Probability in the Engineering and Informational Sciences, **13**, 1999, 55–70. Printed in the U.S.A.

OPTIMALITY OF CONTROL LIMIT MAINTENANCE POLICIES UNDER NONSTATIONARY DETERIORATION

Zvi Benyamini and Uri Yechiali

Department of Statistics and Operations Research
School of Mathematical Sciences, Tel Aviv University
Tel Aviv 69978, Israel

Control limit type policies are widely discussed in the literature, particularly regarding the maintenance of deteriorating systems. Previous studies deal mainly with *stationary* deterioration processes, where costs and transition probabilities depend only on the state of the system, regardless of its cumulative age. In this paper, we consider a *nonstationary* deterioration process, in which operation and maintenance costs, as well as transition probabilities "deteriorate" with both the system's state and its cumulative age. We discuss conditions under which control limit policies are optimal for such processes and compare them with those used in the analysis of stationary models.

Two maintenance models are examined: in the first (as in the majority of classic studies), the only maintenance action allowed is the replacement of the system by a new one. In this case, we show that the nonstationary results are direct generalizations of their counterparts in stationary models. We propose an efficient algorithm for finding the optimal policy, utilizing its control limit form. In the second model we also allow for repairs to better states (without changing the age). In this case, the optimal policy is shown to have the form of a *3-way control limit rule*. However, conditions analogous to those used in the stationary problem do not suffice, so additional, more restrictive ones are suggested and discussed.

1. INTRODUCTION

A great deal of research has been devoted to the maintenance of stochastically deteriorating systems, in particular to the structural form of optimal maintenance policies. Of special interest are policies known as control limit policies (CLPs), in

 0269-9648/99 $12.50

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

which the system operates uninterrupted up to a certain degree of deterioration, and maintenance of some kind is performed whenever this limit is exceeded. Such policies are intuitive in nature and easily implemented in real-life systems. As a result, conditions ensuring the optimality of CLPs for various maintenance models are of great interest.

In the basic models introduced by Derman [1] and Kolesar [6], a system deteriorates according to a Markov process, where different states correspond to different levels of deterioration. In these models, the only possible action is the replacement of the system with a new one. Other models, such as Kijima, Morimura, and Suzuki [5] and Douer and Yechiali [3] also allow for partial repairs, which are cheaper (though less effective) than complete replacement. In all of these models, the optimal policy is shown to be a CLP under intuitively meaningful conditions, namely that costs and failure probabilities increase as the system deteriorates.

In many real-life systems, the deterioration process may change its characteristics (such as transition probabilities or cost functions) as the system ages. In such cases, a Markovian process does not accurately describe the system's deterioration. In order to take the aging process into account in a more general way, Kao [4] developed a semi-Markovian process. He discussed "stage-age" replacement policies, in which the optimal replacement time depends on both the state and the sojourn time in it. Other "state-age" models were developed by Lam and Yeh [7] and So [9], among others.

The use of the term "system age" with respect to the above semi-Markovian process is somewhat misleading: while the sojourn time is measured from the last transition or maintenance action, "true" age accumulates throughout the system's life cycle; while some of the outcomes of this cumulative aging may be overturned through maintenance, others may not (for example, when nonrepairable components exist).

In this paper, therefore, we discuss an extended notion of an aging process: we address the problem of a *nonstationary*, age-dependent deterioration process. In such a process, all parameters—transition probabilities, maintenance and operation costs—depend explicitly on both the system's state and its total cumulative age. We discuss both a "replacement-only" model and a "repair–replacement" model in which repair actions may involve changing the system's state to a better one, but cannot change its age. The system's age can be reduced (to zero) only through replacement by a new system. This property describes, for example, a system which contains a repairable part (whose condition is represented by "states") and a nonrepairable one (whose condition depends on its cumulative age).

The structure of this paper is as follows. In Section 2, we define the concept of a control limit policy in the nonstationary context. In Section 3, we discuss the replacement-only model and the optimality of CLPs under various optimization criteria. The results are shown to be direct generalizations of previous results for stationary models [1,3,6]. In Section 4, we propose an algorithm for finding the optimal policy, utilizing the fact that it has a control limit form. A detailed example, using the proposed algorithm, is provided. Section 5 offers an advanced, more efficient solution algorithm. In Section 6, we discuss the repair–replacement model. We show that, in this case, direct analogies to previous conditions do not ensure optimality of CLPs. More restrictive conditions are suggested and discussed.

2. AN AGE-DEPENDENT CLP

In the classic stationary model, control limit policies may be viewed as bisections of the one-dimensional system state space: The system is replaced if and only if its state i is greater than some control limit i^*. In the nonstationary model, a replacement policy defines the set of state-age pairs in which the system is replaced. Intuitively, one might therefore attempt to define a CLP in the nonstationary case as a bisection of the two-dimensional state-age space. However, certain two-dimensional bisections may not be as simple or intuitive as one would expect of a control limit policy (once defined). For example, a control limit policy would be expected to be monotone in the sense that the extent of maintenance grows as deterioration increases. Hence, the need for a somewhat finer definition emerges.

We therefore define an *age-dependent* CLP as a policy having the following structure: For any fixed age t, it has a control limit structure regarding the state; that is, the system is replaced if and only if the state is above $i^*(t)$. Also, for any fixed state i, it has a control limit form regarding the age; that is, the system is replaced when its age is greater than $t^*(i)$. Such a policy is indeed monotone (in the meaning described above) as long as one of the parameters (state or age) is fixed. We also define a *partial* CLP as a policy which is a CLP regarding state for any fixed age, but not necessarily vice versa. Partial CLPs will be used by the solution algorithm proposed in Section 4.

3. REPLACEMENT-ONLY MODEL

A system deteriorates according to a discrete time, nonstationary Markovian process. The degree of deterioration is denoted by one of a finite number of states $\{0,1,\ldots,N\}$, where state 0 denotes a fully operational system, state N denotes a failed system, and intermediate states denote increasing levels of deterioration. We denote by (i,t) a system at state i and age t.

The transition probability of moving from state i at age $t-1$ to state j at age t is $P_{ij}(t)$ (we assume that the probability of eventually reaching state N from any given state i is nonzero). Immediately after every transition, the system is inspected and a decision is made whether to continue its operation for an additional unit of time (without taking any action) at cost $R_i(t)$ or to replace the system with a new one at cost $B_i(t)$. If the system is replaced, it moves immediately to (0,0) and operates for one unit of time at cost $R_0(0)$. At state N (failure) replacement is mandatory. We also assume that the system is always replaced at its maximal life span $T^* \leq \infty$.

We impose the following conditions on the deterioration process:

Condition 3.1: $R_i(t)$, $B_i(t)$ are increasing functions, both in i and in t (throughout this paper, we consider functions as increasing or decreasing in the weak sense).

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

Condition 3.2: For every $t < T^*$ and $i < N$, $B_i(t) + R_0(0) \geq R_i(t)$.

Condition 3.2 implies that for a one-step horizon, replacement is always more expensive than doing nothing.

Condition 3.3: The function $\sum_{j=k}^{N} P_{ij}(t)$ is increasing in both i and t for any fixed $k = 0, \ldots, N$. It may be shown (see a similar result for the stationary case in [1]) that this condition is equivalent to the following: for any function $h(j,t)$ which is increasing in both j and t, the function $\sum_{j=0}^{N} P_{ij}(t)h(j,t)$ is also increasing in both i and t.

Condition 3.3 may be viewed as a generalized increasing failure rate (IFR) condition [3].

Condition 3.4: The function $B_i(t) - R_i(t)$ is decreasing in both i and t.

It should be noted that Conditions 3.1–3.4 are direct two-dimensional extensions of the conditions used in previous one-dimensional stationary models.

We now show that under the above conditions, and for the discounted cost criterion, for any horizon, the optimal policy is a CLP.

Denote by $\Phi_\alpha^T(i,t)$ the expected total discounted cost under the optimal policy, given that the system is at state i and age t, the discount factor is α and the horizon is $0 < T < \infty$ time units. The following optimality equations may easily be derived:

$$\begin{aligned}
\Phi_\alpha^T(0,0) &= R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) && \\
\Phi_\alpha^T(N,t) &= B_N(t) + R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) && 0 < t \leq T^* \\
\Phi_\alpha^T(i,T^*) &= B_i(T^*) + R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) && 0 \leq i \leq N \\
\Phi_\alpha^T(i,t) &= \min\left\{\begin{array}{l} R_i(t) + \alpha \sum_{j=0}^{N} P_{ij}(t+1)\Phi_\alpha^{T-1}(j,t+1), \\ B_i(t) + R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) \end{array}\right\} && \begin{array}{l} 0 < i < N, \\ 0 < t < T^* \end{array}
\end{aligned} \qquad \textbf{(1)}$$

The initial conditions for $T = 1$ are:

$$\begin{aligned}
\Phi_\alpha^1(i,t) &= R_i(t) && i < N;\ t < T^* \\
\Phi_\alpha^1(N,t) &= B_N(t) + R_0(0) && \text{for all } t \\
\Phi_\alpha^1(i,T^*) &= B_i(T^*) + R_0(0) && \text{for all } i.
\end{aligned} \qquad \textbf{(2)}$$

LEMMA 1: *For fixed α and T, $\Phi_\alpha^T(i,t)$ is an increasing function of both i and t.*

PROOF: The proof follows readily by induction on T, as every expression on the right-hand side of Eq. (1), Eq. (2) is increasing by Conditions 3.1–3.3. ■

THEOREM 1: *Under Conditions* 3.1–3.4, *and for the discounted cost criterion for any horizon* $T > 0$, *the optimal policy is a CLP.*

PROOF: For $T = 1$, the control limit form is immediately evident from Eq. (2). For $1 < T < \infty$, assuming replacement is the optimal action at some (i,t):

$$B_i(t) + R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) \le R_i(t) + \alpha \sum_{j=0}^{N} P_{ij}(t+1)\Phi_\alpha^{T-1}(j,t+1). \tag{3}$$

Rearranging terms results in

$$\begin{aligned} &B_i(t) + R_0(0) - R_i(t) \\ &\quad \le \alpha \left[\sum_{j=0}^{N} P_{ij}(t+1)\Phi_\alpha^{T-1}(j,t+1) - \sum_{j=0}^{N} P_{0j}(1)\Phi_\alpha^{T-1}(j,1) \right]. \end{aligned} \tag{4}$$

The left-hand side of inequality (4) is decreasing by Condition 3.4, while the right-hand side is increasing by Condition 3.3 and Lemma 1. Therefore, inequality (4) holds for any (m,s) with $m \ge i$ and $s \ge t$, thus proving that the optimal policy is a CLP.

The infinite-horizon case is dealt with in a similar manner, using the total cost function $\Phi_\alpha(i,t) \equiv \lim_{T\to\infty} \Phi_\alpha^T(i,t)$. ■

Denote by $\Phi^{(d)}$ the average cost per unit time for policy d. The following well-known lemma connects the value of $\Phi^{(d)}$ with that of the infinite-horizon discounted cost (its proof may be found in Derman [2]). We note that for $T^* = \infty$, additional regularity conditions are required for this result to hold (for example, bounded costs, or other more subtle conditions discussed in Puterman [8] and references therein).

LEMMA 2: *Denote by* $\Phi_\alpha^{(d)}(i,t)$ *the infinite-horizon expected discounted cost under policy* d. *Then*

$$\Phi^{(d)} = \lim_{\alpha\to 1}\{(1-\alpha)\Phi_\alpha^{(d)}(i,t)\}$$

for any i,t.

THEOREM 2: *The optimal policy for the average cost criterion is also a CLP. Furthermore, the average-cost optimal policy is the limit of the discounted-cost optimal policies as the discount factor approaches* 1.

PROOF: Let $\{\alpha_k\}_{k=1}^{\infty}$ be an increasing sequence of distinct discount factors, with $\lim_{k\to\infty}\alpha_k = 1$, having the property that the optimal policy under the infinite-horizon discounted cost criterion—for all discount factors α_k—is the same (such a sequence exists, as there are only finitely many possible replacement policies). Denote this common policy by d^*. For every policy d we therefore have, for every k,

$$\Phi_{\alpha_k}^{(d)}(i,t) \ge \Phi_{\alpha_k}^{(d^*)}(i,t) \tag{5}$$

for all i, t. Using Lemma 2 with inequality (5), we have

$$\Phi^{(d)} = \lim_{k\to\infty}(1-\alpha_k)\Phi_{\alpha_k}^{(d)}(i,t) \geq \lim_{k\to\infty}(1-\alpha_k)\Phi_{\alpha_k}^{(d^*)}(i,t) = \Phi^{(d^*)},$$

thus proving optimality of d^*, which, by Theorem 1, is a CLP. ■

We note the following economic interpretation of Inequalities (3) and (4): the left-hand side of Inequality (4) is the marginal one-step cost of replacement compared with doing nothing, while the right-hand side is the expected saving in future cost as a result of the replacement. These inequalities therefore state, as may intuitively be expected, that replacement at (i, t) is cost-effective if and only if the marginal one-step replacement cost is outweighed by the expected future savings resulting from the replacement. The optimality of CLPs naturally follows from the above intuitive explanation: Conditions 3.1 and 3.3 imply that the expected future savings increase in both i and t, while Condition 3.4 implies that the marginal one-step cost of replacement decreases in both i and t.

Another aspect of Condition 3.4, particularly of its age-related part, should also be noted: it limits the rather general cost deterioration structure characterized by Condition 3.1 and requires that the operation costs "deteriorate" at a higher rate than the replacement costs, for the optimality of CLPs to hold. Without Condition 3.4, the optimal policy is not necessarily a CLP, although such optimal policies are possible. As an example, assume the costs and transition probabilities used in Section 4, with $R_i(t)$ and $P_{ij}(t)$ held constant at their $t = 0$ values (so Condition 3.4 does not hold). It may be seen that as α changes, the optimal policy changes from a CLP to a non-CLP form. For example, for $\alpha = 0.6$, the optimal policy is not a CLP, but, for $\alpha = 0.65$ or $\alpha = 0.55$, it is. For $\alpha = 0.734$, the optimal policy is not a CLP, but it is a CLP for $\alpha = 0.733$ or $\alpha = 0.735$.

4. FINDING THE OPTIMAL POLICY

The nonstationary replacement-only problem may be considered a Markovian decision problem, with each pair (i, t) viewed as a separate Markovian "state," and solved using one of many existing algorithms for such problems. In practice, however, this method (if applied in a straightforward manner) would be quite inefficient, as such general algorithms pay little respect to the specific form of the problem (note the relatively large transition probability matrix, consisting mainly of zeros, as an example) or the solution.

We shall now propose an algorithm for finding the optimal replacement policy. The proposed algorithm is based on the well-known general policy iteration algorithm (cf. Tijms [10]), but it utilizes the known control limit form of the optimal policy. In this section, we describe a basic version of the proposed algorithm. This version is rather inefficient, but it is more clearly described and justified than the advanced, computationally efficient version described in the next section. We discuss only the infinite-horizon discounted cost criterion, noting that the average cost criterion may be handled in a similar manner.

If d is a CLP, denote by $i_d^*(t)$ the control limit for age t under d. That is, under d, a system at age t is replaced if and only if the state is at least $i_d^*(t)$. Clearly, $i_d^*(T^*) = 0$ for any policy d (as replacement is mandatory at T^*). For completeness of notation we also write $i_d^*(0) = N$ for any policy d. The expected total cost under policy d, when starting at (i,t), is denoted by $v_d(i,t)$. We refer to the policy attained after the nth iteration of the algorithm as d_n, denoting its control limit and cost functions by $i_n^*(t)$ and $v_n(i,t)$, respectively.

The following well-known theorem (often referred to as Howard's policy improvement theorem) is the theoretical basis for the policy iteration algorithm. Its proof can be found in [2] and [8]. It is written in terms of a general Markovian decision process (as previously noted, the nonstationary problem may be described in this manner).

THEOREM 3: *Let S be the state space of a (stationary) Markovian decision process. Denote by $d(s)$ the action taken under policy d at state s, by $F(s,a)$ the one-step cost incurred when action a is taken at state s, and by $P(s'|s,a)$ the transition probability of moving to state s' after action a is taken at state s.*

Let d be any policy. If, for some policy g,

$$F(s,g(s)) + \alpha \sum_{s' \in S} P(s'|s,g(s)) v_d(s') \le v_d(s) \tag{6}$$

for every $s \in S$, then $v_g(s) \le v_d(s)$ for all $s \in S$. Furthermore, if for every $s \in S$, $d(s)$ itself minimizes the left-hand side of inequality (6), *then d is optimal.*

We now describe the basic version of the proposed algorithm for the nonstationary replacement-only problem.

1. **Initialization:** The algorithm starts with an arbitrary initial policy d_0, which must be a partial CLP. The algorithm begins at age $t = 1$.
2. **Policy evaluation**: The evaluation of a policy (i.e., determination of the values of $v_n(i,t)$ for every i,t) is performed from T^* backward: at T^* replacement is mandatory, so for all i

$$v_n(i,T^*) = B_i(T^*) + v_n(0,0).$$

Note that the value of $v_n(0,0)$ is yet unknown, and we refer to it as a variable for the time being. Suppose now that $v_n(i,t+1)$ is known for $i = 0,\ldots,N$. $v_n(i,t)$ is evaluated by the following equations:

$$v_n(i,t) = \begin{cases} B_i(t) + v_n(0,0) & i \ge i_n^*(t) \\ R_i(t) + \alpha \sum_{j=0}^{N} P_{ij}(t+1) v_n(j,t+1) & i < i_n^*(t) \end{cases} \tag{7}$$

Note that all expressions on the right-hand side of Eq. (7) are known (in terms of $v_n(0,0)$). Continuing in this manner, we reach the following equation:

$$v_n(0,0) = R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1) v_n(j,1). \tag{8}$$

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

Equation (8) can be solved for $v_n(0,0)$, whose value may be substituted into $v_n(i,t)$ for complete evaluation of d_n. The above method is, in fact, equivalent to the method used in the general policy iteration algorithm, that is, the solution of a set of linear equations. However, it utilizes the special structure of our problem, in which only transitions to very few states are possible (those corresponding to age $t+1$).

3. **Policy improvement:** The algorithm attempts to shift the control limit $i_n^*(t)$ up or down by one state. Using test conditions (9) and (10) below, which are based on Theorem 3, the algorithm tests whether such a shift is improving, that is, whether the cost under the new policy is lower than under d_n (we refer to i^* rather than $i_n^*(t)$ hereafter for simplicity of notation).

$$R_{i^*}(t) + \alpha \sum_{j=0}^{N} P_{i^*,j}(t+1) v_n(j,t+1) \overset{?}{<} B_{i^*}(t) + v_n(0,0) \tag{9}$$

$$B_{i^*-1}(t) + v_n(0,0) \overset{?}{<} R_{i^*-1}(t) + \alpha \sum_{j=0}^{N} P_{i^*-1,j}(t+1) v_n(j,t+1). \tag{10}$$

If the shift is found to be improving, then the new policy d_{n+1} is evaluated (by step 2) and additional improvement in the same direction is attempted. Otherwise, the algorithm moves on to $t+1$, and continues in this manner cyclically. Note that d_{n+1}, like d_n, is a partial CLP.

4. **Stopping criterion:** The algorithm terminates when an entire cycle is completed without any improving actions. d_n is the optimal policy.

We must now prove that our stopping criterion is indeed a valid one. As opposed to the general algorithm, Theorem 3 cannot be used directly in order to establish optimality, as the policy improvement step in the above algorithm is more restrictive: only two actions are tested at each iteration, namely, shifting of $i_n^*(t)$ at some t up or down by one step, rather than the complete set of possible actions (as in the general algorithm). A priori, this could perhaps lead to some local minimum rather than to an optimal policy.

We first prove the following lemma (we suppress the subscript n hereafter).

LEMMA 3: *If, for policy d, no improvement is made throughout an entire cycle, then $v(i,t)$ is increasing in i for every t.*

PROOF: The proof is by backward induction on t. For $t = T^*$,

$$v(i,T^*) = B_i(T^*) + v(0,0),$$

which is clearly an increasing function of i, by Condition 3.1. Assume now that $t < T^*$ and suppose $i > j$. If $i,j \geq i^*(t)$ or $i,j < i^*(t)$, monotony is evident from Eq. (7) using the induction hypothesis.

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

It is left to show that $v(i^*(t),t) \geq v(i^*(t)-1,t)$. As shifting (particularly down) of $i^*(t)$ is not improving,

$$B_{i^*(t)-1}(t) + v(0,0) \geq R_{i^*(t)-1}(t) + \alpha \sum_{j=0}^{N} P_{i^*(t)-1,j}(t+1)v(j,t+1). \qquad \textbf{(11)}$$

Using the fact that $B_{i^*(t)-1}(t) \leq B_{i^*(t)}(t)$ with inequality (11), the proof is complete. ■

THEOREM 4: *If, for policy d, no improvement is made throughout an entire cycle, then d is optimal.*

PROOF: We will show that if no improvements can be made by the algorithm, then any deviation from *d*—even one that does not result in a partial CLP—is not improving, thus proving optimality by Theorem 3. Assume, therefore, that for all *t*,

$$B_{i^*(t)}(t) + v(0,0) < R_{i^*(t)}(t) + \alpha \sum_{j=0}^{N} P_{i^*(t),j}(t+1)v(j,t+1) \qquad \textbf{(12)}$$

$$B_{i^*(t)-1}(t) + v(0,0) > R_{i^*(t)-1}(t) + \alpha \sum_{j=0}^{N} P_{i^*(t)-1,j}(t+1)v(j,t+1). \qquad \textbf{(13)}$$

Assume now that there exist (k,t), such that $k < i^*(t)$, and

$$B_k(t) + v(0,0) < R_k(t) + \alpha \sum_{j=0}^{N} P_{kj}(t+1)v(j,t+1). \qquad \textbf{(14)}$$

Such (k,t) would imply that d is not optimal, by Theorem 3. Using Condition 3.4 with case (14), we have

$$B_{i^*(t)-1}(t) + v(0,0) < R_{i^*(t)-1}(t) + \alpha \sum_{j=0}^{N} P_{kj}(t+1)v(j,t+1).$$

By Condition 3.3 and Lemma 3, we have

$$B_{i^*(t)-1}(t) + v(0,0) < R_{i^*(t)-1}(t) + \alpha \sum_{j=0}^{N} P_{i^*(t)-1,j}(t+1)v(j,t+1),$$

thus contradicting case (12). Therefore, such (k,t) with $k < i^*(t)$ do not exist. In the same manner it is shown that the existence of (k,t) with $k \geq i^*(t)$, in which doing nothing is better than replacement, would contradict case (13), thus completing the proof. ■

We now present a detailed example of a problem solved using the proposed algorithm. We denote a policy d by a vector containing the values of $i_d^*(t)$. Let $N = 4$, $T^* = 4$, and $\alpha = 0.9$. Assume the following operation and replacement costs: $B_i(t) = 5 + 2i + 0.2it$, $R_i(t) = 1 + 2i + 0.5it$. The transition probabilities are:

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

$$P_{i,j<N}(t) = \begin{bmatrix} 0.1 & 0.7 & 0.1 & 0.05 \\ 0 & 0.8 & 0.1 & 0.05 \\ 0 & 0 & 0.5 & 0.25 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \cdot 0.9^t \qquad P_{i,N}(t) = 1 - \sum_{j=0}^{N-1} P_{ij}(t).$$

It is easily verified that Conditions 3.1–3.4 hold. We start with the initial policy, 'replace always':

$$d_0 = [\,4\;0\;0\;0\;0\,].$$

We first evaluate $v_0(i,t)$. It is easily seen that for $t \geq 1$,

$$v_0(i,t) = B_i(t) + v_n(0,0) = 5 + 2i + 0.2it + v_0(0,0).$$

Solving the following equation for $v_0(0,0)$,

$$v_0(0,0) = R_0(0) + \alpha \sum_{j=0}^{N} P_{0j}(1)v_0(j,1) = 1 + 0.9(8.35 + v_0(0,0)),$$

we have $v_0(0,0) = 85.2$. Substituting this into the previous calculations yields the following values for $v_0(i,t)$:

$$v_0 = \begin{bmatrix} 85.2 & 90.2 & 90.2 & 90.2 & 90.2 \\ — & 92.4 & 92.6 & 92.8 & 93.0 \\ — & 94.6 & 95.0 & 95.4 & 95.8 \\ — & 96.8 & 97.4 & 98.0 & 98.6 \\ — & 99.0 & 99.8 & 100.6 & 101.4 \end{bmatrix}.$$

We now test the benefit of shifting the control limit up at $t = 1$:

$$R_0(1) + \alpha \sum_{j=0}^{N} P_{0j}(2)v_0(j,2) = 86.0 < 90.2 = B_0(1) + v_0(0,0).$$

The control limit is therefore shifted up, and $v_1(i,t)$ evaluated. Yet another shift up is found to be successful, after which we have

$$d_2 = [\,4\;2\;0\;0\;0\,], v_2 = \begin{bmatrix} 70.4 & 72.6 & 75.4 & 75.4 & 75.4 \\ — & 75.3 & 77.8 & 78.0 & 78.2 \\ — & 79.8 & 80.2 & 80.6 & 81.0 \\ — & 82.0 & 82.6 & 83.2 & 83.8 \\ — & 84.2 & 85.0 & 85.8 & 86.6 \end{bmatrix}.$$

Testing the benefit of a third shift up, we find that

$$R_2(1) + \alpha \sum_{j=0}^{N} P_{2j}(2)v_2(j,2) = 80.3 > 79.8 = B_2(1) + v_2(0,0).$$

Therefore, the algorithm moves to $t = 2$ without changing the control limit. Again, the first two shifts are found to be improving, while the third is not, so that $d_4 = [\,4\;2\;2\;0\;0\,]$. Moving on to $t = 3$, only one shift up is improving:

$$d_5 = [\,4\;2\;2\;1\;0\,],\ v_5 = \begin{bmatrix} 68.3 & 70.2 & 71.6 & 72.5 & 73.3 \\ — & 72.9 & 74.8 & 75.9 & 76.1 \\ — & 77.7 & 78.1 & 78.5 & 78.9 \\ — & 79.9 & 80.5 & 81.1 & 81.7 \\ — & 82.1 & 82.9 & 83.7 & 84.5 \end{bmatrix}.$$

It can now be seen that there are no further improvements. d_5 is, therefore, the optimal policy. Optimality of d_5 may indeed be directly verified through satisfaction of the optimality equations.

5. IMPROVING COMPUTATIONAL EFFICIENCY

The algorithm described in Section 4 has one major drawback when compared with the general policy iteration algorithm: while the general algorithm evaluates the new policy d_{n+1} only after improving d_n (using Theorem 3) at all states, the above algorithm reevaluates the policy after every shift. As the evaluation step is time-consuming, this appears to cause a significant reduction in the efficiency of the proposed algorithm. However, using the fact that the new policy is only slightly different than the previous one, a more efficient policy evaluation step can be devised, thus greatly improving the algorithm.

Denote by q_n the value of the variable $v_n(0,0)$, and by Q_n the matrix $v_n(i,t)$, evaluated in terms of q_n. We begin with the following lemma, which demonstrates the effects of shifting $i^*(t)$ on Q. Proof is rather straightforward, based on evaluation Eqs. (7).

LEMMA 4: *Suppose the control limit is shifted, and the action at (i,t) changed. Define $X_n(j,s) = Q_{n-1}(j,s) - Q_n(j,s)$ for all j,s. Then:*

1. *For any $s > t$ and for all j: $X_n(j,s) = 0$.*
2. *For all s and for any $j \geq i_n^*(s)$: $X_n(j,s) = 0$.*
3. *For any $j \neq i$: $X_n(j,t) = 0$.*
4. *For any $j < i_n^*(t-1)$: $X_n(j,t-1) = \alpha \cdot P_{ji}(t) \cdot X_n(i,t)$.*
5. *For all $s < t-1$ and for any $j < i_n^*(s)$:*

$$X_n(j,s) = \alpha \cdot \sum_{k<i_n^*(s+1)} P_{jk}(s+1) \cdot X_n(k,s+1).$$

It is evident from Lemma 4 that updating Q_n is considerably easier than reevaluating $v_n(j,s)$ for all j,s using Eqs. (7). Let

$$A_n(q) = B_i(t) + q - R_i(t) - \alpha \sum_{j=0}^{N} P_{ij}(t+1) Q_n(j,t+1).$$

https://doi.org/10.1017/S026996489913105X Published online by Cambridge University Press

Obviously, when the value of q_n is substituted into A_n, we have a test quantity for the benefit of shifting the control limit up at (i,t): if $A_n(q_n) > 0$, the shift is improving, otherwise it is not. Therefore, in order to evaluate the benefit of a shift, only q_n and A_n need be evaluated explicitly. $v_n(i,t)$ need not be evaluated in full, as the benefit can easily be determined in terms of Q_n. Furthermore, if the shift is indeed beneficial, then $A_n = X_n(i,t)$, so that $X_n(i,t)$ need not be evaluated separately. A similar test quantity for evaluation of a shift down can also be established. The entire algorithm can, therefore, be executed in terms of Q_n, without having to evaluate $v_n(i,t)$ (with the exception of $v_n(0,0) = q_n$) at every iteration.

Executing the algorithm using Q_n obviously involves considerably less computation then solving the problem using $v_n(i,t)$. The following improved algorithm may therefore be devised: After each step, Q_n is reevaluated using Lemma 4. Next, $q_n = v_n(0,0)$ is explicitly calculated. Finally, the benefit of the next step is tested using the test quantity A_n.

We demonstrate the improved algorithm using the same example as before. We show only the first iteration of the algorithm, beginning with the same d_0. Evaluation of Q_0 yields:

$$Q_0 = \begin{bmatrix} 8.52 + 0.9q & 5+q & 5+q & 5+q & 5+q \\ — & 7.2+q & 7.4+q & 7.6+q & 7.8+q \\ — & 9.4+q & 9.8+q & 10.2+q & 10.6+q \\ — & 11.6+q & 12.2+q & 12.8+q & 13.4+q \\ — & 13.8+q & 14.6+q & 15.4+q & 16.2+q \end{bmatrix}.$$

Solving the equation $8.52 + 0.9q = q$ yields, as before, $q_0 = v_0(0,0) = 85.2$. In order to evaluate the benefit of a shift up at $t = 1$ we look at the test quantity

$$A_0(q) = B_0(1) + q - R_0(1) - \alpha \sum_{j=0}^{N} P_{0j}(2) Q_0(j,2) = 0.1q - 4.3.$$

Substituting q_0 yields $A_0(85.2) = 4.2 > 0$. Therefore, the control limit is shifted. We know that $X_1(0,1) = A_0 = 0.1q - 4.3$. Using Lemma 4, Q_1 is now evaluated:

$$\begin{aligned} Q_1(0,1) &= Q_0(0,1) - X_1(0,1) = 0.9q + 9.3 \\ X_1(0,0) &= \alpha \cdot P_{00}(1) \cdot X_1(0,1) = 0.01q - 0.35 \\ Q_1(0,0) &= Q_0(0,0) - X_1(0,0) = 0.89q + 8.87. \end{aligned}$$

The rest of Q remains unchanged. It is now easily seen that $q_1 = 82.1$, and so on. The decrease in computational effort is quite evident.

6. REPAIR–REPLACEMENT MODEL

In this section we add another possible maintenance action—repairing the system to any state $j < i$ without changing the age t of the system, at cost $C_{ij}(t)$. After repair, the system operates at state j for one unit of time at cost $R_j(t)$.

In this case, we will show the optimal policy to hold the form of a *3-way* CLP, that is, do nothing at "good" state-age pairs, repair at "medium" ones, and replace at "bad" ones.

We assume the following extensions of Conditions 3.1–3.4:

Condition 6.1: $R_i(t)$, $B_i(t)$, and $C_{ik}(t)$ (for all $k < i$) are increasing in both i and t.

Condition 6.2: $B_i(t) + R_0(0) \geq C_{ik}(t) + R_k(t) \geq R_i(t)$ for every $t < T^*$, $i = 1, \ldots, N-1$, and $k < i$.

Condition 6.3: The deterioration process is IFR (i.e., Condition 3.3 holds).

Condition 6.4: The functions $B_i(t) - R_i(t)$, $B_i(t) - [C_{ik}(t) + R_k(t)]$, $[C_{ik}(t) + R_k(t)] - R_i(t)$ are all decreasing in both i and t.

The following additional conditions are also imposed.

Condition 6.5: The repair and replacement costs satisfy the triangle inequality

$$C_{ik}(t) \leq C_{ij}(t) + C_{jk}(t); \qquad B_i(t) \leq C_{ik}(t) + B_k(t).$$

Condition 6.6: The function

$$\sum_{j=0}^{N} P_{ij}(t)\Phi_\alpha^T(j,t)$$

is superadditive (a function $h(i,t)$ is superadditive if $[h(i,s) - h(i,t)]$ is increasing in i for $s > t$). The meaning of this condition will be discussed later.

THEOREM 5: *Under Conditions* 6.1–6.6, *and for all aforementioned optimization criteria*, *the optimal policy is a* 3-*way CLP.*

PROOF: Optimality equations similar to (1) are easily established. The equivalent of Lemma 1 is also proven along similar lines. We may then show, similar to the proof of Theorem 1, that every pair of actions (replace/repair/do nothing) obeys a CLP, so that the optimal policy is a 3-way CLP. The extension of the discounted result to the average cost criteria is achieved, as before, through Lemma 2. ■

It is also easily seen that, when repair is the optimal action, it is best to repair to a state in which nothing should be done, due to Condition 6.5.

The replacement-only model in Section 3 is a relatively straightforward generalization of the analogous stationary model [3,6]. The repair–replacement model in this section, however, is more complicated: previously used conditions are greatly extended, and new conditions are required. Of special interest is Condition 6.6: not only is it of entirely new nature, it is also of rather complicated form.

We now discuss the intuitive meaning of this condition, as well as the partial results obtained without it. Reviewing the proof of Theorem 1, and, in particular, Inequality (4) and its intuitive interpretation, we note that while the expected future cost if nothing is done increases in i and t (by Lemma 1), the future cost under replacement is independent of the state and age at which replacement took place (we refer to this property of replacement as renewal). The expected saving in future cost, which is the difference between the aforementioned quantities, is therefore increasing—as indeed noted in the proof of Theorem 1. Such renewal also exists when two systems at different states, but the same age, are repaired to the same state. However, it does not exist in the repair of systems at different ages.

The equivalent of inequality (4), when "doing nothing" is compared with "repair to state k," is

$$C_{ik}(t) + R_k(t) - R_i(t) \le \alpha\left[\sum_{j=0}^{N} P_{ij}(t+1)\Phi_\alpha^{T-1}(j,t+1) - \sum_{j=0}^{N} P_{kj}(t+1)\Phi_\alpha^{T-1}(j,t+1)\right]. \quad \textbf{(15)}$$

The right-hand side of inequality (15) is increasing in i, by the same reasoning described above. However, monotony in t is not quite as simple: as both quantities are dependent upon t, both increasing, the difference between them is not necessarily increasing as well. It is therefore not obvious that the future savings do indeed increase in t, as needed for completion of the proof. Condition 6.6 overcomes this difficulty, as it may be rewritten in the following maner:

$$\sum_{j=0}^{N} P_{ij}(s)\Phi_\alpha^{T}(j,s) - \sum_{j=0}^{N} P_{kj}(s)\Phi_\alpha^{T}(j,s) \ge \sum_{j=0}^{N} P_{ij}(t)\Phi_\alpha^{T}(j,t) - \sum_{j=0}^{N} P_{kj}(t)\Phi_\alpha^{T}(j,t) \quad \textbf{(16)}$$

for any $s > t$ and $i > k$. Inequality (16) has the following intuitive meaning: it states that the reduction in future cost due to repair to state k (as opposed to doing nothing) becomes more substantial as the system ages—thus increasing the benefit of repair at greater age, as required.

The present form of Condition 6.6 is quite impractical. It may be shown that Condition 6.6 can be replaced with the following two conditions.

Condition 6.7: $\sum_{j=k}^{N} P_{ij}(s)$ is superadditive for all k.

Condition 6.8: $\Phi_\alpha^T(i,s)$ is superadditive.

Even though these conditions seem more appealing, they are still somewhat problematic, as Condition 6.8 does not depend explicitly on the problem data. We were unable to establish explicit conditions which ensure the satisfaction of Condition 6.8. However, we now present an intuitively simple example, in which Conditions 6.6 and 6.8 do not hold, thus proving that the remaining conditions are insufficient.

Let $N = 4$, $T^* = 5$, and $\alpha = 0.9$. The maintenance costs $B_i(t)$, $C_{ik}(t)$, as well as the transition probabilities $P_{ij}(t)$ are assumed to be constant with age, while only the operating costs $R_i(t)$ deteriorate:

$$B_0 = 36,\ B_1 = 36,\ B_2 = 37,\ B_3 = 37,\ B_4 = 40$$

$$C_{10} = 17,\ C_{20} = 18,\ C_{30} = 20,\ C_{21} = C_{32} = 7,\ C_{31} = 9$$

$$R_0(0) = R_1(0) = 1,\ R_2(0) = 4,\ R_3(0) = 6,\ R_i(t) = R_i(0) + 0.25 \cdot i \cdot t$$

$$P_{i,j} = \begin{bmatrix} 0.1 & 0.7 & 0.1 & 0.05 & 0.05 \\ 0.1 & 0.6 & 0.2 & 0.05 & 0.05 \\ 0.1 & 0.2 & 0.4 & 0.15 & 0.15 \\ 0.1 & 0.2 & 0.4 & 0.15 & 0.15 \end{bmatrix}.$$

It can be verified that conditions 6.1–6.5 and 6.7 are satisfied. The following policy may be shown to satisfy the optimality equations with respect to the infinite-horizon criterion, and is therefore optimal (R stands for doing nothing, C_1 for repair to state 1, and B for replacement).

		age					
		0	1	2	3	4	T^*
	0	R	R	R	R	R	B
state	1	—	R	R	R	R	B
	2	—	C_1	C_1	R	R	B
	3	—	C_1	C_1	C_1	R	B
	N	—	B	B	B	B	B.

Obviously, this policy is not a 3-way CLP. For example, a system at state 2 should be repaired to state 1 if its age is $t = 1,2$, but left uninterrupted at ages $t = 3,4$. Indeed, evaluation of $\Phi_\alpha(i,t)$ and of $\sum_{j=0}^{N} P_{ij}(t)\Phi_\alpha(j,t)$ reveals that neither is superadditive.

In case Condition 6.6 does not hold (or cannot be verified), only a partial result is obtained. The optimal policy in this case may easily be shown to hold the following form (as does the policy described in the above example): it is a CLP regarding replacement, but only a partial CLP regarding repair.

7. CONCLUSIONS

In this paper we extend the classic stationary Markovian deterioration models to a nonstationary deterioration process, which explicitly depends on the system's cumulative age. The concept of control limit policies is defined in a two-dimensional context, and conditions are established for the optimality of such policies. In the replacement-only model, results are shown to be direct extensions of previous stationary models. We propose an efficient algorithm for finding the optimal policy, utilizing its known form. In the repair–replacement model, however, direct extensions of the conditions used in stationary models are not sufficient to ensure optimality of control limit policies in the nonstationary case. A simple counterexample to the optimality of such policies, under the classic conditions alone, is presented.

Additional, more restrictive and complicated conditions, which ensure optimality of control limit policies, are suggested and discussed.

One direction of future research is further investigation of sufficient conditions for the optimality of CLPs. Of particular importance are explicit and less restrictive conditions for the repair–replacement model of Section 6. Also, more subtle conditions replacing Condition 3.4 may be of interest.

References

1. Derman, C. (1963). On optimal replacement rules when changes of state are Markovian. In R. Bellman (ed.), *Mathematical optimization techniques*. Berkeley: University of California Press, pp. 201–210.
2. Derman, C. (1970). *Finite state Markovian decision processes*. New York: Academic Press.
3. Douer, N. & Yechiali, U. (1994). Optimal repair and replacement in Markovian systems. *Communications in Statistics—Stochastic Models* 10: 253–270.
4. Kao, E.P.C. (1973). Optimal replacement rules when changes of state are semi-Markovian. *Operations Research* 21: 1231–1249.
5. Kijima, M., Morimura, H., & Suzuki, Y. (1988). Periodical replacement problem without assuming minimal repair. *European Journal of Operational Research* 37: 194–203.
6. Kolesar, P. (1966). Minimum cost replacement under Markovian deterioration. *Management Science* 12: 694–706.
7. Lam, C.T. & Yeh, R.H. (1994). Optimal replacement policies for multistate deteriorating systems. *Naval Research Logistics* 41: 303–315.
8. Puterman, M.L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. New York: John Wiley & Sons.
9. So, K.C. (1992). Optimality of control limit policies in replacement models. *Naval Research Logistics* 39: 685–697.
10. Tijms, H.C. (1994). *Stochastic models: An algorithmic approach*. New York: John Wiley & Sons.