

Optimal Replacement of Continuously Degrading Systems in Partially Observed Environments

John A. Flory¹ Jeffrey P. Kharoufeh² David Abdul-Malak³

Department of Industrial Engineering
University of Pittsburgh
1048 Benedum Hall
3700 O'Hara Street
Pittsburgh, PA 15261 USA

To appear in *Naval Research Logistics*
Accepted July 7, 2015

Abstract

Motivated by wind energy applications, we consider the problem of optimally replacing a stochastically degrading component that resides and operates in a partially observable environment. The component's rate of degradation is modulated by the stochastic environment process, and the component fails when its accumulated degradation first reaches a fixed threshold. Assuming periodic inspection of the component, the objective is to minimize the long-run average cost per unit time of performing preventive and reactive replacements for two distinct cases. The first case examines instantaneous replacements and fixed costs, while the second considers time-consuming replacements and revenue losses accrued during periods of unavailability. Formulated and solved are mixed state space, partially observable Markov decision process (POMDP) models, both of which reveal the optimality of environment-dependent threshold policies with respect to the component's cumulative degradation level. Additionally, it is shown that for each degradation value, a threshold policy with respect to the environment belief state is optimal if the environment alternates between two states. The threshold policies are illustrated by way of numerical examples using both synthetic and real wind turbine data.

1 Introduction

Rising energy prices, global climate change, escalating demand for electricity, and global energy supply uncertainties have generated enormous interest in clean, renewable energy sources. Wind energy, generated by land-based and offshore wind turbines, is poised to play a prominent role in a global shift towards alternative energy supplies. However, the cost of producing wind energy remains a significant barrier with operating and maintenance costs contributing as much as 20 to 47.5% of the total cost of energy (see, for example, [32]). These significant costs are attributed primarily to the replacement of major, critical components (e.g., gear boxes, generators or turbine blades) that are subjected to randomly-varying loads due to time-varying wind speeds and dynamic atmospheric conditions (e.g., temperature, humidity, etc.). Discerning the exact state of a wind turbine's environment is complicated by the fact that wind conditions depend on the topography of the immediate surrounding land and the effects of adjacent wind turbines. Furthermore, the wind turbine components themselves are housed in a protective nacelle whose internal conditions may differ substantially from those of the ambient atmosphere. Because the environment state is not known with certainty, a model that accounts for a partially-observed environment, and its influence on the degradation of critical components, is needed. In this paper, we consider the problem of optimally replacing a component whose stochastic degradation process is governed by an exogenous

¹Ph: (505) 284-3469; Email: jaflory@sandia.gov

²Corresponding author. Ph: (412) 624-9832; Email: jkharouf@pitt.edu

³Email: dta10@pitt.edu

random environment, which is assumed to be partially-observed, while the component’s degradation level is known with certainty. Considered are two different replacement models. The first model assumes the replacement costs are deterministic, while the second model incorporates stochastic revenue losses (or downtime costs). We formulate stochastic optimization problems to examine the optimal times at which replacements should occur in order to minimize the long-run average cost per unit time of performing preventive replacements (those done prior to component failure) and reactive replacements (those done in response to unanticipated failures).

The optimal replacement or repair of a stochastically-degrading system is a classical applied probability problem that has received considerable attention in the operations research community. Some important surveys of this topic are provided in [25, 28, 31]. For wind turbine applications, partially observable Markov decision process (POMDP) models are especially useful, as maintenance decisions are often made using either incomplete information or imperfect models of degradation. General POMDP models have found wide applicability in multi-state maintenance optimization problems (cf. [11, 20, 21, 33] and references therein). Maintenance models pertaining to wind turbines in particular are fairly sparse with the exception of Byon et al. [4] who formulated a POMDP model to optimally maintain a wind turbine component whose degradation state evolves as a finite, discrete-time Markov chain (DTMC). Their objective is to minimize the expected per period costs, which are related to both maintenance activities and turbine downtime. That model was extended in Byon and Ding [3] to include multiple components and season-dependent weather conditions. They used a discounted POMDP model to minimize the total expected discounted cost of performing corrective maintenance actions. Using a backward dynamic programming algorithm, they solved the POMDP model and illustrated the optimal policies for a wind turbine gearbox application.

In a more general setting, Makis and Jiang [22] consider optimally replacing a component whose degradation evolves as a continuous-time Markov chain (CTMC) on a finite state space and is observed imperfectly at discrete times. Associated with each observation is a probability mass function conditioned on the true degradation state. The replacement model is formulated as an optimal stopping problem, and the long-run expected replacement cost per unit time is minimized. That work was extended in [16] to include multivariate observations that are normally distributed with mean and covariance matrices determined by the degradation state. A similar model was used by Jiang et al. [13] to maximize the component’s long-run expected availability per unit time. Zhou et al. [35] analyzed a partially observable semi-Markov process (POSMDP) model which is continuous in both time and state. In their model, component degradation evolves as a Gamma-based state space model. A Monte Carlo density projection procedure (see [34]) was used to reduce the infinite-dimensional belief space to a finite-dimensional space so as to formulate the problem as a Markov decision process (MDP) model on the lower-dimensional space. The MDP model was solved numerically using policy iteration for both long-run average cost and availability objectives.

Despite the importance of the operating environment, relatively few replacement models take into account the environment’s impact on accelerating (or decelerating) degradation over time. Çekyay and Özekici [5] surveyed condition-based maintenance models that lead to structured control-limit (or threshold) policies. Waldmann [30] was the first to analyze the structure of an optimal replacement policy for a system subjected to stochastic deterioration in a random environment. He considered the effects of uncontrollable internal and external factors on the progression of the system’s deterioration status in a continuous-time shock model and derived sufficient conditions for the optimality of a threshold policy with respect to the cumulative damage of the system. Özekici [23] modeled the deterioration of a system in an uncontrollable environment by its intrinsic age, where the environment was assumed to evolve as a semi-Markov jump process and the intrinsic age of the system is determined by the total cumulative degradation. They showed that, if the system’s

lifetime distribution has an increasing failure rate in each environment state, then a threshold-type policy, with respect to the degradation state, is optimal. Kurt and Maillart [18] examined the optimal replacement of a system that fails due to random shocks arriving according to a Poisson process whose rate is modulated by a discrete-time Markov chain. They consider both controllable and uncontrollable Markovian environments and analyze the structure of the resulting optimal cost function with respect to the shock arrival rate and the cumulative number of shocks received. Kurt and Kharoufeh [17] extended that work by relaxing the fixed cost assumption. Relevant to our work here is a model analyzed by Ulukus et al. [29], who considered the problem of optimally replacing a component whose operating environment evolves as a CTMC on a finite state space \mathcal{S} . When the environment is in state i , the component accumulates degradation at a constant rate r_i ($r_i > 0$) for $i \in \mathcal{S}$. In their model, both the cumulative degradation and the environment state are assumed to be completely observable, and at each inspection time, a decision maker may choose to either preventively replace the component or do nothing until the next inspection time. The problem was formulated using a MDP model with the objective of minimizing the total expected discounted cost of performing preventive and reactive replacements over an infinite time horizon. It was shown that, for each environment state, there exists a degradation threshold above which it is always optimal to preventively replace the component.

In this paper, we consider the problem of optimally replacing a component whose observable degradation is influenced by its operating environment, which is only partially observable. The degradation dynamics are similar to those described in [29], but our optimization problem is complicated by the partial observability of the environment. We formulate and solve two problems using a POMDP modeling framework. The first model assumes that preventive and reactive replacements are performed instantaneously with fixed, deterministic costs, whereas the second model considers time-consuming replacements during which revenue losses are accrued at a rate that depends on the environment state. The objective in both models is to minimize the long-run average cost per unit time of performing preventive and reactive replacements. For each model, we prove the existence and optimality of a threshold-type replacement policy with respect to the cumulative degradation level whose thresholds depend on the decision maker's belief of the current environment state. Furthermore, in the case of deterministic replacement costs, the optimality of a threshold policy, with respect to the environment belief state, is proved for each possible degradation value when the environment alternates between two states. Numerical solution techniques are presented to compute optimal policies, and numerical examples using notional and real wind turbine data are presented to illustrate the replacement policies.

The remainder of the paper is organized as follows. Section 2 describes the stochastic degradation process and the general POMDP framework of both replacement models. In Sections 3 and 4, the replacement models and their associated optimality equations are presented for fixed replacement costs and stochastic revenue losses, respectively. Numerical techniques for approximating the optimal policies of the models are described in Section 5, while Section 6 presents numerical examples to illustrate the replacement policies.

2 Problem Formulation

Before presenting the general problem formulation that forms the basis of the models in Sections 3 and 4, we first review the stochastic degradation model analyzed by Ulukus et al. [29]. Consider a component that is placed into service at time $t = 0$ in new condition. The component degrades over time due to normal usage and the influence of its operating environment. Once the cumulative degradation exceeds a critical threshold value ξ ($\xi > 0$), the component is considered to be failed. For each $t \geq 0$, the environment can occupy one (and only one) state in the set $\mathcal{S} = \{1, \dots, \ell\}$

where $\ell < \infty$. Let $Z(t)$ be the state of the environment at time t , and suppose $\{Z(t) : t \geq 0\}$ is an \mathcal{S} -valued, continuous-time Markov chain (CTMC) with infinitesimal generator matrix $\mathbf{Q} = [q_{ij}]$, $i, j \in \mathcal{S}$. The CTMC is assumed to be irreducible; therefore, it possesses a unique, positive limiting distribution vector π_s that satisfies $\pi_s \mathbf{Q} = \mathbf{0}$ and $\pi_s \mathbf{e} = 1$, where $\mathbf{0}$ is the $1 \times \ell$ zero-vector and \mathbf{e} is an $\ell \times 1$ vector of ones. Let $r : \mathcal{S} \rightarrow (0, \infty)$ be a real-valued, (Borel) measurable function such that whenever $Z(t) = i \in \mathcal{S}$, the component degrades linearly at a unique, constant rate r_i ($r_i > 0$). For convenience, assume $r_1 < r_2 < \dots < r_\ell$ and define $\mathbf{r} = (r_1, \dots, r_\ell)$, the degradation rate vector. We pause here to note that this modeling framework does not assume that the component degrades linearly overall; rather, it is assumed that the rate of degradation is constant within a given environment state. Therefore, the degradation process can be characterized by several mean rates of degradation as a function of time (or usage). Variation about these mean degradation rates certainly exists, but if the times between transitions are short relative to the overall time to failure, or if there is ample evidence to suggest a certain growth pattern, this approach provides a great deal of modeling flexibility, as it allows one to estimate the degradation rate associated with each state. In practice, it may also be necessary to estimate the number of distinct environment states (ℓ) to which the component might be exposed. If the number of environment states is large, aggregation of the states may be necessary. Two approaches have been suggested for aggregating and estimating the number of states. Those include: (1) clustering the degradation rates, as proposed by Kharoufeh and Cox [14], and (2) using the Bayesian inference criterion (BIC), as was done by Flory et al. [8]. Numerical illustrations in those papers demonstrate that only a few states (usually not more than three) are needed to adequately approximate the modulating environment process.

Denote by $X(t)$ the cumulative degradation of the component at time t , and let $T(\xi) \equiv \inf\{t > 0 : X(t) \geq \xi\}$ be the first passage time of the process $\{X(t) : t \geq 0\}$ to the critical threshold ξ . The cumulative degradation at time t is

$$X(t) = X(0) + \int_0^t r_{Z(u)} du, \quad (1)$$

where we assume $\mathbb{P}(X(0) = 0) = 1$ and

$$\mathbb{P}\left(\int_0^t |r_{Z(u)}| du < \infty\right) = 1$$

to ensure $X(t)$ is well-defined. The strict positivity of \mathbf{r} ensures that the sample paths of $\{X(t) : t \geq 0\}$ are monotone increasing with probability 1 (w.p. 1). The following additional notation will be used throughout. All random variables are defined on a common and complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For an event $A \in \mathcal{F}$, denote by $\mathbb{I}(A)$ the indicator function with $\mathbb{I}(A) = 1$ if A occurs and $\mathbb{I}(A) = 0$ otherwise. For any $a, b \in \mathbb{R}$, let $a \wedge b \equiv \min\{a, b\}$ and $a \vee b \equiv \max\{a, b\}$.

The component is inspected (or observed) periodically at times in the set $I = \{k\delta : k \in \mathbb{N}\}$ for some $\delta > 0$, where each time in I represents a decision epoch. A *period* is the time between two consecutive decision epochs. For decision epoch $n \in I$, let X_n be the cumulative degradation of the component and Z_n be the state of the environment at this time. Our models assume that X_n is completely observable (i.e., degradation can be discerned with certainty), but that Z_n is only partially observable through a probability distribution (i.e., the environment state is inferred from an observed degradation increment and represented using the concept of a belief state). For example, some forms of wind turbine degradation are fully observable, such as the length of a crack on a turbine blade, or the oil contamination level in the wind turbine's gearbox, whereas other forms can be discerned through a signal of degradation, such as a vibration signal, as in Gebraeel

et al. [9]. Furthermore, there is ample evidence to suggest that the environment is only partially observable. For instance, it is well known that several factors contribute to the random evolution of the environment, including wind speed, wind turbulence, ambient air temperature, humidity, and the degradation levels of other components within the wind turbine. While many of these environment conditions can be monitored, measurement errors and the inability to measure these conditions at all points spatially and temporally leads to partial observability. Indeed, characterizing the operating environment of a wind turbine is known to be very challenging, as noted by Dueñas-Osorio and Basu [7] and Peinke et al. [24]. Immediately following an inspection, a decision maker chooses one of two feasible actions in the set $\mathcal{A} = \{0, 1\}$, where action 0 means “do nothing,” and action 1 means “preventively replace” the component. Inspections are assumed to be performed instantaneously at a fixed cost c_0 ($c_0 > 0$). If an inspection reveals that the system is not failed, the decision maker can perform a preventive replacement at a fixed cost c_1 where $0 < c_0 \ll c_1 < \infty$. On the other hand, the decision maker may elect to do nothing until the start of the next period. Should the component fail during any period (between two consecutive decision epochs), it is immediately, reactively replaced at a fixed cost $c_1 + c_2$ where $c_2 > 0$ is a penalty cost for unplanned replacements. For the model in Section 3, both preventive and reactive replacements are assumed to be performed instantaneously. However, in Section 4, this assumption is relaxed, and the penalty cost c_2 is replaced by stochastic downtime costs (revenue losses) that are incurred due to time-consuming replacements. Whenever a preventive or reactive replacement occurs, the cumulative degradation level is reset to 0 at the start of the next period. The objective is to minimize the long-run average cost per unit time of performing preventive and reactive replacements.

To address this problem, we use a partially observable Markov decision process (POMDP) model. The states of the POMDP model are vectors of the form (x, π) , where $x \in [0, \xi]$ is the component’s degradation level and $\pi \in \mathbb{R}_+^\ell$ is the belief state of the environment whose belief space is the ℓ -dimensional probability simplex

$$\Pi \equiv \left\{ \left[\pi^{(1)}, \dots, \pi^{(\ell)} \right] : \sum_{i \in \mathcal{S}} \pi^{(i)} = 1 \right\}.$$

The belief space of the POMDP model is $B = [0, \xi] \times \Pi$. At the n th decision epoch, x_n is the observed degradation, and the belief state of the environment is the probability vector $\pi_n = [\pi_n^{(1)}, \dots, \pi_n^{(\ell)}]$, where

$$\pi_n^{(i)} = \mathbb{P}(Z_n = i | \Delta X_n, \pi_{n-1}), \quad i \in \mathcal{S},$$

and $\Delta X_n \equiv X_n - X_{n-1}$ is the (random) degradation increment during the n th period. Immediately following the observation of a degradation increment, the belief state is recursively updated. Let $W_{ij}(x, t) = \mathbb{P}(X(t) \leq x, Z(t) = j | Z(0) = i)$ for $i, j \in \mathcal{S}$ and $t \geq 0$, and define the probability density function

$$w_{ij}(u) \equiv \left. \frac{\partial W_{ij}(x, \delta)}{\partial x} \right|_{x=u}$$

where $w_{ij}(u)$ is defined for all $u \in (r_1 \delta, r_\ell \delta)$ such that $u \neq r_k \delta$ for all $k \in \mathcal{S}$. For $j \in \mathcal{S}$, define

$$T_j(u, \pi) \equiv \mathbb{P}(Z_{n+1} = j | \Delta X_n = u, \pi_n = \pi) = \sum_{i \in \mathcal{S}} \mathbb{P}(Z_{n+1} = j | \Delta X_n = u, Z_n = i) \pi^{(i)},$$

where

$$\mathbb{P}(Z_{n+1} = j | \Delta X_n = u, Z_n = i) = \begin{cases} \frac{w_{ij}(u)}{\sum_{k \in \mathcal{S}} w_{ik}(u)}, & u \in (r_1 \delta, r_\ell \delta) \text{ and } u \neq r_k \delta, \forall k \in \mathcal{S}, \\ \mathbb{I}(j = k), & u = r_k \delta \text{ for some } k \in \mathcal{S}. \end{cases}$$

The quantity $T_j(u, \pi)$ is the probability that the environment is in state j at the next decision epoch, given an initial belief state π and an observed degradation increase $u = x_{n+1} - x_n$ over the period. Given $\pi_n = \pi$ and a realized degradation increment u , it follows that $\pi_{n+1} = \mathbf{T}(u, \pi)$, where $\mathbf{T}(u, \pi) \equiv [T_1(u, \pi), T_2(u, \pi), \dots, T_\ell(u, \pi)]$.

Let the belief state at the n th decision epoch be denoted by $b^{(n)} = (x_n, \pi_n)$, $n \in I$. For two states in B , (x, π) and (x', π') , define the conditional probability

$$K((x, \pi), (x', \pi')) \equiv \mathbb{P}(\Delta X_{n+1} \leq x' - x, \pi_{n+1} = \pi' | \pi_n = \pi).$$

Note that $K((x, \pi), (x', \pi')) = K((0, \pi), (x' - x, \pi'))$, where for $u \geq 0$,

$$K((0, \pi), (u, \pi')) = \begin{cases} \int_0^u \sum_{i \in \mathcal{S}} q(v|i) \pi^{(i)} dv + \sum_{i \in \mathcal{S}} \mathbb{I}(u \geq r_i \delta) \pi^{(i)} \exp(q_{ii} \delta), & \pi' = \mathbf{T}(u, \pi), \\ 0, & \pi' \neq \mathbf{T}(u, \pi), \end{cases}$$

and $q(v|i) \equiv \sum_{j \in \mathcal{S}} w_{ij}(v)$. Denote the transition kernel density between $(x, \pi) \in B$ and $(x', \pi') \in B$ as

$$k(x' - x, \pi) \equiv \left. \frac{\partial}{\partial u} K((0, \pi), (u, \pi')) \right|_{u=x'-x}$$

where for $u \geq 0$,

$$k(u, \pi) = \begin{cases} \sum_{j \in \mathcal{S}} q(u|j) \pi^{(j)}, & u \neq r_i \delta \text{ for all } i \in \mathcal{S}, \\ \pi^{(i)} \exp(q_{ii} \delta), & u = r_i \delta \text{ for some } i \in \mathcal{S}. \end{cases}$$

In Sections 3 and 4, we consider the POMDP model formulation under two different cost structures, respectively. The first model excludes the cost of downtime due to replacements, while the second includes this cost explicitly. For both models, it is shown that a threshold replacement policy is optimal with respect to the component's cumulative level of degradation, and these thresholds depend on the decision maker's assessment of the environment state.

3 Replacement with Fixed Costs

For this first model, as described in Section 2, inspection and preventive and reactive replacements are assumed to be instantaneous and impose fixed costs. All components are assumed to begin operation in the belief state $(0, \pi_s)$. Define a *policy* as a function $a : B \rightarrow \mathcal{A}$, where $a(X_n, \pi_n)$ is the action taken in state $(X_n, \pi_n) \in B$, and let \mathcal{P} denote the set of all possible policies. The objective is to find the policy a^* that minimizes the long-run average cost of replacements per unit time, γ , given by

$$\gamma = \inf_{a \in \mathcal{P}} \mathbb{E}_a \left\{ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N c_0 + c_1 \mathbb{I}\{a(X_n, \pi_n) = 1\} + (c_1 + c_2) \mathbb{I}\{a(X_n, \pi_n) = 0, \widehat{H}(X_n, \pi_n) = 1\} \right\}$$

where $\widehat{H}(X_n, \pi_n)$ is an indicator for the event that the component fails between decision epochs n and $n + 1$, given $(X_n, \pi_n) \in B$.

The optimality equations are now provided. Let $V(x, \pi)$ be the minimum relative cost per unit time, given that a component starts operation in $(x, \pi) \in B$, and define $V_0(x, \pi)$ and $V_1(x, \pi)$ as the relative costs if either no action or preventive replacement, respectively, are taken in $(x, \pi) \in B$.

The expected survival time of the component in the next period, given $(x, \pi) \in B$, is denoted $\tau(x, \pi)$, where

$$\tau(x, \pi) = \sum_{i \in \mathcal{S}} \left[\sum_{j \in \mathcal{S}} \int_0^\delta W_{ij}(\xi - x, t) dt \right] \pi^{(i)}.$$

The optimality equations are

$$V(x, \pi) = \min\{V_1(x, \pi), V_0(x, \pi)\}, \quad (x, \pi) \in B, \quad (2)$$

where for $\mathbb{I}^+(x, u) \equiv \mathbb{I}(x + u \geq \xi)$, $\mathbb{I}^-(x, u) \equiv \mathbb{I}(x + u < \xi)$, and $V_\pi(x, u) \equiv V(x + u, \mathbf{T}(u, \pi))$,

$$V_1(x, \pi) = c_0 + c_1 + V(0, \pi_s) \quad (3)$$

$$V_0(x, \pi) = c_0 + \int_0^\infty \mathbb{I}^+(x, u) [c_1 + c_2 + V(0, \pi_s)] k(u, \pi) du + \int_0^\infty \mathbb{I}^-(x, u) V_\pi(x, u) k(u, \pi) du - \gamma \tau(x, \pi). \quad (4)$$

An exact analytical solution for the optimality equations (2) is attainable only in trivial cases. Nonetheless, it is possible to prove some basic structural results to characterize the optimal replacement policy.

3.1 Structural Results

Here, we examine attributes of the cost function and establish structural results that lead to a characterization of the optimal policy in Section 3.2. For $(x, \pi) \in B$, let

$$H(x, \pi) \equiv \int_0^\xi \mathbb{I}^+(x, u) k(u, \pi) du$$

be the probability that the component fails in the next period starting in state (x, π) , i.e., the probability that the component's cumulative degradation exceeds the failure threshold ξ prior to the next decision epoch, given the current belief state is (x, π) . The first result provides a nonnegative lower bound for the optimal average cost γ .

Lemma 1. *The average cost of an optimal policy is bounded below as follows:*

$$\gamma > \frac{c_0}{\delta}. \quad (5)$$

Proof. The lower bound can be established by considering the average cost of a policy for the case when $c_1 = c_2 = 0$. Note that, immediately following replacement, it is optimal to do nothing; otherwise, $\gamma = \infty$. Set $V(0, \pi_s) = 0$ and observe that $V(0, \pi_s) = V_0(0, \pi_s) = 0$. This implies that

$$c_0 + (c_1 + c_2)H(0, \pi_s) + \int_0^\infty \mathbb{I}^-(0, u) V_\pi(0, u) k(u, \pi) du - \gamma \tau(0, \pi_s) = 0. \quad (6)$$

Solving (6) for γ gives

$$\gamma = \frac{c_0 + (c_1 + c_2)H(0, \pi_s) + \int_0^\infty \mathbb{I}^-(0, u) V_\pi(0, u) k(u, \pi) du}{\tau(0, \pi_s)} > \frac{c_0}{\delta}. \quad \square$$

The lower bound (5) is not tight; however, its existence is useful to establish basic properties of the optimal policy. Lemma 2 bounds $V(x, \pi)$ from above.

Lemma 2. For all $(x, \pi) \in B$,

$$V(x, \pi) \leq C_0 \equiv c_0 + c_1 + V(0, \pi_s).$$

Proof. For an arbitrary state $(x, \pi) \in B$,

$$\begin{aligned} V(x, \pi) &= \min\{V_0(x, \pi), V_1(x, \pi)\} \\ &\leq V_1(x, \pi) \\ &= c_0 + c_1 + V(0, \pi_s). \quad \square \end{aligned}$$

To facilitate the discussion that follows, let the subset of B where preventive replacement is optimal be denoted by $D = \{(x, \pi) \in B : V_1(x, \pi) \leq V_0(x, \pi)\}$, and let D^c be its complement. Additionally, let $D_\pi \equiv \{x : (x, \pi) \in D\}$ and $D_x \equiv \{\pi : (x, \pi) \in D\}$. Characterizing the optimal policy is tantamount to describing the structure of the region D ; however, it is difficult to determine properties of D , such as convexity or even connectedness. Fortunately, Theorem 1 and Corollary 1 provide bounds on the region.

Theorem 1. If the component survives the next period w.p. 1 for $(x, \pi) \in B$, then $(x, \pi) \in D^c$.

Proof. If the component survives in the next period w.p. 1 for $(x, \pi) \in B$, then

$$\begin{aligned} V_0(x, \pi) &= c_0 + \int_0^\infty I^-(x, u) V_\pi(x, u) k(u, \pi) du - \gamma\delta \\ &\leq c_0 + \int_0^\infty I^-(x, u) [c_0 + c_1 + V(0, \pi_s)] k(u, \pi) du - \gamma\delta \quad (\text{by Lemma 2}) \\ &= 2c_0 + c_1 + V(0, \pi_s) - \gamma\delta \\ &< c_0 + c_1 + V(0, \pi_s) \\ &= V_1(x, \pi). \end{aligned}$$

Therefore, $(x, \pi) \in D^c$. □

Intuitively, preventive replacement is never optimal if the likelihood of component failure before the next decision epoch is zero. Theorem 1 leads immediately to Corollary 1, which provides some insights into the structure of the set D .

Corollary 1. Let $D_\delta \equiv \{(x, \pi) \in B : x \in (\xi - r_\ell \delta, \xi]\}$. Then $D \subseteq D_\delta$.

Proof. For all $(x, \pi) \in B$ such that $x \in [0, \xi - r_\ell \delta]$, the component will survive the next period with certainty since the maximum degradation increment cannot exceed $r_\ell \delta$ in the next period. Hence, $D_\delta^c = \{(x, \pi) \in B : x \in [0, \xi - r_\ell \delta]\} \subseteq D^c$. The result follows by taking the complements of both sets. □

Corollary 1 confines D to a subset of B where $x \geq \xi - r_\ell \delta$ so that the probability of failure before the next decision epoch is nonzero. Figure 1 depicts this bound in the case of a two-state environment.

It is tempting to conclude that a sufficient condition for the optimality of preventive replacement in (x, π) is the certainty of failure in the next period; however, a stronger condition is required to account for pathological cases, e.g., the inspection interval is longer than the component lifetime. A sufficient condition for the optimality of preventive replacement is provided in Theorem 2, which requires the next lemma.

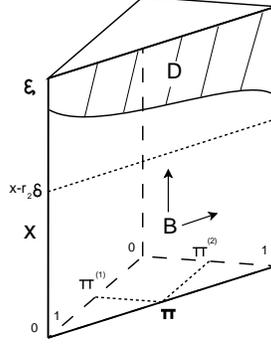


Figure 1: Depiction of boundary for preventive replacement region D .

Lemma 3. *An upper bound for γ is*

$$\gamma \leq c_0 + (c_1 + c_2) \frac{r_\ell}{\xi}. \quad (7)$$

Proof. Applying Lemma 1 of [15], it can be shown that $T(\xi) \geq \xi/r_\ell$ w.p. 1. Therefore, an upper bound for the optimal cost can be obtained by considering the average cost of allowing a component to operate continuously at rate r_ℓ (without intervention) until failure. In such a case,

$$\gamma \leq \frac{\left\lfloor \frac{\xi}{r_\ell} \right\rfloor c_0 + c_1 + c_2}{\frac{\xi}{r_\ell}} \leq c_0 + (c_1 + c_2) \frac{r_\ell}{\xi}. \quad \square$$

Similar to the lower bound (5), the upper bound (7) is not particularly tight; however, its existence is useful to establish a sufficient condition for the optimality of preventive replacement.

Theorem 2. *If the component fails w.p. 1 in the next period for $(x, \pi) \in B$, then it is optimal to preventively replace if*

$$\tau(x, \pi) \leq \frac{c_2 \xi}{c_0 + (c_1 + c_2)r_\ell}. \quad (8)$$

Proof. Suppose failure is certain in the next period. Then

$$\begin{aligned} V_0(x, \pi) &= c_0 + \int_0^\infty \mathbb{I}^+(x, u) [c_1 + c_2 + V(0, \pi_s)] k(u, \pi) du \\ &\quad + \int_0^\infty \mathbb{I}^-(x, u) V_\pi(x, u) k(u, \pi) du - \gamma \tau(x, \pi) \\ &= c_0 + c_1 + c_2 + V(0, \pi_s) - \gamma \tau(x, \pi). \end{aligned}$$

Preventive replacement is optimal if $V_0(x, \pi) \geq V_1(x, \pi)$, or equivalently,

$$c_0 + c_1 + c_2 + V(0, \pi_s) - \gamma \tau(x, \pi) \geq c_0 + c_1 + V(0, \pi_s),$$

which implies $c_2 - \gamma \tau(x, \pi) \geq 0$. From Lemma 3, it follows that

$$c_2 - \gamma \tau(x, \pi) \geq c_2 - \left[c_0 + (c_1 + c_2) \frac{r_\ell}{\xi} \right] \tau(x, \pi).$$

Hence, a sufficient condition for preventive replacement is

$$c_2 - \left[c_0 + (c_1 + c_2) \frac{r_\ell}{\xi} \right] \tau(x, \pi) \geq 0,$$

or equivalently,

$$\tau(x, \pi) \leq \frac{c_2 \xi}{c_0 + (c_1 + c_2) r_\ell}. \quad \square$$

3.2 Optimal Policy for Fixed π

The main result of this section is Theorem 3, which asserts the optimality of a threshold policy with respect to the cumulative degradation level x for each $\pi \in \Pi$. That is, for each belief state of the environment, there exists a control limit above which it is always optimal to preventively replace the component. To help prove this result, the next lemma establishes bounds for $V_0(x, \pi)$.

Lemma 4. *For each $(x, \pi) \in B$, $V_0(x, \pi)$ is bounded as follows:*

$$V_0(x, \pi) \leq V_0(x, \pi) \leq \bar{V}_0(x, \pi), \quad (9)$$

where

$$V_0(x, \pi) = (C_0 - c_1) + (c_1 + c_2)H(x, \pi) - \gamma\tau(x, \pi), \quad (10)$$

$$\bar{V}_0(x, \pi) = c_0 + C_0 + (c_2 - c_0)H(x, \pi) - \gamma\tau(x, \pi). \quad (11)$$

Proof. By Lemma 2, $V_\pi(x, u) \leq C_0$ for all $u \geq 0$. Let $M \equiv c_1 + c_2 + V(0, \pi_s)$, then

$$\begin{aligned} V_0(x, \pi) &= c_0 + \int_0^\infty \mathbb{I}^+(x, u)[c_1 + c_2 + V(0, \pi_s)]k(u, \pi)du \\ &\quad + \int_0^\infty \mathbb{I}(x + u < \xi)V_\pi(x, u)k(u, \pi)du - \gamma\tau(x, \pi) \\ &= c_0 + M \int_0^\infty \mathbb{I}^+(x, u)k(u, \pi)du \\ &\quad + \int_0^\infty \mathbb{I}(x + u < \xi)V_\pi(x, u)k(u, \pi)du - \gamma\tau(x, \pi) \\ &\leq c_0 + H(x, \pi)M + C_0[1 - H(x, \pi)] - \gamma\tau(x, \pi) \\ &= c_0 + C_0 + (M - C_0)H(x, \pi) - \gamma\tau(x, \pi) \\ &= c_0 + C_0 + (c_2 - c_0)H(x, \pi) - \gamma\tau(x, \pi). \end{aligned}$$

Likewise, $V(x, \pi) \geq V(0, \pi_s)$ for all $\pi \in \Pi$. Therefore,

$$\begin{aligned} V_0(x, \pi) &= c_0 + M \int_0^\infty \mathbb{I}^+(x, u)k(u, \pi)du \\ &\quad + \int_0^\infty \mathbb{I}(x + u < \xi)V_\pi(x, u)k(u, \pi)du - \gamma\tau(x, \pi) \\ &\geq c_0 + M \int_0^\infty \mathbb{I}^+(x, u)k(u, \pi)du \\ &\quad + V(0, \pi_s) \int_0^\infty \mathbb{I}(x + u < \xi)k(u, \pi)du - \gamma\tau(x, \pi) \\ &= c_0 + H(x, \pi)M + (C_0 - c_0 - c_1)[1 - H(x, \pi)] - \gamma\tau(x, \pi) \\ &= (C_0 - c_1) + [M - V(0, \pi_s)]H(x, \pi) - \gamma\tau(x, \pi) \\ &= (C_0 - c_1) + (c_1 + c_2)H(x, \pi) - \gamma\tau(x, \pi). \quad \square \end{aligned}$$

Although the monotonicity of $V_0(x, \pi)$ in x , and by extension $V(x, \pi)$, are not guaranteed, the lower bound (10) and upper bound (11) are both monotone increasing in x . Therefore, as x increases, it can be shown that $\bar{V}_0(x, \pi)$ will eventually exceed $V_1(x, \pi)$ (which is constant in x), implying the existence of a threshold of cumulative degradation above which preventive replacement is optimal. A depiction of the relationship between $V(x, \pi)$ and the bounds of Lemma 4 is provided in Figure 2.

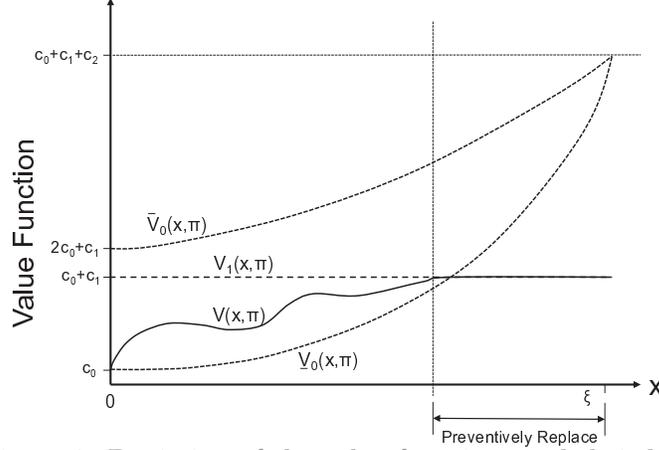


Figure 2: Depiction of the value functions and their bounds.

The next lemma states that if preventive replacement is optimal for some degradation level in a fixed environment state, then preventive replacement is also optimal for all higher degradation levels. The lemma is particularly useful because it establishes that a single interval comprises D_π .

Lemma 5. *For each $\pi \in \Pi$, if $(x, \pi) \in D$, then $(x', \pi) \in D$ for all $x' \geq x$.*

Proof. For π fixed, assume that $(x, \pi) \in D$ and there exists some $x^* > x$ such that $(x^*, \pi) \in D^c$. Let $B_{(x^*, \pi)}^D$ be the subset of B that is reachable from (x^*, π) under the policy defined by D prior to a replacement event. Consider a policy defined by D^* that is obtained by modifying D as follows:

- (i) Set $D^* = D \setminus \{(x, \pi)\}$;
- (ii) For all $(y, \pi') \in B_{(x^*, \pi)}^D$, set $D^* = D^* \cup \{(y + x - x^*, \pi')\}$ if $(y, \pi') \in D$; otherwise, set $D^* = D^* \setminus \{(y + x - x^*, \pi')\}$.

Note that, prior to a replacement event, the evolution of two sample paths starting from (x, π) and (x^*, π) under the policy defined by D^* differ only through a translation of degradation levels by $x^* - x$. For any sample path starting at (x^*, π) that exceeds the threshold at some time t , the corresponding sample path starting from (x, π) will not exceed the threshold at time t . Therefore, $\gamma_{D^*} < \gamma_D$, and the policy defined by D cannot be optimal. This implies that $(x', \pi) \in D$ for all $x' \geq x$. \square

Using the lower bound of (9) and Lemma 5, the following theorem establishes the existence of a threshold policy with respect to x for a given belief state π .

Theorem 3. *For each $\pi \in \Pi$, there exists a threshold x_π ($x_\pi < \xi$) at which it is optimal to preventively replace for all $x \geq x_\pi$.*

Proof. A sufficient condition to replace in $(x, \pi) \in B$ is that $\underline{V}_0(x, \pi) \geq V_1(x, \pi)$. Therefore, $\underline{V}_0(x, \pi) \geq V_1(x, \pi)$ implies that

$$(C_0 - c_1) + (c_1 + c_2)H(x, \pi) - \gamma\tau(x, \pi) \geq C_0,$$

which implies

$$(c_1 + c_2)H(x, \pi) - \gamma\tau(x, \pi) \geq c_1. \quad (12)$$

Note that the left-hand side of inequality (12) is increasing in x as $H(x, \pi)$ is increasing in x and $\gamma\tau(x, \pi)$ is decreasing in x . Furthermore, the inequality is strict at the point (ξ, π) for all $\pi \in \Pi$ since $H(\xi, \pi) = 1$ and $\tau(\xi, \pi) = 0$. Let $x' = \inf\{x : (c_1 + c_2)H(x, \pi) - \gamma\tau(x, \pi) \geq c_1\}$. Then $x' < \xi$ and (12) is satisfied for all $x \geq x'$. Therefore, $\{x : x \geq x'\} \subseteq D_\pi$. Moreover, it follows by Lemma 5 that $D_\pi = \{x : x \geq x_\pi\}$ for some $x_\pi \leq \xi$. However, since $\{x : x \geq x'\} \subseteq D_\pi$, it follows that $x_\pi < \xi$. \square

3.3 Optimal Policy for Fixed x

Subsection 3.2 established the fact that, for each belief state π , there exists a threshold degradation level, x_π , above which preventive replacement is always optimal. It is natural to inquire whether an optimal threshold policy exists with respect to the belief state for each degradation level. That is, for each degradation level x ($x \in [0, \xi)$), is there a threshold belief state π_x for which it is optimal to replace the component whenever the belief state is worse (in some sense) than π_x ? To answer this question definitively, it is necessary to characterize the set $D_x = \{\pi : (x, \pi) \in D\}$, which is difficult to do in general. However, the following proposition provides at least one attribute of D_x – namely that it is an increasing set in x .

Proposition 1. *For all x and x' such that $x \leq x'$, $D_x \subseteq D_{x'}$.*

Proof. Consider $\pi \in D_x$ for $x \geq x_\pi$. Then for $x' \geq x$, $(x', \pi) \in D$ by Theorem 3, which implies that $\pi \in D_{x'}$. Therefore, $D_x \subseteq D_{x'}$. \square

Proposition 1 asserts that the subset of Π in which preventive replacement is optimal for a given degradation level x will tend to grow larger as x increases. However, characterizing this subset is nontrivial.

The main result of this subsection (Theorem 4) establishes the optimality of a threshold policy for each fixed x in a particular case. Specifically, for the case when $\ell = 2$, it is possible to establish a threshold policy with respect to the belief state π , provided that \mathbf{Q} and \mathbf{r} satisfy certain mild conditions. Assume the environment has infinitesimal generator \mathbf{Q} and degradation rate vector \mathbf{r} given by

$$\mathbf{Q} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \quad \mathbf{r} = [r_0 \quad r_1],$$

where $\alpha > \beta > 0$ and $r_1 > r_0 > 0$. In what follows, we parameterize the belief state so that $\pi(\rho) \equiv [1 - \rho, \rho] \in \Pi$ for $0 \leq \rho \leq 1$.

Theorem 4. *For $x \in [0, \xi]$ fixed, if it is optimal to replace in state $(x, \pi(\rho))$, then it is optimal to replace for all $(x, \pi(\rho')) \in B$ such that $\rho \leq \rho' \leq 1$.*

Proof. The detailed proof is provided in the Appendix. \square

Theorem 4 asserts that, for fixed x , if preventive replacement is optimal in a particular belief state, then it is also optimal for any belief state that assigns a higher probability to the more detrimental environment state.

For environments in which $\ell > 2$, it is natural to order the belief states via standard stochastic orders. Two of the most common stochastic orders are the usual stochastic ordering and the likelihood ratio ordering of two belief states π and π' . The following definitions are adopted from Maillart and Zheltova [21].

Definition 1. *Suppose π and π' are two belief states in the ℓ -dimensional simplex Π . Then, π' is larger than π in the usual stochastic order if and only if, for all $m = 1, 2, \dots, \ell$,*

$$\sum_{i=m}^{\ell} \pi^{(i)} \leq \sum_{i=m}^{\ell} \pi'^{(i)}.$$

In such case, we write $\pi \preceq_{\text{st}} \pi'$. Belief state π' is larger than π in the likelihood ratio order if and only if

$$\pi^{(i)} \pi'^{(j)} \geq \pi^{(j)} \pi'^{(i)}$$

for all $j \geq i$. In such case, we write $\pi \preceq_{\text{lr}} \pi'$.

Next, we will show that when $\ell > 2$, the existence of an optimal threshold policy with respect to π , for each fixed degradation level x , cannot be guaranteed. To that end, note that, in either the usual stochastic or likelihood ratio orders, the ℓ -dimensional simplex Π has a largest element, namely $\mathbf{e}_\ell \equiv [0 \ 0 \ \dots \ 1]$, where \mathbf{e}_i denotes the i th unit vector for each $i \in \mathcal{S}$. We formalize this fact in Proposition 2.

Proposition 2. *Under the usual stochastic order and the likelihood ratio order, $\mathbf{e}_\ell \equiv [0 \ 0 \ \dots \ 1]$ is the largest element in the probability simplex*

$$\Pi = \left\{ \left[\pi^{(1)} \ \dots \ \pi^{(\ell)} \right] : \sum_{i=1}^{\ell} \pi^{(i)} = 1 \right\}.$$

More specifically, there does not exist any $\pi \in \Pi$ where $\pi \neq \mathbf{e}_\ell$ such that $\pi \succeq_{\text{st}} \mathbf{e}_\ell$ or $\pi \succeq_{\text{lr}} \mathbf{e}_\ell$.

Proposition 3 establishes conditions under which the replacement region is non-empty but no threshold belief state π exists for at least one $x \in [0, \xi)$. In what follows, $x_{\hat{\pi}}$ and $x_{\mathbf{e}_\ell}$ denote the degradation thresholds corresponding to belief states $\hat{\pi}$ and \mathbf{e}_ℓ , respectively.

Proposition 3. *If for some $\hat{\pi} \in \Pi$ such that $\hat{\pi} \neq \mathbf{e}_\ell$ we have that $x_{\hat{\pi}} < x_{\mathbf{e}_\ell} < \xi$, then*

1. $D_{x_{\hat{\pi}}} \neq \emptyset$;
2. *For degradation level $x_{\hat{\pi}}$, there does not exist a threshold $\pi_{x_{\hat{\pi}}} \in \Pi$ such that, for all $\pi \succeq_{\text{st}} \pi_{x_{\hat{\pi}}}$ or $\pi \succeq_{\text{lr}} \pi_{x_{\hat{\pi}}}$, preventive replacement is optimal.*

Proof. By Theorem 3, for fixed $\hat{\pi}$, it is true that $(x_{\hat{\pi}}, \hat{\pi}) \in D$, which implies that $\hat{\pi} \in D_{x_{\hat{\pi}}}$. Hence, $D_{x_{\hat{\pi}}}$ is non-empty. For part (2), by contradiction, assume there exists a threshold belief state $\pi_{x_{\hat{\pi}}}$ above which preventive replacement is optimal. By Proposition 2, we have that $\pi_{x_{\hat{\pi}}} \preceq_{\text{st}} \mathbf{e}_\ell$ and $\pi_{x_{\hat{\pi}}} \preceq_{\text{lr}} \mathbf{e}_\ell$, which implies $(x_{\hat{\pi}}, \mathbf{e}_\ell) \in D \Rightarrow x_{\hat{\pi}} \in D_{\mathbf{e}_\ell}$. By Theorem 3, $x \in D_{\mathbf{e}_\ell}$ if and only if $x \geq x_{\mathbf{e}_\ell}$. Therefore, $x_{\hat{\pi}} \geq x_{\mathbf{e}_\ell}$, but this is a contradiction. Hence, there does not exist such a threshold. \square

Proposition 3 provides sufficient conditions under which a threshold policy does not exist for a given degradation level $x \in [0, \xi)$. To see that these conditions can, in fact, be met, consider the following numerical illustration. Suppose the environment infinitesimal generator matrix and degradation rate vector are, respectively,

$$Q = \begin{bmatrix} -1 & 4/5 & 1/5 & 0 \\ 0 & -1 & 4/5 & 1/5 \\ 0 & 0 & -1 & 1 \\ 1/3 & 1/3 & 1/3 & -1 \end{bmatrix} \quad \text{and} \quad r = [5 \ 6 \ 7 \ 8].$$

The fixed costs are $c_0 = 1$, $c_1 = 10$, $c_2 = 5$, and the failure threshold is $\xi = 24$ units. In this case, it can be shown that $x_{e_4} = 17.6$ and $x_{e_3} = 17$. Hence, the conditions of Proposition 3 are met. Thus, there does not exist a threshold policy with respect to π when $x = 17$. Moreover, consider any degradation value $\hat{x} \in (17, 17.6)$. Then, $(\hat{x}, e_3) \in D$, so $D_{\hat{x}} \neq \emptyset$; however, $(\hat{x}, e_4) \notin D$ since $\hat{x} < 17.6$. Therefore, there does not exist a threshold belief state $\pi_{\hat{x}}$ above which preventive replacement is optimal, as e_4 is the largest element of Π (in the usual stochastic and likelihood ratio orders).

4 Replacement with Stochastic Downtime Costs

Realistically, wind turbine component replacements are time-consuming and require a shut-down period during which no power is generated and revenue is lost. This section extends the model of Section 3 to consider non-instantaneous replacements and the cost of downtime, which depends explicitly on the state of the random environment. To set the stage for this model, we first discuss the costs. Inspections are performed instantaneously at a fixed cost c_0 . If a preventive replacement is elected (at cost c_1), it is initiated immediately at the start of the period, and if reactive replacement is required (also at cost c_1), it likewise begins immediately upon failure. The deterministic time required to complete a preventive (reactive) replacement is δ_1 (δ_2), and these durations are such that $0 < \delta < \delta_1 < \delta_2$. The assumption that $\delta_1 < \delta_2$ is justified since, in the event of an unplanned replacement, additional time is required to assemble the resources needed to complete the replacement (e.g., a large crane, specialized equipment and personnel). In addition to the fixed replacement cost c_1 , a stochastic downtime cost is accrued for each replacement type at a rate that depends on the environment state. Let d_i denote the downtime cost (or revenue loss) rate for a replacement that commences while the environment is in state $i \in \mathcal{S}$, and define the vector $\mathbf{d} = (d_1, d_2, \dots, d_\ell)$. Let $C_p(\pi)$ be the expected total downtime cost of preventive replacement starting in belief state (π) , and $C_r(x, \pi)$ be the expected total downtime cost of reactive replacement commencing in state $(x, \pi) \in B$.

Let $V(x, \pi)$ again be the minimum relative cost per unit time starting in state $(x, \pi) \in B$, and $V_0(x, \pi)$ and $V_1(x, \pi)$ are the relative costs of doing nothing and performing preventive replacement, respectively, starting in state (x, π) . The optimality equations,

$$V(x, \pi) = \min\{V_1(x, \pi), V_0(x, \pi)\}, \quad (x, \pi) \in B, \quad (13)$$

are identical to those of Section 3 with the exception of additional downtime cost terms in the

relative cost functions. The modified functions are

$$V_1(x, \pi) = c_0 + c_1 + C_p(\pi) + V(0, \pi_s) - \delta_1 \gamma, \quad (14)$$

$$V_0(x, \pi) = c_0 + [c_1 + C_r(x, \pi) - \delta_2 \gamma + V(0, \pi_s)] \int_0^\infty \mathbb{I}^+(x, u) k(u, \pi) du \\ + \int_0^\infty \mathbb{I}^-(x, u) V_\pi(x, u) k(u, \pi) du - \gamma \tau(x, \pi). \quad (15)$$

The next subsection provides expressions for the expected downtime costs $C_p(\pi)$ and $C_r(x, \pi)$, followed by structural results for this model.

4.1 Expected Downtime Costs

Here, we show how to obtain the expected downtime costs $C_p(\pi)$ and $C_r(x, \pi)$. We begin by defining the cumulative downtime cost accrued during a replacement up to time t , $D(t)$, given by

$$D(t) = D(0) + \int_0^t d_{Z(u)} du,$$

where $\mathbb{P}(D(0) = 0) = 1$, i.e., a replacement that consumes no time accrues no downtime costs. For $i, j \in \mathcal{S}$ and $t \geq 0$, let $F_{ij}(y, t) \equiv \mathbb{P}(D(t) \leq y, Z(t) = j | Z(0) = i)$, and define the density function

$$f_{ij}(u, t) \equiv \left. \frac{\partial F_{ij}(y, t)}{\partial y} \right|_{y=u}.$$

The expected total downtime cost accrued during a replacement commencing in environment state i and requiring a duration of t time units, denoted $C_i(t)$, is

$$C_i(t) \equiv \mathbb{E}(D(t) | Z(0) = i) \\ = t d_i \exp(q_{ii} t) + \sum_{j \in \mathcal{S}} \int_{td_1}^{td_e} u f_{ij}(u, t) du.$$

Therefore, the expected total downtime cost of a preventive replacement in belief state (π) is

$$C_p(\pi) = \mathbb{E}(D(\delta_1) | \pi) = \sum_{i \in \mathcal{S}} C_i(\delta_1) \pi^{(i)}. \quad (16)$$

For the downtime cost of reactive replacement, let $Y(x, \pi)$ be the elapsed time from the current inspection until failure in the next period starting in state (x, π) . To simplify notation, we suppress the dependence of Y on (x, π) and write simply Y . For $u \geq 0$, define the row vector $\mathbf{T}^*(u, \pi) \equiv [T_i^*(u, \pi)]$, $i \in \mathcal{S}$ where

$$T_i^*(u, \pi) \equiv \mathbb{P}(Z(u) = i | \pi, Y = u) = \frac{\sum_{j \in \mathcal{S}} p^{(u)}(\xi - x, i | j) \pi^{(j)}}{\sum_{j \in \mathcal{S}} \sum_{k \in \mathcal{S}} p^{(u)}(\xi - x, j | k) \pi^{(k)}},$$

where

$$p^{(u)}(\xi - x, j | i) \equiv \left. \frac{\partial W_{ij}(v, u)}{\partial v} \right|_{v=\xi-x}.$$

Therefore, starting in state (x, π) , the expected total downtime cost of a reactive failure in the next period is

$$\begin{aligned} C_r(x, \pi) &= \mathbb{E}(D(Y + \delta_2) | D(Y) = 0, X(Y) = \xi - x, \pi) \\ &= \int_{\nu_i}^{\nu_1} \sum_{i \in \mathcal{S}} C_i(\delta_2) T_i^*(u, \pi) d\mathbb{P}(Y \leq u), \end{aligned} \quad (17)$$

where $\nu_i = [r_i^{-1}(\xi - x)] \wedge \delta$, $i \in \mathcal{S}$.

As before, we assume all new components begin operation in state $(0, \pi_s)$, and the objective is to minimize the long-run average replacement costs per unit time given by

$$\begin{aligned} \gamma = \inf_{a \in \mathcal{P}} \mathbb{E}_a \left\{ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N c_0 + c_1 \mathbb{I}\{a(X_n, \pi_n) = 1\} \right. \\ \left. + [c_1 + D(Y + \delta_2) - D(Y)] \mathbb{I}\{a(X_n, \pi_n) = 0, Y \leq \delta\} \right\}. \end{aligned}$$

Obtaining structural results for (13) is considerably more difficult than for (2). In the next subsection, we provide some results that help characterize the optimal replacement policy.

4.2 Structural Results

In this section, we present several results that hold under somewhat strong conditions. The main result of this section establishes necessary conditions that ensure the optimality of a threshold-type control policy. The first result bounds the optimal cost from below.

Lemma 6. *The average cost of an optimal policy is bounded below as follows:*

$$\gamma > \frac{c_0}{\delta + \delta_2}.$$

Proof. Set $V(0, \pi_s) = 0$, and note that the optimal action immediately following replacement is $a = 0$, so that $V_0(0, \pi_s) = V(0, \pi_s) = 0$. Substituting equation (15) for $V_0(0, \pi_s)$ and solving for γ yields

$$\gamma = \frac{c_0 + [c_1 + C_r(0, \pi_s)] H(0, \pi_s) + \int_0^\infty \mathbb{I}^-(0, u) V_\pi(0, u) k(u, \pi) du}{\tau(0, \pi_s) + \delta_2 H(0, \pi_s)},$$

which shows that

$$\gamma > \frac{c_0 + [c_1 + C_r(0, \pi_s)] H(0, \pi_s)}{\tau(0, \pi_s) + \delta_2 H(0, \pi_s)} > \frac{c_0}{\delta + \delta_2}. \quad \square$$

In contrast to the model of Section 3, it is not necessarily optimal to delay preventive replacement when the component survives the next period with certainty. This is because it may be advantageous to initiate preventive replacement in an environment state with lower downtime costs (e.g., when wind conditions are mild). But when failure is imminent in the next period, preventive replacement is always optimal if the difference between preventive and reactive downtime costs exceed a certain threshold, as shown in Lemma 7.

Lemma 7. *Starting in state (x, π) , if the component fails w.p. 1 in the next period, then a sufficient condition for the optimality of preventive replacement is*

$$C_r(x, \pi) - C_p(\pi) \geq \frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2}.$$

Proof. Choose any state $(x, \pi) \in B$ for which failure is certain in the next period. Now, preventive replacement is optimal if $V_1(x, \pi) \leq V_0(x, \pi)$ or, equivalently, if

$$c_0 + c_1 + C_p(\pi) + V(0, \pi_s) - \delta_1\gamma \leq c_0 + c_1 + C_r(x, \pi) - \delta_2\gamma + V(0, \pi_s) - \gamma\tau(x, \pi). \quad (18)$$

Rearranging the terms of (18) shows that

$$C_r(x, \pi) - C_p(\pi) \geq \gamma[\delta_2 - \delta_1 + \tau(x, \pi)] > \frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2},$$

where the second inequality follows from Lemma 6. \square

The next three lemmas are needed to prove the optimality of a threshold policy with respect to the degradation level x (for π fixed). The first establishes a lower bound for $V_0(x, \pi)$, and the remaining two assert some properties of the downtime cost function $C_r(x, \pi)$.

Lemma 8. *For each $\pi \in \Pi$,*

$$V_0(x, \pi) \geq \underline{V}_0(x, \pi) \equiv c_0 + [c_1 + C_r(x, \pi)]H(x, \pi) - \gamma[\delta_2H(x, \pi) + \tau(x, \pi)] + V(0, \pi_s). \quad (19)$$

Proof. For $(x, \pi) \in B$ and $u \geq 0$, $V_\pi(x, u) \geq V(0, \pi_s)$. Therefore,

$$\begin{aligned} V_0(x, \pi) &= c_0 + [c_1 + C_r(x, \pi) - \delta_2\gamma + V(0, \pi_s)] \int_0^\infty \mathbb{I}^+(x, u)k(u, \pi)du \\ &\quad + \int_0^\infty \mathbb{I}^-(x, u)V_\pi(x, u)k(u, \pi)du - \gamma\tau(x, \pi) \\ &\geq c_0 + [c_1 + C_r(x, \pi)]H(x, \pi) - \gamma\delta_2H(x, \pi) \\ &\quad + V(0, \pi_s)H(x, \pi) + V(0, \pi_s)[1 - H(x, \pi)] - \gamma\tau(x, \pi) \\ &= c_0 + [c_1 + C_r(x, \pi)]H(x, \pi) - \gamma[\delta_2H(x, \pi) + \tau(x, \pi)] + V(0, \pi_s). \quad \square \end{aligned}$$

Unfortunately, the non-monotonicity of $C_r(x, \pi)$ in x precludes us from asserting the monotonicity of $V_0(x, \pi)$ via (19). By contrast, the properties of the lower bound (10) were transparent.

Lemma 9. *For all $\pi \in \Pi$,*

$$C_r(\xi, \pi) = \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)}.$$

Proof. When $x = \xi$, the remaining lifetime is zero w.p. 1, and $d\mathbb{P}(Y \leq u) = \omega(u)$, where $\omega(u)$ is the Dirac delta function. Therefore,

$$\begin{aligned} C_r(\xi, \pi) &= \int_{\mathbb{R}_+} \sum_{i \in \mathcal{S}} C_i(\delta_2) T_i^*(u, \pi) \omega(u) \\ &= \sum_{i \in \mathcal{S}} C_i(\delta_2) T_i^*(0, \pi) \\ &= \sum_{i \in \mathcal{S}} C_i(\delta_2) \mathbb{P}(Z(0) = i | Y = 0, \pi) \\ &= \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)}. \quad \square \end{aligned}$$

Lemma 10. For $\epsilon > 0$ and $(x, \pi) \in B$ such that $\xi - x < \epsilon$,

$$C_r(x, \pi) \geq \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \exp(q_{ii}\epsilon/r_i).$$

Proof. Choose $\epsilon < \delta r_1$ and x such that $\xi - x < \epsilon$. Then

$$\begin{aligned} C_r(x, \pi) &= \int_{\epsilon/r_\ell}^{\epsilon/r_1} \sum_{j \in \mathcal{S}} C_j(\delta_2) T_j^*(u, \pi) d\mathbb{P}(Y \leq u) \\ &= \int_{\epsilon/r_\ell}^{\epsilon/r_1} \sum_{j \in \mathcal{S}} C_j(\delta_2) \mathbb{P}(Z(u) = j | Y = u, \pi) d\mathbb{P}(Y \leq u) \\ &= \int_{\epsilon/r_\ell}^{\epsilon/r_1} \sum_{j \in \mathcal{S}} C_j(\delta_2) d\mathbb{P}(Z(u) = j, Y \leq u | \pi) \\ &\geq \int_{\epsilon/r_\ell}^{\epsilon/r_1} \mathbb{I}(u \in \{\epsilon/r_i : i \in \mathcal{S}\}) \sum_{j \in \mathcal{S}} C_j(\delta_2) d\mathbb{P}(Z(u) = j, Y \leq u | \pi) \\ &= \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} C_j(\delta_2) d\mathbb{P}(Z(u) = j, Y \leq u | \pi) \Big|_{u=\epsilon/r_i} \\ &\geq \sum_{i \in \mathcal{S}} C_i(\delta_2) d\mathbb{P}(Z(u) = i, Y \leq u | \pi) \Big|_{u=\epsilon/r_i} \\ &= \sum_{i \in \mathcal{S}} \sum_{k \in \mathcal{S}} C_i(\delta_2) d\mathbb{P}(Z(u) = i, Y \leq u | Z(0) = k) \Big|_{u=\epsilon/r_i} \pi^{(k)} \\ &= \sum_{i \in \mathcal{S}} \sum_{k \in \mathcal{S}} C_i(\delta_2) \mathbb{P}(Z(\epsilon/r_i) = i, X(\epsilon/r_i) = \epsilon | Z(0) = k) \pi^{(k)} \\ &\geq \sum_{i \in \mathcal{S}} C_i(\delta_2) \mathbb{P}(Z(\epsilon/r_i) = i, X(\epsilon/r_i) = \epsilon | Z(0) = i) \pi^{(i)} \\ &= \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \exp(q_{ii}\epsilon/r_i). \quad \square \end{aligned}$$

We are now prepared to state the main result of this section, namely the optimality of a threshold policy under certain conditions. Specifically, such a policy is optimal for belief states in which the expected difference between the downtime costs of reactive and preventive replacements is sufficiently large. When these conditions are met, the optimality of preventive replacement can be established for all x sufficiently close to the failure threshold ξ .

Theorem 5. If for some $\pi \in \Pi$,

$$\frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2} < \sum_{i \in \mathcal{S}} [C_i(\delta_2) - C_i(\delta_1)] \pi^{(i)}, \quad (20)$$

then there exists an $\epsilon > 0$ such that preventive replacement is optimal for all $(x, \pi) \in B$ such that $x \in (\xi - \epsilon, \xi]$.

Proof. By Lemma 8, a sufficient condition for the optimality of preventive replacement in state $(x, \pi) \in B$ is $V_0(x, \pi) \geq V_1(x, \pi)$, which implies

$$c_0 + c_1 + C_r(x, \pi) - \delta_2 \gamma + V(0, \pi_s) \geq c_0 + c_1 + C_p(\pi) + V(0, \pi_s) - \delta_1 \gamma. \quad (21)$$

Rearranging the terms of (21) and applying Lemma 7 shows that

$$C_r(x, \pi) - \sum_{i \in \mathcal{S}} C_i(\delta_1) \pi^{(i)} \geq \gamma(\delta_2 - \delta_1) > \frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2},$$

which with further simplification yields

$$C_r(x, \pi) \geq \sum_{i \in \mathcal{S}} C_i(\delta_1) \pi^{(i)} + \frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2} \equiv K. \quad (22)$$

Evaluating (22) at (ξ, π) by applying Lemma 9 and rearranging terms gives condition (20). Now, assuming (20) is satisfied at ξ , it can also be shown to be satisfied for $x \in (\xi - \epsilon, \xi]$ for some $\epsilon > 0$ as follows. Assume condition (20) is satisfied and rearrange terms so that

$$\sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} > \sum_{i \in \mathcal{S}} C_i(\delta_1) \pi^{(i)} + \frac{c_0(\delta_2 - \delta_1)}{\delta + \delta_2} = K,$$

which implies

$$\sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} - K \equiv K_1 > 0. \quad (23)$$

By Lemma 10, for $\epsilon > 0$ and all x such that $\xi - x < \epsilon$,

$$C_r(x, \pi) \geq \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \exp(q_{ii}\epsilon/r_i).$$

Choosing ϵ such that $1 - \exp(q_{ii}\epsilon/r_i) < K_1 / \sum_{j \in \mathcal{S}} C_j(\delta_2)$ for all $i \in \mathcal{S}$ and using (23) yields

$$\begin{aligned} C_r(x, \pi) - K &\geq \sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \exp(q_{ii}\epsilon/r_i) - K \\ &> \left(\sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \right) \left[1 - K_1 \left(\sum_{i \in \mathcal{S}} C_i(\delta_2) \right)^{-1} \right] - K \\ &= K_1 \left[1 - \left(\sum_{i \in \mathcal{S}} C_i(\delta_2) \pi^{(i)} \right) \left(\sum_{i \in \mathcal{S}} C_i(\delta_2) \right)^{-1} \right] \\ &> 0, \end{aligned}$$

where the last inequality holds by (23). Since $C_r(x, \pi) - K > 0$ implies inequality (22) holds, it is optimal to replace for all $x \in (\xi - \epsilon, \xi]$. \square

That a replacement threshold does not necessarily exist for each environment belief state in Theorem 5 indicates that information about the environment is especially critical when downtime costs are incurred as a result of replacement actions.

The next section describes the means by which we obtain optimal policies for the models presented in Sections 3 and 4. Subsequently, we illustrate these policies by way of a few numerical examples.

5 Solution Techniques

This section describes numerical solution techniques to solve the optimal replacement models of Sections 3 and 4. These techniques require discretization of the belief space B and application of either the policy iteration algorithm (cf. Puterman [26]) or the linear programming (LP) approach to solving Markov decision process (MDP) models. We approximate the integrals in equations (4) and (15) as Riemann sums and formulate MDP models on a discrete, finite state space.

To that end, let $\dot{\Pi} = \{\dot{\pi}_1, \dot{\pi}_2, \dots, \dot{\pi}_{L_1}\}$ be a discretization of Π , where $\dot{\pi}_j \in \Pi$ and L_1 is a positive integer. Let $\mathcal{X} = \{0\} \cup [r_1\delta, \xi + r_\ell\delta]$ be the set of possible degradation levels in the numerical procedure, and define b as the step size between the discrete points in $[r_1\delta, \xi + r_\ell\delta]$ where

$$b = [\xi + (r_\ell - r_1)\delta] / L_2,$$

and L_2 ($L_2 \geq 2$) is a positive integer. The ordered set $\dot{\mathcal{X}} = \{0, \dot{x}_1, \dot{x}_2, \dots, \dot{x}_{L_2+1}\}$ is the discretization of \mathcal{X} , where

$$\dot{x}_k = r_1\delta + (k-1)b, \quad k = 1, 2, \dots, L_2 + 1.$$

The complete discretized belief space is denoted $\dot{B} \equiv (\dot{\Pi} \times \dot{\mathcal{X}}) \cup \{(0, \pi_s)\}$, where π_s is the stationary distribution of the environment process. Define $L = |\dot{B}| = L_1(L_2+2)+1$, and let $\mathcal{L} = \{0, 1, \dots, L\}$. The i th belief state in \dot{B} is denoted by $\dot{b}_i \equiv (\dot{x}_i, \dot{\pi}_i)$, where $(\dot{x}_i, \dot{\pi}_i) \in \dot{\Pi} \times \dot{\mathcal{X}}$, $i \in \mathcal{L} \setminus \{0\}$ and $\dot{b}_0 \equiv (0, \pi_s)$. The set of cumulative degradation levels attainable from state \dot{b}_i , given no replacement, is $\dot{\mathcal{X}}_i = \{\bar{x}_{i1}, \bar{x}_{i2}, \dots, \bar{x}_{i,C+1}\}$, where $\bar{x}_{ij} = \dot{x}_i + r_1\delta + (j-1)b$ and $C \equiv |\dot{\mathcal{X}}_i| = \lfloor \delta(r_\ell - r_1)/b \rfloor$, $i \in \mathcal{L}$, $j = 0, 1, \dots, C+1$.

Estimating the transition probabilities between discretized belief states consists of three steps. First, for a given state \dot{b}_i , compute $\bar{\Pi}_i = \{\bar{\pi}_{i1}, \bar{\pi}_{i2}, \dots, \bar{\pi}_{i,C+1}\}$, where $\bar{\pi}_{ij} = \mathbf{T}(\bar{x}_{ij}, \dot{\pi}_i)$ is the updated belief state obtained from \dot{b}_i by observing an increase in degradation of \bar{x}_{ij} in the next decision epoch, $i \in \mathcal{L} \setminus \{0\}$, $j = 1, 2, \dots, C+1$. Next, approximate each $\bar{\pi}_{ij}$ with a discretized $\pi_{ij}^* \in \dot{\Pi}$, where

$$\pi_{ij}^* = \min_{\dot{\pi} \in \dot{\Pi}} \|\bar{\pi}_{ij} - \dot{\pi}\|, \quad i \in \mathcal{L}, \quad j = 0, 1, \dots, C+1,$$

and let $\Pi_i^* \equiv \{\pi_{i1}^*, \pi_{i2}^*, \dots, \pi_{i,C+1}^*\}$. Finally, compute the row vector $\dot{P}_i = (\dot{p}_{i1}, \dot{p}_{i2}, \dots, \dot{p}_{i,C+1})$, where \dot{p}_{ij} is the transition probability between \dot{b}_i and the updated belief state $(\dot{x}_i + \bar{x}_{ij}, \pi_{ij}^*)$, given no replacement. To compute \dot{p}_{ij} , define $R(x; \pi_n, x_n) \equiv \mathbb{P}(X_{n+1} \leq x | x_n, \pi_n)$, the probability that the cumulative degradation during period $n+1$ does not exceed x , given $b^{(n)} = (x_n, \pi_n)$. A simple conditioning argument shows that

$$\begin{aligned} R(x; \pi_n, x_n) &= \mathbb{P}(X_{n+1} \leq x | x_n, \pi_n) \\ &= \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} \mathbb{P}(X_1 \leq x - x_n, z_1 = j | z_0 = i) \pi_n^{(i)} \\ &= \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} W_{ij}(x - x_n, \delta) \pi_n^{(i)}. \end{aligned}$$

The transition probability \dot{p}_{ij} is computed as

$$\dot{p}_{ij} = \begin{cases} R_i(\bar{x}_{ij} + b/2 \wedge \dot{x}_{i,C+1}) - R_i(\dot{x}_{i1} \vee \dot{x}_{ij} - b/2), & i \in \mathcal{L}, \quad j > 0, \\ 0, & i = 0, \quad j = 0, \end{cases} \quad (24)$$

where $R_i(x) \equiv R(x; \dot{\pi}_i, \dot{x}_i)$ is approximated using a technique described in Theorem 4.1 of Bladt et al. [2]. Computing $(\Pi_i^*, \dot{\mathcal{X}}_i, \dot{P}_i)$ for all $\dot{b}_i \in \dot{B}$ provides the transition probabilities between each pair of discretized belief states, given no replacement occurs.

Now the optimality equations are stated on \dot{B} for the model of Section 3. Let $\dot{V}(i) \equiv \dot{V}(\dot{b}_i)$ denote the relative cost, starting in state \dot{b}_i . Then

$$\dot{V}(i) = \min\{\dot{V}_0(i), \dot{V}_1(i)\}, \quad i \in \mathcal{L}, \quad (25)$$

where $\dot{V}_0(i) \equiv \dot{V}_0(\dot{b}_i)$ and $\dot{V}_1(i) \equiv \dot{V}_1(\dot{b}_i)$ denote the relative costs of doing nothing and preventive replacement, respectively, in \dot{b}_i . The relative cost of action $a = 1$ (preventive replacement) in \dot{b}_i is

$$\dot{V}_1(i) = c_0 + c_1 + \dot{V}(0), \quad i \in \mathcal{L}.$$

For action $a = 0$ (do nothing), the relative cost is obtained by conditioning on the event of component survival in the next period. Denote the single-period transition probability between \dot{b}_i and \dot{b}_j , given the component survives, by \tilde{q}_{ij} , where

$$\begin{aligned} \tilde{q}_{ij} &= \mathbb{P}(b^{(n+1)} = \dot{b}_j | b^{(n)} = \dot{b}_i, \dot{x}_j \leq \xi) \\ &= \frac{\dot{p}_{ij}}{R_i(\xi)} \mathbb{I}(\dot{x}_j \leq \xi), \quad i, j \in \mathcal{L}. \end{aligned}$$

Let τ_i be the expected survival time in the next decision period, given \dot{b}_i , where

$$\tau_i = \sum_{j \in \mathcal{S}} \sum_{k \in \mathcal{S}} \int_0^\delta W_{jk}(\xi - \dot{x}_i, t) dt, \quad i \in \mathcal{L}.$$

Then the relative cost of doing nothing starting in state \dot{b}_i is

$$\dot{V}_0(i) = c_0 + (c_1 + c_2 + \dot{V}(0))\bar{R}_i(\xi) + R_i(\xi) \sum_{j=0}^L \dot{V}(j)\tilde{q}_{ij} - \gamma\tau_i, \quad i \in \mathcal{L},$$

where $\bar{R}_i(\xi) \equiv 1 - R_i(\xi)$.

The optimality equations (25) can be solved using the policy iteration algorithm, but for larger problem instances, the LP approach, which facilitates use of a commercial solver, is preferred. By convention, the solution of the LP primal formulation of a MDP model corresponds to the optimal relative cost of each belief state, whereas the solution of the LP dual formulation corresponds to the optimal action for each belief state. The primal and dual objective function values both equal the long-run average cost per unit time under the optimal replacement policy. Before presenting the primal formulation, observe that setting $\dot{V}(0) = 0$ leads to the following set of optimality equations for each discretized belief state $i \in \mathcal{L}$:

$$\dot{V}(i) = \min \left\{ c_0 + c_1, \quad c_0 + (c_1 + c_2)\bar{R}_i(\xi) + R_i(\xi) \sum_{j=0}^L \dot{V}(j)\tilde{q}_{ij} - \gamma\tau_i \right\}. \quad (26)$$

Equation (26) can be expressed as the pair of linear constraints

$$\dot{V}(i) \leq c_0 + c_1, \quad (27a)$$

$$\dot{V}(i) - R_i(\xi) \sum_{j=0}^L \dot{V}(j)\tilde{q}_{ij} + \gamma\tau_i \leq c_0 + (c_1 + c_2)\bar{R}_i(\xi). \quad (27b)$$

Using constraints (27a) and (27b), the primal LP formulation is

$$\max \quad \gamma \tag{28a}$$

$$\text{s.t.} \quad \dot{V}(i) \leq c_0 + c_1, \quad \forall i \in \mathcal{L} \tag{28b}$$

$$\dot{V}(i) - R_i(\xi) \sum_{j=0}^L \dot{V}(j) \tilde{q}_{ij} + \gamma \tau_i \leq c_0 + (c_1 + c_2) \bar{R}_i(\xi), \quad \forall i \in \mathcal{L} \tag{28c}$$

$$\gamma \in \mathbb{R}.$$

The dual LP formulation follows directly from the primal formulation (28). Let x_{ia} denote the limiting probability of being in belief state i and taking action a , $a \in \{0, 1\}$. The dual formulation is

$$\min \quad \sum_{i=0}^L x_{i0} [c_0 + (c_1 + c_2) \bar{R}_i(\xi)] + (c_0 + c_1) \sum_{i=0}^L x_{i1} \tag{29a}$$

$$\text{s.t.} \quad x_{i0} + x_{i1} - \sum_{j=0}^L \tilde{q}_{ji} R_j(\xi) x_{j0} = 0, \quad \forall i \in \mathcal{L} \tag{29b}$$

$$\sum_{i=0}^L \tau_i x_{i0} = 1, \tag{29c}$$

$$x_{ia} \geq 0, \quad \forall i \in \mathcal{L}, \quad a \in \{0, 1\}.$$

From a computational viewpoint, obtaining the optimal policy via formulation (29) can be problematic because the belief space of the POMDP model is comprised of many transient states, and the dual formulation can assign optimal actions only to recurrent states (see Puterman [26]). To circumvent this complication, the optimal policy can be alternatively obtained from the primal solution. Consider a component that begins operation in $\dot{b}_i \in \dot{B}$. Preventive replacement is optimal in state \dot{b}_i only if the expected total bias incurred for immediate preventive replacement, followed by resumption of the optimal policy, does not exceed the expected total bias for doing nothing in a single period, i.e., if

$$c_0 + (c_1 + c_2) \bar{R}_i(\xi) - \gamma \tau_i + R_i(\xi) \sum_{j=0}^L \tilde{q}_{ij} \dot{V}(j) \geq c_0 + c_1,$$

or equivalently,

$$c_2 + R_i(\xi) \left(\sum_{j=0}^L \tilde{q}_{ij} \dot{V}(j) - (c_1 + c_2) \right) - \gamma \tau_i \geq 0. \tag{30}$$

Therefore, the optimal policy is obtained by first solving the primal LP formulation (28) to obtain γ and $\dot{V}(i)$, $i \in \mathcal{L}$. Subsequently, condition (30) is checked for each discretized belief state to obtain the corresponding optimal action.

For the replacement model of Section 4, the optimality equations on \dot{B} follow in a similar manner. Let $\check{V}(i) \equiv \check{V}(\dot{b}_i)$ denote the relative cost given that the component starts operation in i , and let $\check{V}_0(i) \equiv \check{V}_0(\dot{b}_i)$ and $\check{V}_1(i) \equiv \check{V}_1(\dot{b}_i)$ denote the relative costs given no action and preventive replacement, respectively, in state i , $i \in \mathcal{L}$. Then

$$\check{V}(i) = \min\{\check{V}_0(i), \check{V}_1(i)\}, \quad i \in \mathcal{L}, \tag{31}$$

where

$$\begin{aligned}\ddot{V}_0(i) &= c_0 + \left[c_1 + C_r(i) - \delta_2\gamma + \ddot{V}(0) \right] \bar{R}_i(\xi) + R_i(\xi) \sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} - \gamma\tau_i, \\ \ddot{V}_1(i) &= c_0 + c_1 + C_p(i) - \delta_1\gamma + \ddot{V}(0),\end{aligned}$$

and the downtime costs $C_p(i)$ and $C_r(i)$ are obtained via (16) and (17), respectively. Setting $\ddot{V}(0) = 0$, we have for each $i \in \mathcal{L}$,

$$\ddot{V}(i) = \min \left\{ c_0 + c_1 + C_p(i) - \delta_1\gamma, \quad c_0 + (c_1 + C_r(i) - \delta_2\gamma) \bar{R}_i(\xi) + R_i(\xi) \sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} - \gamma\tau_i \right\},$$

which is equivalent to the pair of linear constraints

$$\ddot{V}(i) + \delta_1\gamma \leq c_0 + c_1 + C_p(i), \quad (32a)$$

$$\ddot{V}(i) - R_i(\xi) \sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} + \gamma [\tau_i + \delta_2 \bar{R}_i(\xi)] \leq c_0 + (c_1 + C_r(i)) \bar{R}_i(\xi). \quad (32b)$$

Using constraints (32a) and (32b), the primal LP formulation is

$$\max \quad \gamma \quad (33a)$$

$$\text{s.t.} \quad \ddot{V}(i) + \delta_1\gamma \leq c_0 + c_1 + C_p(i), \quad \forall i \in \mathcal{L} \quad (33b)$$

$$\begin{aligned}\ddot{V}(i) - R_i(\xi) \sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} + \gamma [\tau_i + \delta_2 \bar{R}_i(\xi)] \\ \leq c_0 + (c_1 + C_r(i)) \bar{R}_i(\xi), \quad \forall i \in \mathcal{L}\end{aligned} \quad (33c)$$

$$\gamma \in \mathbb{R}.$$

The dual LP is then formulated directly from the primal as

$$\min \quad \sum_{i=0}^L x_{i0} [c_0 + (c_1 + C_r(i)) \bar{R}_i(\xi)] + (c_0 + c_1 + C_p(i)) \sum_{i=0}^L x_{i1} \quad (34a)$$

$$\text{s.t.} \quad x_{i0} + x_{i1} - \sum_{j=0}^L \tilde{q}_{ji} R_j(\xi) x_{j0} = 0, \quad \forall i \in \mathcal{L}, \quad (34b)$$

$$\begin{aligned}\sum_{i=0}^L [\tau_i + \delta_2 \bar{R}_i(\xi)] x_{i0} + \delta_1 \sum_{i=0}^L x_{i1} = 1, \\ x_{ia} \geq 0, \quad \forall i \in \mathcal{L}, \quad a \in \{0, 1\}.\end{aligned} \quad (34c)$$

Analogous to (30), preventive replacement is optimal in a given belief state only if

$$c_0 + (c_1 + C_r(i) - \delta_2\gamma) \bar{R}_i(\xi) + R_i(\xi) \sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} - \gamma\tau_i \geq c_0 + c_1 + C_p(i) - \delta_1\gamma,$$

or equivalently,

$$\gamma(\delta_1 - \tau_i) + (C_r(i) - \delta_2\gamma) \bar{R}_i(\xi) + R_i(\xi) \left[\sum_{j=0}^L \ddot{V}(j) \tilde{q}_{ij} - c_1 \right] - C_p(i) \geq 0. \quad (35)$$

Therefore, the optimal policy is obtained by solving the primal LP formulation (33) to obtain γ and $\check{V}(i)$, for $i \in \mathcal{L}$. Then condition (35) is checked for each discretized belief state to assign the optimal actions.

6 Numerical Examples

This section presents two numerical examples that illustrate optimal policies obtained from the model of Section 3 and elucidate the advantages of the POMDP policies. The models were solved using the LP approach described in Section 5. The first example illustrates the case in which replacements are instantaneous and only fixed costs are involved.

Example 1. Suppose the component operates in a random environment with $\ell = 3$ distinct states. The generator matrix and degradation rate vector are parameterized so that

$$\mathbf{Q}(q) = \begin{bmatrix} -1 & 0.9 & 0.1 \\ 0.9 & -1 & 0.1 \\ 0.5 + q/2 & 0.5 + q/2 & -1 - q \end{bmatrix} \quad \text{and} \quad \mathbf{r}(w) = [1.0 \quad 2.0 \quad w].$$

We consider parameter values (q, w) in the set $\{(q, w) : q \in \{-0.5, 0, 1\}, w \in \{3, 5, 7\}\}$. For the inspection interval, we set $\delta = 1$ and set the failure threshold at $\xi = 40$ units. The fixed costs of inspection, preventive replacements and reactive replacements are $c_0 = 1$, $c_1 = 10$ and $c_2 = 2.5$, respectively. A discretization interval of 0.5 was used for both Π and $\mathcal{X} \setminus \{0\}$, which corresponds to fixing $L_1 = 21$ and setting $L_2 = 210$, $L_2 = 220$ and $L_2 = 230$ when $w = 3$, $w = 5$, and $w = 7$, respectively. The respective total number of states when $w = 3$, $w = 5$, and $w = 7$ are $L = 4, 453$, $L = 4, 663$, and $L = 4, 875$. The optimal policy costs associated with each environment are shown in Table 1.

Table 1: Policy cost (γ) for various (q, w) .

| Parameter values (q, w) | Policy cost (γ) |
|---------------------------|--------------------------|
| (-0.5,3) | 1.4797 |
| (-0.5,5) | 1.5664 |
| (-0.5,7) | 1.6433 |
| (0,3) | 1.4557 |
| (0,5) | 1.5084 |
| (0,7) | 1.5599 |
| (1,3) | 1.4422 |
| (1,5) | 1.4709 |
| (1,7) | 1.5004 |

Intuitively, the optimal cost γ increases as q decreases and the degradation rate in state 3 (w) increases. As q decreases, the environment spends a greater proportion of time in state 3, thereby reducing the component's expected lifetime. Consequently, the replacement frequency increases, leading to higher policy costs. Likewise, for q fixed, as w increases the component's expected lifetime is reduced as degradation progresses more rapidly, and the resulting replacement frequency increases. Figure 3 plots the optimal replacement thresholds for three environment belief states as a function of r_3 when $q = -0.5$. For a fixed $\hat{\pi}$ and q , the threshold $x_{\hat{\pi}}$ appears to decrease monotonically as r_3 increases. That is, the policy becomes more conservative in order to avoid

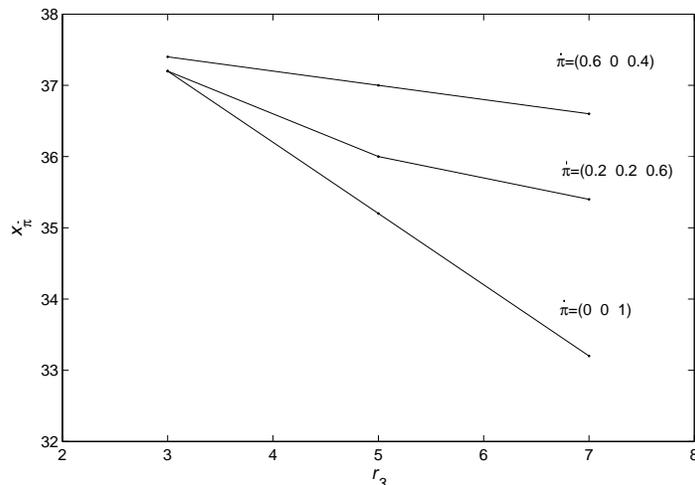


Figure 3: Optimal replacement thresholds (x_{π}^E) for $q = -0.5$.

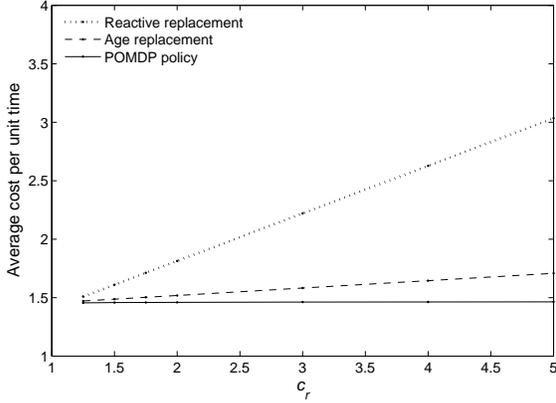
costly reactive replacement penalties. Likewise, the optimal replacement thresholds tend to be lower when there is a stronger belief that the environment is occupying the most detrimental state (state 3).

The performance of the POMDP policies was also evaluated by comparing their costs with those of classical age-replacement and reactive replacement only policies. To estimate the costs of the age- and reactive-replacement policies, the c.d.f. of the first passage time to level ξ was obtained numerically from the parameters \mathbf{Q} and \mathbf{r} using the Laplace transform results of Kharoufeh et al. [15]. Subsequently, we applied classical results from Barlow and Proschan [1] to compute the long-run average costs per unit time for these policies. For the age-replacement policy, the replacement time is the minimum of the failure time or 90% of the component's expected lifetime, whichever occurs first. Policy costs were compared for two cases with $c_0 = 1$, $c_1 = 10$ and $c_2 \in \{2.5, 5, 10, 20, 30, 40\}$. Define

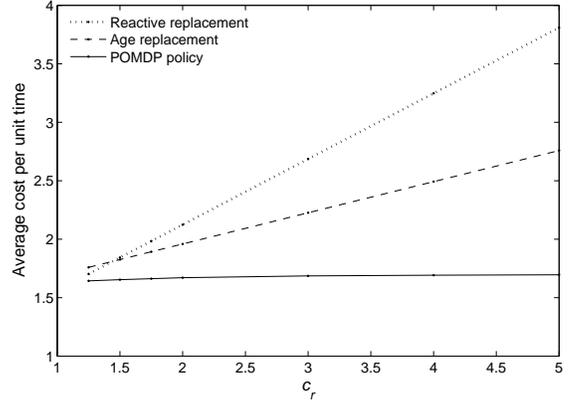
$$c_r \equiv \frac{c_1 + c_2}{c_1}$$

as the ratio of the reactive replacement cost to the preventive replacement cost. Figures 4(a) and 4(b) show a comparison of the policy costs for environments with $(q, w) = (0, 3)$ and $(q, w) = (-0.5, 7)$, respectively, as a function of c_r . In both cases, the performance of the POMDP policies is superior to those of the age- and reactive-replacement policies. As c_r increases and reactive failures become more punitive, the cost of both the age- and reactive-replacement policies increases at a much faster rate than the cost of the POMDP policy. The cost discrepancy is more pronounced in environment $(q, w) = (-0.5, 7)$, indicating that information about the environment can be crucial if there is a significant likelihood that the environment will occupy a very detrimental state for relatively long periods of time.

Example 2. The second example illustrates an optimal replacement policy for a real wind turbine shaft bearing. The effective number of load cycles imposed on the bearing up to a fixed time can be used as a proxy for the cumulative degradation of the bearing. To estimate the environment parameters, the evolution of degradation was simulated up to a failure threshold $\xi = 10^7$ effective load cycles using a physics-based bearing degradation model and empirical temperature and rotor speed observations obtained from a commercial wind turbine. It was assumed that the primary determinant of the degradation rate at a given time is the kinematic viscosity of the bearing



(a) Policy costs for $(q, w) = (0, 3)$.



(b) Policy costs for $(q, w) = (-0.5, 7)$.

Figure 4: Comparison of replacement policy performance.

lubricant. Let $\rho = \nu_0/\nu_1$ be the relative lubricant viscosity, where ν_0 and ν_1 respectively denote the specification and actual lubricant viscosity, and $a(\rho)$ is a life adjustment factor that is a function of ρ . If T_b denotes the bearing's specified lifetime – the time to which the bearing is expected to survive with 90% probability under operational loading conditions – then it is known that $T_b \propto a(\rho)$ (see for example [10]). The relationship between kinematic lubricant viscosity (in mm^2/s) and the lubricant temperature, θ , (in degrees Celsius, C) can be approximated by the empirical Ubbelohde-Walther equation [12], given by

$$\ln[\ln(\nu_1 + 0.8)] = u - v \ln(\theta + 273.15), \quad (36)$$

where u and v are constants that depend on the lubricant type.

Summary data for rotor speed and bearing temperature were available for $N = 25,421$ 10-minute periods (after discarding periods with spurious or missing data). To simulate degradation signals, rotor speed and bearing temperature data were bootstrapped by randomly sampling from 4-hour contiguous blocks of wind turbine data. For the i th 10-minute interval, the relative lubricant viscosity ρ_i is computed as a function of the bearing temperature using equation (36) and assuming $\nu_0 = 150 \text{ mm}^2/\text{s}$. In these calculations, the observed bearing temperature serves as a proxy for the lubricant temperature θ . The constants u and v of equation (36) were obtained using tabulated data from a commercial wind turbine lubricant with the following kinematic viscosity characteristics: $150 \text{ mm}^2/\text{s}$ at 40° C and $20.7 \text{ mm}^2/\text{s}$ at 100° C . Using these values, we obtained $u = 17.76$ and $v = 2.81$. The total number of effective load cycles in the i th 10-minute period, denoted by η'_i , is

$$\eta'_i = 10 a(\rho_i) \omega_i, \quad i = 1, \dots, N,$$

where ω_i is the rotor speed in revolutions per minute (rpm). The life adjustment factor $a(\rho_i)$ is first estimated using data for standard steel bearings [6]. Subsequently, this factor is adjusted to be significantly larger and smaller for relatively small and large values of ρ_i , respectively. This adjustment is intended to represent a material with amplified degradation characteristics. The degradation signal at time t_i (the end of the i th 10-minute period) is given by

$$Y(t_i) = \sum_{j=1}^i \eta'_j, \quad i = 1, \dots, N,$$

with $Y(0) \equiv 0$. Using this physics-based bearing degradation model, we estimated the environment parameters by simulating cumulative bearing degradation up to the failure threshold $\xi = 10^7$

effective load cycles. The environment parameters \mathbf{Q} and \mathbf{r} were inferred from the simulated observations, $\{Y(t_i) : i = 1, \dots, N\}$, by adapting a Markov chain Monte Carlo (MCMC) estimation procedure for switching diffusion processes (see [8, 19]). The estimated number of environment states is $\ell = 2$, and the estimated generator matrix and degradation rate vector are, respectively,

$$\mathbf{Q} = \begin{bmatrix} -0.0089 & 0.0089 \\ 0.0304 & -0.0304 \end{bmatrix} \quad \text{and} \quad \mathbf{r} = 10^3 \times [0.7580 \quad 2.2585].$$

The remaining parameter values were selected as follows: $\delta = 300$ min; $c_0 = 1$; $c_1 = 100$; and $c_2 = 200$. The POMDP model was solved numerically using a discretization interval of 0.2 for Π and approximately 9.0×10^3 for $\mathcal{X} \setminus \{0\}$, which corresponds to setting $L_1 = 6$ and $L_2 = 1,150$ for $L = 6,913$ total belief states. Figure 5 depicts the optimal replacement threshold as a function of the probability $\hat{\pi}^{(2)}$ that the environment occupies the state of rapid degradation. The monotonic-

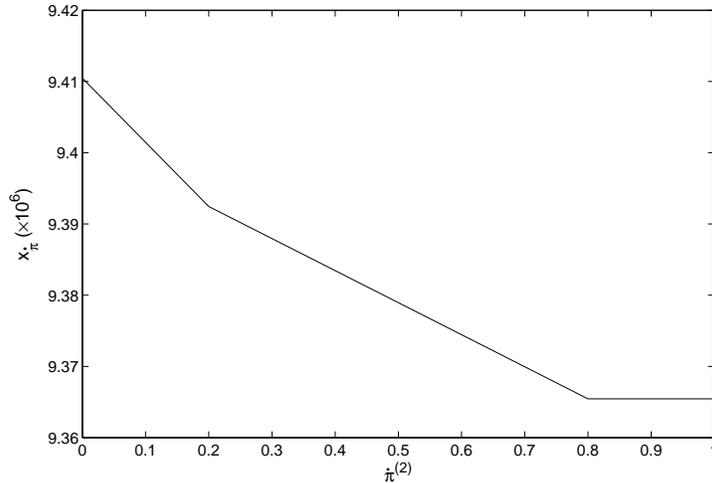


Figure 5: Optimal replacement threshold (x_π) for Example 2.

ity of this curve illustrates the fact that the replacement policy is more conservative if there is a stronger belief that the environment occupies the more detrimental state. The optimal average cost of the POMDP model is $\gamma = 0.0149$, while the costs of the age- and reactive-replacement policies are 0.0158 and 0.0362, respectively. It is noted that the cost of the POMDP policy is approximately 5.7% less than that of the age-replacement policy. This modest improvement can be attributed to the fact that the ratio of reactive to replacement costs is $c_r = 3$. As c_r increases, it is expected that the cost reduction will be more significant, as illustrated in Figure 4(b).

Example 3. The third and final example illustrates the existence and behavior of threshold-type policies for the model of Section 4, which includes stochastic downtime costs. The parameter values for this instance are as follows: $\xi = 40$; $\delta = 1$; $\delta_1 = 0.5$; $c_0 = 1$; and $c_1 = 1.5$. Considered are three distinct environment processes, each with $\ell = 3$ states. The degradation rate vector, which common to each case, is $\mathbf{r} = [1 \ 2 \ 5]$, and the generator matrices are

$$\mathbf{Q}_1 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{Q}_2 = \begin{bmatrix} -1 & 0.5 & 0.5 \\ 0.5 & -1 & 0.5 \\ 0.5 & 0.5 & -1 \end{bmatrix}, \quad \mathbf{Q}_3 = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$$

Matrices \mathbf{Q}_1 and \mathbf{Q}_3 describe the dynamics of cyclic environments for which transitions occur in the sequences $(1 \rightarrow 2 \rightarrow 3 \rightarrow 1)$ and $(3 \rightarrow 2 \rightarrow 1 \rightarrow 3)$, respectively, and generator matrix \mathbf{Q}_2

corresponds to an acyclic environment. In this environment, starting in any one of the states, the process transitions to either of the other two states with equal probability. To examine the effect of the reactive replacement duration (δ_2), optimal replacement thresholds were obtained for each $\delta_2 \in \{0.75, 1, 1.5, 2, 5\}$. Furthermore, it was assumed that the downtime cost rate vector is $\mathbf{d} = [1 \ 1.1 \ 8]$.

The optimality equations were solved numerically by discretizing Π and $\mathcal{X} \setminus \{0\}$ over intervals of length 0.2 so that $L_1 = 21$, $L_2 = 220$, and $L = 4,663$. Figure 6 depicts the optimal replacement thresholds under the downtime cost rate vector \mathbf{d} when the belief state is $\hat{\pi} = [1 \ 0 \ 0]$, i.e., when it is believed that the environment occupies state 1 with certainty. First, it is noted that for each

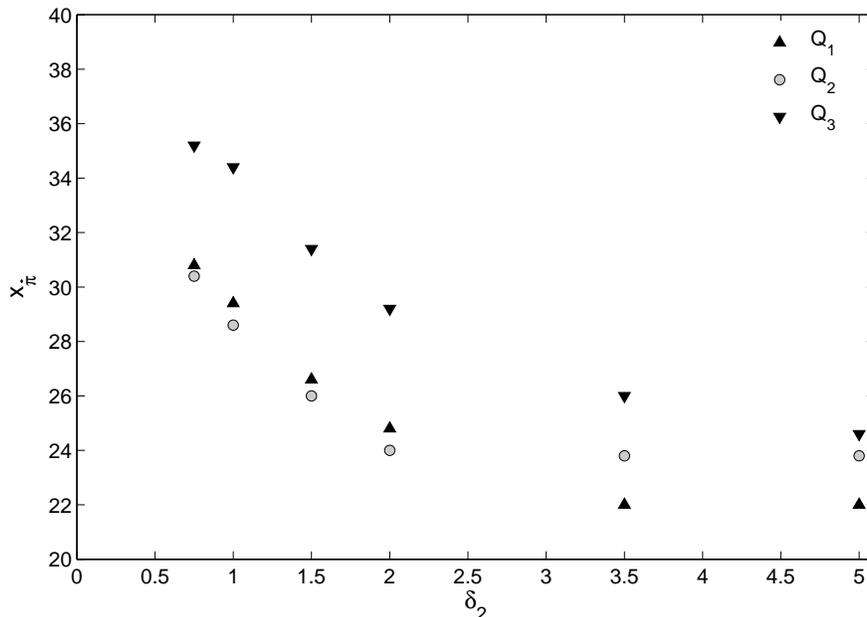


Figure 6: Replacement thresholds as a function of δ_2 ($\hat{\pi} = [1 \ 0 \ 0]$).

of the three cases, the replacement threshold is monotone decreasing in the reactive replacement time. Intuitively, as the time to complete a reactive replacement increases, the optimal policy tends to initiate preventive replacements earlier. Second, the thresholds for the environment described by Q_3 are consistently larger than those of the environment described by generator matrix Q_1 . That is, when it is certain that the environment occupies state 1, the policy is opportunistic in the sense that it prescribes preventive replacement when the downtime cost and degradation rates are lowest, assuming the environment evolves as per Q_1 . By contrast, assuming the environment evolves as per Q_3 , the thresholds are higher, indicating a postponement of replacements. Though not presented here (due to space restrictions), further analysis indicates that, when the belief state is $\hat{\pi} = [0 \ 0 \ 1]$, replacement thresholds are comparatively higher than those of the cases $\hat{\pi} = [1 \ 0 \ 0]$ and $\hat{\pi} = [0 \ 1 \ 0]$. These differences indicate that it is advantageous to avoid initiating replacements in state 3 due to this state's high downtime cost rate. In fact, when $\hat{\pi} = [0 \ 0 \ 1]$ and $\delta_2 \leq 1.5$, preventive replacement is never optimal.

Figure 7 depicts the optimal average costs as a function of the reactive replacement time δ_2 . Interestingly, the costs associated with the cyclic environments (described by generator matrices Q_1 and Q_3) are nearly identical, and those associated with the acyclic environment (described by Q_2) are significantly higher when δ_2 is small to moderate. It is surmised that the average cost is higher in the latter case because it may not be possible to avoid state 3 for at least one period by initiating

a preventive replacement in states 1 or 2. However, the average costs are more comparable as the reactive replacement time increases.

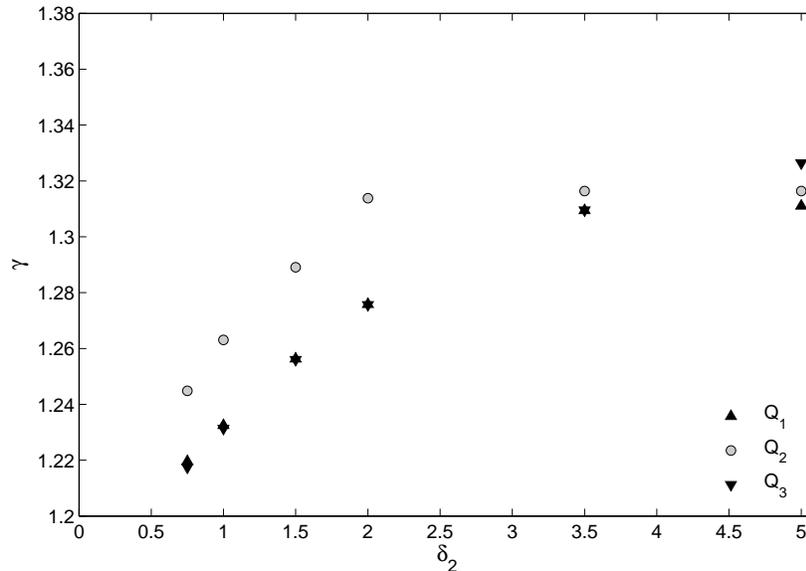


Figure 7: Optimal average cost as a function of δ_2 .

Acknowledgements

The authors are grateful to three anonymous referees and the Associate Editor, whose comments have improved the paper. This research was sponsored, in part, by grants from the U.S. National Science Foundation (CMMI-0856702 and CMMI-1266194).

References

- [1] R. Barlow and F. Proschan. *Mathematical Theory of Reliability*. John Wiley & Sons, Inc., New York, NY, 1965.
- [2] M. Bladt, B. Meini, M. Neuts, and B. Sericola. Distributions of reward functions on continuous-time Markov chains. In *Proceedings of the Fourth International Conference on Matrix Analytic Methods in Stochastic Models*, Adelaide, Australia, 2002.
- [3] E. Byon and Y. Ding. Season-dependent condition-based maintenance for a wind turbine using a partially observed Markov decision process. *IEEE Transactions on Power Systems*, 25(4):1823–1834, 2010.
- [4] E. Byon, L. Ntaimo, and Y. Ding. Optimal maintenance strategies for wind turbine systems under stochastic weather conditions. *IEEE Transactions on Reliability*, 59(2):393–404, 2010.
- [5] B. Çekyay and S. Özekici. Condition-based maintenance under Markovian deterioration. In J. Cochran, A. Cox, P. Keskinocak, J. P. Kharoufeh, and J. C. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*, Hoboken, NJ, 2011. John Wiley & Sons, Inc.

- [6] M. Cutler. Bearing life prediction in paper making machinery, including environmental effects. In *TAPPI Engineering Conference*, Toronto, Ontario, 1990.
- [7] L. Dueñas-Osorio and B. Basu. Unavailability of wind turbines due to wind-induced accelerations. *Engineering Structures*, 30(4):885–893, 2008.
- [8] J. Flory, J. Kharoufeh, and N. Gebraeel. A switching diffusion model for lifetime estimation in randomly-varying environments. *IIE Transactions*, 46(11):1227–1241, 2014.
- [9] N. Gebraeel, M. Lawley, R. Li, and J. Ryan. Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions*, 37(6):543–557, 2005.
- [10] T. Harris and R. Barnsby. Life rating for ball and roller bearings. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 215(6):577–595, 2001.
- [11] J. Ivy and S. Pollock. Marginally monotonic maintenance policies for a multi-state deteriorating machine with probabilistic monitoring and silent failures. *IEEE Transactions on Reliability*, 54(3):489–497, 2005.
- [12] B. Jakoby, M. Scherer, M. Buskies, and H. Eisenschmid. An automotive engine oil viscosity sensor. *IEEE Sensors Journal*, 3(5):562–568, 2003.
- [13] R. Jiang, M. Kim, and V. Makis. Availability maximization under partial observations. *OR Spectrum*, 35(3):691–710, 2013.
- [14] J. Kharoufeh and S. Cox. Stochastic models for degradation-based reliability. *IIE Transactions*, 37(6):533–542, 2005.
- [15] J. Kharoufeh, D. Finkelstein, and D. Mixon. Availability of periodically inspected systems with Markovian wear and shocks. *Journal of Applied Probability*, 43(2):303–317, 2006.
- [16] M. Kim and V. Makis. Optimal control of a partially observable failing system with costly multivariate observations. *Stochastic Models*, 28(4):584–608, 2012.
- [17] M. Kurt and J. Kharoufeh. Monotone optimal replacement policies for a Markovian deteriorating system in a controllable environment. *Operations Research Letters*, 38(4):273–279, 2010.
- [18] M. Kurt and L. Maillart. Structured replacement policies for systems receiving shocks with a Markov-modulated Poisson arrival rate. *Operations Research Letters*, 37(4):280–284, 2009.
- [19] J. Leichty and G. Roberts. Markov chain Monte Carlo methods for switching diffusion models. *Biometrika*, 88(2):299–315, 2001.
- [20] L. Maillart. Maintenance policies for systems with condition monitoring and obvious failures. *IIE Transactions*, 38(6):463–475, 2006.
- [21] L. Maillart and L. Zheltova. Structured maintenance policies on interior sample paths. *Naval Research Logistics*, 54(6):645–655, 2007.
- [22] V. Makis and X. Jiang. Optimal replacement under partial observations. *Mathematics of Operations Research*, 28(2):382–394, 2003.

- [23] S. Özekici. Optimal maintenance policies in random environments. *European Journal of Operational Research*, 2(3):283–294, 1995.
- [24] J. Peinke, S. Barth, F. Böttcher, D. Heinemann, and B. Lange. Turbulence, a challenging problem for wind energy. *Physica A: Statistical Mechanics and its Applications*, 338(1-2):187–193, 2004.
- [25] W. Pierskalla and J. Voelker. A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23(3):353–388, 2006.
- [26] M. Puterman. *Markov Decision Processes*. John Wiley and Sons Inc., New York, NY, 2005.
- [27] B. Sericola. Occupation times in Markov chains. *Communications in Statistics–Stochastic Models*, 16(5):479–510, 2000.
- [28] Y. Sherif and M. Smith. Optimal maintenance models for systems subject to failure. *Naval Research Logistics Quarterly*, 28(1):47–74, 1981.
- [29] M. Ulukus, J. Kharoufeh, and L. Maillart. Optimal replacement policies under environment-driven degradation. *Probability in the Engineering and Informational Sciences*, 26(3):405–424, 2012.
- [30] K. Waldmann. Optimal replacement under additive damage in randomly varying environments. *Naval Research Logistics*, 30(3):377–386, 1983.
- [31] H. Wang. A survey of maintenance policies of deteriorating systems. *European Journal of Operations Research*, 139(3):469–489, 2002.
- [32] R. Wisser and M. Bolinger. *2010 Wind Technologies Market Report*. U.S. Department of Energy, Washington, D.C., June 2011.
- [33] E. Yaylali and J. Ivy. Partially observable MDPs (POMDPs): An introduction. In J. Cochran, L. Cox, P. Keskinocak, J. Kharoufeh, and J. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*, pages 4017–4036, New York, NY, 2011. John Wiley and Sons, Inc.
- [34] E. Zhou, M. Fu, and S. Marcus. Solving continuous-state POMDPs via density projection. *IEEE Transactions on Automatic Control*, 55(5):1101–1116, 2010.
- [35] Y. Zhou, L. Ma, J. Mathew, Y. Sun, and R. Wolff. Maintenance strategy optimization using a continuous-state partially observable semi-Markov decision process. *Microelectronics Reliability*, 51(2):300–309, 2011.

Appendix: Proof of Theorem 4

Proof. Assume $\mathcal{S} = \{1, 2\}$ and the environment process has parameters

$$\mathbf{Q} = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \quad \mathbf{r} = [r_0 \quad r_1],$$

where $\alpha > \beta > 0$ and $r_1 > r_0 > 0$. Define $\lambda = \beta/\alpha$, $\bar{\lambda} = 1 - \lambda$, and $\hat{\mathbf{P}} = \mathbf{I} + \mathbf{Q}/\alpha$, where $\hat{\mathbf{P}}$ is the transition probability matrix of the uniformized environment process given by

$$\hat{\mathbf{P}} = \begin{bmatrix} 0 & 1 \\ \lambda & 1 - \lambda \end{bmatrix}.$$

Recall that $W_{ij}(x, t) = \mathbb{P}(X(t) \leq x, Z(t) = j | Z(0) = i)$ and define the matrix $\mathbf{W}(x, t) = [W_{ij}(x, t)]$, $i, j \in \mathcal{S}$. Sericola [27] obtained a series representation $\mathbf{W}(x, t)$ which, for $\ell = 2$, is

$$\mathbf{W}(x, t) = \sum_{n=0}^{\infty} e^{-\alpha t} \frac{(\alpha t)^n}{n!} \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} \mathbf{C}(n, k), \quad (37)$$

where the matrix $\mathbf{C}(n, k) = [c_{ij}(n, k)]$ is given by

$$c_{ij}(n, k) = \begin{cases} c_{2j}(n-1, k), & i = 1, j \in \{1, 2\}, k \in \{0, 1, \dots, n-1\}, \\ \left(\hat{\mathbf{P}}^n\right)_{1j}, & i = 1, j \in \{1, 2\}, k = n, \\ 0, & i = 2, j \in \{1, 2\}, k = 0, \\ \lambda c_{1j}(n-1, k-1) + \bar{\lambda} c_{2j}(n-1, k-1), & i = 2, j \in \{1, 2\}, k \in \{1, 2, \dots, n\}. \end{cases}$$

For notational brevity, let $c_i(n, k) \equiv c_{i1}(n, k) + c_{i2}(n, k)$ be the i th row sum of $\mathbf{C}(n, k)$, $i \in \{1, 2\}$. It will be shown that a threshold policy exists with respect to π for fixed x if $c_1(n, k) \leq c_2(n, k)$ for all $k \leq n$ and $n \geq 0$. To prove this condition, we employ an induction argument that requires a few preliminary lemmas.

Lemma 11. For $j \in \{1, 2\}$,

$$c_{2j}(n, k) = \begin{cases} \lambda^{n-k} c_{2j}(2k-n, 2k-n) \\ \quad + \bar{\lambda} \left[\sum_{s=1}^{n-k} \lambda^{n-k-s} c_{2j}(2k-n+2s-1, 2k-n+s-1) \right], & k \geq n/2, \\ 0, & k < n/2. \end{cases}$$

Proof. The result can be obtained by recursive substitution of the function definitions. □

Lemma 12. For $n \geq 1$, the n -step transition probability matrix of the uniformized chain is

$$\hat{\mathbf{P}}^n = \begin{bmatrix} 1 - \sum_{k=0}^{n-1} (-1)^k \lambda^k & \sum_{k=0}^{n-1} (-1)^k \lambda^k \\ 1 - \sum_{k=0}^n (-1)^k \lambda^k & \sum_{k=0}^n (-1)^k \lambda^k \end{bmatrix}.$$

Proof. The result follows directly by induction on n . □

Lemma 13. For $n \geq 1$, $c_2(n, n) = \lambda \sum_{k=0}^{n-1} \bar{\lambda}^k$.

Proof. Using Lemmas 11 and 12 and $\mathbf{C}(0, 0) = \mathbf{0}$,

$$\begin{aligned} c_2(n, n) &= \lambda \left[\left(\widehat{\mathbf{P}}^{n-1} \right)_{11} + \left(\widehat{\mathbf{P}}^{n-1} \right)_{12} \right] + \lambda \bar{\lambda} \left[\left(\widehat{\mathbf{P}}^{n-2} \right)_{11} + \left(\widehat{\mathbf{P}}^{n-2} \right)_{12} \right] \\ &\quad + \lambda \bar{\lambda}^2 \left[\left(\widehat{\mathbf{P}}^{n-3} \right)_{11} + \left(\widehat{\mathbf{P}}^{n-3} \right)_{12} \right] + \cdots + \lambda \bar{\lambda}^{n-1} \left[(\mathbf{I})_{11} + (\mathbf{I})_{12} \right] \\ &= \lambda \sum_{k=0}^{n-1} \bar{\lambda}^k. \quad \square \end{aligned}$$

Lemma 14. For all $n \geq 0$, $c_2(n, n) \leq c_1(n, n)$. Moreover, for all $n \geq 1$, $c_2(n, n-1) \leq c_1(n, n-1)$.

Proof. To prove the first part, note that by Lemma 13,

$$c_2(n, n) = \lambda \sum_{k=0}^{n-1} \bar{\lambda}^k \leq 1 = c_1(n, n).$$

For the second part, observe that for $n \geq 1$ and $j \in \{1, 2\}$, Lemma 11 gives the following:

$$\begin{aligned} c_{2j}(n, n-1) &= \lambda c_{2j}(n-2, n-2) + \lambda \bar{\lambda} c_{2j}(n-3, n-3) + \lambda \bar{\lambda}^2 c_{2j}(n-4, n-4) + \cdots + \lambda \bar{\lambda}^{n-3} c_{2j}(1, 1) \\ &\quad + \lambda \bar{\lambda}^{n-2} c_{2j}(0, 0) + \bar{\lambda}^{n-1} c_{2j}(1, 0). \end{aligned}$$

Therefore, we see that

$$\begin{aligned} c_2(n, n-1) &= \lambda^2 \left[\sum_{k=0}^{n-3} \bar{\lambda}^k + \bar{\lambda} \sum_{k=0}^{n-4} \bar{\lambda}^k + \bar{\lambda}^2 \sum_{k=0}^{n-5} \bar{\lambda}^k + \cdots + \bar{\lambda}^{n-3} \right] \\ &\leq \lambda^2 \left[\frac{1}{\lambda} + \bar{\lambda} \frac{1}{\lambda} + \bar{\lambda}^2 \frac{1}{\lambda} + \cdots + \bar{\lambda}^{n-3} \frac{1}{\lambda} \right] \\ &\leq \lambda \sum_{k=0}^{n-2} \bar{\lambda}^k \\ &= c_2(n-1, n-1) = c_1(n, n-1), \end{aligned}$$

where the last equality holds by the first part of the lemma. □

Lemma 15. For all $n \geq 0$ and $k \leq n$, $c_2(n, k) \leq c_2(n-1, k)$.

Proof. By Lemma 11,

$$\begin{aligned}
c_2(n, k) &= \underbrace{\lambda^{n-k} c_2(2k-n, 2k-n)}_{(1)} + \underbrace{\lambda^{n-k-1} \bar{\lambda} c_2(2k-n+1, 2k-n)}_{(2)} \\
&\quad + \underbrace{\lambda^{n-k-2} \bar{\lambda} c_2(2k-n+3, 2k-n+1)}_{(3)} + \cdots + \underbrace{\lambda^2 \bar{\lambda} c_2(n-5, k-3)}_{(n-k-1)} \\
&\quad\quad\quad + \underbrace{\lambda \bar{\lambda} c_2(n-3, k-2) + \bar{\lambda} c_2(n-1, k-1)}_{(n-k)} \\
c_2(n-1, k) &= \underbrace{\lambda^{n-k-1} c_2(2k-n+1, 2k-n+1)}_{(1)} + \underbrace{\lambda^{n-k-2} \bar{\lambda} c_2(2k-n+2, 2k-n+1)}_{(2)} \\
&\quad + \underbrace{\lambda^{n-k-3} \bar{\lambda} c_2(2k-n+4, 2k-n+2)}_{(3)} + \cdots + \underbrace{\lambda^2 \bar{\lambda} c_2(n-6, k-3)}_{(n-k-2)} \\
&\quad\quad\quad + \underbrace{\lambda \bar{\lambda} c_2(n-4, k-2) + \bar{\lambda} c_2(n-2, k-1)}_{(n-k-1)}
\end{aligned}$$

A pairwise comparison of terms $\{(3), \dots, (n-k)\}$ in $c_2(n, k)$ with terms $\{(2), \dots, (n-k-1)\}$ in $c_2(n-1, k)$ shows that each pair is of the form $\alpha c_2(w, v)$ and $\alpha c_2(w-1, v)$, respectively, where $w \leq n$, $v \geq w/2$, and $\alpha > 0$. Therefore, the paired terms in each set have the same relationship as the generating expressions, for which $(\alpha, w, v) = (1, n, k)$. It follows that if the recursion is true for $(\alpha, w, v) = (1, n, k)$, then it holds for all the generated terms by extension. Thus, it is sufficient to show that the sum of terms (1) and (2) in $c_2(n, k)$ is less than or equal to term (1) of $c_2(n-1, k)$. To prove this, we first note that $c_{ij}(n, k) \leq c_{ij}(n, k+1)$ for all $i, j \in \{1, 2\}$, $n \geq 0$, and $k \leq n$ (see [27]). The result follows by observing that the sum of terms (1) and (2) in $c_2(n, k)$ is bounded above by λ^{n-k} , while term (1) of $c_2(n-1, k)$ is bounded below by the same quantity. \square

Lemma 16. *For all $n \geq 0$ and $k \leq n$, $c_2(n, k) \leq c_1(n, k)$.*

Proof. In the case where $k \leq n/2$, $c_2(n, k) = 0 \leq c_1(n, k)$ by Lemma 11. For $k \geq n/2$, Lemma 14 proves that the result holds for all $n \geq 0$ and $k \in \{n-1, n\}$. Assume that $c_2(n, k) \leq c_1(n, k)$ for all $n \geq 0$ and $k \in \{n-(w-1), n-(w-2), \dots, n\}$ for some integer $w \geq n/2$. Then by Lemma 11, it can be shown that

$$\begin{aligned}
c_1(n, n-w) &= K_1 + \bar{\lambda} c_2(n-2, n-w-1) \quad \text{and} \\
c_2(n, n-w) &= K_2 + \bar{\lambda} c_2(n-1, n-w-1),
\end{aligned}$$

where $K_1 > K_2 > 0$. By Lemma 15, $c_2(n-1, n-w-1) \leq c_2(n-2, n-w-1)$ so that $c_2(n, n-w) \leq c_1(n, n-w)$, and the induction holds for w . \square

Proposition 4. *Let $\pi(\rho) = [1-\rho, \rho]$ for some $\rho \in [0, 1]$ and fix $x \in [0, \xi]$. Then $H(x, \pi(\rho))$ is monotone increasing in ρ , and $\tau(x, \pi(\rho))$ is monotone decreasing in ρ .*

Proof. For $\rho \in [0, 1]$ and $(x, \pi(\rho)) \in B$, we have that

$$H(x, \pi(\rho)) = 1 - \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} W_{ij}(x, \delta) \pi^{(i)}(\rho) = \rho \left[\sum_{j \in \mathcal{S}} W_{1j}(x, \delta) - W_{2j}(x, \delta) \right].$$

Therefore, to prove $H(x, \pi(\rho))$ is monotone increasing in ρ , it suffices to show

$$\sum_{j \in \mathcal{S}} (W_{1j}(x, \delta) - W_{2j}(x, \delta)) \geq 0 \quad (38)$$

for all $x \in [0, \xi]$ and $\delta \geq 0$. By Lemma 16, $c_1(n, k) \geq c_2(n, k)$ for all $n \geq 0$ and $k \leq n$. Hence, it follows immediately from (37) that

$$\sum_{j \in \mathcal{S}} W_{1j}(x, \delta) \geq \sum_{j \in \mathcal{S}} W_{2j}(x, \delta).$$

Likewise, to prove that $\tau(x, \pi(\rho))$ is monotone decreasing in ρ , observe that

$$\begin{aligned} \tau(x, \pi(\rho)) &= \sum_{i \in \mathcal{S}} \left[\sum_{j \in \mathcal{S}} \int_0^\delta W_{ij}(\xi - x, v) dv \right] \pi^{(i)}(\rho) \\ &= \int_0^\delta \left[1 - \rho \sum_{j \in \mathcal{S}} (W_{1j}(x, v) - W_{2j}(x, v)) \right] dv. \end{aligned}$$

The integrand is monotone decreasing in ρ as is $\tau(x, \pi(\rho))$. \square

Finally, by Lemma 4, we note that

$$V(x, \pi(\rho)) \geq \underline{V}_0(x, \pi(\rho)) = (C_0 - c_1) + (c_1 + c_2)H(x, \pi(\rho)) - \gamma\tau(x, \pi(\rho)).$$

By Proposition 4, $\underline{V}_0(x, \pi(\rho))$ is monotone increasing in ρ ; therefore, if it is optimal to replace in state $(x, \pi(\rho)) \in B$, it remains optimal to replace for all $(x, \pi(\rho')) \in B$ such that $\rho \leq \rho' \leq 1$. \square