Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 168 (2006) 86–99

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Stochastics and Statistics

Optimal preventive maintenance of a production system with an intermediate buffer

E.G. Kyriakidis *, T.D. Dimitrakos

Department of Statistics and Actuarial Science, University of the Aegean, Karlovassi, 83200 Samos, Greece

Received 30 April 2003; accepted 28 January 2004
Available online 25 June 2004

Abstract

In this paper we consider a model consisting of a deteriorating installation that transfers a raw material to a production unit and a buffer which has been built between the installation and the production unit. The deterioration process of the installation is considered to be nonstationary, i.e. the transition probabilities may depend not only on the working conditions of the installation but on its age as well. The problem of the optimal preventive maintenance of the installation is considered. Under a suitable cost structure it is shown that, for fixed age of the installation and fixed buffer level, the optimal policy is of control-limit type. When the deterioration process is stationary, an efficient Markov decision algorithm operating on the class of control-limit policies is developed. There is strong numerical evidence that the algorithm converges to the optimal policy. Two generalizations of this model are also discussed.

Keywords: Maintenance; Markov processes; Dynamic programming

1. Introduction

During the last three decades a great number of articles have appeared dealing with Markov decision models for the optimal maintenance or replacement of an operating device that is subject to a deterioration. For a survey of the research done on this area and the applicability of the models we refer to the papers by Valdez-Flores and Feldman [19], Scarf [16] and Wang [22]. In many such models (see e.g. [1–3,5,6,10,14,17,21]) it can be shown that the optimal policy initiates a maintenance (or a replacement) of the operating device if the degree of its deterioration is greater than or equal to a critical level. Such a policy

* Corresponding author. Address: P. Zervou 13, P. Psychico, 15452 Athens, Greece.
E-mail address: kyriak@aegean.gr (E.G. Kyriakidis).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2004.01.052

is usually called control-limit policy. Van der Duyn Schouten and Vanneste [20] introduced a finite-state Markov decision model for the optimal preventive maintenance of an installation in a production line with an intermediate buffer. Under some suitable conditions they proved that, for each fixed buffer level, the optimal policy is of control-limit type.

The present paper is mainly concerned with an infinite-state generalisation of Van der Duyn Schouten's and Vanneste's model. In our basic model the deterioration process of the installation is considered nonstationary, in the sense that the transition probabilities depend not only on the working conditions of the installation but on its age as well. The concept of nonstationary deterioration was introduced by Benyamini and Yechiali [1] to describe more accurately the deterioration process of many real-life systems since in many cases the characteristics of the system (such as transition probabilities or cost function) change as the system ages. The state of the system in our model consists of the working condition of the installation, its age and the buffer level. Furthermore, the cost structure of our model is more general than the one introduced in [20] since it includes operating costs of the installation, maintenance costs of the installation, storage costs and shortage costs while in [20] only shortage costs are considered. It is assumed that the operating costs of the installation depend on both the working condition and the age of the installation. The existence of an average-cost optimal policy is proved and then, it is shown that, for fixed age of the installation and fixed buffer level, the optimal policy is of control-limit type.

Note that a similar model was introduced by Meller and Kim [11], in which it was assumed that the failure time of the installation is exponentially distributed. The aim of that study was to determine the optimal buffer inventory level, that triggers preventive maintenance of the installation. A cost model was developed and the average cost was calculated as a function of the critical buffer level.

The rest of the paper is organised as follows. The description of the model is given in Section 2. In Section 3 the existence of an optimal policy is proved and then it is shown that the optimal policy is of control-limit type. Section 4 deals with the special case of stationary deterioration, in which the transition probabilities depend only on the working conditions of the installation. The structure of the optimal policy is given and an efficient Markov decision algorithm is developed that generates a sequence of improving control-limit policies. In Section 5 two generalizations of the model and some numerical results are presented and in Section 6 the main conclusions of the paper are summarised.

2. The model

We consider an installation (I), which supplies a raw material to a subsequent production unit (P). The production unit acts as a "pull"-system with a constant demand rate equal to d (units/time). A buffer (B) has been built between the production unit and its input generating installation to cope with unexpected failures of the installation, which may cause delays in production. It is assumed that the capacity of the buffer is equal to K units. As long as the buffer capacity is not reached the installation operates at a constant rate of p units/time ($p > d$) and the excess output is stored in the buffer. When the buffer is full, the installation reduces its speed from p to d. The three components of this manufacturing system are depicted in Fig. 1.

As mentioned in [20] an example of this production system could be an offshore oil exploration platform which provides the crude oil to onshore refineries. The crude oil is transported by pipelines from the

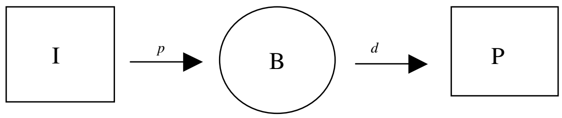

Fig. 1. The three components of the system.

platform to storage tanks, from which is further transported to the refinery. In this case the crude oil, the exploration platform, the refineries and the storage tanks are the raw material, the installation, the production unit and the buffer, respectively. Another possible example (see [11]) of the model is automobile general assembly where the production unit represents the assembly line and the installation represents one of the many parallel operations that directly supply the line.

We suppose that the installation is inspected at discrete, equidistant time epochs $\tau = 0,1,\ldots$ (say every day) and is classified into one of $m+2$ working conditions $0,1,\ldots,m+1$, which describe increasing degrees of deterioration. State 0 denotes a new installation (or functioning as good as new), whereas state $m+1$ denotes a failed (inoperative) installation. The intermediate states $1,\ldots,m$ are operative. If at a time epoch τ the state of the installation is $i<m+1$ and the content of the buffer is $x<K$ then the content of the buffer at the epoch $\tau+1$ will be $\min(x+p-d,K)$. This increase of the buffer content will happen even if the state of the installation at epoch $\tau+1$ is $m+1$. The state of the installation is defined as the pair (i,t), where $i \in \{0,1,\ldots,m+1\}$ is its working condition and $t \in \{0,1,\ldots\}$ is its age. The transition probability of moving from condition i at age t to condition j at age $t+1$ is $p_{ij}(t)$. We assume that the probability of eventually reaching the condition $m+1$ from any given condition i is nonzero.

If the installation at a time epoch is found to be at state $(m+1,t)$, then a corrective maintenance is mandatory. If at a time epoch the installation is found to be at any state (i,t), $i \leqslant m$, a preventive maintenance may be started. Both preventive and corrective maintenance are nonpreemptive, i.e. they cannot be interrupted, and they bring the installation to state (0,0). It is assumed that the preventive and corrective repair times (expressed in time units) are geometrically distributed with probability of success a and b, i.e. the probability that they will last $t \geqslant 1$ time units are equal to $(1-a)^{t-1}a$ and $(1-b)^{t-1}b$, respectively.

The supply of the raw material to the buffer is interrupted during any maintenance (preventive or corrective) of the installation. If during a maintenance the buffer contains some raw material, the production unit operates normally pulling the raw material from the buffer at a constant rate of d units/time. If during a maintenance the buffer is empty, then the operation of the production unit stops.

We suppose that a shortage cost is incurred when a preventive or corrective repair is performed and the buffer is empty. It is assumed that the unit of cost has been chosen in such a way so that, the shortage cost rate is equal to the lost demand d for each unit of time during which a repair (preventive or corrective) is performed.

If at a time epoch the installation is found to be at state $(i,t), 0 \leqslant i \leqslant m, t \geqslant 0$, and no preventive maintenance is initiated, an operating cost is incurred until the next time epoch, which is equal to $c_i(t)$, if the buffer is not full, or to $\tilde{c}_i(t)$, if the buffer is full. When a preventive or a corrective maintenance is performed, a repair cost is incurred, which is equal for each unit of time to $c_p(t)$ or to $c_f(t)$, respectively, where t is the age of the installation. Furthermore, we suppose that the cost of holding a unit of the raw material in the buffer for one time unit is equal to $h>0$.

We introduce the state PM to denote the situation that a preventive maintenance is performed on the installation. Then the state space of the system is the set

$$S = \{0,\ldots,m+1,PM\} \times \{0,1,\ldots\} \times \{0,\ldots,K\},$$

where (i,t,x) denotes the state in which i is the working condition of the installation, t is its age and x is the content of the buffer.

A policy is any rule for choosing actions at each time epoch $\tau = 0,1,\ldots$ There are three actions: 0 (do nothing), 1 (start a preventive maintenance), 2 (start a corrective maintenance). At states $(m+1,t,x), t \geqslant 0, 0 \leqslant x \leqslant K$, action 2 is mandatory. At states (PM,t,x), $t \geqslant 0$, $0 \leqslant x \leqslant K$, the only possible action is 0. At states (i,t,x), $0 \leqslant i \leqslant m$, $t \geqslant 0$, $0 \leqslant x \leqslant K$, actions 0 and 1 are possible.

We consider a discrete-time Markov decision process in which we aim to find a stationary policy which minimises the long-run expected average cost per unit time.

The following conditions on the cost structure and the transition probabilities are assumed to be valid:

Condition 1. For each $t = 0,1,\ldots$, $c_0(t) \leqslant c_1(t) \leqslant \cdots \leqslant c_m(t)$ and $\tilde{c}_0(t) \leqslant \tilde{c}_1(t) \leqslant \cdots \leqslant \tilde{c}_m(t)$. That is, for any given t, as the working condition of the installation deteriorates, the operating cost increases.

Condition 2. For each $t = 0,1,\ldots$, $\tilde{c}_i(t) \leqslant c_i(t)$, $0 \leqslant i \leqslant m$.

That is, the reduction of the speed from p (units/time) to d (units/time) of the installation, as soon as the buffer is filled up, causes a reduction of its operating cost.

Condition 3. $0 < b < a \leqslant 1$. That is, the expected time required for a preventive maintenance is smaller than the expected time required for a corrective maintenance.

Condition 4. For each $t = 0,1,\ldots, c_p(t) \leqslant c_f(t)$. That is, for any given t, the cost rate of a preventive maintenance does not exceed the cost rate of a corrective maintenance.

Condition 5. $c_i(t)$ $(0 \leqslant i \leqslant m)$, $\tilde{c}_i(t)$ $(0 \leqslant i \leqslant m)$, $c_p(t)$, $c_f(t)$ are nondecreasing in t. That is, the operating and maintenance costs of the installation increase as its age increases.

Condition 6. $\sup_{t \geqslant 0} c_m(t) < +\infty$ and $\sup_{t \geqslant 0} c_f(t) < +\infty$. This condition guarantees that the one-step expected costs of the model are bounded.

Condition 7. (an IFR assumption). For each $k = 0, 1,\ldots, m+1$, the function

$$D_k(i,t) = \sum_{j=k}^{m+1} p_{ij}(t)$$

is nondecreasing in both i, $0 \leqslant i \leqslant m$ and $t \geqslant 0$. This condition implies that $I_i(t) \leqslant_{st} I_{i+1}(t), 0 \leqslant i \leqslant m$, where $I_i(t)$ is a random variable representing the next working condition of the installation if its present working condition and age are i and t, respectively.

It may be shown (see [4, pp. 122–123] and [1, p. 58]) that this condition is equivalent to the following one:

Condition 8. For any function $h(j,t)$, $0 \leqslant j \leqslant m+1$, $t \geqslant 0$, which is nondecreasing in both j and t, the function $\sum_{j=0}^{m+1} p_{ij}(t)h(j,t)$ is also nondecreasing in both i and t.

3. The form of the optimal policy

In this section to simplify the analysis we make the assumption $p - d = 1$ (this assumption can be relaxed without problem). Let α $(0 < \alpha < 1)$ be a discount factor. The minimum n-step expected discounted cost $V_n^{\alpha}(i,t,x)$, where (i,t,x) is the initial state, can be found for all $n = 1,2,\ldots$, recursively, from the following equations (see e.g. Chapter 1 in Ross [15]):

$$V_n^{\alpha}(i,t,x) = \min\left\{ c_i(t) + hx + \alpha \sum_{j=0}^{m+1} p_{ij}(t) V_{n-1}^{\alpha}(j,t+1,x+1), V_n^{\alpha}(PM,t,x) \right\},$$
$$0 \leqslant i \leqslant m,\ t \geqslant 0,\ 0 \leqslant x \leqslant K-1,$$

$$V_n^\alpha(i,t,K) = \min\left\{\tilde{c}_i(t) + hK + \alpha\sum_{j=0}^{m+1} p_{ij}(t)V_{n-1}^\alpha(j,t+1,K), V_n^\alpha(PM,t,K)\right\}, \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,$$

$$V_n^\alpha(m+1,t,x) = c_f(t) + hx + (d-x)^+ + \alpha b V_{n-1}^\alpha(0,0,(x-d)^+) + \alpha(1-b)V_{n-1}^\alpha(m+1,t+1,(x-d)^+),$$
$$t \geqslant 0,\ 0 \leqslant x \leqslant K,$$

$$V_n^\alpha(PM,t,x) = c_p(t) + hx + (d-x)^+ + \alpha a V_{n-1}^\alpha(0,0,(x-d)^+) + \alpha(1-a)V_{n-1}^\alpha(PM,t+1,(x-d)^+),$$
$$t \geqslant 0,\ 0 \leqslant x \leqslant K,$$

with initial value

$$V_0^\alpha(i,t,x) = 0, \quad (i,t,x) \in S.$$

Note that $(d-x)^+ = \max(d-x,0)$ and $(x-d)^+ = \max(x-d,0)$ represent the one-period demand lost during maintenance and the buffer content after a one-period maintenance, respectively, when the buffer content equals x at the beginning of that period. The first term in braces in the above equations corresponds to action 0 (do nothing) and the second term corresponds to action 1 (start a preventive maintenance).

Lemma 1. *For each $n = 0,1,\ldots$ we have that*

(i) $V_n^\alpha(i,t,x) \leqslant V_n^\alpha(i+1,t,x), \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,\ 0 \leqslant x \leqslant K,$

(ii) $V_n^\alpha(PM,t,x) \leqslant V_n^\alpha(m+1,t,x), \quad t \geqslant 0,\ 0 \leqslant x \leqslant K,$

(iii) $V_n^\alpha(i,t,x) \leqslant V_n^\alpha(i,t+1,x), \quad 0 \leqslant i \leqslant m+1,\ t \geqslant 0,\ 0 \leqslant x \leqslant K,$

$V_n^\alpha(PM,t,x) \leqslant V_n^\alpha(PM,t+1,x), \quad t \geqslant 0,\ 0 \leqslant x \leqslant K.$

Proof. The lemma holds for $n = 0$, since $V_0{}^\alpha(i,t,x) = 0$ for all $(i,t,x) \in S$. Assume that it holds for $n-1(\geqslant 0)$. We have to show that it also holds for n. For ease of notation we will delete the discount factor α in the notation during the proof of the induction step. First we prove part (ii), then part (i) and finally part (iii).

Part (ii): Let

$$D = V_{n-1}(m+1,t+1,(x-d)^+) - V_{n-1}(0,0,(x-d)^+).$$

Using the expression for $V_n(PM,t,x)$ we have

$$\begin{aligned} V_n(PM,t,x) &= c_p(t) + hx + (d-x)^+ + \alpha a V_{n-1}(0,0,(x-d)^+) + \alpha(1-a)V_{n-1}(PM,t+1,(x-d)^+) \\ &\leqslant c_f(t) + hx + (d-x)^+ + \alpha a V_{n-1}(0,0,(x-d)^+) + \alpha(1-a)V_{n-1}(m+1,t+1,(x-d)^+) \\ &= c_f(t) + hx + (d-x)^+ + \alpha V_{n-1}(m+1,t+1,(x-d)^+) - \alpha a D \leqslant c_f(t) + hx + (d-x)^+ \\ &\quad + \alpha V_{n-1}(m+1,t+1,(x-d)^+) - \alpha b D = V_n(m+1,t,x). \end{aligned}$$

The first inequality follows from Condition 4 and part (ii) of the induction hypothesis. The second inequality follows from Condition 3 and from the inequality $D \geqslant 0$ which is a consequence of parts (i) and (iii) of the induction hypothesis.

Part (i): We have to show that

$$V_n(i,t,x) \leqslant V_n(i+1,t,x), \quad 0 \leqslant i \leqslant m-1,\ t \geqslant 0,\ 0 \leqslant x \leqslant K \tag{1}$$

and

$$V_n(m,t,x) \leqslant V_n(m+1,t,x), \quad t \geqslant 0,\ 0 \leqslant x \leqslant K. \tag{2}$$

The inequality (2) is easily verified for $0 \leqslant x \leqslant K-1$, using part (ii) above:

$$\begin{aligned} V_n(m,t,x) &= \min\left\{c_m(t) + hx + \alpha \sum_{j=0}^{m+1} p_{mj}(t) V_{n-1}(j,t+1,x+1), V_n(PM,t,x)\right\} \\ &\leqslant V_n(PM,t,x) \leqslant V_n(m+1,t,x). \end{aligned}$$

Similarly, we obtain the inequality for $x=K$. For $0 \leqslant x \leqslant K-1$, $t \geqslant 0$ and $0 \leqslant i \leqslant m-1$ we obtain

$$\begin{aligned} V_n(i,t,x) &= \min\left\{c_i(t) + hx + \alpha \sum_{j=0}^{m+1} p_{ij}(t) V_{n-1}(j,t+1,x+1), V_n(PM,t,x)\right\} \leqslant \min\left\{c_{i+1}(t) + hx \right.\\ &\quad \left. + \alpha \sum_{j=0}^{m+1} p_{i+1,j}(t) V_{n-1}(j,t+1,x+1), V_n(PM,t,x)\right\} = V_n(i+1,t,x). \end{aligned}$$

The above inequality follows from Condition 1 and the inequality

$$\sum_{j=0}^{m+1} p_{ij}(t) V_{n-1}(j,t+1,x+1) \leqslant \sum_{j=0}^{m+1} p_{i+1,j}(t) V_{n-1}(j,t+1,x+1),$$

which is implied by parts (i) and (iii) of the induction hypothesis and Condition 8. Hence the inequality (1) has been proved for $0 \leqslant x \leqslant K-1$. Similarly, we obtain the inequality (1) for $x=K$.

Part (iii): From the expressions for $V_n(m+1,t,x)$ and $V_n(PM,t,x)$, from Condition 5 and from part (iii) of the induction hypothesis, it follows that $V_n(m+1,t,x)$ and $V_n(PM,t,x)$ are nondecreasing in $t \geqslant 0$. From Parts (i) and (iii) of the induction hypothesis and Condition 8 it follows that the expression

$$\sum_{j=0}^{m+1} p_{ij}(t) V_{n-1}(j,t+1,x+1)$$

is nondecreasing in $t \geqslant 0$. Hence, from the expression for $V_n(i,t,x)$, the monotonicity of $V_n(PM,t,x)$ in t and the Condition 5, it follows that $V_n(i,t,x)$, $0 \leqslant i \leqslant m$, $0 \leqslant x \leqslant K-1$, is nondecreasing in t. Similarly, it can be proved that $V_n(i,t,K)$, $0 \leqslant i \leqslant m$, is nondecreasing in t.

The proof of the lemma is now complete. □

Let $V^\alpha(s)$ be the minimum expected total α-discounted cost with initial state $s \in S$. From Proposition 3.1 in [15] it follows that $V^\alpha(s)$ is the limit of $V_n^\alpha(s)$ as $n \to \infty$. Hence, from parts (i) and (iii) of Lemma 1 we obtain the following result:

Lemma 2

$$V^\alpha(i,t,x) \leqslant V^\alpha(i+1,t,x), \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,\ 0 \leqslant x \leqslant K,$$

$$V^\alpha(i,t,x) \leqslant V^\alpha(i,t+1,x), \quad 0 \leqslant i \leqslant m+1,\ t \geqslant 0,\ 0 \leqslant x \leqslant K.$$

Let π^* be the stationary policy that minimizes the expected total α-discounted cost for any initial state. The following lemma guarantees the existence of an average-cost optimal stationary policy (see Theorems 2.1 and 2.2 in [15]).

Lemma 3. There exists a positive number B such that

$$|V^\alpha(i,t,x) - V^\alpha(0,0,x^*)| \leqslant B \quad \text{for all } (i,t,x) \in S \quad \text{and}$$
$$\alpha \in (0,1), \text{ where } V^\alpha(0,0,x^*) = \max_{0 \leqslant x \leqslant K} V^\alpha(0,0,x).$$

Proof. Assume that the initial state of the system is $(i,t,x) \in S$. Let X_r, a_r, $C(X_r, a_r)$, $r=0,1,\dots$ denote the state of the system at time r, the action chosen at time r and the corresponding cost incurred at time r, respectively. Let π denote the nonstationary policy that starts a maintenance at the initial state (i,t,x) and, after the completion of the maintenance, chooses the same actions as the optimal policy π^*. Let also $V^\alpha_\pi(i,t,x)$ denote the expected total α-discounted cost under the policy π if (i,t,x) is the initial state. Let $N \in \{1,2,\dots\}$ be the time at which the maintenance is completed. From the definitions of π and N it follows that:

$$\begin{aligned}
V^\alpha(i,t,x) \leqslant V^\alpha_\pi(i,t,x) &= E_\pi\left[\sum_{r=0}^{\infty} \alpha^r C(X_r,a_r) | X_0 = (i,t,x)\right] \\
&= E_\pi\left[\sum_{r=0}^{N-1} \alpha^r C(X_r,a_r) | X_0 = (i,t,x)\right] + E_\pi\left[\sum_{r=N}^{\infty} \alpha^r C(X_r,a_r) | X_0 = (i,t,x)\right] \\
&= E_\pi\left[\sum_{r=0}^{N-1} \alpha^r C(X_r,a_r) | X_0 = (i,t,x)\right] + \sum_{i=1}^{\infty} P(N=i)\alpha^i V^\alpha(0,0,(x-id)^+) \\
&\leqslant E_\pi\left[\sum_{r=0}^{N-1} \alpha^r C(X_r,a_r) | X_0 = (i,t,x)\right] + V^\alpha(0,0,x^*) \\
&\leqslant \max\left\{d + hK + \sup\nolimits_{t \geqslant 0} c_f(t), \sup_{t \geqslant 0} c_m(t) + hK\right\} E_\pi\left[\sum_{r=0}^{N-1} \alpha^r | X_0 = (i,t,x)\right] + V^\alpha(0,0,x^*) \\
&\leqslant \max\left\{d + hK + \sup_{t \geqslant 0} c_f(t), \sup_{t \geqslant 0} c_m(t) + hK\right\} E(N) + V^\alpha(0,0,x^*) (\text{since } \alpha \in (0,1)) \\
&\leqslant B_1 + V^\alpha(0,0,x^*),
\end{aligned}$$

where,

$$B_1 = b^{-1} \max\{d + hK + \sup_{t \geqslant 0} c_f(t), hK + \sup\nolimits_{t \geqslant 0} c_m(t)\}.$$

From parts (i) and (iii) of Lemma 1 it follows that

$$V^\alpha(i,t,x) - V^\alpha(0,0,x^*) \geqslant V^\alpha(0,0,\tilde{x}) - V^\alpha(0,0,x^*), \quad (i,t,x) \in S,$$

where, $V^\alpha(0,0,\tilde{x}) = \min_{0 \leqslant x \leqslant K} V^\alpha(0,0,x)$.

Hence,

$$|V^\alpha(i,t,x) - V^\alpha(0,0,x^*)| \leqslant \max\{B_1, V^\alpha(0,0,x^*) - V(0,0,\tilde{x})\} = B, \quad (i,t,x) \in S. \qquad \square$$

Lemma 3 implies (see Theorem 2.2 in Ross [15, p. 95]) that there exist numbers $v(s)$, $s \in S$, and a constant g such that the following average-cost optimality equations hold:

$$v(i,t,x) = \min\left\{c_i(t) + hx - g + \sum_{j=0}^{m+1} p_{ij}(t) v(j,t+1,x+1), v(PM,t,x)\right\}, \quad 0 \leqslant i \leqslant m, \ t \geqslant 0,$$
$$0 \leqslant x \leqslant K-1,$$

$$v(i,t,K) = \min\left\{\tilde{c}_i(t) + hK - g + \sum_{j=0}^{m+1} p_{ij}(t)v(j,t+1,K), v(PM,t,K)\right\}, \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,$$

$$v(m+1,t,x) = c_f(t) + hx + (d-x)^+ - g + bv(0,0,(x-d)^+) + (1-b)v(m+1,t+1,(x-d)^+),$$
$$t \geqslant 0,\ 0 \leqslant x \leqslant K,$$

$$v(PM,t,x) = c_p(t) + (d-x)^+ + hx - g + av(0,0,(x-d)^+) + (1-a)v(PM,t+1,(x-d)^+),$$
$$t \geqslant 0,\ 0 \leqslant x \leqslant K.$$

The first term in braces in the above equations corresponds to action 0 (do nothing) and the second term to action 1 (start a preventive maintenance). If for some state (i,t,x) the first term is smaller than the second term then the optimal action at this state is the action 0. If the first term is greater than the second term then the optimal action is action 1. If the terms are equal then both actions are optimal. According to Theorem 2.2 in Ross [15] the numbers $v(s), s \in S$, are given by

$$v(s) = \lim_{n\to\infty}[V^{\alpha_n}(s) - V^{\alpha_n}(0,0,x^*)]$$

for some sequence $\alpha_n \to 1 (0 < \alpha_n < 1, n \in \mathrm{N})$.

In view of Lemma 2 the above limiting expressions imply the following result:

Corollary 1

$$v(i,t,x) \leqslant v(i+1,t,x), \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,\ 0 \leqslant x \leqslant K,$$

$$v(i,t,x) \leqslant v(i,t+1,x), \quad 0 \leqslant i \leqslant m+1,\ t \geqslant 0,\ 0 \leqslant x \leqslant K.$$

The proposition below provides a characterization of the structure of the optimal policy.

Proposition 1. *For fixed buffer content $x, 0 \leqslant x \leqslant K$ and for fixed age $t \geqslant 0$ of the installation, there exists a critical working condition $i^*(=i^*(t,x))$ such that the optimal policy initiates a preventive maintenance if and only if the working condition i of the installation is greater than or equal to i^*.*

Proof. Suppose that the optimal policy prescribes action 1 at state $(i,t,x), t \geqslant 0, 0 \leqslant x \leqslant K-1$. This implies that

$$v(PM,t,x) \leqslant c_i(t) + hx - g + \sum_{j=0}^{m+1} p_{ij}(t)v(j,t+1,x+1). \tag{3}$$

To show that the optimal policy prescribes action 1 at state $(i+1,t,x)$ it is enough to verify that

$$v(PM,t,x) \leqslant c_{i+1}(t) + hx - g + \sum_{j=0}^{m+1} p_{i+1,j}(t)v(j,t+1,x+1). \tag{4}$$

From Conditions 1, 8 and Corollary 1 it follows that the right-hand side of (4) is greater or equal to the right-hand side of (3). Hence (3) implies (4). The same result is obtained similarly when $x = K$. □

4. Stationary deterioration

Consider the problem introduced in Section 2 with the following assumptions:

$$p_{ij}(t) = p_{ij}, \quad 0 \leqslant i,\ j \leqslant m+1,\ t \geqslant 0,$$

$$c_i(t) = c_i, \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,$$

$$\tilde{c}_i(t) = \tilde{c}_i, \quad 0 \leqslant i \leqslant m,\ t \geqslant 0,$$

$$c_p(t) = c_p, \quad t \geqslant 0,$$

$$c_f(t) = c_f, \quad t \geqslant 0.$$

The state of the installation is now defined as its working condition $i \in \{0,1,\ldots,m+1\}$ and does not depend on its age. The state space of the system is the finite set

$$S = \{0, 1, \ldots, m+1, PM\} \times \{0, 1, \ldots, K\},$$

where (i,x) denotes the state in which i is the working condition of the installation and x is the content of the buffer. The process under any policy is regenerative with regenerative state (0,0). From known results of Markov decision theory (see e.g. Chapter 3 in [18]) it follows that an optimal stationary policy exists. Proposition 1 of the previous section takes now the following form:

Proposition 2. *For fixed buffer content x, $0 \leqslant x \leqslant K$, there exists a critical working condition $i^*(= i^*(x))$ such that the optimal policy initiates a preventive maintenance if and only if the working condition i of the installation is greater than or equal to i^*.*

Remark 1. Note that if $c_p = c_f = h = 0$, $c_i = \tilde{c}_i = 0$ $(0 \leqslant i \leqslant m)$, $p_{m,m+1} = 1$, $p_{i,i+1} = r_i$, $p_{i,m+1} = 1 - r_i$ $(0 \leqslant i \leqslant m-1)$, where $\{r_i\}$ $(0 \leqslant i \leqslant \mathrm{m})$ is nondecreasing in i then the above model coincides with the one introduced by Van der Duyn Schouten and Vanneste [20]. In this case the average cost under any policy represents the average lost demand of the production unit per unit time. Proposition 2 extends a similar result obtained in [20].

Remark 2. Extensive numerical results provide strong evidence that the critical level $i^*(x)$ is nonincreasing in x. However, an analytical proof of this conjecture seems difficult.

The optimal policy can be computed by implementing any standard algorithm (policy iteration, value iteration, linear programming). However it is possible to develop a computationally tractable algorithm which generates a sequence of improving control-limit policies (i.e. policies that, for each buffer level x, initiate the preventive maintenance if and only if the working condition of the installation is greater than or equal to some critical level $i(x)$). There is strong numerical evidence that the final policy obtained by the algorithm is the optimal one. The design of the algorithm is based on the embedding technique of Tijms (see [18, p. 234]). Similar algorithms have been developed in queueing, inventory, maintenance and pest control models (see [18, pp. 234–248], Nobel and Tijms [12,13], Kyriakidis [7–9]). The description of the algorithm is given below.

4.1. A special-purpose policy iteration algorithm

Consider a particular control-limit policy R characterized by the critical numbers $i(x)$, $0 \leqslant x \leqslant K$, and let $w(s)$, $s \in S$, be the associated relative values (see Tijms [18, p. 188]). We choose

$$w(0,0) = 0. \tag{5}$$

Note that the set of states $E = \bigcup_{x=0}^{K}\{(i,x) : 0 \leqslant i \leqslant i(x)\}$ can be reached from every initial state $s \in S$ if the policy R is employed. The embedding technique can be applied if we take the set E as the embedded set of states. According to relation (3.6.1) in Tijms [18, p. 235] the relative values $w(s)$,$s \in S$, satisfy the following system of linear equations:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), s \in S, \tag{6}$$

where $g(R)$ is the average cost under the policy R, ${p_{s\ell}}^E$ is the probability that the first entry state in the set E equals ℓ given that the policy R is used and the initial state is s and, ${T_s}^E$ and ${C_s}^E$ are the expected time and cost, respectively. It is assumed that, for the initial state $s \in E$, the first entry state in the set E is the state at the next return to the set E.

The quantities ${T_s}^E$, ${C_s}^E$ and ${p_{s\ell}}^E$, $s \in S$, $\ell \in E$ can be easily computed by taking into account the structure of the model under the policy R and by using simple conditional arguments. For example,

$$T_{(i,x)}^E = 1 + a^{-1} \sum_{j=i(x+1)+1}^{m} p_{ij} + b^{-1} p_{i,m+1}, \quad 0 \leqslant i < i(x),\ 0 \leqslant x \leqslant K-1,$$

$$\begin{aligned} C_{(i,x)}^E = {} & a^{-1}\{c_p + (1-a)^{[x/d]}(d - ax + ad[x/d])\} + \frac{h}{2}(1-a)^{[x/d]}\left(1 + \left[\frac{x}{d}\right]\right)\left(2x - \mathrm{d}\left[\frac{x}{d}\right]\right) \\ & + \frac{ha}{2}\sum_{t=1}^{[x/d]}(1-a)^{t-1}(2tx - \mathrm{d}t^2 + \mathrm{d}t), \quad i(x) \leqslant i \leqslant m,\ 0 \leqslant x \leqslant K. \end{aligned}$$

$$p_{(i,K)(j,K)}^E = p_{ij}, \quad 0 \leqslant i < i(K),\ 0 \leqslant j \leqslant i(K),$$

$$p_{(i,x)(0,0)}^E = (1-a)^{[(x+1)/d]} \sum_{j=i(x+1)+1}^{m} p_{ij} + p_{i,m+1}(1-b)^{[(x+1)/d]}, \quad 0 \leqslant i < i(x),\ 0 \leqslant x \leqslant K-1,$$

where $[x/d]$ and $[(x+1)/d]$ denote the integer parts of x/d and $(x+1)/d$, respectively.

The relative values $w(s)$, $s \in E$, and the average cost $g(R)$ can be found by solving the system of linear equations:

$$w(s) = C_s^E - g(R)T_s^E + \sum_{\ell \in E} p_{s\ell}^E w(\ell), \quad s \in E \tag{7}$$

together with the normalization condition (5).

We consider now the so-called policy improvement quantity $Q_R(s;a)$ associated with the policy R defined by

$$Q_R(s;a) = C(s,a) - g(R)T(s,a) + \sum_{\ell \in S} p_{s\ell}(a)w(\ell), \quad s \in \{0,\ldots,m\} \times \{0,\ldots,K\},\ a \in \{0,1\}, \tag{8}$$

where $p_{s\ell}(a)$ is the probability that the next state of the process will be ℓ, given that the present state is s and the action a is chosen, and $T(s,a)$ and $C(s,a)$ are the corresponding one-step expected time and cost, respectively.

Suppose that for some x ($0 \leqslant x \leqslant K$) there exists an integer $\tilde{i}(x)$ such that $0 \leqslant \tilde{i}(x) < i(x)$ and $Q_R((i,x);1) < w(i,x)$, $\tilde{i}(x) \leqslant i < i(x)$. Then, according to Theorem 3.2.1 in Tijms [18, p. 192], the control-limit policy that is characterized by the critical numbers $i(0), \ldots, i(x-1)$, $\tilde{i}(x)$, $i(x+1), \ldots, i(K)$ achieves smaller average cost than $g(R)$.

Similarly, if for some $x(0 \leqslant x \leqslant K)$ there exists an integer $\tilde{i}(x)$ such that $i(x) < \tilde{i}(x) \leqslant m+1$ and $Q_R((i,x);0) < w(i,x)$, $i(x) \leqslant i < \tilde{i}(x)$, then, according to Theorem 3.2.1 in [18], the control-limit policy that is characterized by the critical numbers $i(0), \ldots, i(x-1), \tilde{i}(x), i(x+1), \ldots, i(K)$ achieves smaller average cost than $g(R)$.

The above remarks lead us to design the following algorithm which generates a sequence of strictly improving control-limit policies. The corresponding Matlab program can be found in Web site http://www.actuar.aegean.gr/english/Department/Algorithm.html.

Special-purpose algorithm:

Step 1: (initialization). Choose an initial control-limit policy R characterized by the critical numbers $i(x)$, $0 \leqslant x \leqslant K$.

Step 2: (value-determination step). For the current control-limit policy R compute the average cost $g(R)$ and the associated relative values $w(s)$, $s \in E$, by solving the system of linear equations (5), (7).

Step 3: (policy improvement step). For each x, $0 \leqslant x \leqslant K$

(a) Find, if it exists, the smallest integer $\tilde{i}(x)$ such that $0 \leqslant \tilde{i}(x) < i(x)$ and $Q_R((i,x);1) < w(i,x)$, $\tilde{i}(x) \leqslant i < i(x)$. Otherwise

(b) Find, if it exists, the largest integer $\tilde{i}(x)$ such that $i(x) < \tilde{i}(x) \leqslant m+1$ and $Q_R((i,x);0) < w(i,x)$, $i(x) \leqslant i < \tilde{i}(x)$.

The quantities $Q_R((i,x);1)$ and $Q_R((i,x);0)$ are given by (8), where, if it is necessary, $w(s), s \in S-E$, can be computed from (6).

Replace $i(x)$ by $\tilde{i}(x)$ for those x, $0 \leqslant x \leqslant K$, for which it is possible to find an integer $\tilde{i}(x)$ and go to Step 2.

Step 4: (convergence test). If it is not possible to find any $\tilde{i}(x)$, $0 \leqslant x \leqslant K$, the algorithm is stopped. The final policy is R with average cost $g(R)$.

The algorithm generates a sequence of strictly improving control-limit policies and stops after a finite number of iterations since the set of control-limit policies is finite. There is strong numerical evidence that the final policy obtained by the algorithm is the optimal one. A proof of this conjecture seems to be difficult. The computational time required by the algorithm is considerably smaller than the computational time required by the standard policy-iteration algorithm. This is due to the fact that the number of the unknowns in the value determination step of our algorithm is equal to the number of the elements of the embedded set of states E while the number of the unknowns in the value determination step of the standard policy iteration algorithm is equal to the number of the elements of the entire state space S. Note also, from a great number of examples that we have tested, that there is strong evidence that the number of iterations of the special-purpose policy iteration algorithm is not especially influenced by the initial control-limit policy.

As an illustration we consider the following example.

Example 1. Suppose that $m=50$, $K=10$, $a=0.9$, $b=0.2$, $p=9$, $d=8$, $c_p=0.4$, $c_f=0.8$, $h=0.5$, $c_i=0.1(i+1)$, $\tilde{c}_i = 0.05(i+1)$, $0 \leqslant i \leqslant m$. We assume that the nonzero transition probabilities p_{ij}, $0 \leqslant i, j \leqslant m+1$ are given by

$$p_{ij} = \frac{1}{m+2-i}, \quad i \leqslant j \leqslant m+1.$$

Table 1
The successive control-limit policies generated by the algorithm

The critical values $i(x)$, $0 \leqslant x \leqslant 10$ of the policies	Average cost
$i(0)=\cdots=i(10)=50$	6.416
$i(0)=13,\ i(1)=\cdots=i(10)=0$	4.392
$i(0)=37,\ i(1)=34,\ i(2)=30,\ i(3)=27,\ i(4)=23,$ $i(5)=18,\ i(6)=14,\ i(7)=9,\ i(8)=i(9)=i(10)=0$	3.872
$i(0)=33,\ i(1)=29,\ i(2)=26,\ i(3)=22,\ i(4)=17,\ i(5)=13,$ $i(6)=9,\ i(7)=4,\ i(8)=i(9)=i(10)=0$	3.855

This means that if the present state of the installation is i, then the next state is uniformly distributed in the set $\{i,i+1,\ldots,m+1\}$. These probabilities satisfy Condition 7 since, for each $k=0,\ldots,m+1$, the quantity

$$\sum_{j=k}^{m+1} p_{ij} = \frac{m+2-k}{m+2-i}$$

is increasing in i, $0 \leqslant i \leqslant 50$. As initial control-limit policy we choose the policy that is characterized by the critical numbers $i(x)=50$, $0 \leqslant x \leqslant 10$. In Table 1 we present the successive control-limit policies generated by the algorithm and their average costs.

Note that in this example when the buffer content is quite high (between 8 and 10) the optimal policy initiates a preventive maintenance of the installation even if its working condition is 0. This is intuitively reasonable since the transition probabilities to states with high operating costs are not negligible in this example.

5. Generalizations

In this section we consider two generalizations of the Markov decision model of the previous section.

(i) Suppose that, if the installation at a time epoch is found to be at working condition $i \leqslant m$ and a preventive maintenance is initiated, then the preventive repair time is geometrically distributed with probability of success a_i. It is assumed that $a_0 \geqslant a_1 \geqslant \cdots \geqslant a_m \geqslant b$. This means that, as the working condition of the installation deteriorates, the expected value of the preventive repair time increases.

Extensive numerical results provide strong evidence that the optimal policy, for fixed buffer content x, initiates the preventive maintenance of the installation if and only if its working condition is greater than or equal to a critical value $i^*(x)$. An analytical proof seems to be difficult. As an illustration we consider the following example.

Example 2. Suppose that $m=10$, $K=5$ and the nonzero transition probabilities p_{ij}, $0 \leqslant i,j \leqslant 11$, are given by

$$p_{ij} = \frac{1}{12-i}, \quad i \leqslant j \leqslant 11.$$

Let $c_i=0.1(i+1)$, $\tilde{c}_i = 0.05(i+1)$, $0 \leqslant i \leqslant 10$, $a_i=10(10+i)^{-1}$, $0 \leqslant i \leqslant 10$, $b=0.1$, $c_p=0.4$, $c_f=0.8$, $d=2$, $p=3$, $h=0.2$. The optimal policy is characterized by the critical numbers:

$$i^*(0) = 6, i^*(1) = 5, i^*(2) = 2, i^*(3) = i^*(4) = i^*(5) = 0.$$

Its average cost is 1.51.

(ii) Suppose that, if the installation at a time epoch is found to be at working condition $i \leqslant m$, there is a possibility to initiate a preventive maintenance that will restore the installation to any better state j, where $j \in \{0,1,\ldots,i-1\}$. The notion of partial maintenance was used in [1,5]. It is assumed that the time of such a partial maintenance is geometrically distributed with probability of success a_{i-j}. If the installation is found to be at state $m+1$, a corrective maintenance is mandatory, which restores the installation to state 0 after a geometrically distributed time with expected value b^{-1}. Suppose that $a_0 \geqslant a_1 \geqslant \cdots \geqslant a_m \geqslant b$.

Extensive numerical results provide strong evidence that for this model the optimal policy, for fixed buffer content x, initiates a preventive maintenance if and only if the working condition of the installation is greater than or equal to a critical value $i^*(x)$. An analytical proof seems difficult. As an illustration we consider the following example.

Example 3. Suppose that the state space is $S = \{0,1,2,3,4\} \times \{0,1,2,3\}$.

The matrix P with the underlying transition probabilities p_{ij} for $0 \leqslant i,j \leqslant 4$, is given below:

$$P = \begin{pmatrix} 0.1 & 0.7 & 0.1 & 0.05 & 0.05 \\ 0 & 0.8 & 0.1 & 0.05 & 0.05 \\ 0 & 0 & 0.5 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Let $c_i = 0.1(i+1)$, $\tilde{c}_i = 0.05(i+1)$, $0 \leqslant i \leqslant 3$, $a_i = (3+i)^{-1}$, $0 \leqslant i \leqslant 3$, $b = 0.15$, $c_p = 0.3$, $c_f = 0.6$, $h = 0.4$, $d = 1$, $p = 2$. The optimal policy is characterized by the critical numbers:

$$i^*(0) = 2, i^*(1) = 2, i^*(2) = i^*(3) = 0.$$

Its average cost is 0.72.

6. Conclusions

In this paper we introduced a Markov decision model for the optimal preventive maintenance of an installation that supplies a raw material to a production unit. A buffer has been built between the installation and the production unit to cope with unexpected failures of the installation, which may cause interruptions of the production process. Using suitable techniques from Markov decision theory, we have shown that, for fixed age of the installation and fixed buffer level, the policy that minimizes the expected long-run average cost per unit time is of control-limit type. In the case of stationary deterioration an efficient special-purpose policy iteration algorithm that generates a sequence of stricly improving control-limit policies was developed. Extensive numerical results provide strong evidence that the algorithm converges to the optimal policy.

It would be interesting to examine whether these results are valid in the case in which the preventive and corrective repair times are not geometrically distributed. If these times follow the exponential (or the Erlang) distribution and the expected value of a preventive maintenance does not exceed the expected value of a corrective maintenance it seems again intuitively reasonable that the optimal policy is of control-limit type. However it seems difficult to prove this assertion. When the deterioration process is stationary, a computational approach is possible if we construct a suitable semi-Markov decision model. The results will be presented in a subsequent paper.

In the present paper we assumed that the production unit runs without risk provided that it pulls the raw material from the buffer at a constant rate of d units/time. The construction of a more general model in which the production unit could fail might be a subject of further research.

Acknowledgement

The authors would like to thank Dr. A. N. Yannacopoulos for helpful comments. Special thanks are also due to two anonymous referees whose comments improved the presentation of the paper.

References

[1] Z. Benyamini, U. Yechiali, Optimality of control limit maintenance policies under nonstationary deterioration, Probability in the Engineering and Informational Sciences 13 (1999) 55–70.
[2] M. Berg, A proof of optimality for age replacement policies, Journal of Applied Probability 13 (1976) 751–759.
[3] M. Chen, R.M. Feldman, Optimal replacement policies with minimal repair and age-dependent costs, European Journal of Operational Research 98 (1997) 75–84.
[4] C. Derman, Finite State Markovian Decision Processes, Academic Press, New York, 1970.
[5] N. Douer, U. Yechiali, Optimal repair and replacement in Markovian systems, Stochastic Models 10 (1994) 253–270.
[6] A. Federgruen, K.C. So, Optimal time to repair a broken server, Advances in Applied Probability 21 (1989) 376–397.
[7] E.G. Kyriakidis, A Markov decision algorithm for optimal pest control through uniform catastrophes, European Journal of Operational Research 64 (1993) 38–44.
[8] E.G. Kyriakidis, Optimal control of a truncated general immigration process through total catastrophes, Journal of Applied Probability 36 (1999) 461–472.
[9] E.G. Kyriakidis, Optimal control of a simple immigration-emigration process through total catastrophes, European Journal of Operational Research 155 (2004) 198–208.
[10] C.E. Love, Z.G. Zhang, M.A. Zitron, R. Guo, A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs, European Journal of Operational Research 125 (2000) 398–409.
[11] R.D. Meller, D.S. Kim, The impact of preventive maintenance on system cost and buffer size, European Journal of Operational Research 95 (1996) 577–591.
[12] R.D. Nobel, H.C. Tijms, Optimal control for an Mx/G/1 queue with two services modes, European Journal of Operational Research 113 (1999) 610–619.
[13] R.D. Nobel, H.C. Tijms, Optimal control of a queueing system with heterogeneous servers and set up costs, IEEE Transactions on Automatic Control 45 (2000) 780–784.
[14] S. Ozekici, N.O. Gunlunk, Maintenance of a device with age-dependent exponential failures, Naval Research Logistics 39 (1992) 699–714.
[15] S.M. Ross, Introduction to Stochastic Dynamic Programming, Academic Press, New York, 1983.
[16] P.A. Scarf, On the application of mathematical models in maintenance, European Journal of Operational Research 99 (1997) 493–506.
[17] K.C. So, Optimality of control limit policies in replacement models, Naval Research Logistics 39 (1992) 685–697.
[18] H.C. Tijms, Stochastic Models: An Algorithmic Approach, Wiley, Chichester, 1994.
[19] C. Valdez-Flores, R.M. Feldman, A survey of preventive maintenance models for stochastically deteriorating single-unit systems, Naval Research Logistics 36 (1989) 419–446.
[20] F.A. Van Der Duyn Schouten, S.G. Vanneste, Maintenance optimization of a production system with buffer capacity, European Journal of Operational Research 82 (1995) 323–338.
[21] S.G. Vanneste, A generalized age-replacement model, Probability in the Engineering and Informational Sciences 6 (1992) 525–541.
[22] H. Wang, A survey of maintenance policies of deteriorating systems, European Journal of Operational Research 139 (2002) 469–489.