\beta}{\rho} x_1 x_2\right) \text{ for } t \in [0, t_k - t_{k-1}) \ (\forall k), \tag{8}$$

where $t_0 = 0$ and $t_{n+1} = \Delta t$, respectively refer to the start and end of the planning horizon. We choose the same β for the covariates x_1 and x_2 to ensure that the impact of the cross-term $(x_1 x_2)$ is of the same magnitude as the main effects. The parameter ρ controls the relative impact of the cross-term. The simulation of failure and maintenance costs is not changed from the original set-up. We compare the stratified approach with the data pooling approach but let the data pooling approach fit the following mis-specified model:

$$\widehat{\lambda}(t) = \lambda_0(t) \exp\left(\beta x_1 + \beta x_2\right) \text{ for } t \in [0, t_k - t_{k-1}) \ (\forall k). \tag{9}$$

Thus we can interpret ρ^{-1} as measuring the amount of model mis-specification of the data pooling approach. Note that the stratified approach does not suffer from mis-specification since it makes no assumption of the impact of the covariates on the time-to-failure (nor the failure or maintenance costs). Consequently, the resulting maintenance policies for the pooling approach will result from the mis-specified model and we will be able to test the impact of this mis-specification.

In order to assess the impact of the mis-specification, we compare the costs relative to the oracle of both the pooling approach and the stratified approach for decreasing ρ (increased model mis-specification). The oracle is of course adapted to consider the influence of the cross-term. The results are displayed in Fig. 3. They show that the effect of data pooling outweighs the effect of model mis-specification in all considered settings. This is a strong indication, that, while mis-specification is likely to happen, its negative impact is easily offset by the benefits of data pooling.

4. Conclusion

This paper describes a data-driven approach to optimize the periodic, preventive maintenance policies for a heterogeneous machine population over a finite time horizon. The heterogeneity of the machine population is characterized by observable machine characteristics that induce different machine profiles. Our approach *pools* the available data of failure and maintenance records over the different machine profiles in order to learn as best as possible the failure behaviour and the costs of failure and maintenance for each machine profile. We rely on the assumption of proportional hazards for the failure behaviour and on the assumption of a gamma GLM for the costs to accomplish this data pooling. In conjunction with the estimates for the failure behaviour and the costs, our optimality condition for the number of preventive maintenance interventions delivers tailored maintenance policies for each machine profile. By means of numerical experiments, we compare our *pooling* approach with both a *stratified* approach that splits the data per machine profile and a *uniform* approach that disregards the machine profiles and prescribes the same uniform PM policy for all machine profiles. The *pooling* approach outperforms these benchmarks and additionally has a smaller spread on its performance. We also show how the *pooling* approach is more data-efficient than the *stratified* approach, even under mild model mis-specification. This means that the *pooling* approach obtains better performing maintenance policies for the same amount of data. Finally, we investigate when it is worth differentiating the PM policy given the heterogeneity in the machine population and the failure intensity, motivating the approach introduced in this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.orl.2023.01.006.

References

[1] David T. Abdul-Malak, Jeffrey P. Kharoufeh, Lisa M. Maillart, Maintaining systems with heterogeneous spare parts, Nav. Res. Logist. 66 (6) (2019) 485–501.
[2] Joachim Arts, Rob Basten, Design of multi-component periodic maintenance programs with single-component models, IISE Trans. 50 (7) (2018) 606–615.
[3] Abbas Barabadi, Javad Barabady, Tore Markeset, Application of reliability models with covariates in spare part prediction and optimization–a case study, Reliab. Eng. Syst. Saf. 123 (2014) 1–7.
[4] R.E. Barlow, F. Proschan, Mathematical Theory of Reliability, Wiley, 1965.
[5] Richard Barlow, Larry Hunter, Optimum preventive maintenance policies, Oper. Res. 8 (1) (1960) 90–100.
[6] A. Bobbio, A. Cumani, A. Premoli, O. Saracco, Modeling and identification of nonexponential distributions by homogeneous Markov processes, in: Proc. of the Sixth Advances in Reliability Symposium, 1980.
[7] Richard J. Cook, Jerald Lawless, The Statistical Analysis of Recurrent Events, Springer Science & Business Media, 2007.
[8] David R. Cox, Partial likelihood, Biometrika 62 (2) (1975) 269–276.
[9] Piet De Jong, Gillian Z. Heller, et al., Generalized Linear Models for Insurance Data, Cambridge University Press, Cambridge, 2008.
[10] Bram de Jonge, Arjan S. Dijkstra, Ward Romeijnders, Cost benefits of postponing time-based maintenance under lifetime distribution uncertainty, Reliab. Eng. Syst. Saf. 140 (2015) 15–21.
[11] Łukasz Delong, Mathias Lindholm, Mario V. Wüthrich, Collective reserving using individual claims data, Scand. Actuar. J. (2021) 1–28.
[12] Michel Denuit, Xavier Maréchal, Sandra Pitrebois, Jean-François Walhin, Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems, John Wiley & Sons, 2007.
[13] Laurent Doyen, Olivier Gaudoin, Modeling and assessment of aging and efficiency of corrective and planned preventive maintenance, IEEE Trans. Reliab. 60 (4) (2011) 759–769.
[14] Collin Drent, Stella Kapodistria, Onno Boxma, Censored lifetime learning: optimal bayesian age-replacement policies, Oper. Res. Lett. 48 (6) (2020) 827–834.
[15] İpek Dursun, Alp Akçay, Geert-Jan Van Houtum, Age-based maintenance under population heterogeneity: optimal exploration and exploitation, Eur. J. Oper. Res. 301 (3) (2022) 1007–1020.
[16] Antje Jahn-Eimermacher, Katharina Ingel, Ann-Kathrin Ozga, Stella Preussler, Harald Binder, Simulating recurrent event data with hazard functions defined on a total time scale, BMC Med. Res. Methodol. 15 (1) (2015) 16.
[17] Dhananjay Kumar, Ulf Westberg, Some reliability models for analyzing the effect of operating conditions, Int. J. Reliab. Qual. Saf. Eng. 4 (2) (1997) 133–148.
[18] P.A.W. Lewis, Gerald S. Shedler, Simulation of nonhomogeneous Poisson processes by thinning, Nav. Res. Logist. Q. 26 (3) (1979) 403–413.
[19] Bo Henry Lindqvist, et al., On the statistical modeling and analysis of repairable systems, Stat. Sci. 21 (4) (2006) 532–551.
[20] Chris Metcalfe, Simon G. Thompson, The importance of varying the event generation process in simulation studies of statistical methods for recurrent events, Stat. Med. 25 (1) (2006) 165–179.
[21] Jesper Moller, Rasmus Plenge Waagepetersen, Statistical Inference and Simulation for Spatial Point Processes, CRC Press, 2003.
[22] Toshio Nakagawa, Satoshi Mizutani, A summary of maintenance policies for a finite interval, Reliab. Eng. Syst. Saf. 94 (1) (2009) 89–96.
[23] Esbjörn Ohlsson, Björn Johansson, Non-life Insurance Pricing with Generalized Linear Models, vol. 2, Springer, 2010.
[24] Juliette Pénichoux, Thierry Moreau, Aurélien Latouche, Simulating recurrent events that mimic actual data: a review of the literature with emphasis on event-dependence, arXiv preprint, arXiv:1503.05798, 2015.
[25] W. Wang, P.A. Scarf, M.A.J. Smith, On the application of a model of condition-based maintenance, J. Oper. Res. Soc. 51 (11) (2000) 1218–1227.
[26] Shaomin Wu, A failure process model with the exponential smoothing of intensity functions, Eur. J. Oper. Res. 275 (2) (2019) 502–513.
[27] Shaomin Wu, Ming J. Zuo, Linear and nonlinear preventive maintenance models, IEEE Trans. Reliab. 59 (1) (2010) 242–249.