

Optimally Maintaining a Markovian Deteriorating System with Limited Imperfect Repairs

Murat Kurt¹ and Jeffrey P. Kharoufeh²

Department of Industrial Engineering

University of Pittsburgh

1048 Benedum Hall

3700 O'Hara Street

Pittsburgh, PA 15261 USA

Final version appears in

European Journal of Operational Research, 205 (2010), 368-380.

Abstract

We consider the problem of optimally maintaining a periodically inspected system that deteriorates according to a discrete-time Markov process and has a limit on the number of repairs that can be performed before it must be replaced. After each inspection, a decision maker must decide whether to repair the system, replace it with a new one, or leave it operating until the next inspection, where each repair makes the system more susceptible to future deterioration. If the system is found to be failed at an inspection, then it must be either repaired or replaced with a new one at an additional penalty cost. The objective is to minimize the total expected discounted cost due to operation, inspection, maintenance, replacement and failure. We formulate an infinite-horizon Markov decision process model and derive key structural properties of the resulting optimal cost function that are sufficient to establish the existence of an optimal threshold-type policy with respect to the system's deterioration level and cumulative number of repairs. We also explore the sensitivity of the optimal policy to inspection, repair and replacement costs. Numerical examples are presented to illustrate the structure and the sensitivity of the optimal policy.

Keywords: Reliability, limited repairs, threshold-type policy, Markov decision processes.

¹Corresponding author. Tel: +1 412 624 9807. Fax: +1 412 624 9831. Email: muk7@pitt.edu.

²Tel: +1 412 624 9832. Fax: +1 412 624 9831. Email: jkharouf@pitt.edu.

1 Introduction

In this paper, we consider the problem of optimally maintaining a periodically inspected, repairable system that is subject to stochastic deterioration and has a limit on the number of repairs it may undergo before it must be replaced. Following an inspection, if repair is still feasible, a decision maker must decide whether to repair the system, replace it with a new one, or leave it operating until the next inspection. On the other hand, if repair is no longer- feasible, then the system can only be replaced, or it may be allowed to continue operating if it is not failed. Each repair makes the system more susceptible to future deterioration. The primary objective is to devise a maintenance and replacement strategy that minimizes the total expected discounted cost due to operation, inspection, maintenance, replacement and failure. To this end, we formulate and analyze a sequential decision model.

Stochastically deteriorating systems that can be repaired only a limited number of times can be found in a variety of engineering and service applications. For example, aircraft engine turbine blades operate at high temperatures and experience centrifugal stresses that lead to elongated particles at the microstructure level. This elongation leads to degradation of fatigue strength that can result in voids and crack initiation in the blade tips. The blades can be reworked only a limited number of times before the blade must be scrapped and replaced with a new one to ensure flight safety. Another example is the strategic use of a limited number of available repairs of a system that is purchased with a limited warranty. While most warranties are valid for a fixed time period, some impose a limit on the number of allowable repair actions before the warranty becomes void. Finally, we can consider systems for which each repair exhausts one unit from a finite pool of units needed to repair a critical component in the system. For instance, in military applications, only a finite number of line-replaceable units (LRUs) are held in inventory to perform routine aircraft maintenance between planned missions. Once the LRUs are exhausted, it may be necessary to replace an entire subsystem such as the aircraft's engine.

Maintenance optimization models have been studied extensively in the operations research literature for the past few decades. For a broad overview of several classes of models, the reader is referred to the excellent surveys by Pierskalla and Voelker [19], Sherif and Smith [23], Valdez-Flores and Feldman [25], Pham and Wang [18], Scarf [22], and Wang [26]. Although many researchers have studied systems subject to Markovian deterioration, the case of a limited number of repairs that we examine has been addressed by only a few. A common theme among those approaches is

the development of optimization models for systems that deteriorate stochastically under limited sequential preventive maintenance actions with single or multiple failure modes. In these models, each preventive maintenance action is imperfect in the sense that it affects either the virtual age or the hazard rate function of the system. A representative sample of these works includes [4, 5, 7, 8, 12, 13, 15, 16, 17]. The majority of these models set out to determine the length of a working period between each successive preventive maintenance action and/or the optimal number of repairs that minimizes the average cost rate. To our knowledge, there has not been much emphasis on systems with *predetermined* repair restrictions with the exception of two papers. Goyal et al. [9] studied the lifetime distribution and mean lifetime of a deteriorating unit which can be repaired only a limited number of times. More recently, Lugtigheid et al. [14] developed a dynamic repair/replacement decision model in which the system can be replaced with a new one at any point in time or repaired at failure or preventively, where each repair action increases the system's failure intensity deterministically, and the total number of failure and preventive repairs is limited to fixed number. To minimize the total expected cost of repair and replacement over a finite planning horizon, they formulated a continuous-time Markov decision process (MDP) model and explored several structural properties of the underlying optimal cost function and policy. Specifically, for a system with an exponential lifetime distribution with a failure intensity that is nondecreasing in the cumulative number of completed repairs, they established the non-optimality of preventive repair and replacement actions so that the system can only be repaired or replaced at the time of a failure. They provide a time-dependent control-limit function on the system's cumulative number of repairs to reveal the underlying optimal policy by a switching curve.

In this work, we study optimal maintenance and replacement strategies for a periodically inspected, repairable system that deteriorates stochastically and can be repaired only a limited number of times before it must be replaced. The system's deterioration status is assumed to evolve as a discrete-time Markov chain (DTMC) on a finite state space whose transition probability matrix is influenced by each repair so that the system becomes more susceptible to deterioration as the cumulative number of completed repairs increases. Our main objective is to determine the optimal maintenance or replacement actions in each state that minimize the total expected discounted cost due to operating the system, repairs, inspections, failures, and replacements. To this end, we formulate the problem as a discrete-time, infinite-horizon, discounted MDP model that provides a unifying framework for a broad class of discrete-time, condition-based Markov deterioration models by limiting the number of repairs on the system and incorporating the negative effects of each repair

action into the system’s deterioration process. Therefore, from a modeling point of view, it can be viewed as generalization of a class of optimal replacement models for systems subject to Markovian deterioration.

Our main results include sufficient conditions for the existence of an optimal threshold-type policy and a formal sensitivity analysis of the optimal policy with respect to the major cost parameters of the model. As such, it differs from that of Lugtigheid et al. [14] in a few important ways. Specifically, we derive sufficient conditions that ensure the existence of an optimal policy resulting in three non-overlapping regions separated by switching curves: (i) a region where it is always optimal to do nothing; (ii) a region where it is always optimal to perform maintenance; and (iii) a region for which it is always optimal to replace the system with a new one. We also examine the behavior of these switching curves and their sensitivity to one or two-way changes in inspection, maintenance and replacement costs. With such a sensitivity analysis, we reveal intuitive and interesting insights into the effects of the cost parameters on the optimal policy. Furthermore, we illustrate the structure and the sensitivity of the optimal policy by a few numerical examples. Methodologically, because the proofs of some of our results are difficult to approach by standard dynamic programming techniques, we propose novel methods to establish the main results.

The remainder of the paper is organized as follows: Section 2 defines the problem in detail and provides a formal description of our MDP model. In section 3, we explore several structural properties of the optimal cost function, including its monotonicity, and prove the existence of an optimal threshold-type policy. Section 4 examines the sensitivity of the optimal actions in each state to the major cost parameters. Finally, in section 5, we illustrate the optimal policies graphically through four numerical examples that highlight the behavior of the optimal policies under various conditions.

2 Model Description and Formulation

We consider a repairable system whose deterioration status evolves as a discrete-time Markov chain (DTMC) on a finite state space, $\Psi = \{0, 1, \dots, S\}$, which is structured in order of increasing deterioration levels so that state 0 corresponds to the least amount of deterioration and state S represents failure of the system. The deterioration status of the system is inspected periodically, i.e., it is inspected at time points $\{\tau, 2\tau, 3\tau, \dots\}$ for some $\tau > 0$, and the cost of each inspection is I . Given that the system is found to be in an operational state $s \in \Psi \setminus \{S\}$, the decision maker

can choose among three actions: (i) operate the system for one more period at an immediate cost of $c(s) < \infty$, (ii) repair the system at a cost of M , or (iii) immediately replace the system with a new one at a cost of $R < \infty$ where $R \geq M$. On the other hand, if the system is found to be failed after an inspection, then it must be either repaired or replaced with a new one. Each repair restores the status of the system to the best possible operating condition in a negligible amount of time but affects the evolution of the system's deterioration process by accelerating its movements toward more degraded states. Because the system's susceptibility to deterioration increases by each repair, only a limited number of repairs, N , is allowed. Therefore, following an inspection after N repairs have been completed, given the system is found to be in an operational state, it can either be left in operation for one more period or replaced with a new one; however, if it is detected to be in a failed state then it must be replaced with a new one. Each failure has a fixed penalty cost denoted by ρ ($\rho < \infty$). All costs are discounted at a rate β per period, $0 < \beta < 1$, and the objective is to minimize the total expected discounted cost due to operation, inspection, maintenance, replacement and failure.

We formulate a discrete-time, infinite-horizon, discounted MDP model for this problem. In our model, each decision epoch refers to the instant of time just after an inspection, and the time interval between any two successive inspections is called a *period*. Specifically, we let period t refer to the time interval between decision epochs t and $t + 1$. For notational convenience, we define $\Theta = \{1, 2, \dots, N\}$, $\Theta' = \Theta \setminus \{N\}$ and let $\Phi = \Psi \setminus \{S\}$ denote the set of operational states. Therefore, the state of the MDP is an ordered pair $(s, n) \in \Upsilon \equiv \Psi \times \Theta$, where $s \in \Psi$ represents the deterioration status of the system and $n \in \Theta$ denotes the cumulative number of completed repairs. After each inspection, the decision-maker chooses an action from the set $\{0, 1, 2\}$, where 0 means waiting for one more period, 1 means repairing the system, and 2 means replacing the system immediately. Because each repair makes the system less resistant to deterioration, the system's deterioration process depends on the number of completed repairs, which we model as follows: For a system that has been repaired n times prior to epoch t , we let $P(s'|s, n)$ denote the probability that the system transitions to state s' at epoch $t+1$ given that it starts operating in state s at epoch t . Note that waiting for one more period is not a feasible action if the system is found to be failed at an inspection. Therefore, for each pair $(s, n) \in \Phi \times \Theta$, we let $w(s, n)$ denote the total expected discounted cost corresponding to action 0 when the process starts in state (s, n) . Then, for each $(s, n) \in \Upsilon$, we define $\vartheta(s, n)$ as the minimum total expected discounted cost starting in state (s, n) and denote the corresponding optimal action by $a(s, n)$. Note that in any state $(s, n) \in \Upsilon$, action 2

immediately moves the process into state $(0, 0)$ with probability 1. Likewise, because repairing the system restores its status to the best operating condition, but increments the cumulative number of repairs by one, if feasible, action 1 moves the process into state $(0, n + 1)$ immediately. Then, we can characterize an optimal policy that yields the minimum total expected cost in each state $(s, n) \in \Upsilon$ by solving the following optimality equations:

$$\vartheta(s, n) = \min \{w(s, n), \vartheta(0, n + 1) + M, \vartheta(0, 0) + R\} \quad \text{for } (s, n) \in \Phi \times \Theta', \quad (1a)$$

$$\vartheta(s, N) = \min \{w(s, N), \vartheta(0, 0) + R\} \quad \text{for } s \in \Phi, \quad (1b)$$

$$\vartheta(S, n) = \min \{\vartheta(0, n + 1) + M, \vartheta(0, 0) + R\} \quad \text{for } n \in \Theta', \quad (1c)$$

$$\vartheta(S, N) = \vartheta(0, 0) + R, \quad (1d)$$

where for $(s, n) \in \Phi \times \Theta'$,

$$w(s, n) = c(s) + \beta \left(P(S|s, n)[\vartheta(S, n) + \rho + I] + \sum_{s' \in \Phi} P(s'|s, n) [\vartheta(s', n) + I] \right). \quad (2)$$

To simplify the notation, for $(s, n) \in \Phi \times \Theta'$, we define $\ell(s, n) = c(s) + \beta[P(S|s, n)\rho + I]$ as the *total expected immediate cost* due to operation, failure and inspection. Then the expected waiting cost function of (2) can be rewritten as

$$w(s, n) = \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n)\vartheta(s', n) \quad \text{for } (s, n) \in \Phi \times \Theta. \quad (3)$$

Note that because the immediate operation, replacement and failure penalty costs are finite, by the fact that $\beta < 1$, the optimality equations (1a)-(1d) and (3) admit a unique and finite solution [3]. It is also worth noting that if $N = 0$, i.e., no repairs are allowed, or $M = \infty$, then the optimality equations of our model reduce to those of a classical optimal replacement problem for a Markovian deteriorating system originally introduced by Derman [6], and extended by Klein [10] and Kolesar [11], among others. Therefore, our framework unifies the modeling approaches to the optimal replacement problem under limited imperfect repairs.

In section 3, we explore the structure of the optimal maintenance and replacement policy of the MDP model formulated in this section. Structured policies not only lend themselves to characterization by more efficient solution techniques [3, 21, 24], but they also help facilitate more accurate and efficient approximation architectures when the exact solution of the problem becomes computationally intractable due to the curse of dimensionality [2, 20].

3 Structural Properties

Our results in this section primarily concern the monotonicity of the optimal cost function, ϑ , and its effects on the structure of the optimal policy. Under reasonable assumptions, we show that the optimal cost function is monotonically nondecreasing in both dimensions of the state of the MDP. We exploit the monotonicity of the optimal cost function to characterize an optimal policy in the form of three disjoint regions, one for each action, which are partitioned by two switching curves. We also provide a sufficient condition to examine a monotonic behavior of the switching curve that separates the wait region from repair and replace regions.

Prior to our analysis we make some structural assumptions about the immediate operating cost function and the system's deterioration matrices. First, we assume that it is more costly to operate the system as its condition worsens.

Assumption 1 : *The immediate operating cost function $c(s)$ is nondecreasing in $s \in \Phi$.*

Assumption 2 relates the system's likelihood of moving into a set of states which are worse than a given particular condition. In fact, Assumption 2 can be viewed as the first order stochastic dominance relationship among the rows of the matrix $\mathbf{P}(n) \equiv [P(s'|s, n)]_{s \in \Phi, s' \in \Psi}$, which is also commonly known as the increasing failure rate (IFR) property for stochastic matrices [1]. The IFR property is commonly utilized in the reliability and maintenance optimization literature to analyze the structure of the underlying optimal cost functions and policies [19, 25].

Assumption 2 : *For each $n \in \Theta$, $\sum_{s'=j}^S P(s'|s+1, n) \geq \sum_{s'=j}^S P(s'|s, n)$ for all $s \in \Phi$ and $j \in \Psi$.*

This assumption implies that, given two systems functioning with the same repair history (i.e., with the same cumulative number of completed repairs) in states s and $s+1$, respectively, the system in state $s+1$ is more likely to be worse at the next epoch than the system in state s . Note that Assumption 2 orders the system's rate of deterioration with respect to its current deterioration status. Next, Assumption 3 orders the system's rate of deterioration with respect to its repair history.

Assumption 3 : *For each $s \in \Phi$, $\sum_{s'=j}^S P(s'|s, n+1) \geq \sum_{s'=j}^S P(s'|s, n)$ for all $n \in \Theta'$ and $j \in \Psi$.*

Similar to Assumption 2, Assumption 3 can be viewed as the first order stochastic dominance relationship among the system's deterioration matrices corresponding to different repair histories. More explicitly, given two systems with different repair histories functioning at the same level of

deterioration, the system with the greater cumulative number of completed repairs is more likely to get worse than the other during operation. Along with Assumptions 1–3, by definition, it is clear that the total expected immediate cost function ℓ is nondecreasing in both dimensions of the state of the process. In the remainder of this section, we suppose Assumptions 1–3 hold.

The value iteration algorithm, also known as the successive approximation algorithm, is one of the most commonly used techniques to solve discounted MDP models. Because of its simple structure, it is easy to understand and to implement when solving models numerically. Moreover, it is widely used to derive structural results regarding the optimal value function and the policies of these models. The algorithm starts with an arbitrary, finite initial value vector and updates the values associated with each possible action iteratively until a certain performance guarantee is satisfied. In Lemma 1, we describe a sequence of value iterates in an ordinary value iteration algorithm for our model and analyze its structure.

Lemma 1 : For $(s, n) \in \Upsilon$ and $k \geq 0$, define the function $\vartheta^k(s, n)$ as follows:

$$\vartheta^0(s, n) = 0 \quad \text{for all } (s, n) \in \Upsilon, \quad (4a)$$

$$\vartheta^{k+1}(s, n) = \min \left\{ w^{k+1}(s, n), \vartheta^k(0, n+1) + M, \vartheta^k(0, 0) + R \right\} \quad \text{for } (s, n) \in \Phi \times \Theta' \text{ and } k \geq 0, \quad (4b)$$

$$\vartheta^{k+1}(s, N) = \min \left\{ w^{k+1}(s, N), \vartheta^k(0, 0) + R \right\} \quad \text{for } s \in \Phi \text{ and } k \geq 0, \quad (4c)$$

$$\vartheta^{k+1}(S, n) = \min \left\{ \vartheta^k(0, n+1) + M, \vartheta^k(0, 0) + R \right\} \quad \text{for } n \in \Theta' \text{ and } k \geq 0, \quad (4d)$$

$$\vartheta^{k+1}(S, N) = \vartheta^k(0, 0) + R \quad \text{for } k \geq 0, \quad (4e)$$

where

$$w^{k+1}(s, n) = \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^k(s', n) \quad \text{for } (s, n) \in \Phi \times \Theta' \text{ and } k \geq 0. \quad (5)$$

Then, the following hold for all $k \geq 0$:

- (i) $\vartheta^k(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$,
- (ii) $\vartheta^k(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Psi$.

Proof : We will prove the results by induction on $k \geq 0$.

(i) By construction, $\vartheta^0(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$. As the induction hypothesis, for some $m \geq 0$, suppose $\vartheta^m(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$. By Assumption 2 and Lemma 4.7.2 of Puterman [21], this implies that $\sum_{s' \in \Psi} P(s'|s, n) \vartheta^m(s', n)$ is nondecreasing in

$s \in \Phi$ for all $n \in \Theta$. Then, because $\ell(s, n)$ is nondecreasing in $s \in \Phi$ for all $n \in \Theta$, $w^{m+1}(s, n)$ is also nondecreasing in $s \in \Phi$ for all $n \in \Theta$. By the definition of the function ϑ^{m+1} , this implies that $\vartheta^{m+1}(s, n)$ is nondecreasing in $s \in \Phi$ for all $n \in \Theta$, and the result follows by induction.

(ii) By construction, $\vartheta^0(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Psi$. As the induction hypothesis, for some $m \geq 0$, suppose $\vartheta^m(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Psi$. This implies,

$$\sum_{s' \in \Psi} P(s'|s, n+1)\vartheta^m(s', n+1) \geq \sum_{s' \in \Psi} P(s'|s, n+1)\vartheta^m(s', n) \text{ for all } (s, n) \in \Phi \times \Theta'. \quad (6a)$$

From part (i), recall that $\vartheta^m(s', n)$ is nondecreasing in $s' \in \Psi$ for all $n \in \Theta$. Then, by Assumption 3, Lemma 4.7.2 of Puterman [21] implies

$$\sum_{s' \in \Psi} P(s'|s, n+1)\vartheta^m(s', n) \geq \sum_{s' \in \Psi} P(s'|s, n)\vartheta^m(s', n) \text{ for all } (s, n) \in \Phi \times \Theta'. \quad (6b)$$

By (6a) and (6b),

$$\sum_{s' \in \Psi} P(s'|s, n+1)\vartheta^m(s', n+1) \geq \sum_{s' \in \Psi} P(s'|s, n)\vartheta^m(s', n) \text{ for all } (s, n) \in \Phi \times \Theta'. \quad (7)$$

Note that $\ell(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Phi$. Therefore, by (7), $w^{m+1}(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Phi$. Then, by the definition of the function ϑ^{m+1} and the induction hypothesis, this implies that $\vartheta^{m+1}(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Psi$, and the result follows by induction. \square

By Lemma 1 part (i), we show that, for a particular repair history, the value iteration algorithm preserves the monotonicity of the function ϑ^k with respect to system's deterioration status at each iteration. Then, we exploit this in Lemma 1 part (ii) to show that, at each iteration, the algorithm also preserves the monotonicity of the function with respect to the number of repairs for any particular level of deterioration. Recall that the immediate operation, replacement and failure penalty costs are all finite and $\beta < 1$. Therefore, by Theorem 6.2.1 of Puterman [21], the value iterates $\{\vartheta^k(s, n)\}_{k \geq 0}$ converge to the unique solution of the optimality equations (4a)-(4e) and (5) componentwise, and the monotonicity properties of this sequence in the limit imply the monotonicity of the optimal cost function in both dimensions of the state of the process. More explicitly, by Lemma 1 part (i), Proposition 1 states that the optimal total expected cost does not decrease as the system gets worse with the same repair history.

Proposition 1 : *The function $\vartheta(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$.*

Similarly, by Lemma 1 part (ii), Proposition 2 states that the optimal total expected cost starting in any particular level of deterioration does not decrease as the system gets repaired more.

Proposition 2 : *The function $\vartheta(s, n)$ is nondecreasing in $n \in \Theta$ for all $s \in \Psi$.*

Note that Proposition 2 relies on all Assumptions 1–3, whereas Proposition 1 depends only on Assumptions 1 and 2.

As a consequence of Propositions 1 and 2, Theorem 1 reveals the existence of an optimal policy which is threshold structured in both the system’s deterioration status and the cumulative number of repairs. Specifically, it states that there exists an optimal policy which can be described as the union of three non-overlapping regions, one for each action. The proof of the theorem follows directly from Propositions 1 and 2 and is therefore omitted.

Theorem 1 : *There exists a threshold number of repairs $n^* \in \Theta$, and for each $n \in \Theta$, a threshold operational state $s_n^* \in \Phi$ such that for $n < n^*$*

$$a(s, n) = \begin{cases} 0 & \text{if } s \leq s_n^*, \\ 1 & \text{if } s > s_n^*, \end{cases}$$

and for $n \geq n^*$

$$a(s, n) = \begin{cases} 0 & \text{if } s \leq s_n^*, \\ 2 & \text{if } s > s_n^*. \end{cases}$$

Note that if $M > 0$, then because $\vartheta(0, n)$ is nondecreasing in $n \in \Theta$, by the definition of the function $\vartheta(0, \cdot)$, it is obvious that $a(0, n) \neq 1$ for any $n \in \Theta$. In other words, if maintenance is not free, then it is not optimal to repair a system when it is operating at its best condition. Similarly, when $R > 0$, it is never optimal to replace a unit if it has not been repaired so far and is operating at its best condition.

By Theorem 1, it is clear that the switching curve between repair and replace regions is a straight line. In what follows, for each $n \in \Theta$ we let $s_n^* = \max\{s \in \Psi : a(s, n) = 0\}$ be the associated threshold deterioration status beyond which it is either optimal to repair or replace a system that has been repaired n times. In Theorem 2, we examine the behavior of the thresholds s_n^* . Specifically, we provide a sufficient condition for s_n^* to be nonincreasing in the cumulative number of completed repairs, n . Intuitively, this can be interpreted as follows: As the cumulative number of completed repairs increases, the decision maker becomes less likely to choose action 0 (waiting) at the same deterioration level. For the optimal policy defined in Theorem 1, this implies monotonicity of the switching curve that separates the wait region from the repair and replace regions.

Theorem 2 : *If*

$$\beta\rho[P(S|s, n+1) - P(S|s, n)] \geq (1 - \beta)(R - M) \text{ for all } s \in \Phi \text{ and } n \in \Theta', \quad (8)$$

then, s_n^ is nonincreasing in $n \in \Theta$.*

Note that in Theorem 2, the left-hand side of inequality (8) represents an increase in the expected discounted failure penalty cost by an additional repair at a particular level of deterioration, and the term $R - M$ on the right-hand side represents an upper bound for the increase in the minimum expected cost of repair or replacement by an additional repair. Also note that the discount factor, β , approaches unity as the time between two consecutive inspections, τ , approaches zero. Therefore, condition (8) is weak for systems that are subject to frequent reviews.

Although the result of Theorem 2 is intuitive, it is difficult to establish using standard dynamic programming methods; therefore, we devise a novel technique to prove Theorem 2. Specifically, in Lemma 2, we describe a sequence of value iterates and prove its convergence to the unique solution of the optimality equations (1a)-(1d) and (2). The sequence that we describe differs from an ordinary value iteration algorithm as it updates the associated cost function by iteratively updating only the cost of waiting in each state. Note that this sequence is not appropriate to solve a given problem instance numerically without the value of the optimal cost function in states $\{(0, n) : n \in \Theta\}$.

Before proceeding to Lemma 2, for notational convenience, we restate the optimality equations (1a)-(1d) and (2). For $n \in \Theta$, we define

$$Z(n) = \begin{cases} \min \{ \vartheta(0, n+1) + M, \vartheta(0, 0) + R \} & \text{if } n < N, \\ \vartheta(0, 0) + R & \text{if } n = N. \end{cases} \quad (9)$$

Note that $Z(n)$ represents the minimum expected cost of maintenance or replacement when the system has already been repaired n times. Therefore, using (9), we can rewrite the optimality equations (1a)-(1d) and (2) as follows:

$$\vartheta(s, n) = \min \left\{ \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta(s', n), Z(n) \right\} \text{ for } (s, n) \in \Phi \times \Theta, \quad (10a)$$

$$\vartheta(S, n) = Z(n) \text{ for } n \in \Theta. \quad (10b)$$

Lemma 2 : For $(s, n) \in \Upsilon$ and $k \geq 0$, define the function $\vartheta^k(s, n)$ as follows:

$$\vartheta^0(s, n) = Z(n) \quad \text{for all } (s, n) \in \Upsilon, \quad (11a)$$

$$\vartheta^{k+1}(s, n) = \min \left\{ \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^k(s', n), Z(n) \right\}$$

$$\text{for } (s, n) \in \Phi \times \Theta \text{ and } k \geq 0, \quad (11b)$$

$$\vartheta^{k+1}(S, n) = Z(n) \quad \text{for } n \in \Theta \text{ and } k \geq 0. \quad (11c)$$

Then,

(i) $\vartheta^k(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$ and $k \geq 0$,

(ii) $\lim_{k \rightarrow \infty} \vartheta^k(s, n) = \vartheta(s, n)$ for all $(s, n) \in \Upsilon$.

Proof : (i) By (11a), $\vartheta^0(s, n)$ is nondecreasing in $s \in \Psi$ for all $n \in \Theta$. Then, by induction on $k \geq 0$, the rest of the proof is similar to that of Lemma 1 part (i) and is omitted.

(ii) First, by induction on $k \geq 0$, we will show that $\vartheta^k(s, n)$ is nonincreasing in $k \geq 0$ for all $(s, n) \in \Upsilon$. By definition, $\vartheta^1(s, n) \leq \vartheta^0(s, n)$ for all $(s, n) \in \Upsilon$. As the induction hypothesis, for some $m \geq 1$, suppose $\vartheta^m(s, n) \leq \vartheta^{m-1}(s, n)$ for all $(s, n) \in \Upsilon$. This implies

$$\ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^m(s', n) \leq \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^{m-1}(s', n) \quad \text{for } (s, n) \in \Phi \times \Theta. \quad (12)$$

By the definitions of the functions ϑ^m and ϑ^{m+1} , (12) implies that $\vartheta^{m+1}(s, n) \leq \vartheta^m(s, n)$ for all $(s, n) \in \Upsilon$, and the induction step is complete. Next, we will establish the following:

$$\max_{(s'', n') \in \Upsilon} [\vartheta^k(s'', n') - \vartheta^{k+1}(s'', n')] \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{k-1}(s'', n') - \vartheta^k(s'', n')] \quad \text{for all } k \geq 1. \quad (13)$$

For an arbitrary $m \geq 1$, consider the following possible cases for $\vartheta^m(s, n) - \vartheta^{m+1}(s, n)$. Note that since $\beta > 0$ and $\vartheta^{m-1}(s, n) \geq \vartheta^m(s, n)$ for all $(s, n) \in \Upsilon$,

$$0 \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')]. \quad (14)$$

1. If $s = S$, then by (14) and the fact that $\vartheta^{m+1}(s, n) = \vartheta^m(s, n) = Z(n)$,

$$\vartheta^m(s, n) - \vartheta^{m+1}(s, n) = 0 \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')].$$

2. If $s < S$, then there are two subcases:

(a) If $\vartheta^{m+1}(s, n) = Z(n)$, then because $\vartheta^m(s, n) \leq Z(n)$, by (14) we obtain

$$\vartheta^m(s, n) - \vartheta^{m+1}(s, n) \leq 0 \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')]. \quad (15)$$

(b) If $\vartheta^{m+1}(s, n) = \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^m(s', n)$, then

$$\begin{aligned} \vartheta^m(s, n) - \vartheta^{m+1}(s, n) &\leq \beta \sum_{s' \in \Psi} P(s'|s, n) [\vartheta^{m-1}(s', n) - \vartheta^m(s', n)] \\ &\leq \beta \sum_{s' \in \Psi} P(s'|s, n) \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')] \\ &= \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')], \end{aligned} \quad (16)$$

where the inequality in (16) is implied by $\vartheta^m(s, n) \leq \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^{m-1}(s', n)$.

Thus, $\vartheta^m(s, n) - \vartheta^{m+1}(s, n) \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')]$ for all $(s, n) \in \Upsilon$, implying

$$\max_{(s'', n') \in \Upsilon} [\vartheta^m(s'', n') - \vartheta^{m+1}(s'', n')] \leq \beta \max_{(s'', n') \in \Upsilon} [\vartheta^{m-1}(s'', n') - \vartheta^m(s'', n')]. \quad (17)$$

Recall that we have chosen m arbitrarily. Therefore, by (17), (13) holds. Then, since $\vartheta^k(s, n)$ is nonincreasing in $k \geq 0$ for all $(s, n) \in \Upsilon$, and the optimal cost function ϑ satisfies (10a) and (10b), by the fact that $\beta < 1$, Theorem 6.2.3 of Puterman [21] establishes the result. \square

Note that Lemma 2 part (i) states that the function ϑ^k defined by (11a)-(11c) preserves the monotonicity with respect to the level of deterioration for any particular repair history, and independent from part (i), Lemma 2 part (ii) establishes the convergence of the sequence to the optimal cost function, ϑ . Equipped with Lemma 2, we now proceed to the proof of Theorem 2.

Proof of Theorem 2: First, we will prove the following auxiliary result:

$$Z(n+1) - Z(n) \leq R - M \text{ for all } n \in \Theta'. \quad (18)$$

For the proof of (18), we present only the case for $n = N - 1$ and omit the others since they are similar. Consider the following two possible cases for $Z(N) - Z(N - 1)$. Recall that $Z(N) = \vartheta(0, 0) + R$ and $R \geq M$.

1. If $Z(N - 1) = \vartheta(0, 0) + R$, then $Z(N) - Z(N - 1) = 0 \leq R - M$.
2. If $Z(N - 1) = \vartheta(0, N) + M$, then (by Proposition 2) since $\vartheta(0, 0) \leq \vartheta(0, N)$,

$$Z(N) - Z(N - 1) = \vartheta(0, 0) + R - \vartheta(0, N) - M \leq R - M.$$

Next, we will make use of the relation in (18) to show that

$$\vartheta(s, n+1) - Z(n+1) \geq \vartheta(s, n) - Z(n) \text{ for all } (s, n) \in \Psi \times \Theta'. \quad (19)$$

For $(s, n) \in \Upsilon$, consider the sequence of value iterates $\{\vartheta^k(s, n)\}_{k \geq 0}$ defined by (11a)-(11c). Because $\lim_{k \rightarrow \infty} \vartheta^k(s, n) = \vartheta(s, n)$ for all $(s, n) \in \Upsilon$ (by Lemma 2 part (ii)), to establish (19), it suffices to show that ϑ^k satisfies the following for all $k \geq 0$:

$$\vartheta^k(s, n+1) - Z(n+1) \geq \vartheta^k(s, n) - Z(n) \text{ for all } (s, n) \in \Psi \times \Theta'. \quad (20)$$

By (11a), (20) holds for $k = 0$. As the induction hypothesis, suppose (20) holds for some $k = m \geq 0$. Then, for fixed $s \in \Psi$, choose an arbitrary $n \in \Theta'$, consider the following possible cases for $\vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n)$:

1. If $s = S$, then because $\vartheta^{m+1}(s, n) = Z(n)$ and $\vartheta^{m+1}(s, n+1) = Z(n+1)$,

$$\vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) = Z(n+1) - Z(n).$$

2. If $s < S$, then there are two subcases:

- (a) If $\vartheta^{m+1}(s, n+1) = Z(n+1)$, then since $\vartheta^{m+1}(s, n) \leq Z(n)$,

$$\vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) \geq Z(n+1) - Z(n).$$

- (b) If $\vartheta^{m+1}(s, n+1) = \ell(s, n+1) + \beta \sum_{s' \in \Psi} P(s'|s, n+1) \vartheta^m(s', n+1)$, then since $\vartheta^{m+1}(s, n) \leq \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^m(s', n)$, we have

$$\begin{aligned} & \vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) \\ & \geq \ell(s, n+1) - \ell(s, n) + \beta \sum_{s' \in \Psi} [P(s'|s, n+1) \vartheta^m(s', n+1) - P(s'|s, n) \vartheta^m(s', n)]. \end{aligned} \quad (21)$$

Note that by Lemma 2 part (i), $\vartheta^m(s', n+1)$ is nondecreasing in $s' \in \Psi$. Then, by Assumption 3, Lemma 4.7.2 of Puterman [21] implies

$$\sum_{s' \in \Psi} P(s'|s, n+1) \vartheta^m(s', n+1) \geq \sum_{s' \in \Psi} P(s'|s, n) \vartheta^m(s', n+1). \quad (22)$$

By (21) and (22),

$$\begin{aligned} & \vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) \\ & \geq \ell(s, n+1) - \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) [\vartheta^m(s', n+1) - \vartheta^m(s', n)]. \end{aligned} \quad (23)$$

By the induction hypothesis, $\vartheta^m(s', n+1) - \vartheta^m(s', n) \geq Z(n+1) - Z(n)$ for all $s' \in \Psi$.

By (23), this implies

$$\begin{aligned}
& \vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) \\
& \geq \ell(s, n+1) - \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) [Z(n+1) - Z(n)] \\
& = \ell(s, n+1) - \ell(s, n) + \beta [Z(n+1) - Z(n)] \\
& = \rho\beta [P(S|s, n+1) - P(S|s, n)] + \beta [Z(n+1) - Z(n)]. \tag{24}
\end{aligned}$$

Note that $Z(n+1) - Z(n) \leq R - M$. By satisfying condition (8), this implies

$$\rho\beta [P(S|s, n+1) - P(S|s, n)] \geq (1 - \beta) [Z(n+1) - Z(n)]. \tag{25}$$

Then, by (24), (25) yields $\vartheta^{m+1}(s, n+1) - \vartheta^{m+1}(s, n) \geq Z(n+1) - Z(n)$.

Thus, (20) holds for $k = m + 1$, and the induction step is complete.

Now, to prove the main result, by the definition of s_n^* , it is sufficient to show that for any $s \in \Psi$ and $n \in \Theta'$, $a(s, n) \in \{1, 2\}$ implies $a(s, n+1) \in \{1, 2\}$. For fixed $s \in \Psi$, choose an arbitrary $n \in \Theta'$ and let $a(s, n) \in \{1, 2\}$ implying that $\vartheta(s, n) = Z(n)$. Then, because (19) holds, we must have $\vartheta(s, n+1) \geq Z(n+1)$. By the definition of $\vartheta(s, n+1)$, we also have $\vartheta(s, n+1) \leq Z(n+1)$. These yield $\vartheta(s, n+1) = Z(n+1)$. Therefore, $a(s, n+1) \in \{1, 2\}$. \square

It is clear that the structure of the optimal policy is heavily influenced by the underlying cost structure, i.e., the relative magnitudes of the repair costs, replacement costs, and inspection costs. In the following section, we consider the sensitivity of the optimal policy to these cost parameters.

4 Sensitivity Analysis

Here, we examine the behavior of the optimal actions in each state of the process when we compare the inspection, maintenance and replacement costs of two different systems. Note that our analysis in this section does not impose any particular assumptions on the DTMC describing the system's deterioration. Throughout the section, we consider two systems differing in one or two of the aforementioned cost parameters and let $\mathbf{\Pi}_i$ denote problem instance i , ($i = 1, 2$). We also specify the cost parameters and the cost functions resulting for system i by a subscript i , ($i = 1, 2$). We begin our analyses by a technical lemma. Given the cost of inspection increases by a certain amount, Lemma 3 provides an upper bound for the increase in the optimal cost function in terms of the increase in the cost of inspection and the discount factor.

Lemma 3 : If Π_1 and Π_2 are identical except that $I_1 \geq I_2$, then

$$\vartheta_1(s, n) - \vartheta_2(s, n) \leq \beta[I_1 - I_2]/(1 - \beta) \text{ for all } (s, n) \in \Upsilon.$$

Proof : We will prove the result by induction on the iterates of the value iteration algorithm described by (4a)-(4e) and (5). We will apply the algorithm simultaneously to both problem instances. For problem instance $i = 1, 2$, let $\vartheta_i^k(s, n)$ be the value associated with state $(s, n) \in \Upsilon$ at the k^{th} iteration of the algorithm. By the convergence of the value iteration algorithm, it is sufficient to show that the following holds for all $k \geq 0$.

$$\vartheta_1^k(s, n) - \vartheta_2^k(s, n) \leq \beta[I_1 - I_2]/(1 - \beta) \text{ for all } (s, n) \in \Upsilon. \quad (26)$$

By construction, $\vartheta_1^k(s, n) - \vartheta_2^k(s, n) = 0$ for all $(s, n) \in \Upsilon$. Then, because $I_1 \geq I_2$ and $\beta < 1$, (26) holds for $k = 0$. As the induction hypothesis suppose (26) holds for some $k = m \geq 0$. Then, the following hold for all $(s, n) \in \Phi \times \Theta'$:

$$\begin{aligned} w_1^{m+1}(s, n) - w_2^{m+1}(s, n) &= \ell_1(s, n) - \ell_2(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) [\vartheta_1^m(s', n) - \vartheta_2^m(s', n)] \\ &\leq \beta[I_1 - I_2] + \beta \sum_{s' \in \Psi} P(s'|s, n) \beta \left(\frac{I_1 - I_2}{1 - \beta} \right) = \beta[I_1 - I_2] + \beta^2 \left(\frac{I_1 - I_2}{1 - \beta} \right) = \beta \left(\frac{I_1 - I_2}{1 - \beta} \right). \end{aligned} \quad (27)$$

Note that by the definition of the function ℓ_i for $i = 1, 2$, we have $\ell_1(s, n) - \ell_2(s, n) = \beta[I_1 - I_2]$ for all $(s, n) \in \Upsilon$. By the induction hypothesis this implies the inequality in (27).

Thus, $w_1^{m+1}(s, n) - w_2^{m+1}(s, n) \leq \beta[I_1 - I_2]/(1 - \beta)$ for all $(s, n) \in \Upsilon$. Then, by the induction hypothesis and the definitions of the functions ϑ_1^{m+1} and ϑ_2^{m+1} , this implies that $\vartheta_1^{m+1}(s, n) - \vartheta_2^{m+1}(s, n) \leq \beta[I_1 - I_2]/(1 - \beta)$ for all $(s, n) \in \Upsilon$, and the result follows by induction. \square

Proposition 3 : Suppose Π_1 and Π_2 are identical except that $I_1 \geq I_2$. Then, for any $(s, n) \in \Upsilon$, $a_2(s, n) \in \{1, 2\} \Rightarrow a_1(s, n) \in \{1, 2\}$.

Proof : It is sufficient to show that

$$\vartheta_1(s, n) - Z_1(n) \geq \vartheta_2(s, n) - Z_2(n) \text{ for all } (s, n) \in \Upsilon. \quad (28)$$

For each problem instance $i = 1, 2$, and each pair $(s, n) \in \Upsilon$, consider the sequence of value iterates $\{\vartheta_i^k(s, n)\}_{k \geq 0}$ defined by (11a)-(11c). Because $\lim_{k \rightarrow \infty} \vartheta_i^k(s, n) = \vartheta_i(s, n)$ for all $(s, n) \in \Upsilon$ and $i = 1, 2$

(by Lemma 2 part (ii)), to establish (28), it suffices to show that the following relation holds for all $k \geq 0$.

$$\vartheta_1^k(s, n) - \vartheta_2^k(s, n) \geq Z_1(n) - Z_2(n) \text{ for all } (s, n) \in \Upsilon. \quad (29)$$

By (11a), $\vartheta_i^0(s, n) = Z_i(n)$ for all $(s, n) \in \Upsilon$ and $i = 1, 2$. Therefore, (29) holds for $k = 0$. As the induction hypothesis, suppose (29) holds for some $k = m \geq 0$. Then, choose an arbitrary pair $(s, n) \in \Upsilon$ and consider the following possible cases for $\vartheta_1^{m+1}(s, n) - \vartheta_2^{m+1}(s, n)$. By the definitions of the functions Z_1 and Z_2 , Lemma 3 implies that $Z_1(n) - Z_2(n) \leq \beta[I_1 - I_2]/(1 - \beta)$.

1. If $s = S$, then since $\vartheta_i^{m+1}(s, n) = Z_i(n)$ for $i = 1, 2$, $\vartheta_1^{m+1}(s, n) - \vartheta_2^{m+1}(s, n) = Z_1(n) - Z_2(n)$.
2. If $s < S$, then there are two subcases.

- (a) If $\vartheta_1^{m+1}(s, n) = Z_1(n)$, then since $\vartheta_2^{m+1}(s, n) \leq Z_2(n)$, we have

$$\vartheta_1^{m+1}(s, n) - \vartheta_2^{m+1}(s, n) \geq Z_1(n) - Z_2(n).$$

- (b) If $\vartheta_1^{m+1}(s, n) = \ell_1(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta_1^m(s', n)$, then

$$\begin{aligned} & \vartheta_1^{m+1}(s, n) - \vartheta_2^{m+1}(s, n) \\ & \geq \ell_1(s, n) - \ell_2(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) [\vartheta_1^m(s', n) - \vartheta_2^m(s', n)] \end{aligned} \quad (30a)$$

$$\begin{aligned} & = \beta[I_1 - I_2] + \beta \sum_{s' \in \Psi} P(s'|s, n) [Z_1(n) - Z_2(n)] = \beta[I_1 - I_2] + \beta[Z_1(n) - Z_2(n)] \\ & \geq Z_1(n) - Z_2(n), \end{aligned} \quad (30b)$$

where (30a) is implied by $\vartheta_2^{m+1}(s, n) \leq \ell_2(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta_2^m(s', n)$, and (30b) follows from the fact that $Z_1(n) - Z_2(n) \leq \beta[I_1 - I_2]/(1 - \beta)$.

Thus, (29) holds for $k = m + 1$ and the result follows by induction. \square

Similar to Lemma 2, Lemma 4 describes a special sequence of value iterates and establishes its convergence to the optimal cost function. Specifically, the sequence described in Lemma 4 differs from that of Lemma 2 in that it only keeps the cost associated with replacement action as constant at its optimal value and updates the costs associated with wait and repair actions. Note that this sequence is also not appropriate to compute the optimal cost function without the optimal cost associated with state $(0, 0)$, $\vartheta(0, 0)$. We use Lemma 4 to construct the proofs of some of our sensitivity results.

Lemma 4 : For $(s, n) \in \Upsilon$ and $k \geq 0$, define the function $\vartheta^k(s, n)$ as follows:

$$\vartheta^0(s, n) = \vartheta(0, 0) + R \quad \text{for all } (s, n) \in \Upsilon, \quad (31a)$$

$$\vartheta^{k+1}(s, n) = \min \left\{ \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^k(s', n), \vartheta^k(0, n+1) + M, \vartheta(0, 0) + R \right\} \\ \text{for } (s, n) \in \Phi \times \Theta' \text{ and } k \geq 0, \quad (31b)$$

$$\vartheta^{k+1}(s, N) = \min \left\{ \ell(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta^k(s', n), \vartheta(0, 0) + R \right\} \\ \text{for } s \in \Phi \text{ and } k \geq 0, \quad (31c)$$

$$\vartheta^{k+1}(S, n) = \min \left\{ \vartheta^k(0, n+1) + M, \vartheta(0, 0) + R \right\} \text{ for } n \in \Theta' \text{ and } k \geq 0, \quad (31d)$$

$$\vartheta^{k+1}(S, N) = \vartheta(0, 0) + R \quad \text{for } k \geq 0. \quad (31e)$$

Then, $\lim_{k \rightarrow \infty} \vartheta^k(s, n) = \vartheta(s, n)$ for all $(s, n) \in \Upsilon$.

Proof : By (31b)-(31e), $\vartheta^1(s, n) \leq \vartheta(0, 0) + R$ for all $(s, n) \in \Upsilon$. By (31a), this implies that $\vartheta^1(s, n) \leq \vartheta^0(s, n)$ for all $(s, n) \in \Upsilon$. Then, by induction on $k \geq 0$, it can be shown that $\vartheta^k(s, n)$ is nonincreasing in $k \geq 0$ for all $(s, n) \in \Upsilon$. The remainder of the proof is similar to that of Lemma 2 part (ii) and is therefore omitted. \square

Proposition 4 : Suppose $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ are identical except that $I_1 \geq I_2$. Then, for any $(s, n) \in \Upsilon$, $a_2(s, n) = 2 \Rightarrow a_1(s, n) = 2$.

Proof : We sketch the proof and omit the details. For each problem instance $i = 1, 2$ and pair $(s, n) \in \Upsilon$, consider the sequence of value iterates $\{\vartheta_i^k(s, n)\}_{k \geq 0}$ defined by (31a)-(31e). Because $\vartheta_1(0, 0) - \vartheta_2(0, 0) \leq \beta[I_1 - I_2]/(1 - \beta)$ (by Lemma 3), similar to the proof of Proposition 3, by induction on $k \geq 0$ it can be shown that ϑ_1^k and ϑ_2^k satisfy the following relation at each iteration $k \geq 0$.

$$\vartheta_1^k(s, n) - \vartheta_2^k(s, n) \geq \vartheta_1(0, 0) - \vartheta_2(0, 0) \text{ for all } (s, n) \in \Upsilon. \quad (32)$$

Since $\lim_{k \rightarrow \infty} \vartheta_i^k(s, n) = \vartheta_i(s, n)$ for all $(s, n) \in \Upsilon$ and $i = 1, 2$ (by Lemma 4), satisfying (32) for all $k \geq 0$ implies $\vartheta_1(s, n) - \vartheta_2(s, n) \geq \vartheta_1(0, 0) - \vartheta_2(0, 0)$ for all $(s, n) \in \Upsilon$, and the result follows. \square

Next, we examine the sensitivity of the optimal policy to maintenance and replacement costs. First, we analyze the behavior of the optimal cost function with respect to a change in maintenance and/or replacement costs. Intuitively, Lemma 5 states that the optimal cost function does not decrease as maintenance and/or replacement costs increase.

Lemma 5 : *Suppose Π_1 and Π_2 are identical except that $R_1 \leq R_2$ and/or $M_1 \leq M_2$. Then, $\vartheta_1(s, n) \leq \vartheta_2(s, n)$ for all $(s, n) \in \Upsilon$.*

Proof : We sketch the proof as follows: Suppose we solve problem instances 1 and 2 simultaneously by the value iteration algorithm described by (4a)–(4e) and (5). For each problem instance $i = 1, 2$, if we let $\vartheta_i^k(s, n)$ be the value associated with state $(s, n) \in \Upsilon$ at the k^{th} iteration of the algorithm, then by induction on $k \geq 0$, it can be shown that $\vartheta_1^k(s, n) \leq \vartheta_2^k(s, n)$ for all $(s, n) \in \Upsilon$ and $k \geq 0$, and the result follows from the convergence of the algorithm. \square

Next, in Proposition 5, we use Lemmas 4 and 5 to explore the sensitivity of the optimality of repair/replacement decisions with respect to change(s) in the cost of maintenance and/or replacement. Intuitively, Proposition 5 states that, if the cost of maintenance and/or replacement decreases, then the decision maker becomes less likely to wait in all states.

Proposition 5 : *Suppose Π_1 and Π_2 are identical except that $R_1 \leq R_2$ and/or $M_1 \leq M_2$. Then, for any $(s, n) \in \Upsilon$, $a_2(s, n) \in \{1, 2\} \Rightarrow a_1(s, n) \in \{1, 2\}$.*

Proof : It is sufficient to show that

$$\vartheta_1(s, n) - Z_1(n) \geq \vartheta_2(s, n) - Z_2(n) \text{ for all } (s, n) \in \Upsilon. \quad (33)$$

For each problem instance $i = 1, 2$ and pair $(s, n) \in \Upsilon$, consider the sequence of value iterates $\{\vartheta_i^k(s, n)\}_{k \geq 0}$ defined by (11a)–(11c). By Lemma 2 part (ii), to establish (33) it suffices to show that ϑ_1^k and ϑ_2^k satisfy the following for all $k \geq 0$:

$$\vartheta_1^k(s, n) - Z_1(n) \geq \vartheta_2^k(s, n) - Z_2(n) \text{ for all } (s, n) \in \Upsilon. \quad (34)$$

By construction, for each problem instance $i = 1, 2$, we have $\vartheta_i^0(s, n) = Z_i(n)$ for all $(s, n) \in \Upsilon$. Therefore, (34) holds for $k = 0$. As the induction hypothesis, suppose (34) holds for some $k = m \geq 0$. Then, for a fixed pair $(s, n) \in \Upsilon$, consider the following possible cases for $\vartheta_2^{m+1}(s, n) - \vartheta_1^{m+1}(s, n)$. Note that because the problem instances have identical inspection and failure penalty costs, by the definitions of the functions ℓ_1 and ℓ_2 , we have $\ell_1(s, n) = \ell_2(s, n)$. Also, by the definitions of the functions Z_1 and Z_2 , and Lemma 5, we have $Z_2(n) \geq Z_1(n)$.

1. If $s = S$, then because $\vartheta_i^{m+1}(s, n) = Z_i(n)$ for $i = 1, 2$, we have $\vartheta_2^{m+1}(s, n) - \vartheta_1^{m+1}(s, n) = Z_2(n) - Z_1(n)$.
2. If $s < S$, then consider the following subcases:

(a) If $\vartheta_1^{m+1}(s, n) = Z_1(n)$, then because $\vartheta_2^{m+1}(s, n) \leq Z_2(n)$, we have

$$\vartheta_2^{m+1}(s, n) - \vartheta_1^{m+1}(s, n) \leq Z_2(n) - Z_1(n).$$

(b) If $\vartheta_1^{m+1}(s, n) = \ell_1(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta_1^m(s', n)$, then

$$\vartheta_2^{m+1}(s, n) - \vartheta_1^{m+1}(s, n) \leq \beta \sum_{s' \in \Psi} P(s'|s, n) [\vartheta_2^m(s', n) - \vartheta_1^m(s', n)] \quad (35a)$$

$$\leq \beta \sum_{s' \in \Psi} P(s'|s, n) [Z_2(n) - Z_1(n)] = \beta [Z_2(n) - Z_1(n)] \leq Z_2(n) - Z_1(n), \quad (35b)$$

where (35a) is implied by $\vartheta_2^{m+1}(s, n) \leq \ell_2(s, n) + \beta \sum_{s' \in \Psi} P(s'|s, n) \vartheta_2^m(s', n)$. Then, in (35b), the first inequality follows from the induction hypothesis and because $\beta < 1$, the second inequality is implied by the fact that $Z_2(n) \geq Z_1(n)$.

Thus, (34) holds for $k = m + 1$ and the result follows by induction. \square

Next, Proposition 6 explores the sensitivity of the optimality of a replacement action with respect to a change only in the cost of replacement, R . Intuitively, it states that, if the cost of replacement increases then the decision maker becomes less likely to replace the system in any particular state $(s, n) \in \Upsilon$.

Proposition 6 : *Suppose $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ are identical except that $R_1 \leq R_2$. Then, for any $(s, n) \in \Upsilon$, $a_2(s, n) = 2 \Rightarrow a_1(s, n) = 2$.*

Proof : For the sake of brevity, we again provide only the sketch of the proof. For each problem instance i ($i = 1, 2$) and each pair $(s, n) \in \Upsilon$, consider the sequence of value iterates $\{\vartheta_i^k(s, n)\}_{k \geq 0}$ defined by (31a)-(31e). Because $\vartheta_1(0, 0) \leq \vartheta_2(0, 0)$ (by Lemma 5), similar to the proof of Proposition 3, by induction on $k \geq 0$ it can be shown that ϑ_1^k and ϑ_2^k satisfy the following relation at each iteration $k \geq 0$.

$$\vartheta_1^k(s, n) - \vartheta_2^k(s, n) \geq \vartheta_1(0, 0) + R_1 - \vartheta_2(0, 0) - R_2 \quad \text{for all } (s, n) \in \Upsilon. \quad (36)$$

Since $\lim_{k \rightarrow \infty} \vartheta_i^k(s, n) = \vartheta_i(s, n)$ for all $(s, n) \in \Upsilon$ and $i = 1, 2$ (by Lemma 4), satisfying (36) for all $k \geq 0$ implies $\vartheta_1(s, n) - \vartheta_2(s, n) \geq \vartheta_1(0, 0) + R_1 - \vartheta_2(0, 0) - R_2$ for all $(s, n) \in \Upsilon$, and the result follows. \square

In section 5, we elucidate the optimal policies specified in this section by way of four numerical examples. These depict the structure and the sensitivity of the optimal policies obtained.

5 Numerical Illustrations

In this section, we present a few numerical examples to illustrate the threshold structure and sensitivity of the optimal policy. For each of these numerical examples we assume that inspections are free, i.e., $I = 0$, and let $\Psi = \Theta = \{0, 1, \dots, 9\}$, $c(s) = 4(s + 1)$ for $s \in \Phi$ and $\rho = 2000$. Each example uses the same deterioration process whose transition probability matrix is next defined. Define the functions $g(s) = 0.05 + 0.005s$ for $s \in \Phi$, $h(n) = 1 + 0.05n$ for $n \in \Theta$, and a stochastic matrix $\mathbf{Q} = [Q(s'|s)]_{s,s' \in \Phi}$ as follows:

$$Q(s'|s) = \begin{cases} 0.99 & \text{for } s = s', \\ 0.01 & \text{for } s = 0 \text{ and } s' = 1, \text{ or } s = 8 \text{ and } s' = 7, \\ 0.005 & \text{for } s \in \Phi \setminus \{0, 8\} \text{ and } s' \in \{s - 1, s + 1\}, \\ 0 & \text{otherwise.} \end{cases}$$

By the functional forms g and h , and the stochastic matrix \mathbf{Q} , for $s \in \Phi$, $s' \in \Psi$ and $n \in \Theta$, we define the transition probabilities $P(s'|s, n)$ as follows:

$$P(s'|s, n) = \begin{cases} g(s)h(n) & \text{for } s' = S \text{ and } n \in \Theta, \\ [1 - g(s)h(n)]Q(s'|s) & \text{for } s, s' \in \Phi \text{ and } n \in \Theta, \\ 0 & \text{otherwise.} \end{cases} \quad (37)$$

By (37) and the definitions of the functions g and h , it is clear that the failure probabilities, $P(S|s, n)$, increase proportionally with the system's cumulative number of completed repairs. Finally, for the first three examples we assume the system is inspected daily and consider an annual discount rate 0.99, which is translated into a daily discount factor of $\beta = 0.99^{1/365}$. For the figures that follow, the switching curves that separate the wait regions from repair and replace regions exclude the wait regions, and the vertical lines between repair and replace regions are included in the replace regions.

Example 1: In our first example, we illustrate the threshold structure of the optimal policy. For maintenance and replacement costs we assume $M = 800$ and $R = 5000$, respectively. Along with the other problem parameters, it can be easily verified that these parameter values satisfy Assumptions 1–3 and condition (8). Therefore, in accordance with Theorems 1 and 2, we expect to see an optimal policy that can be described as the union of three regions, one for each action, that are non-overlapping and separated by switching curves. Figure 1 depicts the resulting optimal policy.

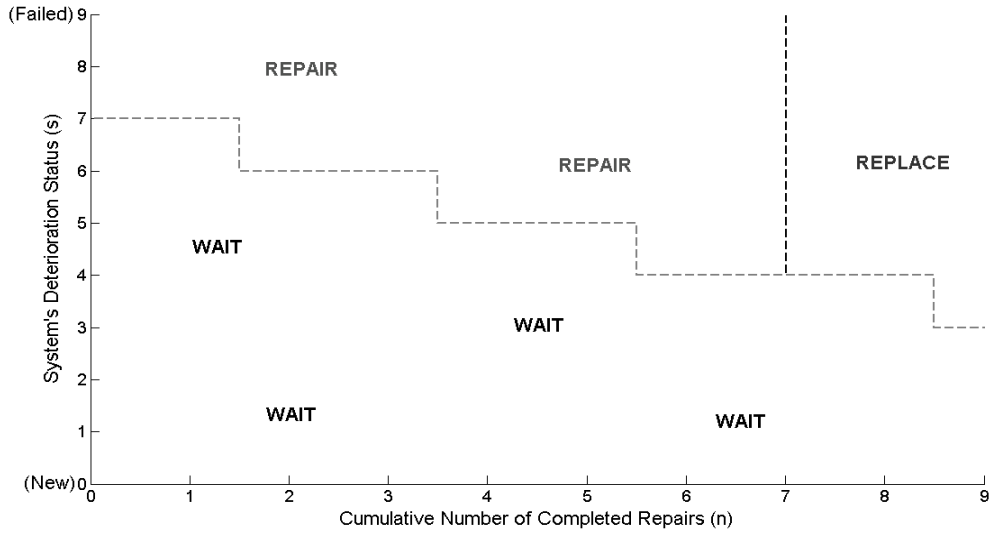


Figure 1: The optimal maintenance and replacement policy for Example 1.

Figure 1 shows that, for each specific cumulative number of repairs, there is a threshold deterioration status beyond which it is either optimal to repair or replace system. Furthermore, as the decision maker repairs the system more, he/she becomes more likely to replace or repair it again. For instance, for a deterioration status corresponding to state 5, waiting is optimal for the system only if it has been repaired three times or less. Interestingly, it is worth noting that the decision maker does not exhaust all repair options and uses, at most, seven of nine allowed repairs. Therefore, a portion of the optimal policy region is useless for a system that has been repaired seven times or less. In other words, for a system starting operation with a current history of no more than seven repairs, given the decision maker behaves optimally, because the system is either left in operation or replaced at seven repairs, the region beyond this threshold number of repairs is not expected to be in use. In Figure 2, the gray-colored region depicts this portion of the optimal policy.

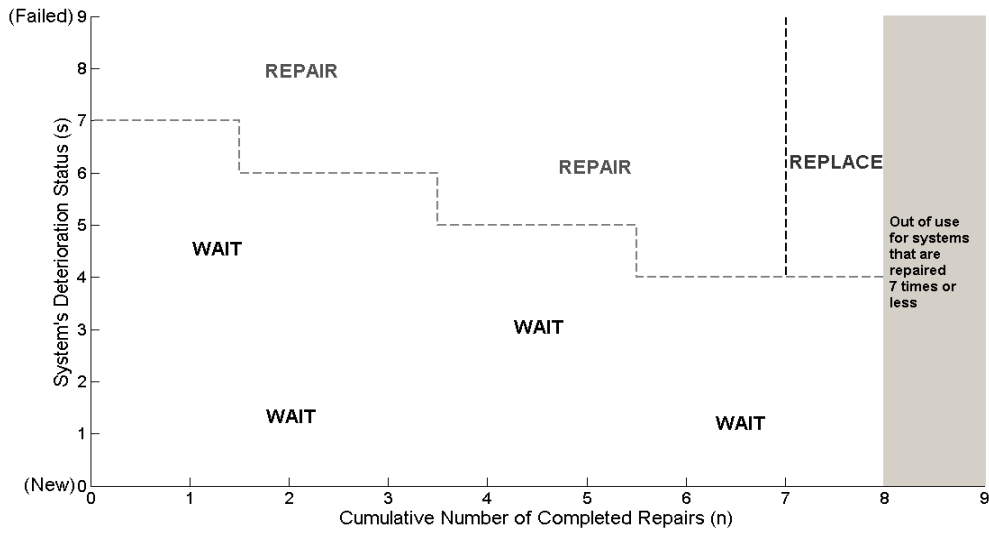


Figure 2: The optimal maintenance and replacement policy for Example 1.

Example 2: In our second example, we illustrate the sensitivity of the optimal policy to a change in the replacement cost. As in Example 1, we assume $M = 800$ but change the cost of replacement to $R = 3000$. Figure 3 depicts the corresponding optimal policy.

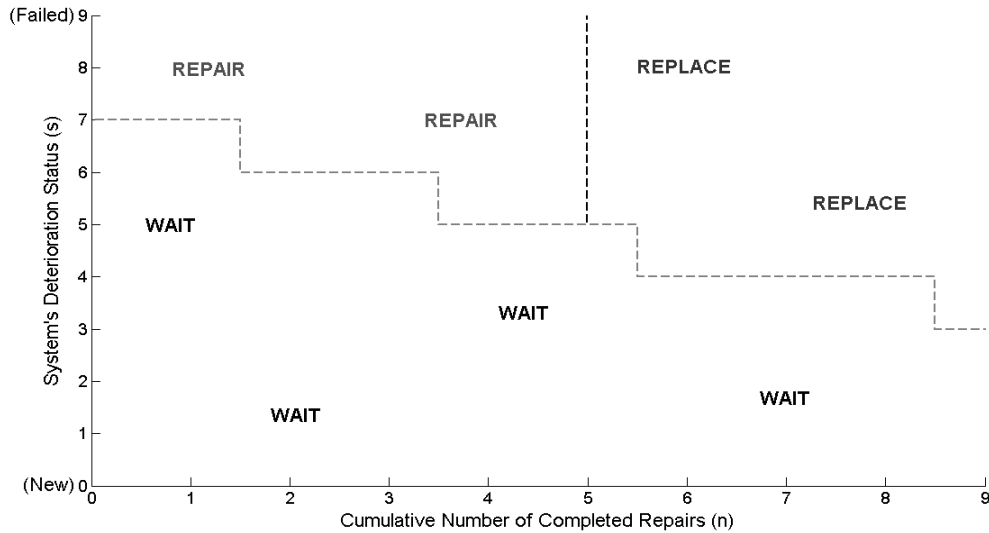


Figure 3: The optimal maintenance and replacement policy for Example 2.

As consistent with Proposition 6, here, the region in which it is optimal to replace the system is at

least as large as that of the optimal policy of Example 1.

Example 3: In our third example, we illustrate the sensitivity of the optimal policy to simultaneous changes in maintenance and replacement costs. We change the cost of maintenance and replacement to $M = 600$ and $R = 3000$, respectively. Figure 4 depicts the resulting optimal policy. As expected by Propositions 5 and 6, here, it is optimal to wait in a region which is at most as wide as that of the optimal policy of Example 1.

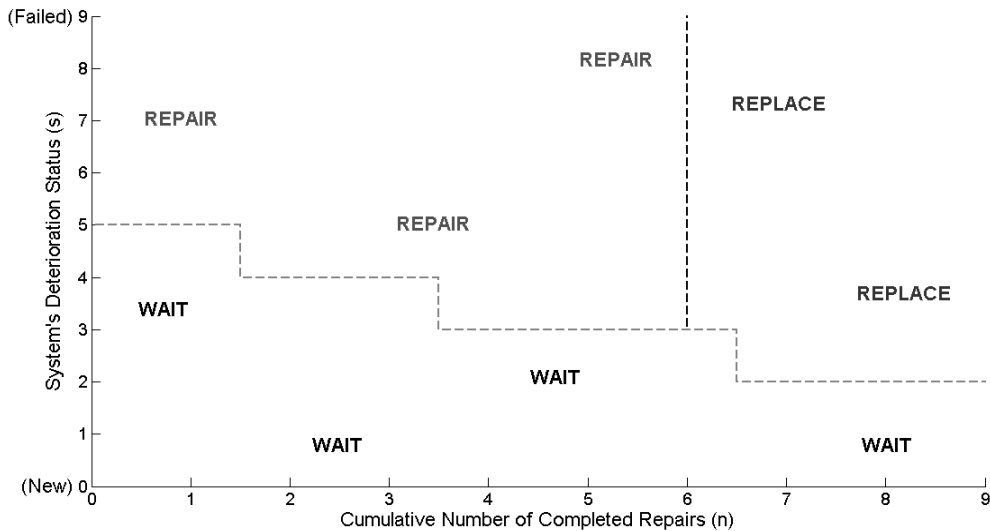


Figure 4: The optimal maintenance and replacement policy for Example 3.

Example 4: In our last example, we illustrate the fact that the switching curve separating the wait region from those of repair and replacement may not be monotone if condition (8) is not satisfied. For maintenance and replacement costs, we assume $M = 500$ and $R = 2000$, respectively. We also assume a periodic discount factor $\beta = 0.9$. It can be easily verified manually that these parameter values do not satisfy condition (8). Figure 5 depicts the resulting optimal policy.

From Figure 5, it is optimal to replace the system only when it fails and the repair action is no longer feasible. It is also clear that the repair region is convex but not monotone, which we can explain as follows: For any particular level of deterioration, higher rates of deterioration induce earlier repairs only up to a particular repair history after which the decision maker's willingness to pay for a further repair decreases.

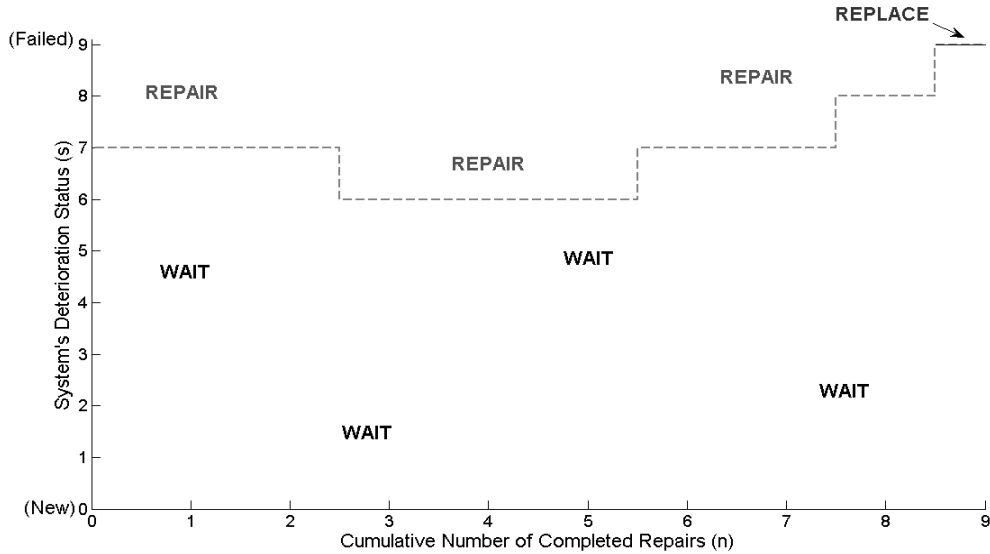


Figure 5: The optimal maintenance and replacement policy for Example 4.

Acknowledgements: The authors wish to express their gratitude to two anonymous referees and the Editor, Professor Jesús Artalejo, whose comments and suggestions led to the refinement of an earlier version of the model. Murat Kurt is especially thankful to Professor Andrew Schaefer for his support during this study.

References

- [1] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. John Wiley & Sons, New York, 1965.
- [2] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Massachusetts, 1996.
- [3] D. P. Bertsekas and J. N. Tsitsiklis. *Dynamic Programming and Optimal Control*. Athena Scientific, Massachusetts, 2001.
- [4] I. T. Castro. A model of imperfect preventive maintenance with dependent failure modes. *European Journal of Operational Research*, 196(1):217 – 224, 2009.
- [5] I. T. Castro and E. L. Sanjuan. An optimal repair policy for systems with a limited number of repairs. *European Journal of Operational Research*, 187(1):84 – 97, 2008.

- [6] C. Derman. On optimal replacement rules when changes of state are Markovian. In R. Bellman, editor, *Mathematical Optimization Techniques*, pages 201–210. University of California Press, 1963.
- [7] S. El-Ferik and M. Ben-Daya. Age-based hybrid model for imperfect preventive maintenance. *IIE Transactions*, 38(4):365 – 375, 2006.
- [8] S. El-Ferik and M. Ben-Daya. Model for imperfect age-based preventive maintenance with age reduction. *Journal of the Operational Research Society*, 59(12):1644 –1651, 2008.
- [9] A. Goyal, V. F. Nicola, A. N. Tantawi, and K. S. Trivedi. Reliability of systems with limited repairs. *IEEE Transactions on Reliability*, 36(2):202 – 207, 1987.
- [10] M. Klein. Inspection-maintenance-replacement schedules under Markovian deterioration. *Management Science*, 9(1):25–32, 1962.
- [11] M. Kolesar. Minimum cost replacement under Markovian deterioration. *Management Science*, 12(9):694–706, 1966.
- [12] D. Lin, M. J. Zuo, and R. C. M. Yam. General sequential imperfect preventive maintenance. *International Journal of Reliability, Quality and Safety Engineering*, 7(3):253 – 266, 2000.
- [13] D. Lin, M. J. Zuo, and R. C. M. Yam. Sequential imperfect preventive maintenance models with two categories of failure modes. *Naval Research Logistics*, 48(2):172 – 183, 2001.
- [14] D. Lugtigheid, X. Jiang, and A. K. S. Jardine. A finite horizon model for repairable systems with repair restrictions. *Journal of the Operational Research Society*, 59(10):1321–1331, 2008.
- [15] V. Makis and A. K. S. Jardine. Optimal preventive maintenance policies for repairable systems. *Operations Research*, 29(6):1181–1194, 1981.
- [16] V. Makis and A. K. S. Jardine. Optimal replacement policy for a general model with imperfect repair. *Journal of the Operational Research Society*, 43(2):111 – 120, 1992.
- [17] T. Nakagawa. Sequential imperfect preventive maintenance policies. *IEEE Transactions on Reliability*, 37(3):295–298, 1998.
- [18] H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

- [19] W. P. Pierskalla and J. A. Voelker. A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23(3):353–388, 1976.
- [20] W. B. Powell. *Approximate Dynamic Programming*. John Wiley and Sons, New Jersey, 2007.
- [21] M. L. Puterman. *Markov Decision Processes*. John Wiley and Sons, New York, 1994.
- [22] P. A. Scarf. On the application of mathematical models in maintenance. *European Journal of Operational Research*, 99(3):493–506, 1997.
- [23] Y. S. Sherif and M. L. Smith. Optimal maintenance models for systems subject to failure. *Naval Research Logistics Quarterly*, 28(1):47–74, 1981.
- [24] H. C. Tijms. *Stochastic Models: An Algorithmic Approach*. Wiley, Chichester, 1994.
- [25] C. Valdez-Flores and R. M. Feldman. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, 36(4):419 – 446, 1989.
- [26] H. Wang. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3):469–489, 2002.