Optimal data pooling for shared learning in maintenance operations

Operations Research Letters 52 (2024) 107056

Contents lists available at ScienceDirect

Operations Research Letters

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

Optimal data pooling for shared learning in maintenance operations

Collin Drent *, Melvin Drent, Geert-Jan van Houtum

Eindhoven University of Technology, School of Industrial Engineering, Eindhoven, the Netherlands

ARTICLE INFO

Article history:
Received 9 February 2023
Received in revised form 6 November 2023
Accepted 29 November 2023
Available online 6 December 2023

Keywords:
Condition-based maintenance
Data pooling
Bayesian learning
Spare parts
Optimal policy

ABSTRACT

We study optimal data pooling for shared learning in two common maintenance operations: condition-based maintenance and spare parts management. We consider systems subject to Poisson input – the degradation or demand process – that are coupled through an unknown rate. Decision problems for these systems are high-dimensional Markov decision processes (MDPs) and are thus notoriously difficult to solve. We present a decomposition result that reduces such an MDP to two-dimensional MDPs, enabling structural analyses and computations. Leveraging this decomposition, we (i) show that pooling data can lead to significant cost reductions compared to not pooling, and (ii) prove that the optimal policy for the condition-based maintenance problem is a control limit policy, while for the spare parts management problem, it is an order-up-to level policy, both dependent on the pooled data.

1. Introduction

Unplanned downtime of advanced technical systems such as aircraft, lithography systems, or rolling stock, is very costly for companies that rely on these systems in their operations. As such, these companies typically have agreements with maintenance service providers to ensure high system availability. Advancements in information technology along with continuous reductions in costs of sensors have led to ample opportunities for service providers to improve their maintenance operations [20]. Indeed, modern systems are increasingly equipped with sensors that relay degradation data of critical components in real-time to decision-makers. This data is useful for inference of degradation behavior; however, the amount of data that each such system generates to predict failures of a particular component is scarce, especially for newly introduced systems.

Maintenance service providers typically maintain several systems of the same type (e.g. similar systems for the same customer at different locations, or similar systems for different customers). At the beginning of the life-cycle of a newly introduced system, the maintenance service provider thus faces a setting where (i) multiple systems of the same type generate a steady stream of real-time degradation data, but at the same time, (ii) each such system alone has not yet generated sufficient amounts of data. A prime example of this can be found in the semiconductor industry. Upon the introduction of a new generation of lithography system in the field, many critical components in this system are also used for the first time, and hence no historical degradation data is available [10]. It is evident that pooling degradation data from multiple systems can lead to cost reductions in the setting described above. However, it remains unclear how to quantify these cost reductions, especially when we are interested in optimal decisions and the state space of the corresponding Markov decision process (MDP) thus becomes large. In this paper, we address this question.

We consider a maintenance service provider that is responsible for maintaining multiple systems of the same type at different locations or customers. We consider a single component that is present in the configuration of all systems. Components face Poisson deterioration with the same but unknown rate, and systems fail when the component's degradation reaches a certain failure threshold. Failures can be prevented by performing preventive maintenance, which is cheaper than replacement upon failure, which generally leads to costly unplanned downtime. We must decide sequentially – based on the data from all systems – for each system to perform preventive maintenance or not, thereby trading off premature interventions with tardy replacements. Systems have the same unknown deterioration rate, but are otherwise heterogeneous (i.e. costs and failure thresholds). We propose a Bayesian procedure to pool all data and jointly learn the rate on-the-fly. We model the problem as a Bayesian MDP for which the optimal policy – in theory – can be computed numerically. However, because the action and state space grow exponentially in the number of systems, this MDP quickly suffers from the curse of dimensionality, making it impossible to assess the value of optimal data pooling. As a remedy, we establish a decomposition result that

* Corresponding author.
E-mail addresses: c.drent@tue.nl (C. Drent), m.drent@tue.nl (M. Drent), g.j.v.houtum@tue.nl (G.-J. van Houtum).

https://doi.org/10.1016/j.orl.2023.11.009

reduces this high-dimensional MDP to multiple two-dimensional MDPs that permit structural analyses and computations.

When components have constant failure rates, maintenance service providers typically replace these components with new spares only correctively upon failure, i.e. they apply repair-by-replacement. The underlying spare parts inventory system responsible for supplying these spares then largely determines the availability of the technical systems. As an extension, we will show that our decomposition result also applies to such a spare parts inventory system consisting of multiple local warehouses that keep spares for the same critical component whose failure rate is unknown.

Sequential Bayesian learning based on sensory data stemming from systems has been used extensively in the maintenance literature to study optimal maintenance decision-making with a-priori unknown parameters [e.g. 11,6,8], but only for single-component systems in isolation (see [4] for an overview of the area). This makes sense when the unknown parameter is unique to the specific system in use. However, as argued above, in practice, parameters may be the same for multiple systems of the same type. In such a setting, which we consider in this paper, it is natural to pool data stemming from all these systems to jointly learn the common parameter.

The benefit of pooling has been studied extensively, yet almost exclusively related to pooling of physical resources. Notable examples include inventory pooling in inventory networks [e.g., 12] and pooling of server capacity in queuing networks [e.g., 19]. Recently, researchers have started to investigate the benefits of pooling data [e.g., 2,14], but so far only two works exist on data pooling in maintenance optimization [5,10]. [5] investigates the benefits of combining data from a set of heterogeneous machines in the context of time-based preventive maintenance. They propose a method to aggregate data from multiple machines such that it can be utilized for selecting a periodic interval at which preventive maintenance is performed for each individual system. [10] pools data to determine whether systems are stemming from a so-called weak or strong population, where the former has lifetimes that are stochastically smaller than the latter. Unlike [5], who proposes a static estimation procedure based on historical pooled data, [10] builds an MDP that sequentially learns as more data becomes available. They numerically show – only for small instances due to the curse of dimensionality – that data pooling can lead to savings of up to 14% compared to not pooling. We also learn from pooled data in a dynamic, sequential way, but beat the curse of dimensionality by leveraging our decomposition result. Both [5] and [10] pool data to learn a time-to-failure model in a time-based maintenance setting, while we focus on learning a degradation model in a condition-based maintenance (CBM) setting.

The contributions of this paper are as follows. First, we formulate the problem of optimally maintaining N systems while pooling data to learn a common unknown deterioration rate as a Bayesian MDP, and establish a novel result to decompose this high-dimensional MDP into N two-state MDPs. Second, using the decomposition, we show that the optimal policy of each individual system has a control limit structure, where the control limit depends on the pooled data. Surprisingly, we show that this optimal control limit is not monotone in general. It typically decreases first but it always increases and converges to the failure level when the pooled data grows very large, implying that preventive maintenance is never optimal in that asymptotic regime. Third, numerical results show that savings due to pooling data can be substantial, even for small values of N. The exact magnitude of these savings largely depends on the degree of the uncertainty in the parameter. Under high uncertainty, savings of close to 57% can be realized on average, while these savings become almost negligible when uncertainty is low. Finally, we apply our decomposition result to a spare parts inventory system consisting of multiple local warehouses where a common failure rate needs to be learned. For this setting, we establish the optimality of monotone order-up-to policies that are non-decreasing in the pooled data.

The remainder is organized as follows. We discuss the model in §2. In §3, we formulate the problem as a Bayesian MDP and we establish the decomposition result. Structural properties of the expected cost and the optimal policy of the alternative MDP are presented in §4. In §5, we report on a numerical study that highlights the benefit of pooling data. In §6, we conclude by applying our decomposition result to a set of spare parts inventory systems. All proofs are relegated to Appendix A.

2. Model description

We consider $N \geq 1$ systems subject to damage accumulation due to random shock arrivals. Random shock degradation is a common assumption in the maintenance literature [see, e.g., 16,18,7] that has been validated by practice-based research [8]. Although data pooling has only value when $N > 1$, the analysis in this paper also holds for $N = 1$. Let $\mathcal{N} \triangleq \{1, \ldots, N\}$ be the set of all systems. Each system has a critical component such that the system breaks down whenever this component fails. The deterioration processes of these components are modeled as independent Poisson processes with the same, but unknown rate λ, denoted with $\{X_i(t), t \geq 0\}$, with $X_i(0) = 0$, for $i \in \mathcal{N}$. A component of system $i \in \mathcal{N}$ deteriorates until its deterioration level reaches or crosses a deterministic failure threshold $\xi_i \in \mathbb{N}_+$, where $\mathbb{N}_+ \triangleq \{1, 2, \ldots\}$, after which the component has failed. This failure threshold ξ_i is essentially the maximum physical capacity of a component to withstand the accumulated damage and under which system $i \in \mathcal{N}$ still adequately performs its function. In most practical situations, components of the same type will have the same failure threshold ξ_i. However, to allow for the general setting in which components have different capacities to withstand deterioration – which is reasonable when some components are of better quality than others – we let the failure threshold ξ_i depend on system $i \in \mathcal{N}$.

Deterioration is monitored at equally spaced decision epochs, though failure moments can happen at any point in time. Replacing only at decision epochs is a reasonable assumption given that critical components in these systems typically have mean lifetimes ranging from 1 to 10 years, while maintenance decisions are made much more frequently, often on a daily to weekly basis [21,17]. For convenience, we re-scale time such that the time between two decision epochs equals 1. If at a decision epoch a component of system $i \in \mathcal{N}$ is failed, it needs to be replaced correctively at costs $c_u^i > 0$. Such a failure can be prevented by performing a preventive replacement, which costs $c_p^i > 0$, with $c_p^i < c_u^i$ for all $i \in \mathcal{N}$. Corrective maintenance is more expensive because it includes costs caused by a component failure in addition to the costs related to the replacement (e.g. unplanned downtime costs). Both replacements involve a new component that starts deteriorating again from level 0 according to a Poisson process with rate λ, that is, $\{X_i(t), t \geq 0\}$ is reset to $X_i(0) = 0$. We assume that replacement times are negligible. This is reasonable given the efficiency of replacing old components with new ones, which usually takes only a few minutes to an hour – significantly shorter than the time between consecutive decision epochs. The systems have a common finite lifespan of length $T \in \mathbb{N}_+$ time units. This lifespan represents the time from their introduction until they are taken out of service, with typical durations of 10 to 30 years [21].

The maintenance service provider, responsible for maintaining the set of N systems, seeks to minimize the total expected maintenance costs over the systems' lifespan. Components used for replacements always have the same rate but this rate is a-priori unknown and needs to be inferred based on the observations of the

deterioration processes throughout this lifespan. Because components have the same rate, we can pool and utilize all accumulated data together in real-time to jointly learn this unknown rate. To this end, we adopt a Bayesian approach and treat the unknown rate λ as a random variable denoted with Λ. Upon the start of operating all systems, at $t = 0$, Λ is modeled by a Gamma distribution with shape parameter α_0 and rate parameter β_0. The subscript notation reflects that this corresponds to $t = 0$; we adopt this notation in the remainder of this paper. Thus, at $t = 0$, the density function of Λ is equal to $f_\Lambda(\lambda; \alpha_0, \beta_0) = \frac{\lambda^{\alpha_0-1} e^{-\beta_0 \lambda} \beta_0^{\alpha_0}}{\Gamma(\alpha_0)}$ for $\lambda > 0$ with $\alpha_0, \beta_0 > 0$, and where $\Gamma(\cdot)$ denotes the Gamma function. Estimation procedures are available in the literature for obtaining the parameters of this initial belief based on expert knowledge or historical data [see, e.g., 1,8]. Suppose that at decision epoch $t \in \mathbb{N}_+$, we observed a cumulative amount of k deterioration increments from all installed components. As degradation is modeled by a Poisson process, which is a non-decreasing, integer-valued process, we know that the degradation increments are non-negative and integer-valued as well. Hence, we know that the cumulative sum of all deterioration increments from all installed components, k, will always be a non-negative integer. Our choice for the Gamma distribution is not only empirically grounded [e.g. 1,8], but also mathematically convenient and therefore quite customary in the literature. Indeed, it is well-known that the Gamma distribution is a conjugate prior for the Poisson distribution, which implies that the new posterior distribution describing our belief of Λ is again a Gamma distribution but with updated parameters [see, e.g., 13, Chapter 2]:

$$\alpha_t = \alpha_0 + k \quad \text{and} \quad \beta_t = \beta_0 + N \cdot t. \tag{1}$$

Observe that from the updating scheme in Equation (1), it is immediately clear that the data stemming from all N systems is pooled for learning the unknown rate λ that the systems have in common. In Bayesian terminology, k is the sufficient statistic (which is thus linear in the observations) and $N \cdot t$ is the total amount of observations at decision epoch t. At each decision epoch, based the current belief of Λ, we wish to predict the future evolution of the deterioration of each component so that we can decide upon potential replacements. This prediction is encoded in the posterior predictive distribution. For this Gamma-Poisson model, it is well-known that the posterior predictive distribution is a Negative Binomial distribution [see, e.g., 13]. Specifically, given parameters α_t and β_t, the deterioration increment (i.e. $X_i(t+1) - x_i(t)$ with $x_i(t)$ the current deterioration at decision epoch t) of a component at system i at the next decision epoch, denoted with Z_i, is Negative Binomially distributed with parameters

$$r = \alpha_t \quad \text{and} \quad p = \frac{\beta_t}{\beta_t + 1}, \tag{2}$$

where r is the number of successes and p is the success probability, so that Z_i can be interpreted as the number of failures until the r^{th} success. In the remainder we use the notation $Z \sim NB(r, p)$ to denote that Z is a Negative Binomially distributed random variable with parameters r and p.

Equations (1) and (2) can be used to sequentially construct an updated posterior predictive distribution in real-time based on the observed data. Since the posterior predictive distributions of the deterioration increments of each system are fully described by only the current decision epoch t and cumulative amount of deterioration increments k, it is a Markov process. This allows us to formulate the optimization problem as a finite horizon (with length T) MDP equipped with the state variable k for Bayesian inference of the unknown rate. Before doing so, we end this section with an important result that establishes a stochastic ordering property of the posterior predictive distribution Z (for brevity we drop the dependence on k, N and t) in the cumulative amount of deterioration increments k when everything else is fixed.

Lemma 1. *The posterior predictive random variable Z is stochastically increasing in k in the usual stochastic order.*

Lemma 1 implies that if the sum of deterioration increments increases, and all else is fixed, then the next random deterioration increments are more likely to take on higher values. This is intuitive since the mean increment $(=\frac{\alpha_t}{\beta_t})$ increases in k.

3. Markov decision process formulation

We will now formulate the problem described in the previous section as an MDP. Let $\mathcal{S} \triangleq \mathbb{N}_0^{N+1}$, where $\mathbb{N}_0 \triangleq \mathbb{N}_+ \cup \{0\}$, be the state space of the MDP. For a state $(\boldsymbol{x}, k) \in \mathcal{S}$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ represents the vector of all deterioration levels, and k denotes the sum of all deterioration increments. Recall that as we are dealing with Poisson degradation, both the deterioration levels and the sum of deterioration increments are non-negative integer-valued. For a given state $(\boldsymbol{x}, k) \in \mathcal{S}$, let $\mathcal{A}(\boldsymbol{x})$ denote the action space. For any action $\boldsymbol{a} = (a_1, a_2, \ldots, a_N) \in \mathcal{A}(\boldsymbol{x})$, a_i represents the action per system, with $a_i \in \{0, 1\}$ when $x_i < \xi_i$ and $a_i = 1$ otherwise. Here, $a_i = 0$ corresponds to taking no action and $a_i = 1$ corresponds to performing maintenance on the component of system i, respectively. This implies that if the critical component of system i is failed (i.e. $x_i \geq \xi_i$), then the maintenance service provider must (correctively) replace it. For all components that have not failed, the maintenance service provider can choose to either preventively replace it, or do nothing and continue to the next decision epoch.

Given the state $(\boldsymbol{x}, k) \in \mathcal{S}$ and an action $\boldsymbol{a} = (a_1, a_2, \ldots, a_N) \in \mathcal{A}(\boldsymbol{x})$, the maintenance service provider incurs a direct cost, denoted by $C(\boldsymbol{x}, \boldsymbol{a})$, equal to

$$C(\boldsymbol{x}, \boldsymbol{a}) \triangleq \sum_{i \in \mathcal{N}} \Big(a_i \big(1 - \mathbb{I}_i(\boldsymbol{x})\big) c_p^i + \mathbb{I}_i(\boldsymbol{x}) c_u^i \Big), \tag{3}$$

where $\mathbb{I}_i(\boldsymbol{x})$ is an indicator function that indicates whether the component of system i has failed in the deterioration vector $\boldsymbol{x}$; that is, $\mathbb{I}_i(\boldsymbol{x}) = 0$ if $x_i < \xi_i$ and $\mathbb{I}_i(\boldsymbol{x}) = 1$ otherwise.

Let $V_t^N(\boldsymbol{x}, k)$ denote the optimal expected total cost over decision epochs $t, t+1, \ldots, T$, starting from state $(\boldsymbol{x}, k) \in \mathcal{S}$, and let the terminal cost, $V_T^N(\boldsymbol{x}, k)$, be equal to the function $C(\boldsymbol{x}) \triangleq \sum_{i \in \mathcal{N}} \mathbb{I}_i(\boldsymbol{x}) c_u^i$ for all k. This terminal cost function essentially assigns a corrective maintenance cost to failed components, while no costs are incurred for non-failed components. Then, by the principle of optimality, $V_t^N(\boldsymbol{x}, k)$ satisfies the following recursive Bellman optimality equations

$$V_t^N(\boldsymbol{x}, k) = \tag{4}$$
$$\min_{\boldsymbol{a} \in \mathcal{A}(\boldsymbol{x})} \left\{ C(\boldsymbol{x}, \boldsymbol{a}) + \mathbb{E}_{\boldsymbol{Z}} \left[V_{t+1}^N \Big(\boldsymbol{x}' + \boldsymbol{Z}, k + \sum_{i \in \mathcal{N}} Z_i \Big) \right] \right\},$$

where $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_N)$ is an N-dimensional random vector with $Z_i \sim NB\Big(\alpha_0 + k, \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\Big)$ (all Z_i's are independent and identically distributed), $\mathbb{E}_{\boldsymbol{Z}}$ denotes that the expectation is taken with respect to $\boldsymbol{Z}$, and $\boldsymbol{x}' = \big(x_1', x_2', \ldots, x_N'\big)$ with $x_i' = x_i$ if $a_i = 0$, and $x_i' = 0$ if $a_i = 1$.

We also refer to $V_t^N(\boldsymbol{x}, k)$ as the value function of the original MDP. The first part between the brackets is the direct costs

while the second part is the expected future costs of taking action $\boldsymbol{a}$ in state $(\boldsymbol{x}, k)$. Specifically, each component's deterioration accumulates further according to the posterior predictive distribution that corresponds to state $(\boldsymbol{x}, k)$, and k increases with the sum of all those increments. Systems that are maintained start with an as-good-as-new component, which is governed by the auxiliary vector $\boldsymbol{x}'$ which ensures that $x_i' = 0$ when $a_i = 1$. The formulation in (4) shows that the learning process about the unknown rate λ is pooled through the evolution of the common state variable k, while the future evolution of all individual deterioration processes depends on all pooled information and the parameter N. The existence of an optimal policy in this setting is guaranteed, see e.g., Proposition 3.4 of [3].

Observe that the minimum total expected cost for N systems over the complete lifespan of length T is given by $V_0^N(\mathbf{0}, 0)$ ($\mathbf{0}$ denotes the N-dimensional zero vector) which can be found by solving Equation (4) via backward induction. It is however clear from the formulation in (4), that as the number of systems grows, the problem will increasingly suffer from the curse of dimensionality: The cardinality of both the action and state space grow exponentially in N. Instead of solving (4) (referred to as the original MDP) directly, we will therefore construct an alternative MDP and show that the original MDP can be decomposed into N of these alternative MDPs: One for each system $i \in \mathcal{N}$. This decomposition is imperative as it allows us to (i) analyze the benefits of pooling of learning when N is relatively large without suffering from the curse of dimensionality, and (ii) establish structural properties of the optimal policy.

To this end, let $\tilde{V}_t^{N,i}(x, k)$ denote the optimal expected total cost for system $i \in \mathcal{N}$, over decision epochs $t, t+1, \ldots, T$, starting from state $(x, k) \in \mathbb{N}_0^2$, and let the terminal cost, $\tilde{V}_T^{N,i}(x, k)$, be equal to the function $C_i(x) \triangleq \mathbb{I}_i(x) c_u^i$ for all k. Then, $\tilde{V}_t^{N,i}(x, k)$ satisfies the following recursive Bellman optimality equations

$$\tilde{V}_t^{N,i}(x,k) = \min_{a \in \mathcal{A}(x)} \left\{ C_i(x,a) + \mathbb{E}_{(Z,K)} \left[\tilde{V}_{t+1}^{N,i}\Big(x \cdot (1-a) + Z, k + Z + K\Big) \right] \right\}, \tag{5}$$

where $Z \sim NB\Big(\alpha_0 + k, \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\Big)$, $K \sim NB\Big((N-1) \cdot (\alpha_0 + k), \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\Big)$, $\mathbb{E}_{(Z,K)}$ denotes that the expection is taken with respect to Z and K, and

$$C_i(x,a) \triangleq a\big(1 - \mathbb{I}_i(x)\big) c_p^i + \mathbb{I}_i(x) c_u^i. \tag{6}$$

The indicator functions and actions (spaces) are as defined before. It is noteworthy to mention that the formulation in (5) resembles a single component optimization problem in isolation, where the transition probabilities depend on N and k. The evolution of state variable k depends on the random deterioration increment of the component (Z) but it also accounts for the evolution of the other components through the random variable K. Below we present the decomposition result, which establishes that the value function of the original MDP is the sum of all N value functions of the alternative MDPs.

Theorem 1. *For each $t \in \{0, 1, \ldots, T\}$, we have $V_t^N(\boldsymbol{x}, k) = \sum_{i \in \mathcal{N}} \tilde{V}_t^{N,i}(x_i, k)$ for all $(\boldsymbol{x}, k) \in \mathcal{S}$.*

The decomposition in Theorem 1 reduces the computational burden of solving (4) significantly. It collapses the original, high-dimensional MDP into N 2-dimensional MDPs with a binary action space, each with their own cost structure and failure threshold, while still taking into account pooled learning across the N systems. The decomposition also eases the process of establishing structural properties on a system level, which is the topic of the next section.

Next to the trivial conditions that the action space and cost function should be decomposable and that actions should not influence the future evolution of the deterioration processes, there are two essential conditions required for our decomposition result to hold. As our proof relies crucially on these two general conditions, we believe they can guide future research on pooled learning in MDPs. Therefore, we discuss these conditions in the remark below.

Remark 1. The decomposition result can be applied to other high-dimensional Bayesian MDPs in which data is pooled to learn a common but unknown parameter. Specifically, the decomposition result necessitates two conditions related to the underlying conjugate pair of the MDP. First, the sufficient statistic in the conjugate pair for learning the parameter should be linear in the observations. This condition enables step (c) in the proof where we extract Z_i from the summation. Secondly, the resulting posterior predictive distribution should be closed under convolutions. This enables step (d) in the proof where we use the fact that the convolution of $N-1$ Negative Binomially distributed random variables is again Negative Binomially distributed. One other conjugate pair – next to the Gamma - Poisson pair used in this paper – that, for instance, satisfies these conditions and which is used very often in the OM literature is the Normal - Normal pair. This pair is generally adopted when the mean of a Normal distribution with known variance is unknown and needs to be learned.

4. Structural properties

In this section we establish structural properties of the alternative MDP, which then carry over to the original MDP through our decomposition result. We first derive monotonicity properties of $\tilde{V}_t^{N,i}(x, k)$, and then use these properties to establish the optimality of a control limit policy. We finally show that the control limit approaches the failure level as the pooled data increases. To that end, we first rewrite (5) into the conventional formulation for single component optimization problems. So, for $x = \xi$, $\tilde{V}_t^{N,i}(x,k) = c_u^i + \mathbb{E}_{(Z,K)}\left[\tilde{V}_{t+1}^{N,i}\Big(Z, k + Z + K\Big)\right]$ because failed components must be replaced correctively at cost c_u^i, and,

$$\tilde{V}_t^{N,i}(x,k) = \min\left\{ c_p^i + \mathbb{E}_{(Z,K)}\left[\tilde{V}_{t+1}^{N,i}\Big(Z, k + Z + K\Big)\right]; \mathbb{E}_{(Z,K)}\left[\tilde{V}_{t+1}^{N,i}\Big(x + Z, k + Z + K\Big)\right] \right\}, \tag{7}$$

when $x < \xi_i$, as we can then either perform a preventive replacement, which costs c_p^i, or leave the component in operation until the next decision epoch at no cost. The terminal costs are as introduced before. The next result establishes the monotonicity of $\tilde{V}_t^{N,i}(x, k)$ in x and k.

Proposition 1. *For each $t \in \{0, 1, \ldots, T\}$ and $i \in \mathcal{N}$, the value function $\tilde{V}_t^{N,i}(x, k)$ is (i) non-decreasing in x, and (ii) non-decreasing in k.*

Proposition 1 implies that (i) if a component is more deteriorated or (ii) when the total amount of deterioration increments is higher, we expect higher costs. This is intuitive: A higher level of deterioration increases the probability of a costly failure and/or the need for preventive replacement, while a higher total amount

of deterioration increments implies that all components are deteriorating relatively fast (i.e. λ is larger). Using, Proposition 1, we may also conclude that $V_t^N(\boldsymbol{x}, k)$ is non-decreasing in the standard component-wise order in $\boldsymbol{x}$, and non-decreasing in k. The former means that for any deterioration vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$ such that $x_i \leq x_i'$ for all $i \in \mathcal{N}$, we have that $V_t^N(\boldsymbol{x}, k) \leq V_t^N(\boldsymbol{x}', k)$. The intuition behind this is similar to the intuition behind Proposition 1.

The next result establishes the optimality of a state-dependent control limit policy for the alternative MDP.

Proposition 2. *For each $t \in \{0, 1, \ldots, T-1\}$, $k \in \mathbb{N}_0$, and $i \in \mathcal{N}$, there exists a control limit $\delta_i^{(k,t)}$, $0 < \delta_i^{(k,t)} \leq \xi_i$, such that the optimal action is to carry out a replacement if and only if $x \geq \delta_i^{(k,t)}$.*

Proposition 2 shows that the control limit at each time of each component, $\delta_i^{(k,t)}$, depends in real-time on the shared learning process across all components through pooling data via the state variable k. The optimality of a control limit policy itself is not only intuitive and convenient for the implementation of this optimal policy in practice, it can also be exploited to further decrease the computational burden of solving the original MDP. That is, existing algorithms that rely on these structural properties such as the modified policy iteration algorithm [see 22, Section 6.5] can be used to efficiently solve the alternative MDP, and hence the original MDP.

Conceivably, one would expect that as we learn from the pooled data that λ is larger (through a higher k) and everything else is fixed, we would impose a lower control limit per component, i.e. $\delta_i^{(k,t)}$ is non-increasing in k. The intuition behind this is that because it is more likely that deterioration increments will take on higher values (see Lemma 1), we should replace a component earlier. Although such a non-increasing $\delta_i^{(k,t)}$ would indeed be intuitive, we found numerically that this is in general not true. Specifically, we found that the control limit usually decreases in k first, as expected, but that it always increases eventually to ξ_i as k grows large, in which case it is never optimal to do preventive maintenance (see Appendix A for an illustration of this behavior). This limiting behavior, which breaks the monotonic behavior of $\delta_i^{(k,t)}$ in k, is formalized in the proposition below.

Proposition 3. *For each $t \in \{0, 1, \ldots, T-1\}$ and $i \in \mathcal{N}$, we have $\lim_{k\to\infty} \delta_i^{(k,t)} = \xi_i$.*

While Proposition 3 may initially appear counter-intuitive, it can be heuristically explained as follows. When k grows very large and everything else is kept fixed, the random deterioration per period grows so large that any component will fail with certainty solely due to the one-period deterioration. In this case, if a component is still working at a certain decision epoch and has deterioration level x ($< \xi_i$), performing preventive maintenance will induce an extra cost of c_p^i because in the next period, the component will fail anyway, regardless of the value of x. To make this argument more explicit, consider a component with deterioration level $x < \xi_i$ at decision epoch $T-1$. As we only have the terminal cost at time T, (7) gives us for the optimality equation:

$$\tilde{V}_{T-1}^{N,i}(x, k) = \min\Bigg\{ \underbrace{c_p^i + \mathbb{P}\big[Z \geq \xi_i\big] \cdot c_u^i + \big(1 - \mathbb{P}\big[Z \geq \xi_i\big]\big) \cdot c_p^i}_{\text{preventive maintenance}};$$

$$\underbrace{\mathbb{P}\big[Z \geq \xi_i - x\big] \cdot c_u^i + \big(1 - \mathbb{P}\big[Z \geq \xi_i - x\big]\big) \cdot c_p^i}_{\text{leave in operation}} \Bigg\}. \tag{8}$$

It is clear that preventive maintenance in this state is not optimal if and only if the following holds:

$$c_p^i + \mathbb{P}\big[Z \geq \xi_i\big] \cdot c_u^i + \big(1 - \mathbb{P}\big[Z \geq \xi_i\big]\big) \cdot c_p^i > \mathbb{P}\big[Z \geq \xi_i - x\big] \cdot c_u^i + \big(1 - \mathbb{P}\big[Z \geq \xi_i - x\big]\big) \cdot c_p^i. \tag{9}$$

Since the random variable Z is increasing in k (see Lemma 1), one can show that $\mathbb{P}\big[Z \geq \xi_i\big] \to 1$ and $\mathbb{P}\big[Z \geq \xi_i - x\big] \to 1$ for each $x < \xi_i$ as $k \to \infty$. This implies, using (8), that at decision epoch $T-1$ as $k \to \infty$, leaving the component in operation costs c_u^i, while performing preventive maintenance costs $c_p^i + c_u^i$ (an extra cost of c_p^i). Since $c_p^i > 0$, Equation (9) will always hold for any $x < \xi_i$ at decision epoch $T-1$ as $k \to \infty$. In the proof of Proposition 3, we formalize this heuristic argument and do so for each decision epoch.

5. Numerical study

This section reports the results of a comprehensive numerical study in which we assess the benefits of pooling data to jointly learn an unknown parameter. Although the results in the previous sections hold for asymmetric – in terms of costs and failure thresholds – systems, we focus on symmetric systems in this numerical study. By doing so, the value function $\tilde{V}_0^N(0,0)$ (we drop the index i as we consider symmetric systems) gives us the cost per system over its lifespan when the data of N systems is pooled. We can use this cost per system to assess the value of pooling learning as a function of N compared to not pooling. To this end, we define the performance measure $\Delta = 100\left[1 - \frac{\tilde{V}_0^N(0,0)}{\tilde{V}_0^1(0,0)}\right]$. These are the relative savings per system over the lifespan when N systems are jointly learning compared to not pooling any data for those systems and learning the unknown rate independently from the other systems.

We first perform an extensive numerical study. Recall that the initial parameter uncertainty is modeled by the random variable Λ which has a Gamma distribution with shape α_0 and rate β_0. By fixing the mean of Λ and subsequently varying its coefficient of variation, we can thus increase or decrease the initial parameter uncertainty. We do so by solving the following set of two equations for the two unknowns α_0 and β_0: $\mathbb{E}[\Lambda] = \frac{\alpha_0}{\beta_0}$, and $cv_\Lambda = \frac{1}{\sqrt{\alpha_0}}$, where cv_Λ is the coefficient of variation of Λ. This allows us to explicitly study the impact of the uncertainty (in terms of its mean and coefficient of variation) on the pooling effects. Our testbed consists of 2268 instances. These are obtained by permuting all parameter values in the second column of Table 1, with the corrective maintenance cost c_u held fixed at 10. These values are representative for the capital goods industry and are in line with the maintenance literature (see, e.g., [23] on typical maintenance costs, and [8] on initial parameter uncertainty). For each instance of the test bed we compute the relative savings $\%\Delta$. The results of the numerical study are summarized in the remaining columns of Table 1. In this table, we present the average and maximum relative savings $\%\Delta$. For each value of N, we first present the average relative savings for subsets of instances with the same value for a given input parameter (row wise), and then present the average results for all instances with that fixed value of N (bottom row), where each average value is accompanied with the maximum value in brackets.

Based on the results in Table 1, we can state the following main observations:

Table 1
Relative savings (%Δ) due to pooled learning.

		N					
Input	Value	2	4	6	8	10	20
ξ	7	2.7 (16.4)	9.8 (37.2)	14.8 (59.2)	18.0 (71.2)	19.7 (77.2)	24.5 (88.5)
	10	3.4 (19.7)	9.9 (36.8)	14.5 (59.4)	17.1 (71.8)	19.1 (79.7)	22.7 (89.2)
T	50	3.0 (19.7)	9.7 (36.3)	14.5 (57.3)	17.4 (70.2)	19.3 (78.4)	23.6 (88.4)
	70	3.1 (19.7)	9.8 (36.6)	14.7 (58.6)	17.6 (71.2)	19.4 (79.2)	23.6 (88.9)
	90	3.1 (19.7)	10.0 (37.2)	14.8 (59.4)	17.6 (71.8)	19.5 (79.7)	23.6 (89.2)
c_p	0.5	4.4 (19.7)	12.7 (37.2)	18.3 (59.4)	21.6 (71.8)	23.7 (79.7)	28.0 (89.2)
	1	2.9 (12.7)	9.5 (33.5)	14.2 (53.9)	17.1 (65.1)	18.9 (72.5)	22.9 (82.8)
	1.5	1.9 (8.7)	7.4 (29.9)	11.5 (48.9)	14.0 (59.4)	15.6 (65.8)	19.9 (77.6)
$\mathbb{E}[\Lambda]$	0.5	3.0 (19.7)	8.3 (33.5)	12.1 (47.5)	14.6 (60.1)	16.2 (69.1)	19.6 (83.6)
	0.75	3.0 (18.7)	10.0 (36.8)	14.9 (54.6)	17.8 (67.6)	19.7 (76.3)	23.8 (87.5)
	1	3.1 (14.3)	11.2 (37.2)	16.9 (59.4)	20.3 (71.8)	22.4 (79.7)	27.3 (89.2)
cv_Λ	0.1	0.0 (0.1)	0.0 (0.2)	0.1 (0.2)	0.1 (0.3)	0.1 (0.3)	0.2 (0.4)
	0.25	0.2 (0.5)	0.4 (0.8)	0.5 (1.0)	0.5 (1.1)	0.6 (1.2)	0.8 (1.4)
	0.5	0.6 (1.3)	1.3 (2.7)	1.8 (3.5)	2.1 (4.1)	2.4 (5.3)	3.1 (6.4)
	1	4.1 (12.0)	7.8 (22.5)	9.4 (26.3)	10.4 (28.3)	11.0 (31.1)	12.4 (35.7)
	2	10.3 (19.7)	22.7 (36.8)	29.5 (44.8)	33.8 (51.5)	36.7 (57.3)	46.0 (73.9)
	4	13.0 (24.6)	27.2 (37.2)	35.3 (59.4)	40.8 (71.8)	44.9 (79.7)	56.9 (89.2)
Total		3.6 (24.6)	9.8 (37.2)	14.0 (59.4)	16.5 (71.8)	18.2 (79.7)	22.3 (89.2)

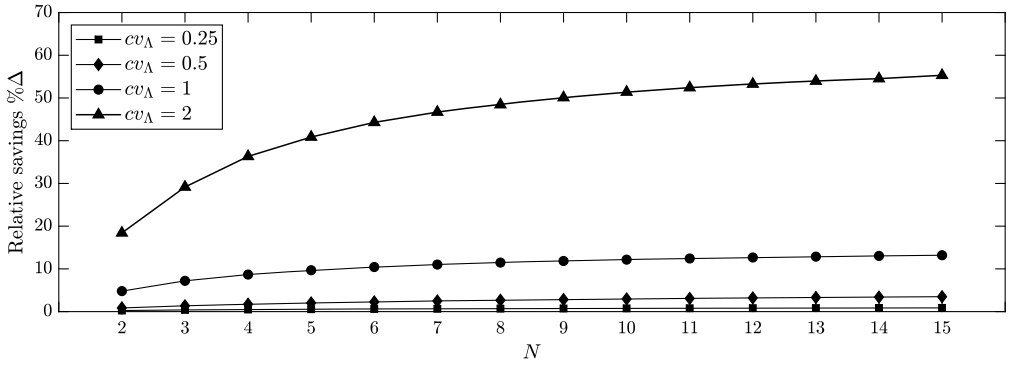

Fig. 1. Relative savings (%Δ) as function of N for various values of cv_Λ for $\mathbb{E}[\Lambda] = 0.75$, $\xi = 10$, $c_p = 0.5$, $c_u = 10$, and $T = 90$.

1. Pooling of data for learning a common unknown parameter can lead to significant savings compared to not pooling data and learning it independently.
2. The magnitude of the savings seems to be inextricably linked with the magnitude of uncertainty in the parameter λ measured by the coefficient of variation of Λ. When cv_Λ is high, savings of up to 56.9% on average (over all instances with $cv_\Lambda = 4$ and $N = 20$) can be achieved, while if cv_Λ is low, savings become almost negligible ($\leq 0.2\%$ on average). This can be explained as follows. When there is high uncertainty in the unknown parameter, pooling data allows the maintenance service provider to faster learn the unknown parameter compared to learning it from data generated by a single system. This result implies that pooling data is especially beneficial for real-life settings where there is high uncertainty in λ through limited knowledge, limited historical data, and/or poor estimation procedures. The opposite is also true. When there is little uncertainty in the unknown parameter, the benefit of data pooling vanishes; a maintenance service provider already has an accurate belief of the unknown parameter that needs little updating.
3. When comparing the average savings for increasing values of N, we find that pooling has already a significant impact for small values of N, and that the marginal savings gradually decrease when N increases.
4. The savings for each value of N tend to decrease as the ratio c_u/c_p decreases (recall that we keep c_u fixed and vary c_p). When this ratio decreases and N is fixed, maintenance decisions have less impact on the resulting costs – simply because their cost difference decreases. Consequently, the benefits of utilizing pooled learning in such maintenance decisions also decrease when c_u/c_p decreases.
5. The savings for each value of N tend to increase as $\mathbb{E}[\Lambda]$ increases. When $\mathbb{E}[\Lambda]$ increases and N is fixed, the expected deterioration increment between two consecutive decision epochs is larger and, as a result, the optimal control limit will be more conservative. The results suggest that in that regime, the choice of the control limit has a higher impact on the resulting costs than when $\mathbb{E}[\Lambda]$ is low and a less conservative control limit is chosen. By pooled learning, one is able to better choose this control limit, and as a result, the relative savings of pooled learning also increase when $\mathbb{E}[\Lambda]$ increases.

Observations 1-3 are also illustrated by Fig. 1. In this figure we plot the relative savings (%Δ) as a function of N for various values of cv_Λ when $\mathbb{E}[\Lambda] = 0.75$, $\xi = 10$, $c_p = 0.5$, $c_u = 10$, and $T = 90$. The plot indeed shows that for a given level of parameter uncertainty, pooling data across a larger number of systems increases the relative savings. The rate at which the savings increase in N increases significantly in the coefficient of variation. This confirms that pooling data can lead to significant cost reductions, especially when the uncertainty surrounding an unknown parameter is high. We further clearly see that the marginal savings due to adding an extra system to the pooled systems decreases as N increases.

6. An application to spare parts inventory systems

We conclude this paper by illustrating that our decomposition result can also be applied to spare parts inventory systems. We first redefine some notation introduced in the previous sections. Again, we consider a maintenance service provider that operates a set of $N \geq 1$ local spare parts warehouses, and we denote this set by $\mathcal{N} = \{1, \ldots, N\}$. Each local warehouse stocks spare parts of the same critical component to serve an installed base of technical systems. As is common in the spare parts inventory literature [e.g. 9], we model demand for spare parts at each local warehouse as an independent Poisson process. The rate of these Poisson processes λ is identical across all local warehouses. This is a reasonable assumption when the installed bases served by each local warehouse are of similar size.

The maintenance service provider is concerned with stocking decisions over a finite horizon of T periods. At the start of each such period, the maintenance service provider decides how many new spare parts are transported to each local warehouse $i \in \mathcal{N}$. Each unit has a transportation cost c_v^i. Since the lead times to the local warehouses are typically much shorter than the duration of a period, we assume that these new spare parts are instantly delivered, after which the period commences. When a component in a technical system fails during the period, local warehouse $i \in \mathcal{N}$ responsible for this system immediately replaces the failed component with a read-for-new one, if it has one available. Otherwise, the part is backordered at unit cost c_b^i, which reflects expensive downtimes or emergency shipments from a central depot or an external supplier. Spare parts on stock that are carried over to the next period cost c_h^i per unit. We account for both backorder and holding costs at the end of each period. We assume that each period lasts 1 time unit so that demand in each period is Poisson distributed with mean λ. We employ the Bayesian approach of Section 2 to infer the a-priori unknown rate λ based on the observed demands at all local warehouses over the entire planning horizon. Observe that in the updating scheme of this approach (cf. Equation (1)), k is now defined as the total cumulative demand at all N local warehouses up to period t. Given k and t, the posterior predictive Z_i now represents the total demands that arrive at local warehouse $i \in \mathcal{N}$ during the next decision epoch.

The state space of the Bayesian MDP corresponding to the decision problem described above is given by $\mathcal{S} \triangleq \mathbb{Z}^N \times \mathbb{N}_0$. For a given state $(\boldsymbol{x}, k) \in \mathcal{S}$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ represents the vector of net inventory levels of all local warehouses before order placement at the start of a period, and k denotes the sum of all observed demands until that period. For a given state $(\boldsymbol{x}, k) \in \mathcal{S}$, the action space $\mathcal{A}(\boldsymbol{x})$ contains all possible net inventory levels after orders are placed and received but before demand is realized, i.e. for any action $\boldsymbol{a} = (a_1, a_2, \ldots, a_N) \in \mathcal{A}(\boldsymbol{x})$, $a_i \in \{x_i, x_i + 1, \ldots\}$ is the net inventory level per local warehouse. As before, we let $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_N)$ denote an N-dimensional random vector with $Z_i \sim NB\left(\alpha_0 + k, \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\right)$. As is customary in inventory theory, the direct cost in a given period accounts for the expected holding and backorder costs of the orders placed in that period. As such, the total transportation, holding, and backorder costs over all local spare parts warehouses is given by $C(\boldsymbol{a}, \boldsymbol{x}, k) \triangleq \sum_{i \in \mathcal{N}} \left(c_v^i (a_i - x_i) + c_h^i \mathbb{E}\left[(a_i - Z_i)^+\right] + c_b^i \mathbb{E}\left[(Z_i - a_i)^+\right]\right)$ with $x^+ \triangleq \max(x, 0)$. While the direct cost now depends on the pooled data through state variable k, we note that it remains decomposable in N direct costs $C_i(a, x, k) \triangleq c_v^i (a_i - x_i) + c_h^i \mathbb{E}[(a_i - Z_i)^+] + c_b^i \mathbb{E}[(Z_i - a_i)^+]$, each associated with a local warehouse $i \in \mathcal{N}$. Let $V_t^N(\boldsymbol{x}, k)$ denote the optimal expected total cost over decision epochs $t, t+1, \ldots, T$, starting from state $(x, k) \in \mathcal{S}$, and let $V_T^N(\cdot, \cdot) \triangleq 0$. By the principle of optimality, $V_t^N(\boldsymbol{x}, k)$ satisfies the recursive Bellman optimality equations $V_t^N(\boldsymbol{x}, k) = \min_{\boldsymbol{a} \in \mathcal{A}(\boldsymbol{x})} \left\{ C(\boldsymbol{a}, \boldsymbol{x}, k) + \mathbb{E}_{\boldsymbol{Z}}\left[V_{t+1}^N\left(\boldsymbol{a} - \boldsymbol{Z}, k + \sum_{i \in \mathcal{N}} Z_i\right)\right]\right\}$. We now formulate the corresponding alternative MDP in which the original MDP can be decomposed. For each $i \in \mathcal{N}$, we let $\tilde{V}_t^{N,i}(x, k)$ denote the optimal expected total cost over decision epochs $t, t+1, \ldots, T$, starting from state $(x, k) \in \mathbb{Z} \times \mathbb{N}_0$. Then, $\tilde{V}_t^{N,i}(x, k)$ satisfies the following recursive Bellman optimality equations

$$\tilde{V}_t^{N,i}(x, k) = \min_{a \in \mathcal{A}(x)} \left\{ C_i(a, x, k) + \mathbb{E}_{(Z,K)}\left[\tilde{V}_{t+1}^{N,i}\left(a - Z, k + Z + K\right)\right]\right\}, \tag{10}$$

where $Z \sim NB\left(\alpha_0 + k, \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\right)$, $K \sim NB\left((N-1) \cdot (\alpha_0 + k), \frac{\beta_0 + N \cdot t}{\beta_0 + N \cdot t + 1}\right)$, and $\tilde{V}_T^{N,i}(\cdot, \cdot) \triangleq 0$. Observe that the alternative formulation in (10) resembles a single spare parts warehouse problem in isolation, but where the dynamics of the system depend on the learned information of all warehouses through k.

In the result below, we present our decomposition result applied to the spare parts inventory system setting. Its proof is almost verbatim the proof of Theorem 1 and therefore omitted.

Theorem 2. *For each $t \in \{0, 1, \ldots, T\}$, we have: $V_t^N(\boldsymbol{x}, k) = \sum_{i \in \mathcal{N}} \tilde{V}_t^{N,i}(x_i, k)$ for all $(\boldsymbol{x}, k) \in \mathcal{S}$.*

As before, the above decomposition result motivates us to establish structural properties of the alternative 2-dimensional MDP, which then carry over to the original, high-dimensional MDP. To that end, we first establish convexity of the value function in the inventory level before order placement.

Proposition 4. *For each $t \in \{0, 1, \ldots, T\}$, $k \in \mathbb{N}_0$, and $i \in \mathcal{N}$, the value function $\tilde{V}_t^{N,i}(x, k)$ is convex in x.*

The above result also implies that the optimal policy of the decomposed MDP is characterized by an order-up-to structure, in which we place orders such that the inventory level after ordering reaches a certain target level (if needed). Our next result formalizes the optimality of order-up-to levels, together with their non-decreasing monotonic behavior in the state variable k.

Proposition 5. *For each $t \in \{0, 1, \ldots, T-1\}$, $k \in \mathbb{N}_0$, and $i \in \mathcal{N}$, there exists a single target level $\delta_i^{(k,t)} \in \mathbb{Z}$ such that the optimal action is $a_i = \delta_i^{(k,t)}$ if $x < \delta_i^{(k,t)}$ and $a_i = x$ otherwise. The optimal target level $\delta_i^{(k,t)}$ is non-decreasing in k.*

Proposition 5 shows that the optimal target levels depend in real-time on the shared learning process across all local warehouses via the state variable k in a monotonic way. This is intuitive: As we learn from the pooled data that λ is higher (through a higher k) and everything else fixed, the next demand will likely take on higher values. Therefore, we should increase the target level to which we raise our spare part inventories. This monotonicity result stands in contrast to the limiting result of the optimal control limits in the CBM setting, as described in Proposition 3. The key difference is that, unlike in the CBM setting, there is a cost incentive that is proportional to demand realizations. This cost incentive remains proportional, even if the demand becomes very large because of a very large k, so that ordering quantities are monotonically non-decreasing in k. We conclude by noting that

similar monotonicity results for Bayesian inventory systems exist in the literature, but only for single inventory systems in isolation and without any data pooling considerations [e.g. 15].

Acknowledgements

The authors are grateful to the editorial team whose comments greatly improved the paper.

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.orl.2023.11.009.

References

[1] K.-P. Aronis, I. Magou, R. Dekker, G. Tagaras, Inventory control of spare parts using a Bayesian approach: a case study, Eur. J. Oper. Res. 154 (3) (2004) 730–739.
[2] H. Bastani, D. Simchi-Levi, R. Zhu, Meta dynamic pricing: transfer learning across experiments, Manag. Sci. 68 (3) (2022) 1865–1881.
[3] D.P. Bertsekas, S.E. Shreve, Stochastic Optimal Control: The Discrete Time Case, Academic Press, 1978.
[4] B. de Jonge, P.A. Scarf, A review on maintenance optimization, Eur. J. Oper. Res. 285 (3) (2020) 805–824.
[5] L. Deprez, K. Antonio, J.J. Arts, R.N. Boute, Data-driven preventive maintenance for a heterogeneous machine portfolio, Oper. Res. Lett. 51 (2) (2023) 163–170.
[6] C. Drent, S. Kapodistria, O. Boxma, Censored lifetime learning: optimal Bayesian age-replacement policies, Oper. Res. Lett. 48 (6) (2020) 827–834.
[7] C. Drent, M. Drent, J.J. Arts, Condition-based production for stochastically deteriorating systems: optimal policies and learning, arXiv:2308.07507, 2023.
[8] C. Drent, M. Drent, J.J. Arts, S. Kapodistria, Real-time integrated learning and decision making for cumulative shock degradation, Manuf. Serv. Oper. Manag. 25 (1) (2023) 235–253.
[9] M. Drent, J.J. Arts, Expediting in two-echelon spare parts inventory systems, Manuf. Serv. Oper. Manag. 23 (6) (2021) 1431–1448.
[10] İ. Dursun, A. Akçay, G.-J. van Houtum, Data pooling for multiple single-component systems under population heterogeneity, Int. J. Prod. Econ. 250 (2022) 108665.
[11] A.H. Elwany, N.Z. Gebraeel, L.M. Maillart, Structured replacement policies for components with complex degradation processes and dedicated sensors, Oper. Res. 59 (3) (2011) 684–695.
[12] G.D. Eppen, Effects of centralization on expected costs in a multi-location newsboy problem, Manag. Sci. 25 (5) (1979) 498–501.
[13] A. Gelman, J. Carlin, H. Stern, D. Rubin, Bayesian Data Analysis, Chapman and Hall, 1995.
[14] V. Gupta, N. Kallus, Data pooling in stochastic optimization, Manag. Sci. 68 (3) (2022) 1595–1615.
[15] D.L. Iglehart, The dynamic inventory problem with unknown demand distribution, Manag. Sci. 10 (3) (1964) 429–440.
[16] M. Kurt, J.P. Kharoufeh, Monotone optimal replacement policies for a Markovian deteriorating system in a controllable environment, Oper. Res. Lett. 38 (4) (2010) 273–279.
[17] D.P.T. Lamghari-Idrissi, A new after-sales service measure for stable customer operations, PhD thesis, Eindhoven University of Technology, 2021.
[18] C. Li, B. Tomlin, After-sales service contracting: condition monitoring and data ownership, Manuf. Serv. Oper. Manag. 24 (3) (2022) 1494–1510.
[19] A. Mandelbaum, M.I. Reiman, On pooling in queueing networks, Manag. Sci. 44 (7) (1998) 971–981.
[20] T. Olsen, B. Tomlin, Industry 4.0: opportunities and challenges for operations management, Manuf. Serv. Oper. Manag. 22 (1) (2020) 113–122.
[21] K.B. Öner, A. Scheller-Wolf, G.-J. Van Houtum, Redundancy optimization for critical components in high-availability technical systems, Oper. Res. 61 (1) (2013) 244–264.
[22] M. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, John Wiley & Sons, 2005.
[23] C. Van Oosterom, H. Peng, G.-J. Van Houtum, Maintenance optimization for a Markovian deteriorating system with population heterogeneity, IISE Trans. 49 (1) (2017) 96–109.