Available online at www.sciencedirect.com
ScienceDirect
Int. J. Production Economics 111 (2008) 752–762
international journal of production economics
www.elsevier.com/locate/ijpe

A semi-Markov decision algorithm for the maintenance of a production system with buffer capacity and continuous repair times

T.D. Dimitrakos[a], E.G. Kyriakidis[b,*]

[a]*Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Karlovassi 83200, Samos, Greece*
[b]*Department of Financial and Management Engineering, University of the Aegean, 31 Fostini Str., 82100 Chios, Greece*

Received 19 January 2005; accepted 27 March 2007
Available online 27 April 2007

Abstract

We consider a model consisting of a deteriorating installation (I) that transfers a raw material to a production unit and a buffer, which has been built between the I and the production unit to cope with unexpected failures of the I that may cause delays in production. The problem of the optimal preventive maintenance of the I is considered. It is assumed that the repair times follow some known continuous distributions. It seems intuitively reasonable that, for fixed buffer level, the optimal policy is of control-limit type, i.e. it initiates the preventive maintenance of the I if and only if the degree of its deterioration exceeds a critical level. An efficient semi-Markov decision algorithm, which operates on the class of control-limit policies, is developed. Numerical examples provide strong evidence that the algorithm converges to the optimal policy.

Keywords: Maintenance; Buffer; Control-limit policies; Semi-Markov decision processes

1. Introduction

The Markov decision process is a mathematical model, which is used to describe a stochastic process controlled by a sequence of actions. Many papers have appeared dealing with Markov decision models for the optimal maintenance or replacement of a device, which operates in time and is subject to deterioration. The papers of Scarf (1997) and Wang (2002) give a summary of the research done in this area.

In many such models (see e.g. Federgruen and So (1989), Douer and Yechiali (1994), Chen and Feldman (1997), Love et al. (2000)) it can be shown that the optimal policy initiates a repair or a replacement of the device if the degree of its deterioration is greater than or equal to a critical level. Such a policy is usually called control-limit policy and the critical level the control limit.

Van Der Duyn Schouten and Vanneste (1995) considered a finite-state Markov decision process for the optimal preventive maintenance of an installation (I) in a production line with an

*Corresponding author. P. Zervou 13, P. Psychico 15452, Athens, Greece. Tel.: +22710 35464; fax: +22710 35499.
E-mail address: kyriak@fme.aegean.gr (E.G. Kyriakidis).

0925-5273/$ - see front matter
doi:10.1016/j.ijpe.2007.03.010

intermediate buffer. They assumed that a shortage cost is incurred when a repair of the I is performed and the buffer is empty and they represented the state of the production system by the age of the I and the buffer level. Under the assumption that the repair times of the I are geometrically distributed, they proved that, for each fixed buffer level, the average-cost optimal policy is of control-limit form.

In a previous paper (see Sections 4 and 5 in Kyriakidis and Dimitrakos, 2006), we considered a generalization of Van Der Duyn Schouten's and Vanneste's model. We represented the state of the production system by the working condition of the I and the buffer level and we introduced operating costs and repair costs of the I, shortage costs and storage costs. Under the assumption that the repair times are geometrically distributed we proved that the average-cost optimal policy has the same structure as in Van Der Duyn Schouten's and Vanneste's model.

In the present paper, we study the case in which the repair times of the I follow some known continuous distributions (e.g. Exponential, Gamma, Weibull). It seems again intuitively reasonable that, for any fixed buffer level, the optimal policy is of control-limit type. A proof of this conjecture seems to be difficult. However, a computational approach is possible. We formulate an approximate semi-Markov decision process with discrete state space and, then, we design an efficient tailor-made policy iteration algorithm that generates a sequence of improving control-limit policies. There is strong numerical evidence that the final policy obtained by the algorithm is the optimal one. Furthermore, the computational time required by our algorithm is considerably smaller than any standard Markov decision algorithm.

We point out that four similar models have been studied by Meller and Kim (1996), Salameh and Ghattas (2001), Ribeiro et al. (2007) and Charlot et al. (2007). In Meller's and Kim's model the failure time of the I was assumed to be exponentially distributed. The aim of that study was to determine the optimal buffer level that triggers preventive maintenance of the I. A cost model was developed and the average cost was calculated as a function of the critical buffer level. In Salameh's and Ghattas's model the optimal just-in-time buffer level was determined by minimizing the sum of the holding cost per unit of time and the shortage cost per unit of time. The objective of the work of Ribeiro et al. was to jointly optimize the maintenance of the I, the production unit and the buffer size, by developing a suitable mixed integer linear programming model. No stochastic elements were contained in this model. In the model of Charlot et al., it was assumed that there are two different kinds of repairs of the I: repairs requiring prior lockout/tagout and repairs carried out without lockout/tagout. The objective was to control the transition rates between the different modes of the system so as to minimize the expected total discounted cost.

The rest of the paper is organized as follows. The description of the model and the specification of the parameters are given in Section 2. In Section 3, we present the special-purpose policy iteration algorithm. Three numerical examples are presented in Section 4 and the effect of varying the buffer capacity is studied in Section 5. In the last section, the conclusions of this work are given.

2. The model

We consider a deteriorating I, which supplies a raw material to a subsequent production unit (P). A buffer (B) has been built between the production unit and its input generating I to cope with unexpected failures of the I which may cause interruptions in production. The capacity of the buffer is equal to K units. The production unit pulls the raw material from the buffer with a constant demand rate equal to d units per unit of time. As long as the buffer capacity is not reached, the I operates at a constant rate of p units per unit of time ($p > d$) and the excess output is stored in the buffer. When the buffer is full, the I reduces its speed from p to d. Henceforth, we assume that $p-d = 1$. This assumption can be relaxed without problem. The three components of the production system are depicted in Fig. 1.

As mentioned in Van Der Duyn Schouten and Vanneste (1995), 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 platform to storage tanks, from which is further transported to the refineries. In this case the

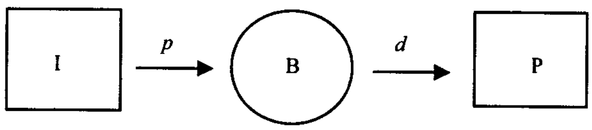

Fig. 1. The three components of the system.

crude oil, the exploration platform, the refineries and the storage tanks are the raw material, the I, the production unit and the buffer, respectively.

We suppose that the I deteriorates as time evolves and it is monitored at discrete, equidistant time epochs $\tau = 0,1,\ldots$ (say every day) and a decision must be made at each epoch. There are three possible actions $a \in \{0,1,2\}$, which are selected at each time epoch. The possible actions are the action of doing nothing ($a = 0$) the action of starting a preventive maintenance of the I ($a = 1$) and the action of starting a corrective maintenance of the I ($a = 2$). A policy is any rule for choosing actions at each decision epoch. A policy is said to be stationary, if at each decision epoch it chooses one action, which depends only on the current state of the process.

The state of the I at each decision epoch is classified into one of the $m+2$ working conditions $0,1,\ldots,m+1$, which represent increasing degrees of deterioration. State 0 denotes a new I (or functioning as good as new). State $m+1$ denotes a failed (inoperative) I. The intermediate states $1,\ldots,m$ are operative. If at a decision epoch the I is found to be at the working condition i, $0 \leqslant i \leqslant m$, and the action of doing nothing is selected, the working condition of the I at the next decision epoch is r, $i \leqslant r \leqslant m+1$ with probability p_{ir}. We further assume that the I can eventually reach the working condition $m+1$ from any working condition i with non-zero probability.

If at a decision epoch the working condition of the I is $i<m+1$, the content of the buffer is $x<K$ and the action of doing nothing is chosen, then the content of the buffer at the next decision epoch will be $\min(K, x+1)$. This increase of the buffer will happen even if the working condition of the I at the next decision epoch is $m+1$.

If at a decision epoch the I is found to be at the working condition $m+1$ then the action of starting a corrective maintenance is compulsory. If at a decision epoch the I is found to be at any state i, $0 \leqslant i \leqslant m$, either the action of doing nothing or the action of starting a preventive maintenance may be selected. We suppose that both actions of preventive and corrective maintenance are non-preemptive, i.e. they cannot be interrupted and bring the I to the working condition 0. It is assumed that the preventive and the corrective repair times are continuous random variables with probability density functions $f_1(x)$ and $f_2(x)$, respectively.

We describe the state of the production system in terms of two variables (i, x), where i denotes the working condition of the I at a decision epoch and x the content of the buffer at that decision epoch. The state space S of the production system is the following set:

$$S = \{(i,x), \quad 0 \leqslant i \leqslant m+1, \quad 0 \leqslant x \leqslant K\}.$$

Note that S is a two-dimensional state space with one discrete state variable, the working condition i of the I, and one continuous state variable, the content of the buffer x. In order to discretize the continuous state variable x the parameter ξ is introduced and the interval $[0,K]$ is divided into K/ξ slices such that the content of the buffer at each decision epoch is represented by the variable $j\xi$, where $j = 0, 1, \ldots, K/\xi$. Thus, the state space of the production system can be rewritten approximately as follows:

$$S = \{(i,j\xi), \quad 0 \leqslant i \leqslant m+1, \quad 0 \leqslant j \leqslant K/\xi\}.$$

The approximation becomes better as the parameter ξ takes very small values (e.g. $\xi = 0.001$).

We suppose that, during any maintenance (preventive or corrective) of the I, the supply of the raw material to the buffer is interrupted. 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. A shortage cost is incurred when a preventive or a corrective maintenance is performed and the buffer is empty. The unit of cost has been chosen in such a way so that the shortage cost is equal to the lost demand d for each unit of time during which a preventive or a corrective repair is performed. We also suppose that the cost of holding a unit of the raw material in the buffer for one unit of time is equal to $h>0$.

If at a decision epoch the I is found to be at the working condition i, $0 \leqslant i \leqslant m$, and the action of doing nothing is selected an operating cost is incurred until the next decision epoch which is equal to c_i, if the buffer is not full, or to $\tilde{c}_i$, if the buffer is full. If the action of preventive or corrective maintenance is selected, a repair cost is incurred which is equal for each unit of time to c_p or to c_f, respectively.

Let m_{PM} and m_{CM} be the expected times required for a preventive maintenance and a corrective maintenance, respectively. The following conditions on the cost structure and the transition probabilities are assumed to be valid:

Condition 1. $c_0 \leqslant c_1 \leqslant \cdots \leqslant c_m$, $\tilde{c}_0 \leqslant \tilde{c}_1 \leqslant \cdots \leqslant \tilde{c}_m$. That is, as the working condition of the I deteriorates, the operating cost increases.

Condition 2. $\tilde{c}_i \leqslant c_i$, $0 \leqslant i \leqslant m$. That is, the reduction of the speed from p (units/time) to d (units/time) of the I, as soon as the buffer is filled up, causes a reduction of its operating cost.

Condition 3. $m_{\mathrm{PM}} < m_{\mathrm{CM}}$. That is, the expected time required for a preventive maintenance is smaller than the expected time required for a corrective maintenance.

Condition 4. $c_p \leqslant c_f$. That is, the cost rate of a preventive maintenance does not exceed the cost rate of a corrective maintenance.

Condition 5. (*An increasing failure rate assumption*). For each $k = 0, 1, \ldots, m+1$, the function

$$D_k(i) = \sum_{r=k}^{m+1} p_{ir}$$

is non-decreasing in i, $0 \leqslant i \leqslant m$.

This condition implies that $I_i \leqslant_{\mathrm{st}} I_{i+1}$, $0 \leqslant i \leqslant m$, where I_i is a random variable representing the next working condition of the I if its present working condition is i and "$\leqslant_{\mathrm{st}}$" means "stochastically smaller than or equal to" (see, for example, Ross (1983, p. 153)).

If at a decision epoch the production system is at a state $(m+1, j\xi)$, $0 \leqslant j\xi \leqslant K$, the action of corrective maintenance ($a = 2$) brings the system either to the state (0, 0) or to one of the states $(0, j'\xi)$, where $1 \leqslant j' \leqslant j$. If at a decision epoch the production system is at a state $(i, j\xi)$, $0 \leqslant i \leqslant m$, $0 \leqslant j\xi \leqslant K$, then either the action of preventive maintenance ($a = 1$) or the action of doing nothing ($a = 0$) may be chosen. If the action $a = 1$ is chosen, the production system makes a transition either to the state (0, 0) or to one of the states $(0, j'\xi)$, where $1 \leqslant j' \leqslant j$. If the action $a = 0$ is chosen and $0 \leqslant j\xi \leqslant K-1$, then the production system makes a transition to one of the states $(r, j\xi+1)$ where $i \leqslant r \leqslant m+1$. If the action $a = 0$ is chosen and $j\xi > K-1$, then the production system makes a transition to one of the states (r, K), $i \leqslant r \leqslant m+1$.

We consider a semi-Markov decision process with state space S in which we aim to find a policy that minimizes the long-run expected average cost per unit time. The relevant theory can be found in Chapter 3 in Tijms (1994). The expected long-run average cost per unit time of a policy π is defined as the limit as $t \to \infty$ of the expected cost incurred in the time interval $[0,t]$ divided by t, given that the policy π is employed. The decision epochs in our problem are all time epochs $\tau = 0,1,\ldots$ at which the system enters a state in S. Note that under any policy the state (0, 0) can be reached from every initial state in S. Hence, since S is finite, it follows (see Ross (1983, p. 98)) that there exists an average-cost optimal stationary policy.

Let $p_{su}(a)$ be the probability that the next state of the system will be $u \in S$, if the present state is $s \in S$ and the action a is chosen and let $\tau_s(a)$ and $c_s(a)$ be the corresponding expected transition time and cost, respectively.

The transition probabilities of the production system for the three possible actions are given in Eqs. (1)–(6) below:

$$p_{(m+1,j\xi)(0,0)}(2) = \int_{j\xi/d}^{\infty} f_2(t)\mathrm{d}t, \quad 0 \leqslant j\xi \leqslant K, \tag{1}$$

$$p_{(m+1,j\xi)(0,j'\xi)}(2) = \int_{(\xi(j-j')-(\xi/2))/d}^{(\xi(j-j')+(\xi/2))/d} f_2(t)\mathrm{d}t, \quad \xi \leqslant j\xi \leqslant K, \quad 1 \leqslant j' \leqslant j. \tag{2}$$

$$p_{(i,j\xi)(0,0)}(1) = \int_{j\xi/d}^{\infty} f_1(t)\,\mathrm{d}t, \quad 0 \leqslant j\xi \leqslant K, \tag{3}$$

$$p_{(i,j\xi)(0,j'\xi)}(1) = \int_{(\xi(j-j')-\xi/2)/d}^{(\xi(j-j')+\xi/2)/d} f_1(t)\mathrm{d}t, \quad \xi \leqslant j\xi \leqslant K, \quad 1 \leqslant j' \leqslant j, \quad 0 \leqslant i \leqslant m, \tag{4}$$

$$p_{(i,j\xi)(r,j\xi+1)}(0) = p_{ir}, \quad i \leqslant r \leqslant m+1, \quad 0 \leqslant j\xi \leqslant K-1, \quad 0 \leqslant i \leqslant m, \tag{5}$$

$$p_{(i,j\xi)(r,K)}(0) = p_{ir}, \quad i \leqslant r \leqslant m+1, \quad K-1+\xi \leqslant j\xi \leqslant K, \quad 0 \leqslant i \leqslant m. \tag{6}$$

The expected transition times for the three possible actions are given in Eqs. (7)–(9) below:

$$\tau_{(m+1,j\xi)}(2) = \int_0^{\infty} tf_2(t)\mathrm{d}t, \quad 0 \leqslant j\xi \leqslant K, \tag{7}$$

$$\tau_{(i,j\xi)}(1) = \int_0^{\infty} tf_1(t)\mathrm{d}t, \quad 0 \leqslant i \leqslant m, \quad 0 \leqslant j\xi \leqslant K, \tag{8}$$

$$\tau_{(i,j\xi)}(0) = 1, \quad 0 \leqslant i \leqslant m, \quad 0 \leqslant j\xi \leqslant K. \tag{9}$$

The expressions for the expected transition costs $c_s(a)$, which correspond to the actions of corrective ($a = 2$) and preventive ($a = 1$) maintenance are more complicated. When the maintenance actions $a \in \{1, 2\}$ are performed on the I, the expected transition costs $c_s(a)$ consist of expected maintenance costs $c_s^{\mathrm{M}}(a)$, expected shortage costs $c_s^{\mathrm{SH}}(a)$ and expected holding (storage) costs $c_s^{\mathrm{H}}(a)$. These quantities are given in Eqs. (10)–(12) below for the case of corrective maintenance $a = 2$. They can be derived by conditioning on the time that the maintenance lasts. By substituting $m+1$ by i, $0 \leqslant i \leqslant m$, c_f by c_p and f_2 by f_1 in Eqs. (10)–(12) we obtain the corresponding expressions for the action of preventive maintenance $a = 1$:

$$c^{\mathrm{M}}_{(m+1,j\xi)}(2) = c_f \int_0^{\infty} t f_2(t) \mathrm{d}t, \quad 0 \leqslant j\xi \leqslant K, \tag{10}$$

$$c^{\mathrm{SH}}_{(m+1,j\xi)}(2) = \int_{j\xi/d}^{\infty} (td - j\xi) f_2(t) \mathrm{d}t, \quad 0 \leqslant j\xi \leqslant K, \tag{11}$$

$$\begin{aligned} c^{\mathrm{H}}_{(m+1,j\xi)}(2) = &\int_0^{j\xi/d} \left[\int_0^t h(j\xi - sd)\,\mathrm{d}s \right] f_2(t)\,\mathrm{d}t \\ &+ \int_{j\xi/d}^{\infty} \left[\int_0^{j\xi/d} h(j\xi - sd)\,\mathrm{d}s \right] f_2(t)\,\mathrm{d}t, \\ &0 \leqslant j\xi \leqslant K. \end{aligned} \tag{12}$$

The expected transition costs $c_s(a)$ for the three possible actions $a \in \{0, 1, 2\}$ are given in Eqs. (13)–(16) below:

$$c_{(m+1,j\xi)}(2) = c^{\mathrm{M}}_{(m+1,j\xi)}(2) + c^{\mathrm{SH}}_{(m+1,j\xi)}(2) + c^{\mathrm{H}}_{(m+1,j\xi)}(2), \quad 0 \leqslant j\xi \leqslant K, \tag{13}$$

$$c_{(i,j\xi)}(1) = c^{\mathrm{M}}_{(i,j\xi)}(1) + c^{\mathrm{SH}}_{(i,j\xi)}(1) + c^{\mathrm{H}}_{(i,j\xi)}(1), \quad 0 \leqslant i \leqslant m, \quad 0 \leqslant j\xi \leqslant K, \tag{14}$$

$$c_{(i,j\xi)}(0) = c_i + hj\xi, \quad 0 \leqslant i \leqslant m, \quad 0 \leqslant j\xi \leqslant K - \xi, \tag{15}$$

$$c_{(i,K)}(0) = \tilde{c}_i + hK, \quad 0 \leqslant i \leqslant m. \tag{16}$$

Note that in expressions (15) and (16), which correspond to the action $a = 0$, the expected cost $c_s(a)$ consists of operating costs and holding costs.

In Section 3, we will present an efficient semi-Markov decision algorithm to compute the optimal stationary policy.

3. The algorithm

Kyriakidis and Dimitrakos (2006) assumed that the maintenance times of the I are geometrically distributed. Under the Conditions 1–5 they have shown that, for fixed buffer level, the policy that minimizes the expected long-run average cost per unit time is of control-limit form. That is, for fixed buffer content x, $0 \leqslant x \leqslant K$, there exists a critical working condition $i^*(x)$ such that the optimal policy initiates the preventive maintenance if and only if the working condition i of the I is equal to or greater than $i^*(x)$.

When the maintenance times of the I are continuous random variables, it seems intuitively reasonable that, if Conditions 1–5 hold, the optimal policy, for any fixed buffer content, is again of control-limit form. A rigorous proof of this conjecture seems to be difficult. The optimal policy can be computed by implementing the value iteration algorithm, the standard policy iteration algorithm and the linear programming algorithm. We refer to Chapter 3 of Tijms's (1994) book for the description of these algorithms and for various applications.

The standard policy-iteration algorithm generates a sequence of strictly improving stationary policies, which converges to the optimal one. Each iteration of the algorithm consists of the value-determination step and the policy-improvement step. At the value-determination step, the average cost and the relative values that correspond to the current stationary policy are computed, by solving a system of linear equations in which the number of unknowns is equal to the number of the elements of the state space. At the policy-improvement step, a strictly better policy is determined. The relative values of the various states play a key role in improving the policy and are defined as follows. If the stationary policy π is employed and r is a state that can be reached from any initial state of the system, then the relative value $w(s)$ that correspond to the state s is defined as the number $C_{sr} - g(\pi)T_{sr}$, where $g(\pi)$ is the average cost of the policy π and T_{sr}, C_{sr} are the expected time and cost, respectively, until the process reaches the state r if the initial state is s. It can be proved that $w(s) - w(l)$ is the difference in total expected costs over an infinite planning horizon by starting in state s rather than in state l when π is used.

It is possible to develop for our problem a computationally tractable special-purpose policy-iteration algorithm, which generates a sequence of improving control-limit policies (i.e. policies that, for some buffer level $j\xi$, $j=0,1,\ldots,K/\xi$, initiate the preventive maintenance if and only if the working condition of the I is equal to or greater than some critical level $i(j\xi)$). The design of the algorithm is based on the embedding technique of Tijms (1994, p. 234), which reduces the system of linear equations for the average cost and the relative values of a particular control-limit policy to a considerably smaller system of linear equations on an embedded set of states. Similar algorithms have been developed in various queueing, inventory, maintenance and pest control models. We refer to the book of Tijms (1994, pp. 234–248) and the papers of Tijms and Van Der Duyn Schouten (1984), Kyriakidis (1993) and Nobel and Tijms (1999). Note that the embedding technique is especially useful in problems in which the state space is infinite since in this case, the standard policy-iteration algorithm is not applicable (see e.g. Nobel and Tijms, 1999).

The description of the algorithm follows. Consider a particular control-limit policy R which is characterized by the critical numbers $i(j\xi)$, $j\in\{0,1,\ldots,K/\xi\}$. We define the set of states E as follows:

$$E=\bigcup_{j=0}^{K/\xi}\{(i,j\xi):\quad 0\leqslant i\leqslant i(j\xi)\}.$$

Note that the set E 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. Let $g(R)$ be the long-run expected average cost per unit time under the policy R and T_s^E, C_s^E be the expected time and the expected cost, respectively, until the first entry in the set E when the initial state of the production system is s. Let also p_{sr}^E be the probability that the first entry state in the set E equals r given that the policy R is employed and the initial state of the production system is s. 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.

According to relation (3.6.1) in Tijms (1994, p. 235) the relative values $w(s)$, $s\in S$, associated with the policy R (see Tijms, 1994, p. 188) satisfy the following system of linear equations:

$$w(s)=C_s^E-g(R)T_s^E+\sum_{r\in E}p_{sr}^E w(r),\quad s\in S, \tag{17}$$

together with the equation

$$w(0,0)=0. \tag{18}$$

The quantities T_s^E, C_s^E and p_{sr}^E, $s\in S$, $r\in E$, can be computed by using simple conditional arguments and by taking into account the transitions of the production system under the policy R. We give below some of these expressions:

$$T_{(i,j\xi)}^E=1+\sum_{r=i(j\xi+1)+1}^{m}p_{ir}\left[\int_0^\infty tf_1(t)\mathrm{d}t\right]+p_{i,m+1}\int_0^\infty tf_2(t)\mathrm{d}t,$$
$$0\leqslant i<i(j\xi),\quad 0\leqslant j\xi\leqslant K-1,$$

$$C_{(i,j\xi)}^E=c_i+hj\xi+\sum_{r=i(j\xi+1)+1}^{m}p_{ir}\left[c_{(i,j\xi+1)}^{\mathrm{M}}(1)+c_{(i,j\xi+1)}^{\mathrm{SH}}(1)+c_{(i,j\xi+1)}^{\mathrm{H}}(1)\right]+p_{i,m+1}\left[c_{(m+1,j\xi+1)}^{\mathrm{M}}(2)+c_{(m+1,j\xi+1)}^{\mathrm{SH}}(2)+c_{(m+1,j\xi+1)}^{\mathrm{H}}(2)\right],$$
$$0\leqslant i<i(j\xi),\quad 0\leqslant j\xi\leqslant K-1,$$

$$C_{(i,j\xi)}^E=c_{(i,j\xi)}^{\mathrm{M}}(1)+c_{(i,j\xi)}^{\mathrm{SH}}(1)+c_{(i,j\xi)}^{\mathrm{H}}(1),$$
$$i(j\xi)\leqslant i\leqslant m,\quad 0\leqslant j\xi\leqslant K,$$

$$p_{(i,j\xi)(r,j\xi+1)}^E=p_{ir},\quad 0\leqslant i<i(j\xi),\quad 0\leqslant r\leqslant i(j\xi+1),$$
$$0\leqslant j\xi\leqslant K-1,$$

$$p_{(i,j\xi)(0,0)}^E=\int_{j\xi/d}^\infty f_1(t)\,\mathrm{d}t,\quad i(j\xi)\leqslant i\leqslant m,\quad 0\leqslant j\xi\leqslant K.$$

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

$$w(s)=C_s^E-g(R)T_s^E+\sum_{r\in E}p_{sr}^E w(r),\quad s\in E, \tag{19}$$

together with the Eq. (18). Once this system of linear equations has been solved we can compute each relative value $w(s)$, $i\in S-E$, from (17) by a single-pass calculation. The so-called policy improvement quantity $Q_R(s;a)$ associated with the policy R will be used in the policy-improvement step of our algorithm. It is defined as follows:

$$Q_R(s;a)=c_s(a)-g(R)\tau_s(a)+\sum_{r\in S}p_{sr}(a)w(r),$$
$$s\in\{(i,j\xi),\ 0\leqslant i\leqslant m,\ 0\leqslant j\leqslant K/\xi\},\ a\in\{0,1\}, \tag{20}$$

where $p_{sr}(a)$ is the probability that the next state of the production system will be r given that the present state is s and the action a is chosen and $c_s(a)$ and $\tau_s(a)$ are the corresponding expected cost and time, respectively. The transition probabilities $p_{sr}(a)$ are given by Eqs. (1)–(6), the expected times $\tau_s(a)$ are given by Eqs. (7)–(9) and the expected costs $c_s(a)$ are given by Eqs. (13)–(16).

Suppose that for some buffer content $j\xi$, $0\leqslant j\xi\leqslant K$, there exists an integer $\tilde{i}(j\xi)$ such that $0\leqslant\tilde{i}(j\xi)<i(j\xi)$ and $Q_R((i,j\xi);1)<w(i,j\xi)$, $\tilde{i}(j\xi)\leqslant i<i(j\xi)$. Then, according to Theorem 3.2.1 on p. 192 in Tijms (1994), the control-limit policy that is characterized by the critical numbers $i(0),\ldots, i(j\xi-\xi), \tilde{i}(j\xi), i(j\xi+\xi),\ldots,i(K)$ achieves smaller average cost than $g(R)$.

Similarly, if for some buffer content $j\xi$, $0\leqslant j\xi\leqslant K$, there exists an integer $\tilde{i}(j\xi)$ such that $i(j\xi)<\tilde{i}(j\xi)\leqslant m+1$ and $Q_R((i,j\xi);0)<w(i,j\xi)$, $i(j\xi)\leqslant i<\tilde{i}(j\xi)$, then, according again to Theorem 3.2.1 in Tijms (1994), the control-limit policy that is characterized by the critical numbers $i(0),\ldots, i(j\xi-\xi)$, $\tilde{i}(j\xi)$, $i(j\xi+\xi)$, …, $i(K)$ achieves smaller average cost than $g(R)$.

The above remarks lead us to develop the following special-purpose policy iteration algorithm which generates a sequence of strictly improving control-limit policies.

Special-purpose algorithm:

Step 1 (*initialization*): Choose an initial control-limit policy R characterized by the critical numbers $i(j\xi), j=0,1,\ldots,K/\xi$.

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 (18) and (19).

Step 3 (*policy improvement step*): For each buffer content $j\xi, j=0,1,\ldots,K/\xi$:

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

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

The quantities $Q_R((i,j\xi);1)$ and $Q_R((i,j\xi);0)$ are given by (20), where, if it is necessary, $w(s)$, $s\in S-E$, can be computed from Eq. (17).

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

Step 4 (*convergence test*): If it is not possible to find any $\tilde{i}(j\xi)$, $0\leqslant j\xi\leqslant K$, the algorithm is stopped. The final policy is the policy 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 algorithm converges to the optimal policy, since, in all examples that we have tested, the final policy obtained by the algorithm coincides with the final policy obtained by the standard policy iteration algorithm. 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 elements of the entire state space S. From a great number of examples that we have tested, 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.

4. Numerical examples

As illustrations of the algorithm we present three examples. In Examples 1–3 below, we assume that the maintenance times of the I follow the Exponential, the Gamma and the Weibull distribution, respectively.

Example 1. Suppose that $m=50$, $K=30$, $p=16$, $d=15$, $c_f=0.8$, $c_p=0.4$, $c_i=0.1(i+1)$ and $\tilde{c}_i=0.05(i+1)$, $0\leqslant i\leqslant m$. We assume that the non-zero transition probabilities p_{ir}, $0\leqslant i$, $r\leqslant m+1$ are given by $p_{ir}=(m+2-i)^{-1}$, $i\leqslant r\leqslant m+1$. This means that if the present state of the I is i then the next state is uniformly distributed in the set $\{i, i+1,\ldots,m+1\}$. These probabilities satisfy Condition 5 since, for each $k=0,\ldots,m+1$, the quantity

$$\sum_{r=k}^{m+1}p_{ir}=\frac{m+2-\max(k,i)}{m+2-i}$$

is increasing in i, $0\leqslant i\leqslant 50$. We assume that the preventive repair time and the corrective repair time of the I are exponentially distributed with mean 0.125 and 0.25, respectively. We choose $\xi=0.05$ so that the interval $[0,K]$ is divided into $K/\xi=600$ slices. As initial control-limit policy we choose the policy that is characterized by the critical numbers $i(j\xi)=50$, $0\leqslant j\leqslant 600$.

In Table 1, we present the critical numbers $i^*(j\xi)$, $0 \leqslant j \leqslant 600$, of the final control-limit policy obtained by our special-purpose algorithm for various values of the holding cost h. This policy is the optimal one since it coincides with the final policy obtained by the standard policy iteration algorithm.

From Table 1, it can be seen that for fixed buffer level $j\xi$, the critical number $i^*(j\xi)$ decreases as h increases.

In Table 2, we present the average cost of the final control-limit policy obtained by the algorithm, the number of iterations required by the special-purpose algorithm and the CPU time (in s) of the corresponding Matlab program that we run on a PC Acer Aspire 1605DLC for $h \in \{0.2, 0.8, 1.4, 2\}$. In parentheses we present the CPU time required by the standard policy iteration algorithm.

Note that in all cases the CPU time required by the special-purpose policy iteration algorithm is considerably smaller than the CPU time required by the standard policy iteration algorithm. It can also be seen that, as expected, the average cost of the final policy increases as h increases.

Example 2. Suppose that $m = 30$, $K = 15$, $p = 11$, $d = 10$, $c_f = 1.2$, $c_p = 0.7$ and $h = 0.4$. We also suppose that the operating costs c_i, $\tilde{c}_i$, $0 \leqslant i \leqslant m$, and the non-zero transition probabilities p_{ir}, $0 \leqslant i, r \leqslant m+1$, are the same as in Example 1. We assume that the preventive and the corrective repair times follow the Gamma distribution with parameters $a_1 > 0$, $b_1 > 0$ and $a_2 > 0$, $b_2 > 0$, respectively. Their probability density functions are given by

$$f_1(t) = \frac{1}{\Gamma(a_1) b_1^{a_1}} t^{a_1 - 1} \exp(-t b_1^{-1}) \text{ and}$$

$$f_2(t) = \frac{1}{\Gamma(a_2) b_2^{a_2}} t^{a_2 - 1} \exp(-t b_2^{-1}), \quad t \geqslant 0,$$

respectively, where, $\Gamma(a) = \int_0^\infty e^{-t} t^{a-1}\, dt$, is the Gamma function. We assume that $a_1 = 3$ and $a_2 = 2$, $b_2 = 8$. We choose $\xi = 0.05$ so that the interval $[0, K]$ is divided into $K/\xi = 300$ slices. As initial control-limit policy we choose the policy that is characterized by the critical numbers $i(j\xi) = 30$, $0 \leqslant j \leqslant 300$.

In Table 3, we present the critical numbers $i^*(j\xi)$, $0 \leqslant j \leqslant 300$, of the final control-limit policy obtained by our algorithm for $b_1 \in \{3,4,5\}$. The final policy is also the optimal one since it coincides with the final policy obtained by the standard policy-iteration algorithm. Note that Condition 3 is satisfied for the above values of b_1 since the mean of the Gamma distribution with parameters $a > 0$ and $b > 0$ is ab. By increasing the parameter b_1 of the Gamma distribution we increase the expected time required for a preventive maintenance.

From the Table 3, it can be seen that for fixed buffer level $j\xi$, the critical number $j^*(j\xi)$ increases as the expected time of the preventive maintenance increases.

In Table 4, we present the average cost of the final control-limit policy obtained by the algorithm, the number of iterations required by our special-purpose algorithm and the CPU time (in s) of the corresponding Matlab program for $b_1 \in \{3,4,5\}$. In parentheses we present the CPU time required by the standard policy iteration algorithm.

Table 1
The final policy as h varies

h	Critical numbers $i^*(j\xi)$, $0 \leqslant j \leqslant 600$: j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$
0.2	0	18	14–17	13	36–40	8	61–66	3
	1	17	18–22	12	41–44	7	67–73	2
	2–5	16	23–26	11	45–49	6	74–82	1
	6–9	15	27–30	10	50–55	5	83–600	0
	10–13	14	31–35	9	56–60	4		
0.8	0	18	6–7	13	13–14	8	22–23	3
	1	17	8	12	15–16	7	24–26	2
	2	16	9	11	17–18	6	27–28	1
	3–4	15	10–11	10	19	5	29–600	0
	5	14	12	9	20–21	4		
1.4	0	18	5	12	10	7	16–17	2
	1	17	6	11	11–12	6	18	1
	2	15	7	10	13	5	19–600	0
	3	14	8	9	14	4		
	4	13	9	8	15	3		
2.0	0	17	5	11	10	5		
	1	16	6	9	11	3		
	2	15	7	8	12	2		
	3	13	8	7	13	1		
	4	12	9	6	14–600	0		

Table 2
The effect of varying h

Holding cost h	Average cost	Number of iterations	CPU times
0.2	0.9621	4	208.6 (520.7)
0.8	1.3053	5	240.4 (538.1)
1.4	1.4734	4	205.2 (493.1)
2.0	1.5714	4	220.3 (528.7)

Note that in all cases the CPU time required by the special-purpose policy iteration algorithm is considerably smaller than the CPU time required by the standard policy iteration algorithm. It can also be seen, as expected, that the average cost of the final policy increases as the expected time for the preventive maintenance increases.

Example 3. Suppose that $m = 20$, $K = 10$, $p = 9$, $d = 8$, $c_f = 2.5$ and $h = 0.3$. We also assume that the operating costs c_i, $\tilde{c}_i$, $0 \leqslant i \leqslant m$, and the non-zero transition probabilities p_{ir}, $0 \leqslant i$, $r \leqslant m+1$, are the same as in the Examples 1 and 2 above. We assume that the preventive and the corrective repair times of the I follow the Weibull distribution with parameters $\alpha_1 > 0$, $\lambda_1 > 0$ and $\alpha_2 > 0$, $\lambda_2 > 0$, respectively. Their probability density functions are given by

$$f_1(t) = \alpha_1 \lambda_1 (\lambda_1 t)^{\alpha_1 - 1} \exp[-(\lambda_1 t)^{\alpha_1}] \text{ and}$$
$$f_2(t) = \alpha_2 \lambda_2 (\lambda_2 t)^{\alpha_2 - 1} \exp[-(\lambda_2 t)^{\alpha_2}], \quad t \geqslant 0,$$

respectively. We assume that $\alpha_1 = 1$, $\lambda_1 = 3$ and $\alpha_2 = 0.5$, $\lambda_2 = 5$. Note that Condition 3 is satisfied since the mean of the Weibull distribution with parameters $\alpha > 0$ and $\lambda > 0$ is $\lambda^{-1}\Gamma(1+\alpha^{-1})$, where, $\Gamma(x)$ is the Gamma function. We again choose $\xi = 0.05$ so that the interval [0, K] is divided into $K/\xi = 200$ slices. As initial control-limit policy we choose the policy that is characterized by the critical numbers $i(j\xi) = 20$, $0 \leqslant j \leqslant 200$.

Table 3
The final policy as b_1 varies

Critical numbers $i^*(j\xi)$, $0 \leqslant j \leqslant 300$								
b_1	j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$
3	0–4	27	36–40	20	66–69	13	94–97	6
	5–11	26	41–44	19	70–73	12	98–101	5
	12–16	25	45–48	18	74–77	11	102–105	4
	17–21	24	49–52	17	78–81	10	106–109	3
	22–26	23	53–56	16	82–85	9	110–114	2
	27–31	22	57–61	15	86–89	8	115–119	1
	32–35	21	62–65	14	90–93	7	120–300	0
4	0–5	30	50–54	22	84–87	14	117–120	6
	6–15	29	55–58	21	88–91	13	121–124	5
	16–23	28	59–62	20	92–95	12	125–128	4
	24–29	27	63–66	19	96–99	11	129–133	3
	30–35	26	67–71	18	100–103	10	134–137	2
	36–40	25	72–75	17	104–108	9	138–142	1
	41–45	24	76–79	16	109–112	8	143–300	0
	46–49	23	80–83	15	113–116	7		
5	0–34	30	69–72	22	102–105	14	134–137	6
	35–40	29	73–76	21	106–109	13	138–142	5
	41–45	28	77–80	20	110–113	12	143–146	4
	46–50	27	81–84	19	114–117	11	147–150	3
	51–55	26	85–88	18	118–121	10	151–155	2
	56–59	25	89–92	17	122–125	9	156–159	1
	60–64	24	93–96	16	126–129	8	160–300	0
	65–68	23	97–101	15	130–133	7		

Table 4
The effect of varying b_1

b_1	Average cost	Number of iterations	CPU time
3	2.6347	4	139.2 (386.9)
4	3.0621	5	152.4 (420.8)
5	3.3903	4	135.1 (380.6)

In Table 5, we present the critical numbers $i^*(j\xi)$, $0 \leqslant j \leqslant 200$, of the final control-limit policy obtained by the algorithm for different values of the cost rate c_p of a preventive maintenance. The final policy is also the optimal one since it coincides with the final policy obtained by the standard policy iteration algorithm.

From the Table 5, it can be seen that for fixed buffer level $j\xi$, the critical number $i^*(j\xi)$ increases as c_p increases. In Table 6, we present the average cost of the final control-limit policy obtained by the algorithm, the number of iterations required by the special-purpose algorithm and the CPU time (in s) of the corresponding Matlab program for $c_p \in \{0.8, 1.5, 2, 2.5\}$. In parentheses, we present the CPU time required by the standard policy iteration algorithm.

Note that in all cases the CPU time required by the special-purpose policy iteration algorithm is considerably smaller than the CPU time required by the standard policy iteration algorithm. It can also be seen, as expected, that the average cost of the final policy increases as c_p increases.

Remark. There is strong evidence from the above three examples and from all the examples that we have tested that the critical number $i^*(j\xi)$ that corresponds to the optimal policy is non-increasing in $j \in \{0, 1, \ldots, K/\xi\}$.

5. The effect of varying the buffer size

In Sections 2–4, we assumed that the buffer capacity is fixed. However, it would be interesting to examine the effect of the variation of K on the average cost of the optimal policy. The result of this

Table 5
The final policy as c_p varies

Critical numbers $i^*(j\xi)$, $0\leqslant j\leqslant 200$								
c_p	j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$	j	$i^*(j\xi)$
0.8	0–2	18	19–22	13	38–41	8	59–63	3
	3–6	17	23–26	12	42–45	7	64–68	2
	7–10	16	27–29	11	46–50	6	69–74	1
	11–14	15	30–33	10	51–54	5	75–200	0
	15–18	14	34–37	9	55–58	4		
1.5	0–2	19	19–22	14	37–41	9	58–62	4
	3–6	18	23–25	13	42–45	8	63–66	3
	7–10	17	16–29	12	46–49	7	67–72	2
	11–14	16	30–33	11	50–53	6	73–78	1
	15–18	15	34–37	10	54–57	5	79–200	0
2.0	0	20	21–24	14	44–47	8	70–74	2
	1–5	19	25–28	13	48–51	7	75–80	1
	6–9	18	29–32	12	52–55	6	81–200	0
	10–13	17	33–35	11	56–59	5		
	14–17	16	36–39	10	60–64	4		
	18–20	15	40–43	9	65–69	3		
2.5	0–3	20	24–27	14	46–49	8	72–76	2
	4–7	19	28–30	13	50–53	7	77–82	1
	8–12	18	31–34	12	54–57	6	83–200	0
	13–15	17	35–38	11	58–62	5		
	16–19	16	39–42	10	63–66	4		
	20–23	15	43–45	9	67–71	3		

Table 6
The effect of varying c_p

c_p	Average cost	Number of iterations	CPU times
0.8	1.3923	4	134.3 (288.9)
1.5	1.5125	4	149.2 (313.1)
2.0	1.5967	4	142.6 (325.5)
2.5	1.6794	4	152.7 (340.4)

Table 7
The effect of varying K

K	$g(K)$	K	$g(K)$
1	8.6340	11	8.6318
2	8.6338	12	8.6315
3	8.6335	13	8.6320
4	8.6333	14	8.6325
5	8.6331	15	8.6328
6	8.6329	16	8.6332
7	8.6327	17	8.6337
8	8.6324	18	8.6344
9	8.6321	19	8.6351
10	8.6319	20	8.6356

investigation could be useful for the determination of the optimal buffer capacity in the design phase of the system.

For example, let $m = 20$, $p = 16$, $d = 15$, $c_f = 30$, $c_p = 10$ and $h = 100$. The operating costs c_i, $\tilde{c}_i$, $0\leqslant i\leqslant m$ and the non-zero transition probabilities p_{ir}, $0\leqslant i$, $r\leqslant m+1$ are the same as in Examples 1–3 of the previous section. We assume that the preventive repair time and the corrective repair time of the I are exponentially distributed with means 0.25 and 1, respectively. We choose $\xi = 0.05$ so that the interval [0, K] is divided into K/ξ slices. In Table 7, we give the minimum average cost $g(K)$, for various values of K.

From the Table 7, we see that $g(K)$ is decreasing with respect to K when $1\leqslant K\leqslant 12$, and $g(K)$ is increasing with respect to K when $K\geqslant 12$. Hence, it is deduced that the optimal buffer capacity is 12.

6. Conclusions

A great number of replacement/maintenance models have been introduced and analyzed in the literature. The optimal maintenance of a system can result in substantial saving in operation and, also, in increased availability of the system. In this article, we consider a production system in which a buffer has been built between a production unit and its input generating I. The purpose of the buffer is to avoid frequent interruptions of the production process due to failures of the I. A semi-Markov decision model is constructed for the optimal preventive maintenance of the I. It is assumed that the repair times of the I are continuous random variables with known distributions. An intuitively appealing class of policies consists of those policies that, for fixed buffer level, initiate the preventive maintenance of the I if and only if its degree of deterioration exceeds some critical level. An efficient algorithm that generates a sequence of this kind of policies is developed. There is strong numerical evidence that the final policy obtained by the algorithm is the average-cost optimal policy.

In the present work, we assumed that the production unit operates without risk provided that it pulls the raw material from the buffer at a constant rate of d units/time. The construction and analysis of a more general model in which the production unit could fail might be a subject for future research. In this case, the main problem would be the determination of the policy that optimizes jointly the maintenance of the I and the production unit.

Acknowledgments

We would like to thank two anonymous referees for their helpful comments on the previous version of this paper.

References

Charlot, E., Kenne, J.P., Nadeau, S., 2007. Optimal production, maintenance and lockout/tagout control policies in manufacturing systems. International Journal of Production Economics, in press, doi:10.1016/j.ijpe.2007.03.010.

Chen, M., Feldman, R.M., 1997. Optimal replacement policies with minimal repair and age-dependent costs. European Journal of Operational Research 98, 75–84.

Douer, N., Yechiali, U., 1994. Optimal repair and replacement in Markovian systems. Stochastic Models 10, 253–270.

Federgruen, A., So, K.C., 1989. Optimal time to repair a broken server. Advances in Applied Probability 21, 376–397.

Kyriakidis, E.G., 1993. A Markov decision algorithm for optimal pest control through uniform catastrophes. European Journal of Operational Research 64, 38–44.

Kyriakidis, E.G., Dimitrakos, T.D., 2006. Optimal preventive maintenance of a production system with an intermediate buffer. European Journal of Operational Research 168, 86–99.

Love, C.E., Zhang, Z.G., Zitron, M.A., Guo, R., 2000. A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs. European Journal of Operational Research 125, 398–409.

Meller, R.D., Kim, D.S., 1996. The impact of preventive maintenance on system cost and buffer size. European Journal of Operational Research 95, 577–591.

Nobel, R.D., Tijms, H.C., 1999. Optimal control for an $M^X/G/1$ queue with two services modes. European Journal of Operational Research 113, 610–619.

Ribeiro, M.A., Silvera, J.L., Qassim, R.Y., 2007. Joint optimisation of maintenance and buffer size in a manufacturing system. European Journal of Operational Research 176, 405–413.

Ross, S.M., 1983. Introduction to Stochastic Dynamic Programming. Academic Press, New York.

Salameh, M.K., Ghattas, R.E., 2001. Optimal just-in-time buffer inventory for regular preventive maintenance. International Journal of Production Economics 74, 157–161.

Scarf, P.A., 1997. On the application of mathematical models in maintenance. European Journal of Operational Research 99, 493–506.

Tijms, H.C., 1994. Stochastic Models: An Algorithmic Approach. Wiley, Chichester.

Tijms, H.C., Van Der Duyn Schouten, F.A., 1984. A Markov decision algorithm for optimal inspections and revisions in a maintenance system with partial information. European Journal of Operational Research 21, 245–253.

Van Der Duyn Schouten, F.A., Vanneste, S.G., 1995. Maintenance optimization of a production system with buffer capacity. European Journal of Operational Research 82, 323–338.

Wang, H., 2002. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 139, 469–489.