Vlerick Repository

The effect of multi-sensor data on condition-based maintenance policies

Authors	van Staden, Heletjé E.;Boute, Robert
DOI	10.1016/j.ejor.2020.08.035
Publisher	Elsevier
Journal	European Journal of Operational Research
Download date	2025-04-23 10:32:29
Item License	https://creativecommons.org/licenses/by-nc-nd/4.0/
Link to Item	http://hdl.handle.net/20.500.12127/6529

The effect of multi-sensor data on condition-based maintenance policies

Heletjé E. van Staden[a,*], Robert N. Boute[a,b]

[a]KU Leuven, Faculty of Economics and Business, Naamsestraat 69, Box 3555, Leuven, 3000, Belgium
[b]Vlerick Business School, Technology and Operations Management, Vlamingenstraat 83, Leuven, 3000, Belgium
heletje.vanstaden@kuleuven.be, robert.boute@vlerick.com

Abstract

Industry 4.0 promises reductions in maintenance costs through access to digital technologies such as the Internet of Things, cloud computing and data analytics. Many of the promised benefits to maintenance are, however, dependent on the quality of the data obtained through sensors and related technologies. In this work, we consider the effect of access to different levels of deterioration data quality, resulting in partial information about the underlying state of the system being monitored, by means of sensors, on condition-based maintenance policies. The sensors may be either internal company sensors, or more informative external sensors of which access is obtained at a cost. We analyze the structure of the optimal policy, where the actions are either to perform maintenance, to pay for external sensor information or to continue system operation with internal sensor information only. We show that the optimal policy consists of at most four regions based on the believed deterioration state of the system. The analysis allows us to numerically investigate the decision maker's willingness to pay for more informative external sensor information with respect to the level of external sensor informativeness, when compared to that of the internal sensor, and the cost thereof.

Keywords: maintenance, multi-sensor data, condition-based maintenance, partially observable Markov decision process

1. Introduction

Industry 4.0 and the Internet of Things bring with it increasing data and data acquiring opportunities. Access to system data by means of sensors and other Industry 4.0 technologies allow for condition-based maintenance (CBM). CBM is a maintenance policy whereby systems are preventively maintained according to their perceived condition. The complete true deterioration state of the system being monitored, however, may not always be inferred from sensor observations due to factors such as outside interference, inherent limitations in sensor quality and misspecification in the system deterioration model. Information obtained from sensors therefore typically only provide the decision maker with a partial representation of the real state of the system. To overcome incomplete information, consulting firms such as PwC (Haarman et al., 2018) and McKinsey (Manyika et al., 2015) report that, for CBM to deliver its maximum value, it should

[*]Corresponding author
Email address: heletje.vanstaden@kuleuven.be (Heletjé E. van Staden)

take a maximum of external sources into account. While increasing the number of data sources on which the state prediction of a system is based may improve the prediction quality of an imminent system failure, the increased information quality may still be stochastically related to the true state of the system and may not necessarily justify the cost involved in acquiring and processing such information.

Saghafian et al. (2018) investigated costly information sharing between autonomous sensors to improve sensor prediction quality by analyzing a general information-quality sensor communication framework. Their work inspired us to understand the effect of such a setup in a CBM environment. In this work, we consider a system whose deterioration is continuously measured by means of internal system sensors. Given the Internet of Things, a number of other, external sensors may be connected and operating in the system environment. These sensors may measure, for instance, the system's production quality or the deterioration of equipment connected to the system. Observations from both the internal and external sensors are stochastically related to the exact deterioration state of the system, such that the true state is unknown to the decision maker. We refer to the system as a partially observable system. The sensor observations are used to predict the unknown deterioration state of the system. Observations from the external sensors may offer additional insights to that of the internal sensors and, when combined, result in a more accurate system deterioration state prediction. Access to external sensor data comes at a cost, however. Such costs may include acquisition costs if sensors are owned by external entities or the setup cost of processing the additional sensor data. Depending on the cost of the external sensor information and the difference in prediction accuracy, it may be cost optimal to only acquire such costly external sensor information once the system has reached a particular believed deterioration state. Delayed payment for external sensor information takes advantage of the internal sensor at a time when the system is believed to be in a healthy state and not at a high risk of failure. The problem may also be seen as one where the external information is obtained by means of a technician dispatched to perform manual inspections on the system. While a technician visit may provide a more informative view on the state of the system than the internal sensor, such information may still only partially relate to the true state of the system due to inherent human errors. Technician visits are costly. Considering the prediction quality of the internal sensor against technician and related costs, it may be cost optimal to compromise on prediction accuracy to save on technician fees when the system is predicted to be in a good working condition.

With the objective to minimize the average long-run cost for the system being monitored, we determine policies advising when it is cost optimal to obtain more informative yet costly external sensor information. For ease of exposition, our definition of external sensor information may be seen as the system state prediction indicator after combining the costless continuous internal sensor information with the external information. Our use of the term external sensor information therefore includes the combination of both internal and external sensor information, such that external sensor information is always at least as informative as internal information. Similarly, external sensor information may encompass a set of different external sensors, whose separate observations are combined into a single, more informative system state prediction indicator, with costs adjusted accordingly. In what follows, we will focus on the effect and value of such external sensor information on CBM policies. We refer the interested reader to Mitchell (2007) and Khaleghi et al. (2013) for methods concerning sensor information fusion.

We adopt a partially observable problem setting for a single component that transitions between two unobservable working states, a healthy and a warning state. The decision maker bases decisions on observations from the sensors that provide a probability that the system may be in one of the two working states at any given moment. At any moment, a decision maker may decide to either perform maintenance, to acquire external sensor information or to continue normal system operation with internal sensor information only. For the case where the decision maker may choose between costless internal sensor information and costly, but more informative, external sensor information, we show that the optimal policy consists of at most four regions. The obtained policies result in easily interpretable control rules for choosing between sensor information and maintenance decision making. We present numerical analyses and visualizations of the implications of different levels of informativeness on the optimal maintenance decisions. The analysis allows for the determination of a decision maker's willingness to pay for external sensor information.

We finally provide an extension of the problem setting to allow for multiple external sensor options. The extension demonstrates the cost improvements that may be achieved by policies resulting from such complex problems. Nonetheless, there exist settings for which the cost improvements may be minor and may not outweigh the increased complexity of the optimal policy.

2. Literature review

Some of the earliest work on informativeness in the partial information setting dates back to Monahan (1980, 1982) and Lovejoy (1987). Monahan (1980) investigated the expected value of information for varying levels of informativeness in a partial information environment. Monahan (1982) found that, with increasing informativeness, the optimal value function for his model based on partial information is non-decreasing and convex. Lovejoy (1987) obtained monotonicity results for informativeness in a partial information environment, providing insights into when one observation set is more informative than another. The informativeness of the sensor observations are further motivated by Lévesque & Maillart (2008) and Van Oosterom et al. (2017). In their work, Lévesque & Maillart (2008) investigated the effect of costly, partial information observations (not by means of sensors) of varying informativeness on the assessment of business opportunities, such as diamond mining. The authors showed that the optimal policy consists of a pair of threshold values governing the acceptance or rejection of such an opportunity. Decreasing sensor informativeness over time was investigated by Van Oosterom et al. (2017). The authors jointly optimized and analyzed the maintenance of a system and its deteriorating sensor. In a multi-sensor environment where the sensors' informativeness is known and unique, but the underlying inference models for the sensors may differ, Saghafian et al. (2018) found that each sensor will eventually acquire information from a constant sensor set. The authors showed that the set consists of a single unique sensor when sensor qualities are asymmetric.

While our work touches upon the value of information in maintenance decision making, a wealth of related literature exists in fields other than operations management, most notably in economics of information. Considering a payoff maximising decision maker facing uncertainty given access to observations from a defined information structure, Radner & Stiglitz (1984), Chade & Schlee (2002) and De Lara & Gilotte (2007) determine conditions for the non-concavity in the value of information and their effect on informativeness and payoff. Weibull et al. (2007) provide non-monotonicity results in which it may not always be optimal for a decision maker to choose observations from the most informative option in a given set.

De Jonge & Scarf (2020) recently conducted a comprehensive review on maintenance optimization, providing a classification framework for both single- and multi-component systems. Typically, maintenance optimization seeks to find optimal threshold policies so as to allow for simplified maintenance decisions. In their work, Poppe et al. (2018) numerically optimized a CBM policy of multi-components by using a two threshold structure. Due to the sequential decision making nature of the maintenance problem, dynamic programming or Markov decision processes (MDPs) are commonly used for maintenance optimization (see for instance Olde Keizer et al. (2016)). When the true state of the system being monitored is only stochastically related to the monitoring process, a partially observable Markov decision process (POMDP) formulation is applied. In contrast to an MDP, where the system state is known with certainty, a POMDP depends on the belief that the system is in a particular deterioration state. We refer the interested reader to the book by Powell & Ryzhov (2012) for a comprehensive discussion on the different decision making methodologies in the presence of partial information. Considering POMDPs, the work closest to ours is the seminal work by Ross (1971). The author showed that the optimal POMDP maintenance policy for a system with partial information updating and perfect information from inspections consists of at most four regions. In the results, the optimal maintenance and inspection regions are separated by regions where it is cost-optimal for the decision maker to continue system operations without interference. Ohnishi et al. (1986) and Kim & Makis (2013) similarly determined that the optimal information sampling and maintenance optimization problem consists of at most four regions. Ohnishi et al. (1986) consider the problem where a decision maker can choose between three actions, which are either to continue system operation with incomplete information, to perform a perfect inspection at a cost, or to replace the system. In the work of Kim & Makis (2013), the decision maker can choose between either imperfect system state sampling, perfect system inspection and maintenance if necessary or continue system operations with no new system state information. Our work differs to that of Ohnishi et al. (1986) in that we have that external sensor information improves on the informativeness of the internal sensor, but which is still (typically) imperfect. Our work differs to that of Kim & Makis (2013) in that they only receive system state information when the action is to sample (imperfect information) or to inspect the system (perfect information), whereas we have constant system updating by means of the internal sensor.

A POMDP may be translated into a fully observable MDP by translating the belief states of the POMDP into the actual states of the MDP. Advantage may then be taken of existing MDP solution algorithms. Upon a POMDP translation, the MDP is no longer partially observable as the belief state is known at any decision epoch. The resulting state space of the belief MDP will be continuous, however, since the POMDP beliefs are defined on the interval $[0, 1]$. This continuous state space increases the complexity of solving the MDP. Early work in translating partially observable problems into MDPs are Monahan (1980, 1982). The author modeled a POMDP as an optimal stopping problem, which may be interpreted as determining the optimal threshold at which to stop the system to perform maintenance. Monahan highlighted that the main difference between the maintenance problem and the stopping problem is its time horizon, given positive costs. While the stopping problem will end in a finite length of time, the maintenance problem typically has an infinite horizon. The model of Lévesque & Maillart (2008) is similarly an application of optimal stopping in a POMDP. Makis & Jiang (2003) transformed their POMDP into an MDP by means of a smooth semi-martingale process. They proved optimality of the maintenance policy and used an algorithm based on λ-minimization (Aven & Bergman, 1986) and the convergence properties of the value functions to solve the problem. Kim & Makis (2013) and Maillart et al. (2018) used a renewal process to model the maintenance optimization problem as an optimal stopping problem. As in our work, the authors translated their POMDP into an MDP by means of the λ-minimization technique (Aven & Bergman, 1986) and showed that the optimal policy over the infinite horizon consists of at most four and at most three regions, respectively. Our modeling approach is inspired by that of Kim & Makis (2013), Naderkhani & Makis (2015) and Maillart et al. (2018).

3. Problem description and formulation

We aim to find a set of decision rules governing the actions taken by the decision maker, referred to as a policy, that minimizes the expected long-run average maintenance cost per running hour. System deterioration may usually be separated into two distinct phases: a longer, gradual degradation phase and a shorter degradation phase with a clear non-negligible increase in the failure rate of the system. We refer to the shorter phase as the warning state (state 1) as it is in this phase that unexpected system failure is most likely. The longer phase is referred to as the healthy state (state 0). Such a two-state model is similar to a delay time model used in engineering (Christer, 1999). The true state of the system cannot be directly inferred from the sensors used, such that we model the system operating states, healthy and warning, as unobservable operating states. The failure state is observable. We formulate the problem as a POMDP. We have that the state process $\{X_t, t \in R_+\}$ is a continuous time homogeneous Markov chain with 2 unobservable working states, $\mathcal{S}_X = \{0, 1\}$, where 0 refers to the healthy state and 1 to the warning state. We refer to the observable failure state as state 2, such that the state space is $\mathcal{S}_X \cup \{2\}$. For ease of reference, we provide a list of the frequently used notation in Appendix A.

Decisions about maintenance and external sensor information acquisition are made at discrete time points, referred to as decision epochs. Since external sensor information is solicited from external sources, we may assume that there is a delay in the acquisition of such information. The decision epoch length may be set accordingly such that observations from either the internal or the external sensor becomes available at the start of the next decision epoch. We let τ be the running hours that elapse between two consecutive decision epochs. At every decision epoch, the decision maker may choose to stop system operations and perform preventive maintenance that restores the system to an as-good-as-new state, or to continue system operations and possibly acquire external sensor information in addition to the internal system sensor information. The acquired external sensor information becomes available at the next decision epoch. The external sensor information provides a more accurate estimation of the actual deterioration state of the system, healthy or warning, when compared to the internal system sensor information. The decisions are made based on a univariate probability that the system is in the warning state. Given that n decision epochs have passed since the last preventive or corrective maintenance action, and given all the available sensor information up to decision epoch n, we define the univariate probability as $T_n = \pi$. We also refer to π as the belief state.

The univariate belief state, π, resulting from our two working states model (excluding the observable failure state), allows for maintenance policies that may be easily visualized and interpreted by decision makers. While the problem setting with more than two working states, one healthy and one warning state, may be studied, the univariate belief state will increase in dimensionality accordingly. Visualization and interpretation will therefore become more complex which, in turn, decreases the odds of buy-in and adoption of the model in practice. Additionally, such a multi-dimensional dynamic program is known to be PSPACE hard (Papadimitriou, 1995). As a result, analytically determining the relevant decision making thresholds becomes extremely difficult, if possible at all. Examples of work including only two working states in partially observable environment include Kim & Makis (2013), Van Oosterom et al. (2017) and Maillart et al. (2018).

We refer to the three actions available to the decision maker as either to continue system operations for another epoch with internal sensor information only (action one). Action two is to acquire information from external sensors at a cost of $C_S > 0$. The final action is to perform preventive maintenance at a cost of $C_M > 0$ (action three). The system is immediately replaced upon failure at a cost of $C_F > 0$. We have that $C_F > C_M > C_S > 0$. While this is a realistic assumption, we show in Lemma 2 that, should $C_F > C_M$ not hold, preventive maintenance will never be performed in the optimal policy.

We refer to the sensor information obtained at a decision epoch, from either the internal system sensor or the external sensor, as an observation. The system observation obtained at decision epoch n for the internal sensor is denoted as $Y_n \in \{1, 2, \ldots, M\} = \mathcal{S}_Y$ and $Z_n \in \{1, 2, \ldots, N\} = \mathcal{S}_Z$ for the external sensors. We have that N and M are not necessarily equal. The larger the observation in Y_n and Z_n at decision epoch n, the likelier it is that the system is in the warning state (state one). Each possible observation from the internal and external sensors is stochastically related to the true underlying state of the system, which is either healthy or warning. We assume that this relationship is known and refer to it as the state-observation matrix. For examples of how to use data to determine the state-observation matrices, we refer the interested reader to Uit Het Broek et al. (2019) and Maillart et al. (2018). We use

$$\mathcal{D} = (d_{iy})_{2\times M} \tag{1}$$

to denote the state-observation matrix for the internal sensor, where $d_{iy} = P(Y_n = y | X_n = i)$, for $i \in S_X, y \in S_Y$ is the probability of observing y from the internal sensor while the system is in the unobservable true state i. Likewise, we denote the state-observation matrix for the external sensors as

$$\mathcal{G} = (g_{iz})_{2\times N} \tag{2}$$

where $g_{iz} = P(Z_n = z|X_n = i)$, for $i \in S_X, z \in S_Z$ is the probability of observing z from the external sensor while the system is in the unobservable true state i. The smaller the observation received from either sensor type, the more likely it is that the system is in the healthy state (state 0). This relationship implies that the state-observation matrices are totally positive of order 2 (TP2). State-observation matrix $\mathcal{G}$ ($\mathcal{D}$) is said to be TP2 if $g_{1z^+}g_{0z^-} \geq g_{1z^-}g_{0z^+}$ ($d_{1y^+}d_{0y^-} \geq d_{1y^-}d_{0y^+}$) holds for $z^+ \geq z^-$ ($y^+ \geq y^-$) (Karlin, 1968).

We model the system deterioration as a continuous process. We define the instantaneous system state transition rates, with the instantaneous rate of the system transitioning between the unobservable healthy and warning states and the observable failure state, as $q_{ij} = \lim_{\tau\to 0^+} \frac{P(X_\tau = j|X_0=i)}{\tau} < +\infty, i \neq j \in \mathcal{S}_X \cup \{2\}$, with resulting matrix

$$Q = (q_{ij})_{3\times 3}. \tag{3}$$

Note that

$$q_{ii} = -\sum_{j=0}^{2} q_{ij} \qquad i \neq j.$$

The state sojourn times are assumed to be exponential (Tijms, 1994) with parameter $\nu_i = \sum_j q_{ij}$, for all $i, j \in \mathcal{S}$ and $i \neq j$ such that $\nu_0 = q_{01} + q_{02}$ and $v_1 = q_{12}$. The assumption ensures non-decreasing and monotonic deterioration. Additionally, with the exponential assumption, the state transition rate matrix Q is upper triangular with $q_{ij} = 0$ for all $j < i$ and $q_{i2} \leq q_{j2}$ for all $i < j$. From the instantaneous state transition matrix Q and (Tijms, 1994, Sec. 2.8), we can obtain the transition probability matrix, P, governing the transitions between the unobservable healthy and warning states and the observable failure state between decision epochs with duration τ, where $p_{ij}(t) = P(X_t = j|X_0 = i)$, for all $i, j \in \mathcal{S}_X \cup \{2\}$ follows as

$$P(\tau) = p_{ij}(\tau) = \begin{pmatrix} e^{-v_0\tau} & \frac{q_{01}(e^{-q_{12}\tau} - e^{-v_0\tau})}{v_0 - q_{12}} & 1 - e^{-v_0\tau} - \frac{q_{01}(e^{-q_{12}\tau} - e^{-v_0\tau})}{v_0 - q_{12}} \\ 0 & e^{-q_{12}\tau} & 1 - e^{-q_{12}\tau} \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

For any $t \in [0, \tau]$, the system's reliability is given as

$$R(t|\pi) = (1 - p_{02}(t))(1 - \pi) + (1 - p_{12}(t))\pi, \tag{5}$$

and may be interpreted as the probability that the system will not fail within the next t time units given that the system has not yet failed with probability $1 - p_{02}(t) - p_{12}(t)$ and conditional on the probability, π, of being in the warning state (state 1).

At the beginning of every decision epoch n, a new observation, Y_n or Z_n, from either the internal or external sensors, respectively, is made available. The observation is used to determine, by means of Bayes' rule, the posterior probability that the system is in the warning state given the current belief state, $T_n = \pi$.

We update the belief state similarly for actions one and two (respectively continue normal system operations with internal sensor observations, and acquire external sensor information), as

$$T_{n+1}(\pi|Y_{n+1}) = \frac{d_{1Y_{n+1}}(p_{01}(\tau)(1-\pi)+p_{11}(\tau)\pi)}{d_{0Y_{n+1}}p_{00}(\tau)(1-\pi)+d_{1Y_{n+1}}(p_{01}(\tau)(1-\pi)+p_{11}(\tau)\pi)} \tag{6}$$

and

$$T_{n+1}(\pi|Z_{n+1}) = \frac{g_{1Z_{n+1}}(p_{01}(\tau)(1-\pi)+p_{11}(\tau)\pi)}{g_{0Z_{n+1}}p_{00}(\tau)(1-\pi)+g_{1Z_{n+1}}(p_{01}(\tau)(1-\pi)+p_{11}(\tau)\pi)}. \tag{7}$$

In line with Bayes' rule, we have that the numerator of Eq. (6) is the product of the probability of observing Y_{n+1} given that the system is in the unobservable warning state, $d_{1Y_{n+1}}$, and the probability of being in the warning state at a belief state value of π, which is $p_{01}(\tau)(1-\pi)+p_{11}(\tau)\pi$. The denominator of Eq. (6) is the probability of observing Y_{n+1} regardless of the underlying state of the system. Eq. (7) follows similarly.

The updated belief states are also used to determine a threshold $a \in [0,1]$ for which it is possible to distinguish sensors according to informativeness. We provide an intuitive description of distinguishing between informativeness, after which we present a less restrictive definition which we will use in our analysis in Section 4. We let "high" sensor observations correspond to observations for which the updated posterior probability that the system is in the warning state is above such a threshold a. Recall that external sensor observations are more informative than internal system sensor observations, encouraging the willingness to pay for the external information. Using Blackwell (1951), Monahan (1980) and Lévesque & Maillart (2008), we define external state-observation matrix $\mathcal{G}$ to be more informative than internal state-observation matrix $\mathcal{D}$, denoted as $\mathcal{G} \succ \mathcal{D}$, if and only if, for all values of the belief state π and a threshold $a \in [0,1]$,

$$\sum_{z=1}^{z(\pi,a)} g_{0z} \geq \sum_{y=1}^{y(\pi,a)} d_{0y} \tag{8}$$

and

$$\sum_{z=z(\pi,a)+1}^{N} g_{1z} \geq \sum_{y=y(\pi,a)+1}^{M} d_{1y}. \tag{9}$$

Observation $y(\pi,a) = \max\{y : T(\pi|y) \leq a\}$, is the maximum internal sensor observation y for which the posterior probability that the system is in the warning state, given y, is at most as great as threshold a. Similarly, $z(\pi,a) = \max\{z : T(\pi|z) \leq a\}$. From Eqs. (8) and (9), state-observation matrix $\mathcal{G}$ is more informative than $\mathcal{D}$, $\mathcal{G} \succ \mathcal{D}$, if, for all thresholds a, the following two statements hold. Firstly, given that the system is in the warning state, the decision maker is more likely to observe "high" observations under $\mathcal{G}$ than under $\mathcal{D}$. Secondly, given that the system is in the healthy state, the decision maker is more likely to observe "low" observations under $\mathcal{G}$ than under $\mathcal{D}$. A less restrictive definition,

$$a(1-\pi)\sum_{z=1}^{z(\pi,a)} g_{0z} + \pi(1-a)\sum_{z=z(\pi,a)+1}^{N} g_{1z} \geq a(1-\pi)\sum_{y=1}^{y(\pi,a)} d_{0y} + \pi(1-a)\sum_{y=y(\pi,a)+1}^{M} d_{1y}, \tag{10}$$

describes the relationship $\mathcal{G} \succ \mathcal{D}$ in one equation.

From POMDP theory (Bertsekas & Shreve, 1978; Monahan, 1980) and our two-state formulation, we have that T_n, the univariate probability that the system is in the warning state given that it has not failed up to decision epoch n and given all the acquired sensor observations up to decision epoch n, represents sufficient information for decision making at the nth decision epoch. We let Δ be the class of all stationary policies. Any stationary policy $\delta \in \Delta$ with $\delta(\pi) \in \{1, 2, 3\}$, $\delta(\pi)$ indicates the action to be taken for $T_n = \pi$.

Recall that the objective is to minimize the expected long-run maintenance cost of the system. We define $C(\delta)$ and $L(\delta)$ as the expected total cost incurred in one maintenance cycle and the expected length of a maintenance cycle, respectively. A maintenance cycle consists of the time between any two consecutive maintenance actions, be it corrective or preventive. Using renewal theory, the stopping problem

$$\frac{E_{T_0}[C(\delta)]}{E_{T_0}[L(\delta)]}, \tag{11}$$

is minimized to obtain the optimal stationary policy $\delta^* \in \Delta$ that minimizes the expected long-run cost per running hour, if it exists. We have that E_{T_0} is the conditional expectation given T_0, the univariate initial state distribution. Since a maintenance action restores the system to an as-good-as-new state, we have that $T_0 = [0]$. We apply the λ-minimization technique developed in Aven & Bergman (1986) to transform Eq. (11) into an additive function, allowing for the POMDP belief states to be translated into the actual states of an MDP, with parameter $\lambda > 0$. The parameter λ is the expected cost per running hour. The value function

$$V^\lambda(\pi) = \inf_{\delta \in \Delta} E_\pi[C(\delta) - \lambda L(\delta)], \tag{12}$$

results, in which we aim to minimize the difference between the expected cost incurred in a maintenance cycle and the product of the expected length of such a cycle with λ. We use the additive formulation in Eq. (12) to simplify the calculation of λ and, by minimizing λ, we minimize $V^\lambda(\pi)$. The value functions therefore depend on the value of λ. The optimal stationary policy $\delta^* \in \Delta$ follows as the minimization of Eq. (12) for $\lambda = \lambda^*$, with $\lambda^* = \inf\{\lambda > 0 : V^\lambda(T_0) \leq 0\}$, the optimal expected average cost per running hour. To simplify notation, we denote $V(\pi)$ instead of $V^\lambda(\pi)$ throughout the remainder of the paper. Instead, we use $V^1(\pi)$, $V^2(\pi)$ and $V^3(\pi)$ to denote the value function obtained when either of the three allowable actions are taken.

For the finite horizon maintenance problem, Eq. (12) satisfies the dynamic equations

$$V_m(\pi) \quad = \quad \min\{V_m^1(\pi), V_m^2(\pi), V_m^3(\pi)\}, \tag{13}$$

where $V_m^{\cdot}(\pi)$ denotes the state-action value functions for the three separate actions with at most m decision epochs remaining before preventive maintenance is performed. From Eq. (13), we have, for all values of π and a fixed λ, that Eq. (12) is the minimum value function over the three different actions. The terminal value function, $V_0(\pi)$, follows as

$$V_0(\pi) \quad = \quad C_M. \tag{14}$$

In Eq. (13), action one, to continue normal system operations for another epoch with internal sensor information only, is described by $V_m^1(\pi)$. Using the difference between the expected cycle costs and the expected cycle length of Eq. (12), action one results in the cost of corrective maintenance should the system fail between decision epochs, the mean time to failure of the system during that decision epoch, and the continuation cost should the system not fail between epochs. The value function obtained from action two, $V_m^2(\pi)$, is described similarly to that of action one. The difference between $V_m^1(\pi)$ and $V_m^2(\pi)$ is that $V_m^2(\pi)$ includes the cost of acquiring information from the external sensor and that the continuation cost is updated using the more informative state-observation matrix $\mathcal{G}$. Finally, action three is described by $V_m^3(\pi)$ and results in the cost incurred upon maintenance of the system. Recall that, without loss of generality, we assume that a maintenance action is instantaneous. The result is that the expected cycle time for action three is negligible. The value functions are formulated as

$$V_m^1(\pi) = C_F(1 - R(\tau|\pi)) - \lambda \int_0^\tau R(u|\pi)du + R(\tau|\pi) \sum_{y \in \mathcal{S}_Y} V_{m-1}(T(\pi|y))\phi(y, \pi), \tag{15}$$

$$V_m^2(\pi) \quad = \quad C_S + C_F(1 - R(\tau|\pi)) - \lambda \int_0^\tau R(u|\pi)du + R(\tau|\pi) \sum_{z \in \mathcal{S}_Z} V_{m-1}(T(\pi|z))\phi(z, \pi) \tag{16}$$

and

$$V_m^3(\pi) \quad = \quad C_M. \tag{17}$$

Since an observation is only partially related to the true underlying state of the system, the function $\phi(., \pi)$ in Eqs. (15) and (16) is required. The function describes the probability of obtaining observation $Y_{n+1} = y$ for $\phi(y, \pi)$ or $Z_{n+1} = z$ for $\phi(z, \pi)$, given a belief state value π and is given as

$$\phi(y, \pi) = \frac{d_{0y}p_{00}(\tau)(1 - \pi) + d_{1y}(p_{01}(\tau)(1 - \pi) + p_{11}(\tau)\pi)}{R(\tau|\pi)} \tag{18}$$

for the internal sensor observations. For the external sensors, we have

$$\phi(z, \pi) = \frac{g_{0z}p_{00}(\tau)(1 - \pi) + g_{1z}(p_{01}(\tau)(1 - \pi) + p_{11}(\tau)\pi)}{R(\tau|\pi)}. \tag{19}$$

4. Model analysis

In order to determine whether an optimal stationary policy, $\delta^* \in \Delta$, governing the decision maker's actions which minimizes the expected average long-run maintenance cost, exists, we characterize the form of the infinite horizon value function, $V(\pi)$, in this section. We begin with properties describing the finite horizon value function, $V_m(\pi)$, of the m-decision epoch problem. (We relegate the proofs of all Lemmas, Theorems and Corollaries to Appendix B.)

Lemma 1. *The finite horizon value functions, $V_m(\pi)$, are concave in π for every $m \geq 0$ and are uniformly bounded from below by*

$$V_m(\pi) \geq -\frac{\lambda\tau}{1 - R(\tau|0)}.$$

From $V_0(\pi) = V_m^3(\pi) = C_M$ and Eq. (13), we have that $C_M \geq V_m(\pi)$. Moreover, with m epochs remaining until a preventive maintenance action is performed, it follows that $V_m(\pi) \geq V_{m+1}(\pi)$. Together with Lemma 1, we have that the infinite horizon value function can be obtained as the limit $V(\pi) = \lim_{m\to\infty} V_m(\pi)$. We next provide a condition for which, given an optimal policy, the decision maker will never stop the system to perform preventive maintenance.

Lemma 2. *For $0 < C_F < C_M$, any policy that stops the system to initiate preventive maintenance is not optimal.*

Since the external sensor is more informative than the internal sensor, it follows from Eqs. (15) - (17) that, for costless sensor information, $C_S = 0$, $V^2(0) < V^1(0) < V^3(0)$. For such a case, the optimal action for a belief state of $\pi = 0$ is to acquire external sensor information. We next show that, for some $C_S > 0$, the optimal action is also to acquire external sensor information when the belief state $\pi = 0$.

Lemma 3. *There exists a $C_S > 0$ for which $V^2(0) < V^1(0)$.*

While the exact value of the external sensor cost, C_S, for which it holds that $V^2(0) < V^1(0)$, cannot be exactly derived from Lemma 3, the unconventional possibility of $V^2(0) < V^1(0)$ has an impact on the optimal maintenance policy. For the conventional setting of $V^1(0) < V^2(0) < V^3(0)$, Lemmas 1-3 allow a description of the optimal policy for the infinite horizon problem setting:

Theorem 1. *For $V^1(0) < V^2(0) < V^3(0)$, the optimal infinite horizon policy is characterized by three thresholds, π_L^*, π_U^* and π_M^*, where $0 \leq \pi_L^* \leq \pi_U^* \leq \pi_M^* \leq 1$.*

From Theorem 1, we have that the optimal policy for our POMDP has an at most four region structure, in line with literature (Ross, 1971; Kim & Makis, 2013). The three thresholds may be interpreted as follows. We refer to the smallest (largest) belief state for which action two, to solicit more informative external sensor information to increase the prediction accuracy of the system belief state, is optimal as the lower (upper) sensor threshold, π_L^* (π_U^*). The smallest belief state for which action three, to perform preventive maintenance, is optimal is referred to as the preventive maintenance threshold, π_M^*. For any other values of π, action one, to continue normal system operations with internal sensor information only, is optimal.

Figure 1 illustrates how the optimal policy may be represented by means of a control chart in the belief state of the system, governing the maintenance decisions. The decision maker should perform preventive maintenance (action three) when the belief state π is in the region $\pi \in [\pi_M^*, 1]$. The decision maker should acquire costly, but more informative, external sensor information (action 2) when the belief state π is in the region $\pi \in [\pi_L^*, \pi_U^*)$. For all other values of the belief state π, the decision maker "does nothing," i.e. does not perform maintenance on the system nor acquires external sensor information.

While, in Theorem 1, we have shown the existence of the thresholds at which the optimal actions change, it is not possible to determine closed-form expressions due to the partial information from the system updating. The thresholds are therefore solved for numerically using an algorithm presented in Appendix C. The smallest

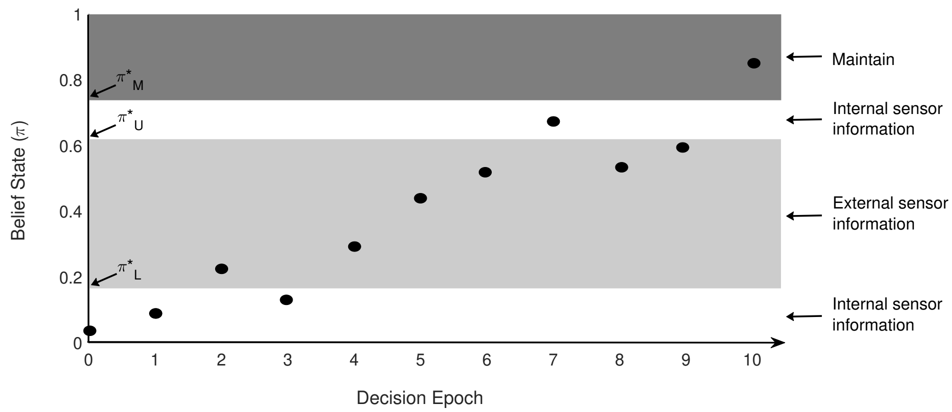

Figure 1: The optimal policy may be represented as a control chart in the belief state of the system, governing the maintenance decisions according to at most three thresholds: π_L^*, after which to acquire external sensor information; π_U^*, after which to continue system operations with only internal sensor information; and π_M^*, after which to perform preventive maintenance.

belief state at which a decision maker is willing to pay for external sensor information, π_L^*, is dependent on both the difference in informativeness between the internal and external sensor, as well as the cost of acquiring the external information. A significant improvement in informativeness due to the external sensor, especially when internal sensor informativeness is unreliable, a low sensor information cost, or a combination of both, leads to a low threshold level π_L^*. While paying for external information at a low belief state that carries a low risk of unexpected failures may seem counter-intuitive, such information may provide a more reliable view on the actual state of the system for maintenance planning purposes in the long run. The greater the improvement in informativeness and the lower the cost of external sensor information, the lower the belief state will be at which a decision maker is willing to pay for external sensor information.

From Theorem 1, it is possible that a second, also counter-intuitive, region exists where the decision maker is not willing to pay for external sensor information. This region is found between the external sensor information region and the preventive maintenance region, as shown in Figure 1 between π_U^* and π_M^*. As the belief state approaches the maintenance threshold π_M^*, it may be optimal for the decision maker to not pay for external sensor information. By definition, a high belief state refers to a high probability that the system is in the warning state, and intuition would suggest to pay for external sensor information so as to obtain an as-reliable-as-possible belief state about the system. However, in this second region between π_U^* and π_M^*, the marginal gain of improved belief state predictions resulting from more informative, but costly, external sensor information is outweighed by the marginal cost of paying for such information.

A sample belief state path may also be seen in Figure 1. According to this path, at decision epochs zero, one, three and seven, the optimal decision is to continue normal system operations with internal sensor information only. For decision epochs two, four to six, eight and nine, the optimal decision is to acquire more informative external sensor information. Finally, at decision epoch ten, the system is believed to be sufficiently deteriorated such that the optimal action is to perform maintenance. From the figure it may also be seen that a decision maker does not necessarily continue to pay for external sensor information once the lower sensor threshold, π_L^*, is reached. Instead, as seen from decision epochs two and three, a more informative external sensor observation may result in a belief state that falls below the lower sensor threshold. The result is that the optimal action is again to continue system operations with internal sensor information only.

Recall, for $C_S = 0$, that $V^2(0) < V^1(0) < V^3(0)$. Given that $C_S = 0$, the decision maker will not make use of internal sensor information only. The problem then reduces to a two region optimal maintenance policy with one threshold, where $\pi_L^* = 0$ and $\pi_U^* = \pi_M^*$. Here, the allowable actions are either to acquire external sensor information or to perform maintenance. Similarly, when the cost of acquiring external sensor information exceeds the cost of corrective maintenance, $C_S > C_M$, we know from Eqs. (15) - (17) that $V^1(0) < V^3(0) < V^2(0)$ and the decision maker will never solicit external sensor information. The optimal policy then reduces to two regions with one threshold, where $\pi_L^* = \pi_U^* = \pi_M^*$. Here, the allowable actions are either to continue normal system operations with internal sensor information only and run the system for another decision epoch or to stop the system to perform maintenance.

From Lemma 3, we have that, for some positive value of C_S, $V^2(0) < V^1(0) < V^3(0)$. In the next result, we use Lemma 3 to show that the optimal policy consists of at most three regions when, for some $0 < C_S < C_M$, we have $V^2(0) < V^1(0) < V^3(0)$.

Theorem 2. *For $V^2(0) < V^1(0) < V^3(0)$, the optimal maintenance policy consists of at most three regions and two thresholds, π_U^* and π_M^*.*

From Theorem 2 we have, given $V^2(0) < V^1(0) < V^3(0)$ holds, that, for $0 \leq \pi_U^* \leq \pi_M^* \leq 1$, the optimal action is to pay for external sensor information if $\pi < \pi_U^*$, to continue normal system operations with internal sensor information only if $\pi_U^* \leq \pi < \pi_M^*$ and to perform preventive maintenance if $\pi \geq \pi_M^*$. For most problem instances it may be, however, that the optimal maintenance policy consists of two regions. Specifically, given a sufficiently large difference between corrective maintenance cost, C_F, and preventive maintenance cost, C_M, the normal system operations with internal information only region may fall away such that there is only one threshold, $0 \leq \pi_M^* \leq 1$. In such instances, the optimal action is to pay for external sensor information if $\pi < \pi_M^*$ and to perform preventive maintenance if $\pi \geq \pi_M^*$.

Recall that the external sensor observations from $\mathcal{G}$ are more informative than that of the internal sensor observations from $\mathcal{D}$. We next investigate how a change in the informativeness of the external sensor state-observation matrices $\mathcal{G}$ impacts the optimal policy. To do so, we first show that the belief state updating of the system is non-decreasing in the observations from either state-observation matrix $\mathcal{G}$ or $\mathcal{D}$.

Lemma 4. *The posterior probability functions, $T(\pi|y)$ and $T(\pi|z)$, are non-decreasing in y and z, respectively, for all $\pi \in [0,1]$ if and only if the observation probability matrices $\mathcal{D}$ and $\mathcal{G}$ are TP2.*

Using Lemma 4, the following result provides insights into the optimal regions under different levels of external sensor informativeness:

Theorem 3. *For identical problem instances but with different TP2 external sensor state-observation probability matrices, $\mathcal{G}$ and $\tilde{\mathcal{G}}$, where $\mathcal{G} \succ \tilde{\mathcal{G}}$ and the conventional $V^1(0) < V^2(0) < V^3(0)$, we have*

1. $V(\pi) \leq \tilde{V}(\pi)$, $\forall \pi$.
2. $\pi_L^* \leq \tilde{\pi}_L^*$, $\tilde{\pi}_U^* \leq \pi_U^*$ *and* $\tilde{\pi}_M^* \leq \pi_M^*$.

From $V(\pi) \leq \tilde{V}(\pi), \forall \pi$ in Theorem 3, we have that the optimal average long-run cost per running hour, λ^*, is greater for the less informative matrix $\tilde{\mathcal{G}}$ than for matrix $\mathcal{G}$. Theorem 3 also shows that, should the external sensor informativeness increase with no change in cost and given that the optimal policy consists of four regions with thresholds $0 \leq \pi_L^* \leq \pi_U^* \leq \pi_M^* \leq 1$ as described in Theorem 1, three changes in the optimal policy regions will be observed. Firstly, as π_L^* decreases and π_U^* increases, the two normal system operations with internal sensor information only regions will become smaller and may eventually disappear. Secondly, the external sensor information region will become larger. An increase in the external sensor observation interval will result in more accurate belief state predictions, reducing the risk of unexpected failures. However, due to the stochastic nature of the belief state updating, it does not necessarily mean that the system running time spent in this interval will increase. Thirdly, as π_M^* increases, the preventive maintenance region will become smaller, such that preventive maintenance is postponed to a larger belief state π. Again, an increase in the preventive maintenance threshold does not necessarily mean that the system running time between preventive maintenance actions will increase, but rather that belief state updating accuracy will improve.

The problem setting may be extended to include multiple different external sensors for which informativeness is known to hold according to Eq. (10) between all the observation matrices for the different sensors. For the setting where all external sensor costs are equal, it is easy to show, using Theorem 3, that the optimal policy will still consist of at most four regions and the most informative external sensor will always be chosen when it is optimal to acquire external information. The result aligns with the findings in Saghafian et al. (2018) that a single external source will be chosen in the long run and under specific informativeness conditions. While our problem setting is more specific than the one studied by Saghafian et al. (2018), their work considers not only the quality of the sensors, but also the level of understanding of the underlying inference models applied to the sensor information. In the problem studied by Saghafian et al. (2018), the most informative sensor may therefore not be the sensor source chosen in the long run.

For the case where the decision maker has access to $W+1$ external sensors with different acquisition costs and increasing as external sensor informativeness increases, we let $\mathbb{G} = \{\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2, ..., \mathcal{G}_W\}$, be the set of external sensor state-observation matrices, ordered according to informativeness. Note that $\mathcal{G}_0$ refers to the original state observation matrix $\mathcal{G}$. We let the cost for acquiring information from a sensor in the updated set, $\mathcal{G}_w \in \mathbb{G}$, be $C_{S,\mathcal{G}_w}$. As the matrices are ordered according to informativeness, $C_{S,\mathcal{G}_{w-1}} < C_{S,\mathcal{G}_w}$ holds for all the matrices in $\mathbb{G}$. We let $V^{2,\mathcal{G}_w}$ be the value function for consulting external sensor information (action 2) with state-observation matrix $\mathcal{G}_w \in \mathbb{G}$. From Theorems 1 and 3, we can derive the following result:

Corollary 1. *For $V^1(0) < V^{2,\mathcal{G}_w}(0) < V^3(0)$, where $\mathcal{G}_w \in \mathbb{G}$, the optimal infinite horizon three threshold* *policy is extended by up to $W-1$ thresholds, $\pi^*_{\mathcal{G}_w}$, for every $\mathcal{G}_w \in \mathbb{G}$. The thresholds are in addition to the thresholds, π^*_U and π^*_M, where $0 \le \pi^*_{\mathcal{G}_w} \le \pi^*_U \le \pi^*_M \le 1$ and for every $\mathcal{G}_w \in \mathbb{G}$.*

Note that the threshold $\pi^*_{\mathcal{G}_0}$ is analogous to π^*_L from Theorem 2. From Corollary 1 we have that, when the cost of obtaining external sensor information increases with informativeness, such that $C_{S,\mathcal{G}_{w-1}} < C_{S,\mathcal{G}_w}$ $\forall \mathcal{G}_w \in \mathbb{G}$, and when $V^1(0) < V^{2,\mathcal{G}_w}(0) < V^3(0)$ holds $\forall \mathcal{G}_w \in \mathbb{G}$, there exist thresholds $\pi^*_{\mathcal{G}_w}$ in $[0, \pi^*_U)$ at which it is optimal to acquire information from external sensors $\mathcal{G}_w \in \mathbb{G}$. However, even if costs are chosen such that $C_{S,\mathcal{G}_{w-1}} < C_{S,\mathcal{G}_w}$ holds for all the state-observation matrices in $\mathbb{G}$, the resulting thresholds may not be of the form $\pi^*_{\mathcal{G}_{w-1}} < \pi^*_{\mathcal{G}_w}$. The exact relationship depends on the trade off between the informativeness of the sensors and the cost thereof. We already note from our numerical analysis that an increase in the number of sensors to solicit information from may increase the value of external information, thereby reducing the average long-run cost per running hour. However, the added value of such a complex policy may, in some cases, be quite small and may therefore not be worth the effort required to implement in practice.

5. Numerical example

In this section, we numerically illustrate the working of our model. While the main goal of the section is to illustrate our analytical results in Theorems 1 - 3, we extend the numerical analysis of Theorem 3 to include a decision maker's willingness to pay for external sensor information in Section 5.4. In Section 5.5, we extend the numerical analysis of Corollary 1 to consider a case when informativeness does not hold between multiple external sensors. We implement the model and obtain the threshold values using an algorithm based on λ-minimization (Kim & Makis, 2013; Maillart et al., 2018) and refer to Appendix C for the pseudocode.

5.1. Preliminary setup

Recall that the observed deterioration observations $Y_n \in \{1, 2, \ldots, M\}$ in Section 3 and $Z_n \in \{1, 2, \ldots, N\}$ are assumed to be discrete. For most systems, however, the deterioration values may be continuous. Several approaches may be applied to discretize the sensor observations. In their work, Uit Het Broek et al. (2019) partitioned their continuous deterioration observations into a pre-specified number of equally sized intervals between the as-good-as-new state and the failure state. Alternatively, and in order to take advantage of the system deterioration pattern, cluster analysis may be used. Maillart et al. (2018) used the K-means cluster analysis algorithm to partition their continuous deterioration observations into discrete observations.

To determine the instantaneous transition rate matrix for the true system states as well as the state-observation matrix that generated the internal sensor observation sequences, hidden Markov model techniques such as the expectation-maximization algorithm may be used. If historical external sensor data is available, a similar approach may be followed to determine the state-observation matrix for external sensor observations.

5.2. Numerical problem setup

After the data pre-processing, our model may be used to determine the optimal policy that minimizes λ, the long-run expected cost per running hour for the system given internal and external sensor observations.

For our stylized example, we assume the following cost structure. We let the preventive maintenance cost be $C_M = 700$ and the corrective maintenance cost be $C_F = 2.5C_M$. We vary the external sensor cost in increments of 5 between $C_S = [0, 100]$ to illustrate Theorems 1 and 3, and to investigate the effect on the optimal policy. For $C_S > 100$, external sensor information is too expensive for the chosen state-observation matrices such that the decision maker will make use of internal sensor information only. To illustrate Theorem 2, we instead use $C_F = 3.5C_M$ and $C_S = 15$.

For the state-observation matrices of the internal and external sensors, we let

$$\mathcal{D} = \begin{pmatrix} 0.46 & 0.30 & 0.24 \\ 0.27 & 0.27 & 0.46 \end{pmatrix} \quad \text{and} \quad \mathcal{G} = \begin{pmatrix} 0.51 & 0.49 & 0.00 \\ 0.13 & 0.14 & 0.73 \end{pmatrix}.$$

We have, for instance, that obtaining observation one while the system is in the healthy state is likelier for matrix $\mathcal{G}$, the external sensor state-observation matrix, with a probability of 0.51, than for matrix $\mathcal{D}$, the internal sensor matrix, with a probability of 0.46. Likewise, the probability of obtaining observation three while the system is in the healthy state is smaller for $\mathcal{G}$, at 0.00, than for $\mathcal{D}$, at 0.24. Using Eq. (10), matrix $\mathcal{G}$ is more informative than matrix $\mathcal{D}$ for all values of π and $a \in [0, 1]$.

The time that elapses between decision epochs, τ, is set as 48 running hours and we let the transition probability matrix between the healthy, warning and failure state at each decision epoch, be

$$P(48) = \begin{pmatrix} 0.79 & 0.17 & 0.04 \\ 0.00 & 0.68 & 0.32 \\ 0.00 & 0.00 & 1.00 \end{pmatrix}.$$

From the transition probability matrix, the probability that the system will transition from the healthy or the warning state to the failure state between two consecutive epochs is 0.04 and 0.32, respectively. Recall from Section 3 that the belief state value, π, is continuous in $[0, 1]$. We discretize the belief state into intervals of 0.001 and find that smaller interval step sizes do not significantly affect the expected cost rate, λ. For increased accuracy, but increased computation time, the interval step size may be decreased.

5.3. Implementation of the optimal policy varying external sensor information cost

In this section we show the optimal policy and the respective thresholds for different values of external sensor information cost. We first consider the case for costless external sensor information, $C_S = 0$. Recall that, for $C_S = 0$, the optimal policy consists of only two regions and one threshold, as $\pi_L^* = 0$ and $\pi_U^* = \pi_M^*$. The optimal actions are then to either acquire external sensor information, or to stop the system to perform maintenance. We illustrate the resulting policy in Figure 2 with the value functions of the three actions. For all values of the belief state, up to the preventive threshold $\pi_M^* = 0.772$, it is optimal to solicit external sensor information. When the belief state exceeds the preventive threshold, it is optimal to perform preventive maintenance. The average long-run cost per running hour is $\lambda^* = 4.19$.

As the cost for external sensor information increases, the decision to pay for external information is postponed to a larger belief state, π, and the optimal policy becomes a three region policy. For the problem

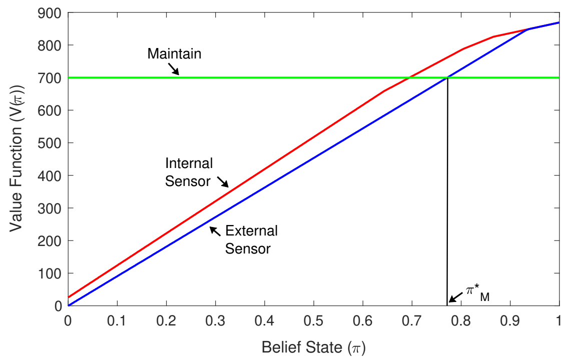

Figure 2: The optimal policy for $C_S = 0$ consists of two regions. Maintenance is performed at or after the cross-over of the external sensor and the maintain value functions.

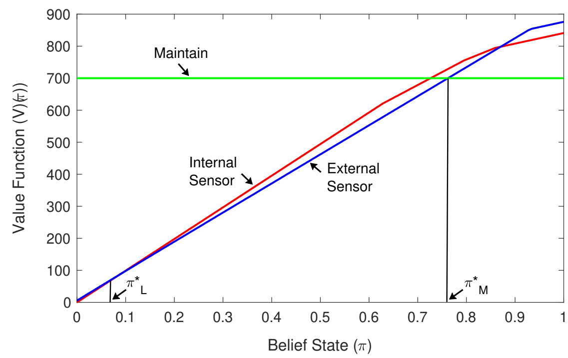

Figure 3: The optimal policy for $C_S = 35$ consists of three regions. The policy is determined as the minimum of the three different value functions.

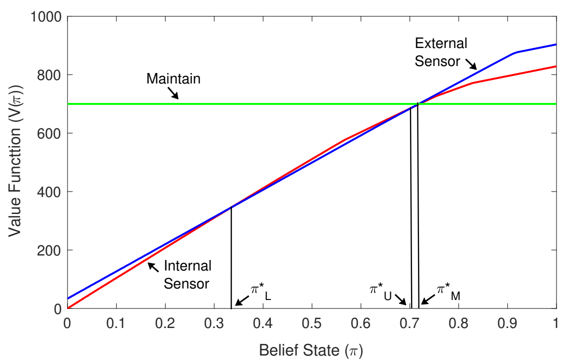

Figure 4: The optimal policy for $C_S = 75$ consists of four regions. The second internal sensor region is obtained for belief state values between 0.703 and 0.720.

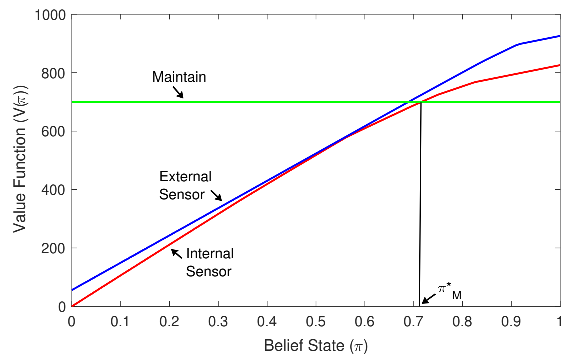

Figure 5: The optimal policy for $C_S = 100$ consists of two regions. A similar policy is obtained for the problem instance when external information is not available.

instance with $C_S = 35$, we show the obtained value functions and the resulting thresholds in Figure 3. The optimal policy consists of three regions. From the figure, it is optimal to continue system operation with internal sensor information only (i.e., action one has the lowest value function) for a belief state π in $0 \leq \pi < 0.081$. For a belief state π in $0.081 \leq \pi < 0.763$, the optimal action is to pay for external sensor information, action two. Finally, for a belief state of $\pi \geq 0.763$, the optimal action is to perform preventive maintenance, action three. The average long-run cost per running hour is $\lambda^* = 4.89$.

For the problem instance with $C_S = 75$, we show the obtained value functions and the resulting thresholds in Figure 4. It may be seen that the optimal policy has a four region structure with three thresholds, $\pi_L^* = 0.335$, $\pi_U^* = 0.703$ and $\pi_M^* = 0.720$. The average long-run cost per running hour is $\lambda^* = 5.20$. It may be seen from Figure 4 that there is little difference in the obtained value functions for the same belief states, π, between action one, to continue system operation as normal, and action two, to acquire external sensor information, in the interval $[\pi_L^*, \pi_M^*)$. In fact, the average long-run cost per running hour for the policy where it is never optimal to utilize external sensor information because it is too expensive, as seen in

Figure 5, and all else kept constant, is $\lambda^* = 5.27$. The cost is an increase of 1.35% when compared to that of Figure 4, and the resulting optimal policy is shown in Figure 5. Also, the preventive maintenance threshold is approximately the same, at $\pi_M^* = 0.717$.

Figure 5 is similar to the instance where only internal sensor information is available and the problem reduces to determining the optimal policy in the absence of any additional external sensor information. From the increase in the average long-run cost per running hour in Figures 2 to 5, as the external sensor information becomes more costly, our model always performs at least as good as the optimal CBM policy where external sensor information is not available, although the cost benefit becomes negligible as C_S increases.

Furthermore, the effect of an increase in external sensor information cost on the optimal policy structure may be seen in Figures 2 to 5. Initially, for $C_S = 0$, the optimal policy consists of two regions where the actions are to either use external sensor information or to perform maintenance. As the external sensor information cost increases, the optimal policy changes from two regions to three regions (Figure 3), where the optimal action now also includes to continue normal system operation with internal sensor information only. As the external sensor information cost continues to increase, the decision maker becomes less likely to pay for external sensor information and a fourth region, the second internal sensor information region, may appear, as in Figure 4. Finally, we know from Theorem 3 that the upper threshold π_U^*, should it exist, becomes smaller while the lower threshold, π_L^*, becomes larger as external sensor informativeness increases. We observe a similar effect on the thresholds as external sensor cost increases, until the optimal policy reduces to two regions again (Figure 5). The two regions now consist of only using internal sensor information (due to the high cost of external sensor information) or performing maintenance.

From Theorem 2, we have the special case where $V^2(0) < V^1(0) < V^3(0)$, such that the optimal policy reduces to at most three regions. Given that $V^2(0) < V^1(0)$ holds, it is optimal to pay for external sensor information when $\pi = 0$. Keeping our model parameters constant, but using $C_F = 3.5C_M$ and $C_S = 15$, we obtain the optimal policy as shown in Figure 6. From the figure it may be seen that, for a combination of a high cost of corrective maintenance and a relatively low cost of external sensor information, it is always optimal to pay for external sensor information. Once the preventive maintenance threshold, $\pi_M^* = 0.612$, is reached, maintenance should be performed. The average long-run cost per running hour is now $\lambda^* = 5.33$.

5.4. The value of external sensor information

We know from Sections 4 and 5.3 that, as the acquisition cost of external sensor information increases, the decision maker is less likely to use such information in the system belief state prediction. Similarly, as the informativeness of the external sensor information increases, the decision maker is more likely to pay for such information. We now numerically investigate the value of external sensor information, and the willingness to pay for it, for different levels of external sensor informativeness and cost.

We consider four different external sensor state-observation matrices of different informativeness. We use

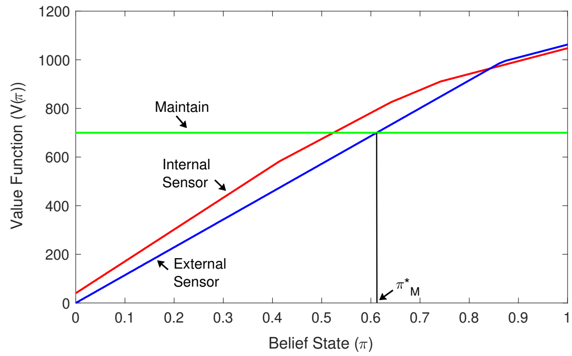

Figure 6: The optimal policy for the problem instance where $C_S = 15$ and $C_F = 3.5C_M$ consists of two regions. It is never optimal to utilize internal sensor information.

the original external sensor state-observation matrix, $\mathcal{G}$, and construct three other matrices. These are

$$\mathcal{G}_2 = \begin{pmatrix} 0.58 & 0.42 & 0.00 \\ 0.03 & 0.10 & 0.87 \end{pmatrix}, \quad \mathcal{G}_3 = \begin{pmatrix} 0.63 & 0.37 & 0.00 \\ 0.00 & 0.01 & 0.99 \end{pmatrix} \text{ and } \overline{\mathcal{G}} = \begin{pmatrix} 0.57 & 0.41 & 0.02 \\ 0.01 & 0.06 & 0.93 \end{pmatrix}.$$

Eq. (10) holds for all values of the belief state π and threshold $a \in [0, 1]$ between matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\mathcal{G}_3$. While matrices $\mathcal{G}_2$ and $\mathcal{G}_3$ are more informative than matrix $\mathcal{G}$, matrix $\overline{\mathcal{G}}$ is chosen in such a way that a decision maker may be indifferent between matrix $\overline{\mathcal{G}}$ and the other three matrices depending on the cost of the sensor information. In other words, according to the definition for informativeness, Eq. (10) does not hold for all $a \in [0, 1]$ and for all $\pi \in [0, 1]$ for which a classification of informativeness may be made. Our choice of indifference between $\overline{\mathcal{G}}$ and matrices $\mathcal{G}$ and $\mathcal{G}_2$ allows us to illustrate how, in such a scenario and for our problem setup, a decision maker's preferences are dependent not only on the state-observation matrix, but may actually change as the external sensor information cost change.

State-observation matrix $\mathcal{G}_3$ represents approximately perfect external sensor information. For $\mathcal{G}_3$, the system is almost certainly in the healthy state when either observation one or two is obtained and almost certainly in the warning state when observation three is obtained. Should observations from an external sensor with state-observation matrix $\mathcal{G}_3$ be available and costless, it will never be optimal to choose action one, to continue system operations with internal sensor information only. The resulting value functions are plotted in Figure 7. From the figure, it may be seen that the preventive maintenance threshold increases to $\pi_M^* = 0.798$, in comparison to that of Figure 2 (0.772), due to the increased deterioration prediction accuracy of $\mathcal{G}_3$ over $\mathcal{G}$. The average long-run cost per running hour for the setting is $\lambda^* = 4.00$, a decrease of 4.53% from $\lambda^* = 4.19$ in Figure 2. The decrease in cost per running hour when a decision maker has access to information from $\mathcal{G}_3$, as opposed to $\mathcal{G}$, illustrates the value of utilizing more informative sensor information.

We numerically determine the willingness to pay for the approximately perfect external sensor with state-observation matrix $\mathcal{G}_3$. We find that the decision maker will be indifferent between acquiring external sensor information from $\mathcal{G}_3$ and not soliciting it for $C_S = 139$. We plot the resulting value functions in Figure 8. The average long-run cost per running hour is $\lambda^* = 5.27$, which is equivalent to the setting when only internal

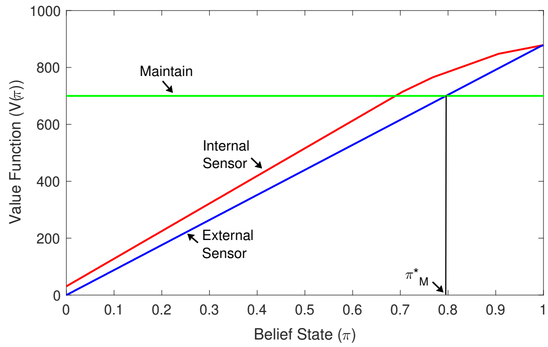

Figure 7: The optimal policy for the problem instance where external sensor information is approximately perfect and costless consists of two regions. Maintenance should be performed at or after the cross-over of the external sensor and the preventive maintenance value functions.

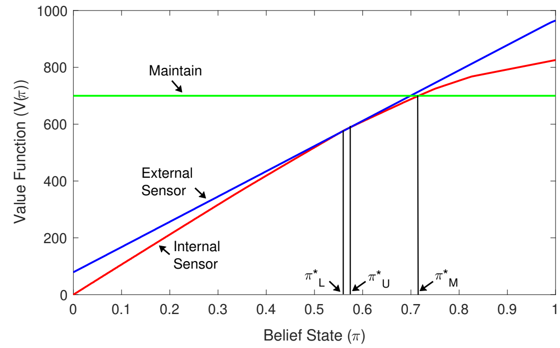

Figure 8: The optimal policy for the problem instance where external sensor information is approximately perfect at a cost of 139 per observation consists of four regions. The external sensor region, $[\pi_L^*, \pi_U^*)$ is small in comparison to the two internal sensor regions, $[0, \pi_L^*)$ and $[\pi_U^*, \pi_M^*)$.

sensor information is available. The decision maker should therefore pay, at most, 139 per approximately perfect external sensor observation from matrix $\mathcal{G}_3$ when the belief state is in $[\pi_L^*, \pi_U^*)$.

We numerically determine the effect on the optimal policy of changes in the cost of observations, C_S, for each of the different external sensor state-observation matrices $\mathcal{G}$, $\mathcal{G}_2$, $\mathcal{G}_3$ and $\overline{\mathcal{G}}$. The results are shown in Figures 9 and 10. In Figure 9, the cost, C_S, is shown on the vertical axis and the four different state-observation matrices are shown on the horizontal axis. The figure shows the maximum cost that a decision maker is willing to pay for an observation from each of the four different external sensor state-observation matrices. For instance, we see that an observation from matrix $\mathcal{G}_3$ at a cost greater than 139, as represented by the white region in Figure 9, will never be solicited and thus have no value to the decision maker. Figure 10 visualizes how the optimal policies change for the four different matrices as the cost, C_S, increases from 0 to 145. Similarly to Figure 1, the white area represents the action to continue normal system operations with internal sensor information only. The light grey area represents the action to pay for external sensor information. The dark grey area represents the action of performing maintenance.

In Figure 9, the optimal policy in the white region consists of two actions, to continue normal system operations with internal sensor information only, or to perform preventive maintenance. Such a two region policy is visualized in Figure 10 for $\mathcal{G}$ at $C_S \geq 105$, for and $\mathcal{G}_2$ and $\overline{\mathcal{G}}$ at $C_S \geq 125$, and for $\mathcal{G}_3$ at $C_S \geq 145$.

For sensor observations from matrix $\mathcal{G}_3$ at a cost between approximately 35 and 139, the light grey area in Figure 9, it is optimal to sometimes make use of external sensor observations. The optimal policy will consist of all three actions with at least three regions and at most four. The four region optimal policies are clearly observed in Figure 10 at $C_S = 105$ for $\mathcal{G}_2$ and at $C_S = 125$ for $\mathcal{G}_3$. For sensor observations from matrix $\mathcal{G}_3$ at a cost of less than 35, the dark grey area in Figure 9, it is always optimal to acquire external sensor information until the preventive maintenance threshold is reached. The optimal policy will consist of two regions and the allowable actions are to either acquire external sensor information, or to perform preventive

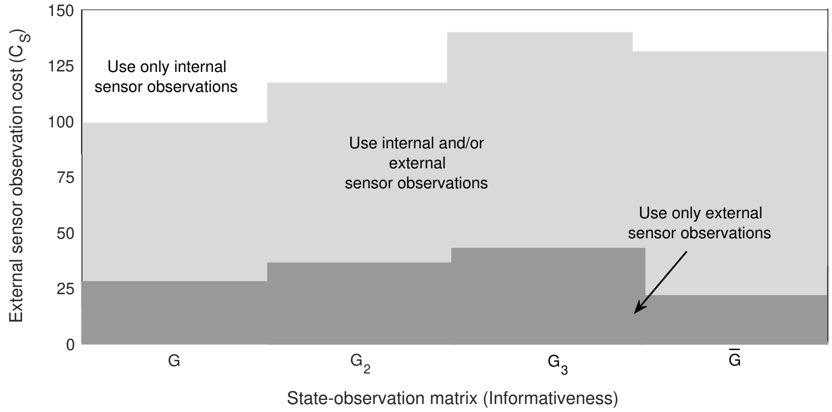

Figure 9: The white area indicates when the cost of external sensor information is too high to make use of such information given the external sensor informativeness. Acquiring external sensor information is sometimes optimal in the light grey area and always optimal during normal system operations in the dark grey area.

maintenance. The policy is visualized for $C_S = 0$ in Figure 10.

Similarly, from Figure 9, sensor information should not be acquired from state-observation matrix $\mathcal{G}_2$ at a cost of more than approximately 120, as visualized in Figure 10 for $C_S = 125$ and 145. Additionally, information should always be acquired from state-observation matrix $\mathcal{G}_2$ at a cost of approximately 30 or less, as visualized in Figure 10 for $C_S = 0$. For a sensor observation cost between about 30 and 120, it is sometimes optimal to acquire external sensor information, as visualized in Figure 10 for $C_S = 50$ and 105.

Note that the indifference in informativeness between matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\overline{\mathcal{G}}$ is also visualized in Figures 9 and 10 and explained as follows. From Figure 9, it may be seen that the maximum a decision maker is willing to pay for an external sensor observation from matrix $\overline{\mathcal{G}}$ (visualized in Figure 10 for $C_S \leq 125$) is more than for one from either $\mathcal{G}_2$ (approximately $C_S \leq 120$ and visualized in Figure 10 for $C_S \leq 125$)) or $\mathcal{G}$ (approximately $C_S \leq 100$ and visualized in Figure 10 for $C_S \leq 105$). However, according to the dark grey area of Figure 9, when it is always optimal to acquire external sensor information during normal system operations, the willingness to pay for an external sensor observation from $\mathcal{G}$ or $\mathcal{G}_2$, approximately 25 and 30, respectively, is higher than that for $\overline{\mathcal{G}}$, approximately 20. Information from $\overline{\mathcal{G}}$ therefore becomes more valuable as C_S increases such that the decision maker begins to prefer $\overline{\mathcal{G}}$ over $\mathcal{G}_2$ or $\mathcal{G}$, confirming the indifference between $\overline{\mathcal{G}}$ and $\mathcal{G}_2$ or $\mathcal{G}$. Finally, for $C_S = 0$, the average cost for $\overline{\mathcal{G}}$ is greater than for either $\mathcal{G}$ or $\mathcal{G}_2$ while, for $C_S = 105$ the average cost for $\overline{\mathcal{G}}$ is less than for either $\mathcal{G}$ or $\mathcal{G}_2$ (Figure 10). The result is that the definition for threshold a, and therefore informativeness according to Eq. (10), does not hold between the matrices.

5.5. The effect of multiple external sensors

In Section 5.4, we considered the impact of external sensor informativeness when only one external sensor is available. In this section, we numerically illustrate the effect on the optimal policy when the decision maker has the option to choose an observation from a predefined set of external sensors. We consider two different

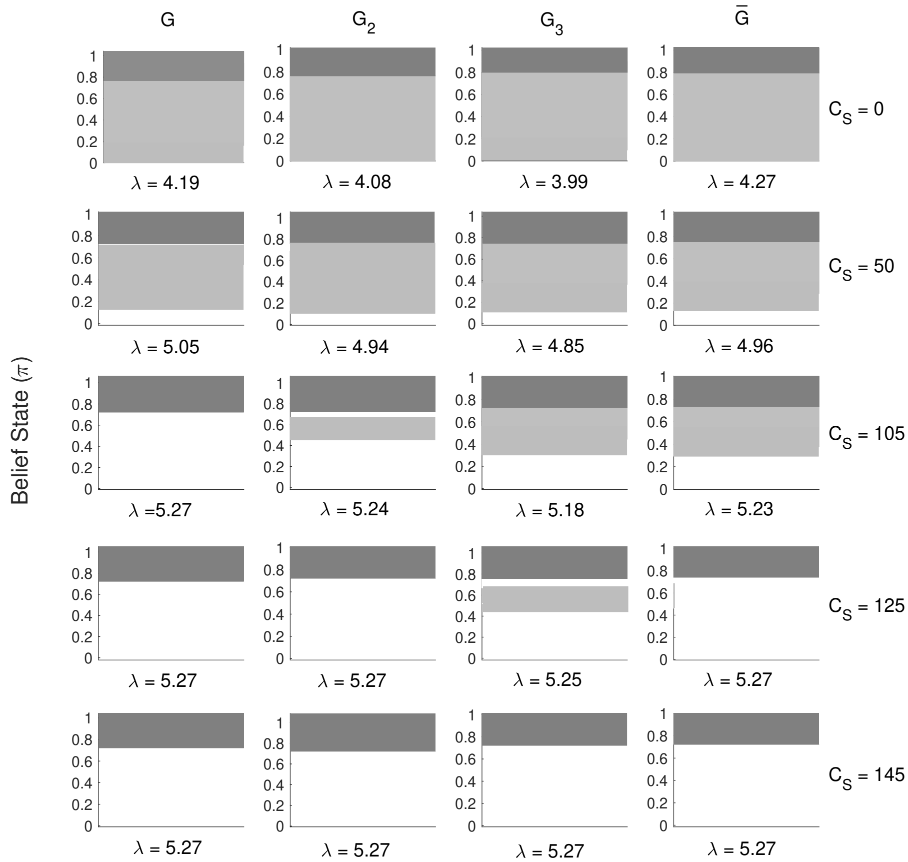

Figure 10: The optimal policy changes as the cost of external sensor information C_S increases and is dependent on the state-observation matrices, $\mathcal{G}$, $\mathcal{G}_2$, $\mathcal{G}_3$ and $\overline{\mathcal{G}}$. At $C_S = 105$ and $C_S = 125$, state-observation matrices $\mathcal{G}_2$ and $\mathcal{G}_3$ both result in four region policies, respectively. As C_S increases up to 145, the optimal policy for all state-observation matrices consists of two regions and will never utilize external sensor information.

settings under both equal and distinct acquisition costs. In the first setting (case one), the decision maker has access to three external sensors with increasing levels of informativeness. These could be observations from three separate sensors, or the third sensor could be a combination of the first two external sensors which, when combined, generates a higher level of informativeness. In the second setting (case two), the decision maker has access to three external sensors for which strict increasing levels of informativeness does not necessarily hold. We describe the relevant optimal policies in greater detail in Appendix D.

For case one, we use state-observation matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\mathcal{G}_3$ from Section 5.4, each at a cost of 105. We find that the optimal policy consists of three regions and only makes use of the most informative external sensor option, $\mathcal{G}_3$. In line with Corollary 1, we obtain two thresholds, $\pi^*_{\mathcal{G}_3} = 0.368$ and $\pi^*_M = 0.734$. The average long-run cost per running hour for this setting is $\lambda^* = 5.18$. The setting is similar to the problem where the decision maker only has access to one external sensor, $\mathcal{G}_3$, at a cost of 105. For case two, we use matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\overline{\mathcal{G}}$ from Section 5.4, each also at a cost of 105. The optimal policy now consists of four regions with three thresholds. Information will be acquired from $\overline{\mathcal{G}}$ with thresholds $\pi^*_{\overline{\mathcal{G}}} = 0.432$, $\pi^*_U = 0.708$ and $\pi^*_M = 0.722$. The average long-run cost per running hour for this setting is $\lambda^* = 5.23$.

We now consider distinct acquisition costs of 35, 45 and 55, respectively, for case one and 80, 100 and 105, respectively, for case two. Recall that, even though a decision maker is indifferent between the sensors $\overline{\mathcal{G}}$ and $\mathcal{G}$ or $\mathcal{G}_2$ in terms of informativeness, sensor $\overline{\mathcal{G}}$ provides a more reliable prediction of when the system is in a warning state. Information from $\overline{\mathcal{G}}$ is therefore more valuable in determining when to stop the system for maintenance when compared to $\mathcal{G}$ or $\mathcal{G}_2$, and hence we set information from $\overline{\mathcal{G}}$ to be more expensive. For case one, our results are as described by Corollary 1, which governs the maximum number of thresholds for an optimal policy. The obtained thresholds are $\pi^*_{\mathcal{G}} = 0.093$, $\pi^*_{\mathcal{G}_2} = 0.185$, $\pi^*_{\mathcal{G}_3} = 0.279$ and $\pi^*_M = 0.776$. We find that information from $\mathcal{G}$ is acquired first. Once a general belief about the system belief state is determined, information from $\mathcal{G}_2$ is chosen. When the system is believed to have sufficiently deteriorated to be close enough to a transition into the warning state, information from the most expensive external sensor, but most informative regarding the warning state, $\mathcal{G}_3$, should be acquired. The average long-run cost per running hour is $\lambda^* = 4.88$. For case two, we obtain similar results, except the optimal policy now consists of a second region where the optimal action is to continue normal system operations with internal sensor information only. The thresholds are obtained as $\pi^*_{\mathcal{G}} = 0.376$, $\pi^*_{\mathcal{G}_2} = 0.505$, $\pi^*_{\overline{\mathcal{G}}} = 0.593$, $\pi^*_U = 0.709$ and $\pi^*_M = 0.721$. The average long-run cost per running hour is $\lambda^* = 5.22$.

For case one, the incremental cost advantage of access to multiple external sensors, at distinct costs, over the single sensor problem is only 0.20% for each of the matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\mathcal{G}_3$. For case two, the incremental cost advantage is 0.19% for $\overline{\mathcal{G}}$ and 0.00%, for both $\mathcal{G}$ and $\mathcal{G}_2$. We have that access to additional sensors may increase the number of thresholds in the optimal policy, such that each additional threshold increases the cost savings. We note that these cost savings, as in the case for our investigated instances, may, however, be quite small. It may therefore be that the minor cost savings achievable from multiple external sensors may not outweigh the increased complexity of soliciting information from such sensors.

6. Conclusions

The Internet of Things allows for improved system deterioration state predictions in maintenance optimization through combining observations from different sensors. In this work, we investigate the benefit of the combined use of costly external sensor observations with existing internal sensor observations on CBM policies for single component partially observable systems. While access to the external sensor observations comes at a cost, such observations are more informative than internal sensor observations, providing more accurate deterioration predictions.

We make two main contributions. Firstly, we analyze the structure of the optimal maintenance policy and show that it consists of at most four regions. Secondly, we show the effect of increasing external sensor informativeness on the optimal policy. The analysis allows us to investigate numerically whether, and how much, a decision maker should be willing to pay in order to acquire external sensor observations of different informativeness. We also investigate the effect of access to multiple external sensors on the optimal policy. We find that increasing the number of external sensors may reduce the average long-run cost per running hour, while increasing the complexity of the optimal policy. The benefit in reduced cost may, however, not be significant enough to justify investment in a more complex maintenance policy.

Our work provides an application for the acquisition of multi-sensor information in a sequential decision making problem. We acknowledge that our model is limited to only two working deterioration states. Nonetheless, we believe that it provides insight into the use and value of multi-sensor techniques in maintenance. It can be beneficial for companies considering investing into Industry 4.0 sensor technologies for maintenance optimization. Future work may include extending the model to include multiple components with deterioration interactions to provide additional insights into the value of multi-sensor data.

Appendix A. Notation list

Table A.1: The list of frequently used notation abbreviations.

Symbol	Meaning	Symbol	Meaning
$\{X_t, t \in R_+\}$	State process	τ	Decision epoch length
$\mathcal{S}_X \cup \{2\}$	State space	n	Epochs since maintenance
$T_n = \pi$	Probability in warning state	C_S	External sensor cost
$Y_n \in \{1, 2, \ldots, M\} = \mathcal{S}_Y$	Internal sensor observation	C_F	Corrective maintenance cost
$Z_n \in \{1, 2, \ldots, N\} = \mathcal{S}_Z$	External sensor observation	C_M	Preventive maintenance cost
$\mathcal{D} = (d_{iy})_{2\times M}$	Internal sensor state-observation matrix	λ	Average long run cost rate
$\mathcal{G} = (g_{iz})_{2\times M}$	External sensor state-observation matrix	π_L	Lower sensor threshold
$Q = (q_{ij})_{3\times 3}$	Instantaneous transition rate matrix	π_U	Upper sensor threshold
$P(\tau) = p_{ij}(\tau)$	Transition probability matrix	π_M	Maintenance threshold
$R(t\|\pi)$	System reliability	$\phi(y, \pi)$	Conditional probability of y
$\mathbb{G} = \{\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2, ..., \mathcal{G}_W\}$	External state-observation matrix set	$\phi(z, \pi)$	Conditional probability of z

Appendix B. Proofs of Lemmas and Theorems

Proof of Lemma 1. The proof follows from Lemmas 1 and 2 in Kim & Makis (2013) and Lemmas 1, 2 and 3 in Maillart et al. (2018). □

Proof of Lemma 2. For $C_F < C_M$, it is easy to show that $V^3(\pi) > V^2(\pi) > V^1(\pi)$ for all $0 \leq \pi \leq 1$, such that it is never optimal to perform preventive maintenance.

□

Proof of Lemma 3. For $C_S \geq 0$,

$$\begin{aligned} V^2(0) - V^1(0) &= C_S + \big(1 - p_{02}(\tau)\big)\big[\sum_z V(T(0|z))\phi(z,0) - \sum_y V(T(0|y))\phi(y,0)\big] \\ &< 0, \end{aligned}$$

for $0 \leq C_S < \big(1 - p_{02}(\tau)\big)\big[\sum_z V(T(0|z))\phi(z,0) - \sum_y V(T(0|y))\phi(y,0)\big]$. □

Proof of Theorem 1. We prove the theorem by showing that the optimal policy consists of four convex subsets. For $\pi = 0$, we know that $V^1(0) < V^2(0) < V^3(0)$. For $\pi = 1$, we now show that $V^3(1) < V^1(1) < V^2(1)$ by first proving the first part of the inequality. The inequality is equivalent to $V^3(1) = V(1)$. We know that $\lim_{m\to\infty} V_m(\pi) = V(\pi)$ exists. Since $V^1(\pi)$ and $V^2(\pi)$ are non-decreasing in π, $V^3(\pi)$ is a constant and $C_F > C_M$, there exists an optimal stopping time such that $V(1) = \lim_{j\to\infty} V_j(1) = V^3(1) < V^1(1)$.

Next, we prove the last part of the inequality, $V^1(1) < V^2(1)$. By Equations (15) and (16),

$$\begin{aligned} V^1(1) - V^2(1) &= R(\tau|1)\sum_{y\in\mathcal{S}_Y} V(T(y|1))\phi(y,1) - R(\tau|1)\sum_{z\in\mathcal{S}_Z} V(T(z|1))\phi(z,1) - C_S \\ &= R(\tau|1)V(1)\sum_{y\in\mathcal{S}_Y}\phi(y,1) - R(\tau|1)V(1)\sum_{z\in\mathcal{S}_Z}\phi(z,1) - C_S \\ &= -C_S \\ &< 0, \end{aligned}$$

which implies that $V^1(1) < V^2(1)$. Thus for $\pi = 1$, $V^3(1) < V^1(1) < V^2(1)$.

The above inequalities and Equation (13) imply that the region $\{\pi : V(\pi) = V^3(\pi)\}$ is a convex subset of $[0,1]$ of the form $[\pi_M^*, 1]$, for some $0 \leq \pi_M^* \leq 1$, and the region $\{\pi : V(\pi) = V^2(\pi)\}$ is a convex subset of $[0,1]$ of the form $[\pi_L^*, \pi_U^*]$, for some $0 \leq \pi_L^* \leq \pi_U^* \leq \pi_M^*$, which completes the proof.

□

Proof of Theorem 2. The proof follows directly from Lemma 3 and Theorem 1.

□

Proof of Lemma 4. The proof follows from Theorem 4 in Whitt (1979).

□

Proof of Theorem 3. For condition 1, we use induction on the value function. Without loss of generality, let $V_0(\pi) = \tilde{V}_0(\pi) = 0$, $\forall\pi$. Assume it holds for some k such that $V_k(\pi) \leq \tilde{V}_k(\pi)$, $\forall\pi$. If $V_{k+1}(\pi) = V_{k+1}^3(\pi)$ or $V_{k+1}(\pi) = V_{k+1}^1(\pi)$, then $V_{k+1}(\pi) \leq \tilde{V}_{k+1}(\pi)$. For the case where $V_{k+1}(\pi) = V_{k+1}^2(\pi)$, we need to show that

$$\begin{aligned} \tilde{V}_{k+1}(\pi) - V_{k+1}(\pi) &\geq \tilde{V}_{k+1}^2(\pi) - V_{k+1}^2(\pi) \\ &= R(\tau|\pi)\Big[\sum_{\tilde{z}\in\mathcal{S}_{\tilde{z}}} \tilde{V}_{k+1}(T(\pi|\tilde{z}))\phi(\tilde{z},\pi) - \sum_{z\in\mathcal{S}_Z} V_{k+1}(T(\pi|z))\phi(z,\pi)\Big] \\ &\geq 0. \end{aligned} \tag{B.1}$$

We first let

$$T_{n+1}(\pi|\emptyset) = \frac{p_{01}(\tau)(1-\pi) + p_{11}(\tau)\pi}{R(\tau|\pi)}$$

be the posterior probability that, given the current decision epoch is n, the system is in the warning state should no sensor observation, either internal or external, be available at the next decision epoch. We have

$$\sum_{z \in \mathcal{S}_Z} T(\pi|z)\phi(z,\pi) = \sum_{\tilde{z} \in \mathcal{S}_{\tilde{z}}} T(\pi|\tilde{z})\phi(\tilde{z},\pi) = T_{n+1}(\pi|\emptyset). \tag{B.2}$$

Using Equation (B.2) and since $V(\pi)$ is concave, we can show that

$$T(\pi|\tilde{z}) \geq_c T(\pi|z),$$

where $\geq_c$ refers to greater in the sense of second-order stochastic dominance (Shanthikumar & Shaked, 1994), holds. Second-order stochastic dominance holds if, for all π and a in $[0,1]$,

$$\int_a^1 \Pr(T(\pi|\tilde{z}) \geq x)dx \leq \int_a^1 \Pr(T(\pi|z) \geq x)dx. \tag{B.3}$$

From Lemma 4, $T(\pi|z)$ is non-decreasing in z, $\forall\ \pi \in [0,1]$, such that Equation (B.3) is equivalent to

$$\sum_{\tilde{z}=1}^{\tilde{z}(\pi,a)} \Pr(\tilde{Z} = \tilde{z})[T(\pi|\tilde{z}) - a] \leq \sum_{z=1}^{z(\pi,a)} \Pr(Z = z)[T(\pi|z) - a]. \tag{B.4}$$

We substitute Equation (7) into Equation (B.4) and obtain

$$\pi(1-a)\sum_{\tilde{z}=1}^{\tilde{z}(\pi,a)} \tilde{g}_{1\tilde{z}} - a(1-\pi)\Big(1 - \sum_{\tilde{z}=\tilde{z}(\pi,a)+1}^{\tilde{M}} \tilde{g}_{0\tilde{z}}\Big) \leq \pi(1-a)\sum_{z=1}^{z(\pi,a)} g_{1z} - a(1-\pi)\Big(1 - \sum_{z=z(\pi,a)+1}^{M} g_{0z}\Big),$$

for all π and $a \in [0,1]$, which is Eq. (10), the informativeness condition. It follows that $T(\pi|\tilde{z}) \geq_c T(\pi|z)$. From $T(\pi|\tilde{z}) \geq_c T(\pi|z)$, the fact that $V(\pi)$ is concave and Theorems 4 and 5 in Blackwell (1951), we have that Inequality (B.1) holds, giving

$$\sum_{\tilde{z} \in \mathcal{S}_{\tilde{z}}} \tilde{V}_{k+1}(T(\pi|\tilde{z})) - \sum_{z \in \mathcal{S}_Z} V_{k+1}(T(\pi|z)) \geq 0, \qquad \forall \pi$$

and the proof for condition 1 follows.

For condition 2, since $\tilde{V}^2(\pi) \geq V^2(\pi)$ and, for values of $\pi \leq \tilde{\pi}_L$, we have that $\tilde{V}^1(\pi) = V^1(\pi)$, it follows that $\tilde{\pi}_L^* \geq \pi_L^*$ and $\tilde{\pi}_U^* \leq \pi_U^*$ holds. Since $\tilde{V}^3(\pi) = V^3(\pi)$, the inequality $\tilde{\pi}_M^* \leq \pi_M^*$ follows directly from condition 1, and the proof follows. □

Proof of Corollary 1. Using Theorem 1 and, since $C_{S,\mathcal{G}_{w-1}} < C_{S,\mathcal{G}_w}$ holds for all the matrices in $\mathbb{G}$, it's easy to show that $V^3(1) < V^1(1) < V^2(1) \leq V^{2,\mathcal{G}_1}(1) \leq \ldots \leq V^{2,\mathcal{G}_W}(1)$, for $\pi = 1$. Since we have that $V^1(0) < V^2(0) < V^3(0)$ and $V^1(0) < V^{2,\mathcal{G}_w}(0) < V^3(0)$, where $\mathcal{G}_w \in \mathbb{G}\backslash\{\mathcal{G}\}$, it is easy to show, using Theorem 1, that $0 \leq \pi^*_{\mathcal{G}_w} \leq \pi^*_U$ for every $\mathcal{G}_w \in \mathbb{G}\backslash\{\mathcal{G}\}$.

□

Appendix C. Solution algorithm

We use a binary search algorithm combined with a dynamic program to find an optimal policy for the minimum λ. The algorithm is used to determine the thresholds for our numerical analysis in Section 5 and is adapted from Kim & Makis (2013) and Maillart et al. (2018).

Step 0. Initialization

Choose $\epsilon > 0, \lambda_L$ and λ_U. Set $m = 1$.

Step 1. Value Initialization

Put $\lambda = \frac{\lambda_L + \lambda_U}{2}, V_0^{\lambda}(\pi) = C_M$.

Step 2. Value Improvement

Calculate $V_m^{\lambda}(\pi)$.

Step 3. Evaluation

If $\max_{\pi} \left\{ |V_m^{\lambda}(\pi) - V_{m-1}^{\lambda}(\pi)| \right\} \le \epsilon$, go to Step 4.

Else, put $m = m + 1$ and go to Step 2.

Step 4. Stopping Criteria Test

If $V_m^{\lambda}(0) < -\epsilon$, put $\lambda_U = \lambda$ and go to Step 1.

If $V_m^{\lambda}(0) > \epsilon$, put $\lambda_L = \lambda$ and go to Step 1.

If $|V_m^{\lambda}(0)| \le \epsilon$, put $\lambda^* = \lambda$, then, for each π, choose $\delta_{\epsilon}(\pi) \in \mathrm{argmin}_{\mathrm{a} \in \mathrm{A}} \left\{ \mathrm{V}_{\mathrm{m}}^{\lambda}(\pi) \right\}$ and Stop.

Appendix D. Investigation into multiple external sensor sources

We describe the analysis from Section 5.5, utilizing information from multiple external sensors, in more detail in this Appendix. We investigate two cases. For case one, the decision maker has access to state-observation matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\mathcal{G}_3$. For case two, the decision maker has access to $\mathcal{G}$, $\mathcal{G}_2$ and $\overline{\mathcal{G}}$.

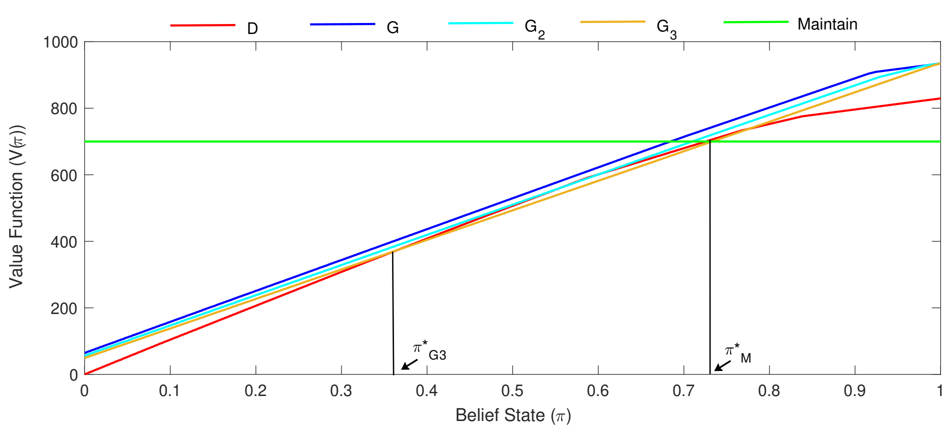

Figure D.11: The optimal policy for the problem instance with external sensor matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\mathcal{G}_3$ at a cost of 105 each. External sensor information is acquired from the most informative matrix, $\mathcal{G}_3$, only.

Figure D.11 illustrates case one at equal sensor costs of 105. From the figure, the thresholds are obtained as $\pi^*_{\mathcal{G}_3} = 0.368$ and $\pi^*_M = 0.734$. The average long-run cost per running hour is $\lambda^* = 5.18$. In Figure D.12 we show the results of case two at equal costs of 105. Observations from $\overline{\mathcal{G}}$ are chosen for $0.432 \le \pi < 0.708$. For $0.708 \le \pi < 0.722$,

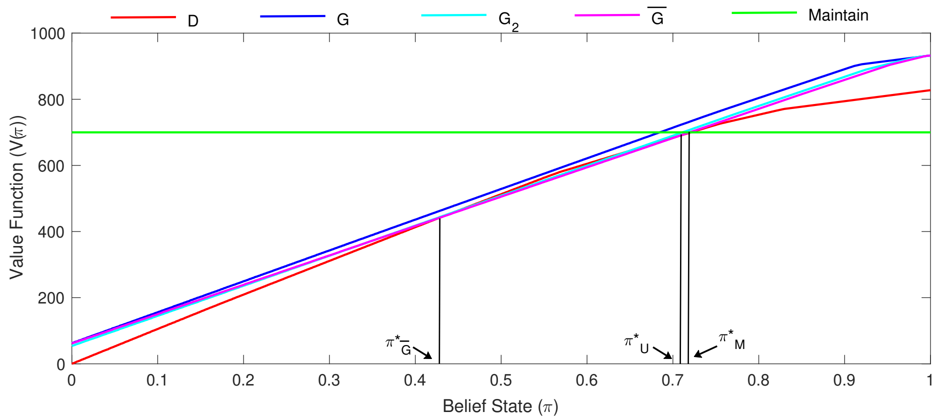

Figure D.12: The optimal policy for the problem instance with external sensor matrices $\mathcal{G}$, $\mathcal{G}_2$ and $\overline{\mathcal{G}}$ at a cost of 105 each. Information from matrix $\overline{\mathcal{G}}$, which is the most informative of the three matrices with respect to the probability of the system being in the warning state, is advised for the majority of the belief state.

it is optimal to not solicit external information from any of the sensors. Maintenance is performed for $\pi \geq 0.722$. The average long-run cost per running hour for the setting is $\lambda^* = 5.23$. In Figure D.13, we show results for case one at distinct costs of 35, 45 and 55, respectively. For Figure D.14, we show results for case two at distinct costs of 80, 100 and 105, respectively. In Figure D.13, information is solicited from $\mathcal{G}_3$ for the majority of the belief state values. In Figure D.14, we obtain a second region, $0.709 \leq \pi < 0.722$, where it is optimal to continue system operations with internal sensor information only. The costs of Figures D.13 and D.14 are, respectively, $\lambda^* = 4.88$ and $\lambda^* = 5.22$.

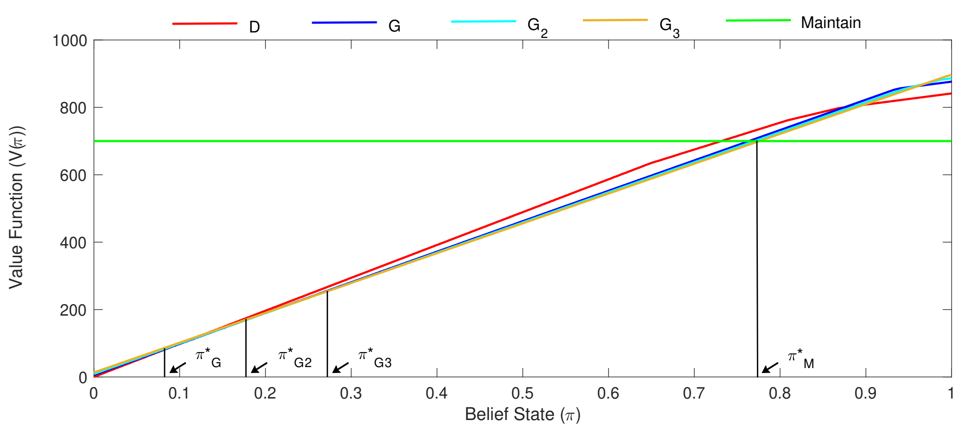

Figure D.13: The optimal policy for the problem instance of Figure D.11, but with sensor costs of 35, 45 and 55, respectively. The most informative matrix, $\mathcal{G}_3$, is used for the majority of the belief state.

References

Aven, T., & Bergman, B. (1986). Optimal replacement times: A general set-up. *Journal of Applied Probability*, *23*, 432–442.

Bertsekas, D. P., & Shreve, S. E. (1978). *Stochastic optimal control: The discrete time case*. Academic Press, New York.

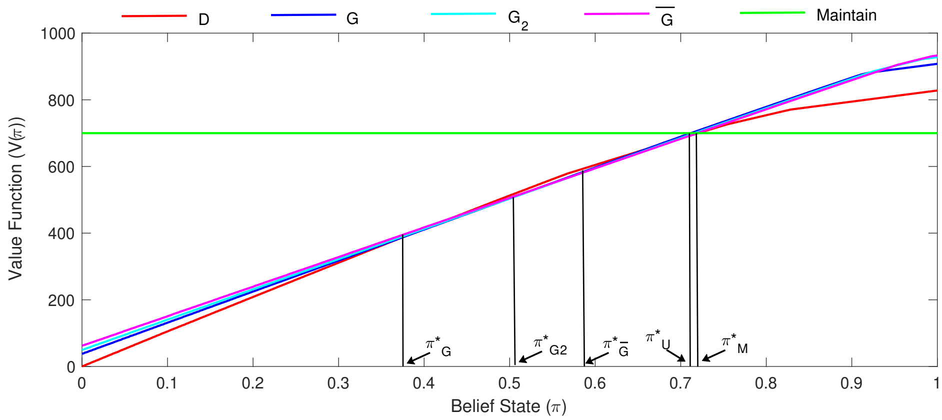

Figure D.14: The optimal policy for the problem instance of Figure D.12, but with sensor costs of 80, 100 and 105, respectively.

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of Second Berkeley Symposium on Mathematical Statistics and Probability* (pp. 93–102). University of California Press.

Chade, H., & Schlee, E. E. (2002). Another look at the radner–stiglitz nonconcavity in the value of information. *Journal of Economic Theory*, *107*, 421–452.

Christer, A. H. (1999). Developments in delay time analysis for modelling plant maintenance. *Journal of the Operational Research Society*, *50*, 1120–1137.

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

De Lara, M., & Gilotte, L. (2007). A tight sufficient condition for radner–stiglitz nonconcavity in the value of information. *Journal of Economic Theory*, *137*, 696–708.

Haarman, M., de Klerk, P., Decaigny, P., Mulders, M., Vassiliadis, C., & Sijtsema, H. (2018). *Predictive maintenance 4.0. beyond the hype: PdM 4.0 delivers results*. Technical Report PricewaterhouseCoopers and Mainnovation.

Karlin, S. (1968). *Total Positivity*. Stanford University Press, Stanford, California.

Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, *14*, 28–44.

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

Lévesque, M., & Maillart, L. M. (2008). Business opportunity assessment with costly, imperfect information. *IEEE Transactions on Engineering Management*, *55*, 279–291.

Lovejoy, W. S. (1987). Some monotonicity results for partially observed Markov decision processes. *Operations Research*, *35*, 736–743.

Maillart, L. M., Mishra, N., & van Staden, H. E. (2018). Data-driven optimal policies for opportunistic maintenance of partially observable systems with economic dependencies. In *Proceedings of Manufacturing and Service Operations Management Conference 2018*. https://www.semanticscholar.org/paper/MSOM-2018-Conference-Data-driven-optimal-policies/d4940d1a2133b7fe74d6096379241f40f225cdff.

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

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The Internet of Things: Mapping the value beyond the hype*. Technical Report McKinsey Global Institute.

Mitchell, H. B. (2007). *Multi-Sensor Data Fusion: An Introduction*. Springer Science & Business Media.

Monahan, E. (1980). Optimal stopping in a partially observable Markov process with costly information. *Operations Research*, *28*, 1319–1334.

Monahan, G. E. (1982). Optimal stopping in a partially observable binary-valued Markov chain with costly perfect information. *Journal of Applied Probability*, *19*, 72–81.

Naderkhani, Z. G., & Makis, V. (2015). Optimal condition-based maintenance policy for a partially observable system with two sampling intervals. *International Journal of Advanced Manufacturing Technologies*, *78*, 795–805.

Ohnishi, M., Kawai, H., & Mine, H. (1986). An optimal inspection and replacement policy under incomplete state information. *European Journal of Operational Research*, *27*, 117–128.

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

Papadimitriou, C. H. (1995). *Computational Complexity*. Addison Wesley.

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

Powell, W. B., & Ryzhov, I. O. (2012). *Optimal learning*. Wiley & Sons.

Radner, R., & Stiglitz, J. E. (1984). A nonconcavity in the value of information. In M. Boyer, & R. E. Kihlstrom (Eds.), *Bayesian Models in Economic Theory* chapter 3. (pp. 33–52). Amsterdam: Eelsevier Science Publishers.

Ross, S. M. (1971). Quality control under Markovian deterioration. *Management Science*, *17*, 587–596.

Saghafian, S., Tomlin, B., & Biller, S. (2018). The Internet of Things and information fusion: Who talks to who? Tuck School of Business Working Paper No. 3124417; HKS Working Paper 18-009. Available at SSRN: https://ssrn.com/abstract=3124417.

Shanthikumar, J. G., & Shaked, M. (1994). *Stochastic Orders and Their Applications*. Boston, MA: Academic.

Tijms, H. C. (1994). *Stochastic models: An algorithmic approach*. Wiley, Chichester.

Uit Het Broek, M., Teunter, R. H., De Jonge, B., Veldman, J., & Van Foreest, N. D. (2019). Condition-based production planning: Adjusting production rates to balance output and failure risk. *Manufacturing and Service Operations Management*, (pp. 1–20).

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

Weibull, J. W., Mattsson, L.-G., & Voorneveld, M. (2007). Better may be worse: Some monotonicity results and paradoxes in discrete choice under uncertainty. *Theory and Decision*, *63*, 121–151.

Whitt, W. (1979). A note on the influence of the sample on the posterior distribution. *Journal of the American Statistical Association*, *74*, 424–426.