IIE Transactions (2002) **34**, 423–435

Integrated maintenance and production control of a deteriorating production system

SEYED M.R. IRAVANI[1] and IZAK DUENYAS[2]*

[1]*Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA*
E-mail: iravani@iems.nwu.edu
[2]*School of Business, University of Michigan, Ann Arbor, MI 48109-1234, USA*
E-mail: duenyas@umich.edu

Received February 2000 and accepted April 2001

We consider a make-to-stock production/inventory system consisting of a single deteriorating machine which produces a single item. We formulate the integrated decisions of maintenance and production using a Markov Decision Process. The optimal dynamic policy is shown to have a rather complex structure which leads us to consider more implementable policies. We present a double-threshold policy and derive exact and approximate methods for evaluating the performance of this policy and computing its optimal parameters. A detailed numerical study demonstrates that the proposed policy and our approximate method for computing its parameters perform extremely well. Finally, we show that policies which do not address maintenance and production control decisions in an integrated manner can perform rather badly.

1. Introduction

Complex and high-tech machinery in advanced production systems constitute a large majority of most industries capital. These production systems are more reliable than their predecessors; however, they are still subject to deterioration with usage and age. The deterioration causes lower production rates (therefore higher production cost per item) and lower product quality. Preventive maintenance is one of the tools to increase the reliability of the production system. Without an effective maintenance program, the production system fails more often, and depending on the magnitude of repair times, the system might be down for significant amounts of time. This means that the effective production rate decreases significantly and the system might not be able to cope with demand. One way of dealing with this scenario is to keep enough inventory in order to satisfy demand during the time that the production facility is down. But, as always, the main question is "how much inventory is enough?" Clearly, how much inventory one should keep should depend on the deterioration rate of the machine as well as the particular maintenance policy used. On the other hand, the maintenance policy to be used should take into account the fact that some inventory can be kept to protect against downtimes. Therefore, there is an intimate relationship between the maintenance/repair and production/inventory policies used in a facility.

This paper deals with the problem of joint maintenance/repair and production/inventory policy in a multiple-state make-to-stock system where the machine deteriorates with usage. Such systems are very common in practice. For example, Berk and Moinzadeh (2000) give examples of such systems in tooling, and semiconductor industries.

We assume that the produced items are held in a finished goods inventory and consumed by exogenous demand. Demand that cannot be met from the finished goods inventory is lost and the system incurs lost sales penalties. The machine has several operational states $(1, 2, \ldots, I-1)$ and one failed state I. The system is assumed to be deteriorating as random shocks take the system to worse states and the production rate is nonincreasing in system state. Performing repair or maintenance operation in state i takes a random time and costs m_i per unit time and changes the state of the system to operational state 1 (as good as new). The system incurs holding costs for each unit held in the finished goods inventory. The objective is to find the best joint production/inventory and repair/maintenance policy in order to minimize the total average cost per unit time.

Most of the existing literature on maintenance policies does not consider the interactions between production/

*Corresponding author

inventory and repair/maintenance decisions. Comprehensive reviews and analysis in this area can be found in Pierskalla and Voelker (1976), Sherif and Smith (1981), McCall (1985), Valdez-Flores and Feldman (1989) and Dekker (1996). These interactions have also received little attention in the production and inventory control literature which typically assumes perfectly reliable machines. Recent literature which attempts to fill this gap can be classified into two groups: (i) literature on the effects of machine failures on production and inventory decisions; and (ii) literature which focuses on developing new integrated production/maintenance policies.

The first group of papers do not attempt to develop integrated production/maintenance policies but rather focus on how failures or a fixed maintenance policy would affect well-known production and inventory policies. Maintenance costs are ignored as the maintenance decisions are assumed to be fixed. For example, Groenevelt *et al.* (1992a, 1992b) focus on the effects of machine breakdowns and corrective maintenance on economic lot sizing decisions. Gallego (1990, 1994) extends the classical Economic Lot Scheduling Problem (ELSP) by providing an algorithm for scheduling the facility after disruptions. Sharafali (1984) shows the effects of a machine failure on the performance measures of a single-machine, single-product production system in which the machine output replenishes the items according to an (s, S) policy. As in the classical (s, S) policy, the machine is shut off as soon as the inventory level reaches S and is put back into operation when the inventory goes below s. Repair starts immediately as soon as the machine fails. Sharifnia (1988) considers a multiple-state machine which produces a single item to satisfy a constant demand rate. The machine changes its state according to a continuous time Markov chain. The objective is to find the optimal production rate with respect to the machine state and inventory level in order to minimize the total average inventory cost. Other examples of the papers which focus on the effects of specific maintenance policies on inventory policies include Meyer *et al.* (1979), Posner and Berg (1989) and Berg *et al.* (1994).

The second group of papers include the repair/maintenance costs in their analysis and introduce policies which integrate the optimal production/inventory and repair/maintenance policies. Lee and Rosenblatt (1987) added the maintenance by inspection feature to the economic lot sizing problem, where the inspections help to determine whether the equipment is in-control or out-of-control. If the equipment is out-of-control, a maintenance operation is required to restore it to an in-control state. The decision variables are the production lot size and the number of inspections per cycle, and the objective is to minimize the total average inventory and inspection/maintenance costs. Srinivasan and Lee (1996) added a preventive maintenance option to an (s, S) policy similar to Sharafali (1984). In their model when the inventory level reaches S, preventive maintenance is undertaken and the machine becomes as good as new. If the system fails before the preventive maintenance is scheduled, the repair process starts. Demand is assumed to be Poisson and the costs involved are preventive maintenance and repair costs and also back order, holding and production setup costs. They obtained the optimal parameters s and S in order to minimize the total average cost. In their model, the production policy depends on the maintenance costs; however, the maintenance policy is fixed and does not depend on inventory/production costs.

Das and Sarkar (1999) considered a similar model to that of Srinivasan and Lee (1996); however, in their model the decision to perform a preventive maintenance depends on the inventory level, as well as the number of items produced since the last repair/maintenance operation. In both models, the production/inventory policy follows the (s, S) policy and the facility idles when the inventory reaches S. Both Srinivasan and Lee (1996) and Das and Sarkar (1999) consider a single (operating) state production facility in which the production rate does not change with usage and the repair/maintenance cost is independent of the facility's age. Our model is different in the sense that we study a multiple (operating) state production system where the production and repair/maintenance characteristics of the system change with usage. Furthermore, in our model the production/inventory policy is not fixed. In fact, we investigate how the structure of the integrated production/maintenance policy changes as the system enters different operating states.

The paper closest to ours is by Van der Duyn Schouten and Vanneste (1995) who considered a single production facility with an increasing failure rate lifetime distribution. In their model the facility aging process does not affect the production rate and the facility produces items either at constant rate p, if the downstream buffer is not full, or at constant rate d (equal to the demand rate), if the downstream buffer is full. The buffer has finite capacity K (exogenously given) and satisfies constant demand rate d. Upon failure the facility goes under repair and becomes as good as new. The option of preventive maintenance exists which takes less time than repair and also puts the facility back into as good as new condition. The objective is to decide when to perform preventive maintenance. This decision is made based on the age of the facility and the downstream buffer level. The criteria is to minimize the total inventory-related measures such as average inventory level, average number of lost sales or backorders. Van der Duyn Schouten and Vanneste introduce a suboptimal policy which prescribes preventive maintenance actions either in age (state) n or N. If the buffer is full, preventive maintenance is undertaken at age n. If the buffer is not full, but has at least k items, preventive maintenance is undertaken at age N. Maintenance is never performed unless the system has at least k items. They develop analytical models to obtain the best values for n, N and k and show that their proposed policy performs well.

In our paper, we also assume a production facility with an increasing failure rate and our decisions are based on the state of the machine and the inventory on hand. However, we assume a stochastic demand and production process. Also in our model the deterioration process (which is a function of machine usage) does affect the performance of the production facility which in turn influences the production/maintenance decisions. More specifically, we consider a multiple-state production facility which produces a single item and deteriorates with usage. The deterioration process affects the production capacity and repair and maintenance operations. In other words, in our model, as the facility deteriorates, its production rate decreases and its maintenance operation becomes more time consuming and costly. Produced items are kept in the Finished Goods (FG) inventory and consumed by exogenous stochastic demand. The cost structure consists of the inventory holding and lost sales cost and also repair/maintenance costs. We look at the problem of finding the best joint production and maintenance policies in order to minimize the total average holding, lost sales and maintenance/repair costs.

The remainder of this paper is organized as follows. We start with a dynamic programming formulation of the problem in Section 2 and we investigate the structure of the optimal policy through some numerical examples. As we show in Section 2, the optimal policy is extremely complex and impractical. This leads us to consider a simpler double-threshold policy which we introduce in Section 3. We develop an exact model for our double-threshold policy; however this exact model requires solving large systems of equations and does not have closed-form. The special case of systems with three states can however be solved easily and its solution is provided in Section 4. We also use the solution for systems with three states in a simple heuristic we develop for systems with any number of states in Section 5. Finally, Section 6 provides a comprehensive numerical study showing that: (i) the proposed double-threshold policy performs very close to the optimal policy; and (ii) the heuristic we propose for the double-threshold policy based on aggregating N states into three states performs very well. The paper concludes in Section 7.

2. Problem formulation

We consider a single-machine make-to-stock manufacturing system producing a single product. Finished items are stored in finished goods inventory at a cost of h per item per unit time. Demand for these items arrives according to a Poisson process with rate λ. The demand that cannot be met from the finished goods inventory is lost and a penalty (lost sales cost) of C per item is incurred.

The machine has several operational states $(1, 2, \ldots, I-1)$ and failed state I. At any operational state i, the machine processing time is a random variable with an exponential distribution having mean processing time $1/\mu_i$. (It is possible to extend the analysis to more complicated situations such as Erlang distributions. However, our focus is on the types of joint maintenance/inventory policies which work well and to gain insight into the interactions between the two decisions. Non-exponential distributions would add tremendous additional complexity with fewer new insights.) It is reasonable to assume that μ_i is non-increasing in i. This can be due to the machine producing a larger ratio of defective units as it deteriorates so that the time between production of two successive good units increases. Maintenance (repair) in state i takes an exponentially distributed amount of time with rate r_i and costs m_i per unit time. The objective is to minimize the total average inventory and maintenance costs by finding the optimal production and maintenance policy. In other words, at any operational state, based on the inventory available, the optimal policy determines whether the machine should produce one more item, stay idle or be maintained. Assuming that pre-emptions are allowed, the problem of determining the optimal policy can be formulated as a semi-Markov decision process:

- The *decision epochs* are: (i) production completion epochs; (ii) demand arrival epochs; (iii) repair completion epochs; and (iv) the epochs when machine changes state.
- The *state* of the system at any decision epoch is presented by a vector (n, i), where $n \in \mathcal{Z}^+$ is the number of items in finished goods inventory, and i is the state of the machine $(i = 1, 2, \ldots, I)$.
- The *actions* are: (i) producing an item; (ii) idling; or (iii) maintenance (repair if failed).

Following Lippman (1975), and defining the indicator function I_x such that

$$I_x = \begin{cases} 1 & \text{when } x = 0, \\ 0 & \text{otherwise,} \end{cases}$$

the optimality equation for the semi-Markov decision process with the objective of minimizing the total average holding, penalty and repair cost is:

$$\begin{aligned} \frac{g}{\Lambda} + V(n,i) = \frac{1}{\Lambda}[\{ & nh + \lambda[V(n-1,i)(1-I_n) \\ & + (V(n,i)+C)I_n] \\ & + \min\{\tau V(n,i),\ \mu_i V(n+1,i) \\ & + \phi_i V(n,i+1) + (\tau - \phi_i - \mu_i)V(n,i), m_i \\ & + r_i V(n,1) + (\tau - r_i)V(n,i)\}], \end{aligned} \tag{1}$$

where $\tau = \text{Max}_i\{r_i, \phi_i + \mu_i\}$, $\Lambda = \lambda + \tau$ and $V(n,i)$ is the relative value of being in state (n,i).

In Equation (1), the first term is the holding cost until the next decision epoch. The terms with λ denote the states after the arrival of demand. If demand arrives and there is at least one unit of inventory in stock (then $I_n = 0$ since $n > 0$), the demand is satisfied and inventory decreases by one. Otherwise, the demand is lost and a penalty of C is incurred. The final expression in Equation (1) denotes the choice between idling, producing or repairing the machine.

We note that the above formulation can be easily extended to the case where machines are replaced instead of maintained. Typically, machine replacement involves a fixed cost (capital expenditure) and a variable cost which is a function of how long the installation of the new equipment takes. Therefore, the only thing that would change in the above formulation would be the addition of a fixed cost term whenever a decision is made to replace the machine with a new one.

In order to investigate the structure of the optimal policy in (1), we solved numerous examples using value iteration and found that the most common structure for the optimal policy is similar to those shown in Fig. 1 (a and b). Figure 1(a and b) shows the optimal solution for a 10-state machine problem with $\lambda = 8$, $\mu_i = 10, \phi_i = 1$; $i = 1, 2, \ldots, 10$, $h = 1$, $C = 50$, and $m_i = 20$; $i = 1, 2, \ldots, 10$, where Fig. 1(a) is for $r_i = 5$; $i = 1, 2, \ldots, 10$ and Fig. 1(b) is for $r = [10\ 10\ 7\ 7\ 5\ 5\ 2\ 2\ 1\ 1]$.

As Fig. 1(a and b) shows, the optimal policy divides the state of the system into three different sets *PI*, *PR* and *R*. In each state in *PI*, the machine continues to produce until the inventory reaches a certain level (which depends on the state), whereafter the machine becomes idle. However, in each state in *PR*, the machine goes under maintenance when inventory reaches a state-dependent threshold. Finally, in states in *R*, no production is ever undertaken, and repair operations start immediately. In Fig. 1(a), $PI = \{1, 2, 3, 4\}$, $PR = \{5, 6, 7, 8, 9\}$, $R = \{10\}$. Note the complexity of the policies in Fig. 1(a and b). Firstly, the production threshold is different for every state. Second, as Fig. 1(b) clearly demonstrates, the switching curves do not even have to be monotonic. The complexity of the optimal policies make their practical implementation unlikely.

In order to get to a simpler and more applicable policy, we introduce a double-threshold policy as shown in Fig. 2. Our double-threshold policy also consists of three sets of states with similar properties to those in Fig. 1(a and b). However, our proposed double-threshold policy has only one production threshold M for states in *PI* where the machine becomes idle when the inventory reaches M. Similarly, there is a single-threshold N for states in *PR* where maintenance is undertaken when inventory reaches N. Thus, implementation of this policy requires stating the values of M and N as well as the two state thresholds that differentiate states *PI*, *PR* and *R*.

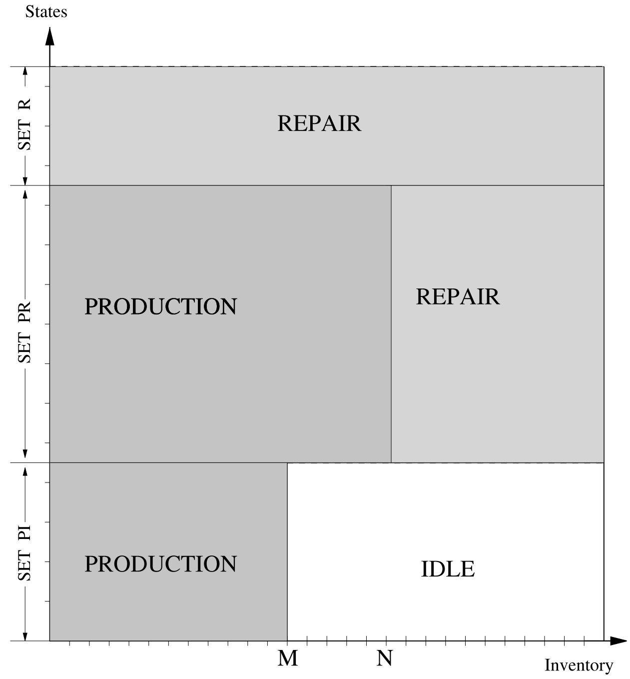

Fig. 2. Double-threshold policy.

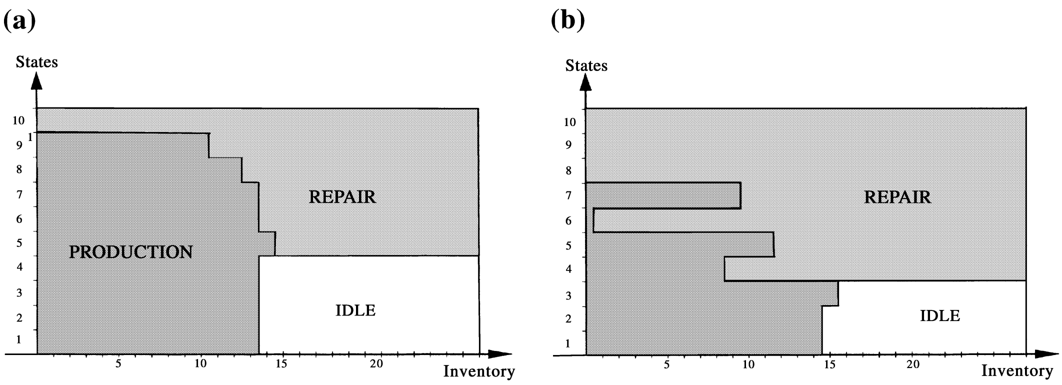

Fig. 1. Examples of the optimal policy when: (a) $r_i = 5$, $i = 1, 2 \ldots, 10$; (b) $r = [10\ 7\ 7\ 5\ 5\ 2\ 2\ 1\ 1]$.

Note that this policy is more practical than the optimal policy for the following reasons: (i) the existence of a single production threshold for all of the states in PI and PR makes the production policy simpler to implement; and (ii) the fact that the production threshold remains the same for a large group of states also makes it easier to keep track of the "state" of the machine. For example, in the case of cutting tools, the tools potentially have a very large number of states. However, the implementation of this policy requires the worker only to recognize two critical states, a much easier task than keeping track of all tool states and measuring the cutting tool all the time. (Ivy and Pollock (1999) focus on the problem of recognition of states based on machine monitoring in cases where machine monitoring may not be perfect. Although this is beyond the focus of this paper, we would like to note that when monitoring is imperfect, recognizing whether the system has deteriorated beyond two given states is much easier than recognizing the state of the system at all times.) We therefore next focus on the double-threshold policy and develop an exact and heuristic analysis of system performance under this policy.

3. An exact analysis of the systems under the double-threshold policy

This section presents an exact performance evaluation of a system which uses a double-threshold policy with parameters M and N. We let

$$PI = \{1, 2, \ldots, K\}, \quad PR = \{K+1, K+2, \ldots, L-1\}, \quad R = \{L, L+1, \ldots, I\}.$$

In the next section we present the exact analysis of the double-threshold policy when $N \geq M$. The analysis for the case where $N < M$ is similar, and is presented in Appendix B of Iravani and Duenyas (1999). (Note as shown in Fig. 1(a and b) that the production thresholds for states in PI may be higher *or lower* than those in states in PR.)

3.1. *Exact analysis when $N \geq M$*

To analyze the number of items in inventory, let $\pi_{n,i}$ be the steady-state probability that the machine is in state i and there are n items in inventory,

$$n = \begin{cases} 1, 2, \ldots, M, & i = 2, 3, \ldots, K, \\ 1, 2, \ldots, N, & i = 1, K+1, K+2, \ldots, L-1, \\ 1, 2, \ldots, N-1, & i = L. \end{cases}$$

Then, we can formulate the model as a continuous-time Markov chain (the balance equations are provided in Iravani and Duenyas (1999)) and letting

$$\Pi_i(z) = \begin{cases} \sum_{n=0}^{M} \pi_{i,n} z^n, & i = 2, 3, \ldots, K, \\ \sum_{n=0}^{N} \pi_{i,n} z^n, & i = 1, K+1, K+2, \ldots, L-1, \\ \sum_{n=0}^{N-1} \pi_{i,n} z^n, & i = L, \end{cases} \tag{2}$$

then we have the following

Lemma 1.

$$\Pi_L(z) = \frac{[\mathcal{C}_L(z)\mathcal{A}_{L-1}(z) + z\phi_{L-1}\mathcal{C}_{L-1}(z)] \prod_{i=1}^{L-2} \mathcal{A}_i(z) + \sum_{j=1}^{L-2} z^{L-j} \mathcal{C}_j(z) \prod_{i=j}^{L-1} \phi_i \prod_{k=1}^{j-1} \mathcal{A}_i(z)}{\prod_{i=1}^{L} \mathcal{A}_i(z) - z^L r_L \prod_{i=1}^{L-1} \phi_i}, \tag{3}$$

and for $n = 1, 2, \ldots L-1$

$$\Pi_i(z) = z^i r_L \Pi_L(z) \left(\frac{\prod_{k=1}^{i-1} \phi_k}{\prod_{k=1}^{i} \mathcal{A}_k(z)} \right) + \sum_{j=1}^{i-1} z^{i-j} \frac{\mathcal{C}_j}{\mathcal{A}_j} \left(\frac{\prod_{k=j}^{i-1} \phi_k}{\prod_{k=j+1}^{i} \mathcal{A}_k(z)} \right) + \frac{\mathcal{C}_i(z)}{\mathcal{A}_i(z)}, \tag{4}$$

where

$$\mathcal{A}_i(z) = \begin{cases} (\mu_i z - \lambda)(1-z) + \phi_i z, & i = 1, 2, \ldots, L-1, \\ r_L z - \lambda(1-z), & i = L, \end{cases} \tag{5}$$

Since the state of the system (machine) only changes when the machine is producing parts, the system will never enter states $L+1, L+2, \ldots, I$. This is because of the fact that as soon as the machine gets to state L, its repair operation starts and when maintenance is completed, the system is back in state 1. Therefore, we only need to consider a machine with states $1, 2, \ldots, L$ $(L \leq I)$, where in states $1, 2, \ldots, K$ the machine produces items until the inventory level reaches M, whereafter it stays idle. However, the machine continues producing items in states $K+1, K+2, \ldots, L-1$ until there are N items in inventory. At this point, the maintenance operation starts and puts the system back in state 1.

$$C_i(z) = \begin{cases} \alpha_1(z) + (\mu_1 z(1-z) + \phi_1 z)\Phi_{M+1}^{N}(z) + \Gamma Z^{N+1}, & i = 1, \\ \alpha_2(z) - \phi_1 z \Phi_{M}^{N}(z), & i = 2, \\ \alpha_i(z) - \phi_{i-1}\pi_{M,i-1}z^{M+1}, & i = 3,4,\ldots,K, \\ \alpha_{K+1}(z) - \phi_K \pi_{M,K} z^{M+1} - r_{K+1}\pi_{N,K+1}z^{N+1}, & i = K+1, \\ \alpha_i(z) - \phi_{i-1}\pi_{N,i-1}z^{N+1} - r_i \pi_{N,i} z^{N+1}, & i = K+2,\ldots,L-1, \\ \alpha_L(z) - \phi_{L-1}\pi_{N,L-1}z^{N+1}, & i = L, \end{cases} \tag{6}$$

and

$$\alpha_i(z) = \begin{cases} (\mu_i z(1-z) + \phi_i z)\pi_{M,i} z^M - \lambda(1-z)\pi_{0i}, & i = 1,2,\ldots,K, \\ (\mu_i z(1-z) + \phi_i z)\pi_{N,i} z^N - \lambda(1-z)\pi_{0i}, & i = K+1, K+2, \ldots, L-1, \\ -\lambda(1-z)\pi_{0L}, & i = L, \end{cases} \tag{7}$$

$$\Phi_{M}^{N}(z) = \sum_{n=M}^{N} \pi_{n1} z^n \qquad \Gamma = \sum_{j=K+1}^{L-1} r_j \pi_{Nj}.$$

Proof. See Appendix A of Iravani and Duenyas (1999) for the outline of the proof. ∎

The denominator of (3) is a polynomial of order $2L-1$ which has $2L-1$ roots ξ_j, $j = 1,2,\ldots,2L-2$, where $\xi_{2L-1} = 1$. Substituting the first $2L-2$ roots into numerator (3) yields the following system of $2L-2$ equations with $2L+N-M-1$ unknowns π_{0i}; $1 \le i \le L$, π_{n1}; $M \le n \le N$ and π_{Mi}; $2 \le i \le K$, π_{Ni}; $K+1 \le i \le L-1$.

$$[\mathcal{C}_L(\xi_j)\mathcal{A}_{L-1}(\xi_j) + \xi_j \phi_{L-1}\mathcal{C}_{L-1}(\xi_j)] \prod_{i=1}^{L-2} \mathcal{A}_i(\xi_j) + \sum_{j=1}^{L-2} \xi_j^{L-j} \mathcal{C}_j(\xi_j) \prod_{i=j}^{L-1} \phi_i \prod_{k=1}^{j-1} \mathcal{A}_i(\xi_j) = 0. \tag{8}$$

Using (4) when $z = 1$, we get

$$\Pi_i(1) = \frac{1}{\phi_i}\left[r_L \Pi_L(1) + \sum_{j=1}^{i} \mathcal{C}_j(1) \right], \quad i = 1,2,\ldots,L-1,$$

and considering $\sum_{i=1}^{L} \Pi_i(1) = 1$, we will have;

$$r_1 \Pi_L(1) \sum_{i=1}^{L-1} \frac{1}{\phi_i} + \sum_{i=1}^{L-1} \sum_{j=1}^{i} \frac{\mathcal{C}_j(1)}{\phi_i} + \Pi_L(1) = 1. \tag{9}$$

Equation (9) along with $\Pi_L(1)$ which can be obtained using L'Hospital's rule in (3) adds one more equation to the system of linear equations (8). However, in order to compute the $2L+N-M-1$ unknowns, we need to obtain more equations. This can be done by considering the balance equations provided in Iravani and Duenyas (1999) where the process of creating a system of $L(N-M+1) - (K-1)(N-M-1)$ linear equations with the same number of unknowns, which can be solved using numerical techniques is described in detail. However, if the number of states is large, then this can potentially be time consuming and this is why we develop a heuristic approach in Section 5.

The analysis of the case where $N < M$ is similar to the case with $N \ge M$ and therefore is omitted. (It is given in Appendix B of Iravani and Duenyas (1999).)

It should be noted that $\Pi_L(1)$ can be obtained by setting $z = 1$ after using L'Hospital's rule in (3). However, the result does not have a nice closed-form and this problem does not only exist for $\Pi_L(1)$. As the structure of the balance equations and the generating functions show, the results for the average number of items in inventory and the probability of having zero inventory are even more complicated. Therefore, although we can obtain exact results by solving the above equations for any system, we next focus on systems with three states. The analysis of the system with three states will be the basis for our heuristic in Section 5 which we develop for systems with any number of states. Furthermore, the case with three states is of interest on its own because there are many situations in which machines are classified into three states such as: (i) *as good as new*; (ii) *deteriorated*; and (iii) *failed.*

4. A systems with three states

In this section we consider a machine which has three states: (i) as good as new; (ii) deteriorated; and (iii) failed. We will use the exact results developed here in our heuristic in the next section. When the machine is in state 3, it

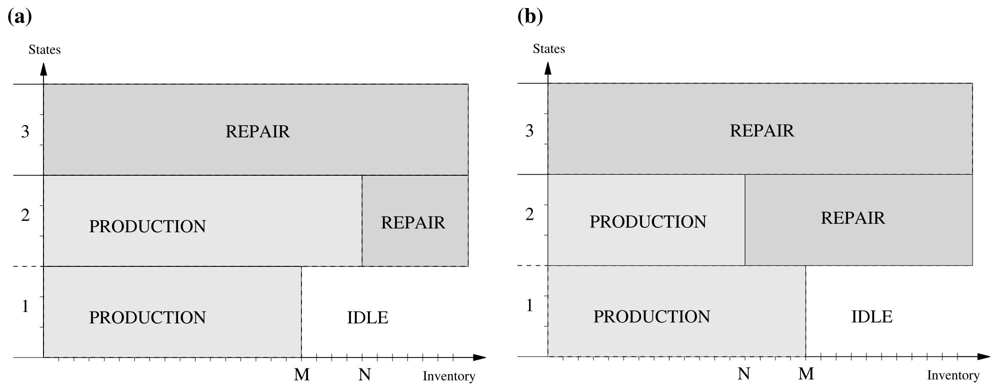

Fig. 3. A three-state system with preventative maintenance when: (a) $N \geq M$; and (b) $N < M$.

has to be repaired as the machine cannot produce any units in this state at all. However, in state 2, when the inventory level reaches a threshold N, the preventive maintenance operation starts and puts the machine back in state 1. The machine becomes idle in state 1 when there are M items in inventory. Figure 3(a and b) show this policy for $N \geq M$ and $N < M$, respectively. We give the analysis for $N \geq M$; the analysis for $N < M$ is similar and is omitted for brevity.

Letting π_{ni} equal the steady-state probability that the machine is in state i and that there are n items in inventory, the balance equations for the system are:

$$\begin{aligned}
\pi_{01}(\mu_1 + \phi_1) &= \lambda\pi_{11} + r_3\pi_{03},\\
\pi_{n1}(\lambda + \mu_1 + \phi_1) &= \lambda\pi_{n+1,1} + \mu_1\pi_{n-1,1} + r_3\pi_{n3},\\
&\qquad n = 1, 2, \ldots, M-1,\\
\pi_{M1}(\lambda) &= \lambda\pi_{M+1,1} + \mu_1\pi_{M-1,1} + r_3\pi_{M3},\\
\pi_{n1}(\lambda) &= \lambda\pi_{n+1,1} + r_3\pi_{n3},\\
&\qquad n = M+1, \ldots, N-1,\\
\pi_{N1}(\lambda) &= r_2\pi_{N2},\\
\pi_{02}(\mu_2 + \phi_2) &= \lambda\pi_{12} + \phi_1\pi_{01},\\
\pi_{n2}(\lambda + \mu_2 + \phi_2) &= \lambda\pi_{n+1,2} + \phi_1\pi_{n1} + \mu_2\pi_{n-1,2},\\
&\qquad n = 1, 2, \ldots, M-1,\\
\pi_{n2}(\lambda + \mu_2 + \phi_2) &= \lambda\pi_{n+1,2} + \mu_2\pi_{n-1,2},\\
&\qquad n = M, M+1, \ldots, N-1,\\
\pi_{N2}(\lambda + r_2) &= \mu_2\pi_{N-1,2},\\
\pi_{03}(r_3) &= \lambda\pi_{13} + \phi_2\pi_{02},\\
\pi_{n3}(\lambda + r_3) &= \lambda\pi_{n+1,3} + \phi_2\pi_{n2},\\
&\qquad n = 1, 2, \ldots, N-2,\\
\pi_{N-1,3}(\lambda + r_3) &= \phi_2\pi_{N-1,2}.
\end{aligned}$$

Therefore we will have:

$$\Pi_1(z) = \frac{\mathcal{C}_1(z)}{\mathcal{A}_1(z)} + \frac{r_3 z}{\mathcal{A}_1(z)}\Pi_3(z), \tag{10}$$

$$\Pi_2(z) = \frac{\mathcal{C}_2(z)}{\mathcal{A}_2(z)} + \frac{\mathcal{C}_1(z)\phi_1}{\mathcal{A}_1(z)\mathcal{A}_2(z)}z + \frac{\phi_1 r_3 z^2}{\mathcal{A}_1(z)\mathcal{A}_2(z)}\Pi_3(z), \tag{11}$$

$$\Pi_3(z) = \frac{\mathcal{C}_1(z)\phi_1\phi_2 z^2 + \mathcal{C}_2(z)\mathcal{A}_1(z)\phi_2 z + \mathcal{C}_3(z)\mathcal{A}_1(z)\mathcal{A}_2(z)}{\mathcal{A}_1(z)\mathcal{A}_2(z)\mathcal{A}_3(z) - r_3\phi_1\phi_2 z^3}, \tag{12}$$

where

$$\begin{aligned}
\mathcal{C}_1(z) &= [\mu_1 z(1-z) + \phi_1 z]\Phi_M^N(z)\\
&\quad + r_2\pi_{N2}z^{N+1} - \lambda(1-z)\pi_{01},\\
\mathcal{C}_2(z) &= [\mu_2 z(1-z) + \phi_2 z]\pi_{N2}z^N\\
&\quad - r_2\pi_{N2}z^{N+1} - \phi_1 z\Phi_M^N(z) - \lambda(1-z)\pi_{02},\\
\mathcal{C}_3(z) &= -\phi_2\pi_{N2}z^{N+1} - \lambda(1-z)\pi_{03},
\end{aligned}$$

and $\mathcal{A}_i(z)$ and $\Phi_M^N(z)$ are as in Lemma 1 for a three-state problem.

Using five roots $\xi_j, j = 1, 2, \ldots, 5$ (except $\xi_5 = 1$) of denominator (12) in the numerator (12) creates the following system of four linear equations with $5 + N - M$ unknowns, π_{01}, π_{02}, π_{03}, π_{N2} and π_{n1} for $M \leq n \leq N$.

$$\begin{aligned}
&\mathcal{C}_1(\xi_j)\phi_1\phi_2\xi_j^2 + \mathcal{C}_2(\xi_j)\mathcal{A}_1(\xi_j)\phi_2\xi_j\\
&\quad + \mathcal{C}_3(\xi_j)\mathcal{A}_1(\xi_j)\mathcal{A}_2(\xi_j) = 0, \quad j = 1, 2, 3, 4.
\end{aligned} \tag{13}$$

On the other hand,

$$\begin{aligned}
\Pi_1(1) &= \Phi_M^N(1) + \frac{r_2}{\phi_1}\pi_{N,2} + \frac{r_3}{\phi_1}\Pi_3(1),\\
\Pi_2(1) &= \pi_{N,2} + \frac{r_3}{\phi_2}\Pi_3(1).
\end{aligned}$$

Therefore considering $\sum_{i=1}^{3}\Pi_i(1) = 1$, we will have

$$\begin{aligned}
&\Pi_3(1)[V+1]\\
&\quad + \pi_{N,2}\left[\frac{r_2}{\phi_1} + 1\right] + \Phi_M^N(1) = 1,
\end{aligned} \tag{14}$$

where $v_i = r_3/\phi_i$ and $V = v_1 + v_2$. Using L'Hospital's rule in (12) we get

$$\Pi_3(1) = \frac{\lambda[\pi_0 - \Phi_M^N(1)] - (r_2 w_1 + \lambda)\pi_{N2}}{r_3 W + \lambda}, \tag{15}$$

where $\pi_0 = \pi_{01} + \pi_{02} + \pi_{03}$, $w_i = (\lambda - \mu_i)/\phi_i$ and $W = w_1 + w_2$.

Now, in order to find unknowns π_{01}, π_{02}, π_{03}, π_{N2} and π_{n1} for $M \le n \le N$, similar to Section 3, the system of linear equations consists of the selected balance equations of the system along with Equations (14), (15) and (13) must be solved.

Finally, the average number of items in the inventory, $E[N] = \sum_{i=1}^{3} \Pi'_i(1)$ can be obtained, where

$$\Pi'_1(1) = \frac{\lambda[\pi_{01} - \Phi^N_M(1)] + \phi_1 \Phi'^N_M(1) + r_2(N - w_1)\pi_{N2}}{\phi_1} + v_1\Pi'_3(1) - v_1 w_1 \Pi_3(1),$$

$$\Pi'_2(1) = \frac{\lambda[\pi_{01} + \pi_{02} - \Phi^N_M(1)] + [N\phi_2 - \lambda + r_2 w_1]\pi_{N2}}{\phi_2} + v_2\Pi'_3(1) - v_2 W \Pi_3(1).$$

On the other hand, obtaining the first derivation of $\Pi_3(z)$, we get

$$\Pi'_3(1) = \frac{\mathcal{Q} - \lambda[\phi'_{M1}(1) + 2\Phi^N_M(1)] + \mathcal{L}\pi_{N2} - \mathcal{P}\Pi_3(1)}{r_3 W + \lambda}, \tag{16}$$

where

$$\mathcal{P} = (2r_3 + \lambda)W + r_3 w_1 w_2 + 2\lambda - v_1\mu_1 - v_2\mu_2,$$
$$\mathcal{Q} = \lambda[2\pi_0 + w1\pi_{02} + W\pi_{03}],$$
$$\mathcal{L} = \frac{r_2\mu_1}{\phi_1} - \phi_2 w_2(N + 2 + w_1) + r_2 w_1(N+2) + \mu_2[w_1 + N + 2].$$

Therefore the optimal limits M and N can be found by searching for the M^* and N^* which minimize the total average inventory and repair cost, $E[TC(M,N)]$, where

$$E[TC(M,N)] = h\,E[N] + C\lambda\pi_0 + m_2\pi_{N2} + m_3\Pi_3(1). \tag{17}$$

5. Heuristic approach for double-threshold policy

In Section 3, we presented an exact analysis of systems with any number of states where the double-threshold policy is used. However, the analysis requires the solving of a potentially high number of simultaneous equations (depending on the number of states). For this reason, we provide a heuristic approach for computing the performance of a system under the double-threshold policy and also for calculating the approximately optimal threshold levels.

Our heuristic converts problems with more than three states to a three-state problem and uses the results in Section 4 to approximate thresholds M^* and N^* and optimal sets PI, PR and R for the optimal double-threshold policy. For each of the given sets $PI = \{1, 2, \ldots, K\}$, $PR = \{K+1, \ldots, L-1\}$ and $R = \{L, L+1, \ldots, I\}$ in an I-state problem, our heuristic defines a three-state problem with states $\hat{1}$, $\hat{2}$ and $\hat{3}$, where state $\hat{1}$ has the following parameters

$$\hat{\mu}_1 = \frac{1}{K}\sum_{i\in PI}\mu_i, \quad \hat{\phi}_1 = \left[\sum_{i\in PI}\frac{1}{\phi_i}\right]^{-1},$$
$$\hat{r}_1 = \frac{1}{K}\sum_{i\in PI} r_i, \quad \hat{m}_1 = \frac{1}{K}\sum_{i\in PI} m_i.$$

In other words, the heuristic considers an aggregated state $\hat{1}$ to represent states in $PI = \{1, 2, \ldots, K\}$. This aggregated state has aggregated production and repair rate $\hat{\mu}_1$ and $\hat{r}_1$ and the aggregated failure rate $\hat{\phi}_1$. The aggregated failure rate actually reflects the average rate of entering state $K+1$ from state 1. Thus, in the three-state problem, it takes on average

$$\sum_{i\in PI}\frac{1}{\phi_i},$$

for the machine to leave state $\hat{1}$ and enter state $\hat{2}$.

Similarly the parameters for state $\hat{2}$ are

$$\hat{\mu}_2 = \frac{1}{L-K-1}\sum_{i\in PR}\mu_i, \quad \hat{\phi}_2 = \left[\sum_{i\in PR}\frac{1}{\phi_i}\right]^{-1},$$
$$\hat{r}_1 = \frac{1}{L-K-1}\sum_{i\in PR} r_i, \quad \hat{m}_2 = \frac{1}{L-K-1}\sum_{i\in PR} m_i.$$

Finally for state $\hat{3}$, the heuristic considers

$$\hat{\mu}_3 = \mu_L, \quad \hat{\phi}_3 = \phi_L, \quad \hat{r}_3 = r_L, \quad \hat{m}_3 = m_L.$$

The heuristic analyzes the three-state problem using the results from Section 4 to obtain the optimal threshold $\hat{M}^*$ and $\hat{N}^*$, which is then used as an approximation for the optimal values M^* and N^* in the original I-state problem.

We summarize our heuristic for an I-state problem in the following algorithm:

Step 0. Set $L = I$, $K = 1$ and go to Step 1.

Step 1. Set $PI = \{1, 2, \ldots, K\}$ and $PR = \{K+1, K+2, \ldots, L-1\}$ and compute

$$\hat{\mu}_1 = \frac{1}{K}\sum_{i\in PI}\mu_i, \quad \hat{\phi}_1 = \left[\sum_{i\in PI}\frac{1}{\phi_i}\right]^{-1},$$
$$\hat{r}_1 = \frac{1}{K}\sum_{i\in PI} r_i, \quad \hat{\mu}_2 = \frac{1}{L-K-1}\sum_{i\in PR}\mu_i,$$
$$\hat{\phi}_2 = \left[\sum_{i\in PR}\frac{1}{\phi_i}\right]^{-1}, \quad \hat{r}_1 = \frac{1}{L-K-1}\sum_{i\in PR} r_i,$$
$$\hat{\mu}_3 = \mu_L \quad \hat{\phi}_3 = \phi_L \quad \hat{r}_3 = r_L$$

Step 2. Find the optimal thresholds $\hat{M}^*$ and $\hat{N}^*$ for the three-state problem defined in Step 1 using both models introduced in Section 4 for $N < M$ and $N \ge M$. Set $TC_{L,K}(\hat{M}^*, \hat{N}^*)$ as the optimal total average cost per unit time for the optimal thresholds $\hat{M}^*$ and $\hat{N}^*$ and go to Step 3.

Step 3. Set $K \leftarrow K + 1$. If $K = L - 1$, set $L \leftarrow L - 1$, and $K = 1$. If $L = 2$, go to Step 4, otherwise, go to Step 1.

Step 4. Find $TC_{L^*,K^*}(M_s^*, N_s^*) = \text{Min}_{L,K}\{TC_{L,K}(\hat{M}^*, \hat{N}^*)\}$. Then the suboptimal inventory thresholds for the I-state problem are M_s^* and N_s^* with suboptimal state sets

$$\hat{P}I = \{1, 2, \ldots, K^*\},$$
$$\hat{P}R = \{K^* + 1, K^* + 2, \ldots, L^* - 1\},$$
$$\hat{R} = \{L^*, L^* + 1, \ldots, I\}.$$

The total number of iterations in our heuristic is $(I - 1)I/2$ where in each iteration a three-state problem is analyzed to obtain the optimal values $\hat{M}^*$ and $\hat{N}^*$. In a Pentium II computer, finding the optimal thresholds for a problem with 10 different machine states takes under 30 seconds.

6. Numerical study

In this section, we report the results of a numerical study we conducted. The purpose of the study was to explore: (i) whether the double-threshold policy is nearly as good as the optimal control policy; (ii) if the heuristic we described performs well in estimating the double-threshold levels; and (iii) how well or badly the policies which ignore interactions between maintenance and production/inventory perform. Finally, we conduct a numerical study to explore the significance of explicitly taking into account machine deterioration information.

6.1. *Evaluation of the double-threshold policy and heuristic*

We studied a set of problems for a machine with 10 states and compared the optimal control policy (obtained by solving (1)) with the optimal double-threshold policy where: (i) the exact model in Section 3 is used to obtain the optimal parameters of the double-threshold policy; and (ii) the heuristic in Section 5 is used to obtain the approximately optimal parameters of the double-threshold policy. Thus, our study tells us both how well the double-threshold policy performs and also how well our heuristic performs in estimating the policy's parameters.

We initially created 32 problems using demand rate $\lambda = 8$, holding cost $h = 1$ and, lost sale cost $C = \{10, 50\}$ per unit, repair cost $m_i = m$; $\forall\, i$ where $m = \{20, 200\}$ per unit time. For production rates, we either used the array μ_{I} or μ_{II}. Similarly, for failure rates, we used ϕ_{I} and ϕ_{II} and for repair rates we used r_{I} and r_{II} given by

$$\mu_{\text{I}} = \{10, 10, 10, 10, 10, 10, 10, 10, 10, 10\},$$
$$\mu_{\text{II}} = \{10, 10, 8, 8, 6, 6, 4, 4, 2, 2\},$$
$$\phi_{\text{I}} = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1\},$$
$$\phi_{\text{II}} = \{0.1, 0.1, 0.25, 0.25, 0.5, 0.5, 0.75, 0.75, 1, 1\},$$
$$r_{\text{I}} = \{5, 5, 5, 5, 5, 5, 5, 5, 5, 5\},$$
$$r_{\text{II}} = \{10, 10, 7, 7, 5, 5, 2, 2, 1, 1\}.$$

Since we had two choices each for production, repair and failure rates, and two choices for C and m, an exhaustive combination resulted in 32 cases displayed in Table 1.

We also analyzed an additional 16 examples (displayed in Table 2) with $h = 1$ and $C = 10$ and variable repair costs as follows:

$$m_1 = \{20, 30, 40, 50, 60, 70, 80, 90, 100, 110\},$$
$$m_2 = \{20, 25, 35, 50, 70, 90, 125, 160, 200, 245\}.$$

Note that whereas the first 32 examples we created have constant repair costs, the next 16 cases have linear and non-linear increasing repair costs as a function of machine state. The criteria for our evaluation in all examples was the relative errors ϵ_{D} and ϵ_{H} which were defined as follows:

$$\epsilon_{\text{D}} = \frac{TC_{\text{D}}(M^*, N^*) - TC(\text{opt.})}{TC(\text{opt.})} \times 100$$

$$\epsilon_{\text{H}} = \frac{TC_{\text{D}}(\hat{M}^*, \hat{N}^*) - TC(\text{opt.})}{TC(\text{opt.})} \times 100$$

where $TC(\text{opt.})$ is the total average cost under optimal control policy; $TC_{\text{D}}(M^*, N^*)$ is the total average cost under the double-threshold policy with parameters (M^*, N^*) where M^* and N^* are obtained using the exact model, and $TC_{\text{D}}(\hat{M}^*, \hat{N}^*)$ is the total average cost under the double-threshold policy with parameters $(\hat{M}^*, \hat{N}^*)$ where $\hat{M}^*$ and $\hat{N}^*$ are obtained using the heuristic.

Tables 1 and 2 summarize our results. The average relative error for the double-threshold policy is about 0.5%, and if heuristic approach is used to obtain the optimal thresholds for the double-threshold policy, the average relative error compared to the optimal dynamic policy is about 0.83%. This implies that the difference between using exact double-threshold levels and approximated ones obtained by our heuristic is about 0.3%. We conclude that: (i) the double-threshold policy with its simple structure is a very good policy which performs close to the much more complex optimal dynamic policy; and (ii) our heuristic is an efficient and accurate tool for approximating the optimal thresholds of the double-threshold policy. In fact, in almost a third of the problems we looked at, our heuristic policy yielded the same costs as the optimal dynamic policy.

Table 1. A comparison of the double-threshold policy and the heuristic with the optimal policy

Number	*(λ,μ,φ,r)*	*(h,C, m)*	*TC(opt.)*	*Double-threshold* (K^*, L^*, M^*, N^*)	*Cost*	ϵ_D	*Heuristic* (K^*, L^*, M^*, N^*)	*Cost*	ϵ_H
1	$(8, \mu_I, \phi_I, r_I)$	(1,10, 20)	9.111	(6,10,8,8)	9.113	0.02	(7,10,8,8)	9.137	0.28
2	$(8, \mu_I, \phi_I, r_I)$	(1,10, 200)	12.306	(9,10,8,–)	12.346	0.32	(9,10,8,–)	12.346	0.32
3	$(8, \mu_I, \phi_I, r_I)$	(1,50, 20)	15.296	(4,10,14,14)	15.300	0.03	(6,10,14,14)	15.356	0.39
4	$(8, \mu_I, \phi_I, r_I)$	(1,50, 200)	18.692	(9,10,15,15)	18.768	0.41	(9,10,15,15)	18.768	0.41
5	$(8, \mu_I, \phi_I, r_{II})$	(1,10, 20)	9.630	(3,6,8,8)	9.674	0.45	(3,6,8,8)	9.671	0.45
6	$(8, \mu_I, \phi_I, r_{II})$	(1,10, 200)	15.387	(5,6,8,–)	15.387	0	(5,6,8,–)	15.387	0
7	$(8, \mu_I, \phi_I, r_{II})$	(1,50, 20)	16.367	(3,6,15,10)	16.396	0.18	(3,6,15,12)	16.411	0.27
8	$(8, \mu_I, \phi_I, r_{II})$	(1,50, 200)	22.320	(5,6,15,–)	22.392	0.32	(5,6,16,–)	22.404	0.38
9	$(8, \mu_I, \phi_{II}, r_I)$	(1,10, 20)	8.608	(4,10,8,8)	8.610	0.02	(4,10,8,8)	8.610	0.02
10	$(8, \mu_I, \phi_{II}, r_I)$	(1,10, 200)	9.411	(7,10,8,9)	9.434	0.24	(7,10,8,9)	9.434	0.24
11	$(8, \mu_I, \phi_{II}, r_I)$	(1,50, 20)	14.348	(3,10,14,13)	14.350	0.01	(3,10,14,13)	14.350	0.01
12	$(8, \mu_I, \phi_{II}, r_I)$	(1,50, 200)	15.268	(6,10,14,13)	15.286	0.12	(6,10,14,14)	15.307	0.12
13	$(8, \mu_{II}, \phi_I, r_I)$	(1,10, 20)	11.794	(2,5,10,2)	11.852	0.49	(2,5,9,2)	11.854	0.51
14	$(8, \mu_{II}, \phi_I, r_I)$	(1,10, 200)	18.966	(4,5,10,–)	19.552	3.09	(4,5,9,–)	19.611	3.40
15	$(8, \mu_{II}, \phi_I, r_I)$	(1,50, 20)	20.791	(2,3,19,–)	20.843	0.25	(2,3,19,–)	20.843	0.25
16	$(8, \mu_{II}, \phi_I, r_I)$	(1,50, 200)	30.353	(4,5,22,–)	31.207	2.81	(4,5,21,–)	31.221	2.86
17	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,10, 20)	8.829	(1,3,8,8)	8.829	0	(1,3,8,8)	8.829	0
18	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,10, 200)	10.233	(2,5,8,1)	10.233	0	(2,5,8,1)	10.233	0
19	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,50, 20)	14.698	(1,3,14,14)	14.698	0	(1,3,14,14)	14.698	0
20	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,50, 200)	16.275	(2,3,14,–)	16.275	0	(2,3,14,–)	16.275	0
21	$(8, \mu_{II}, \phi_I, r_{II})$	(1,10, 20)	10.776	(1,3,9,9)	10.781	0.05	(1,3,9,9)	10.781	0.05
22	$(8, \mu_{II}, \phi_I, r_{II})$	(1,10, 200)	18.092	(3,4,9,–)	18.159	0.37	(3,4,9,–)	18.159	0.37
23	$(8, \mu_{II}, \phi_I, r_{II})$	(1,50, 20)	18.444	(1,3,16,17)	18.447	0.02	(1,3,16,17)	18.447	0.02
24	$(8, \mu_{II}, \phi_I, r_{II})$	(1,50, 200)	26.959	(3,4,19,–)	27.778	3.04	(3,4,19,–)	27.778	3.04
25	$(8, \mu_I, \phi_{II}, r_{II})$	(1,10, 20)	8.611	(2,8,8,7)	8.620	0.10	(3,7,8,7)	8.642	0.36
26	$(8, \mu_I, \phi_{II}, r_{II})$	(1,10, 200)	9.464	(3,6,8,5)	9.466	0.02	(3,5,8,4)	9.486	0.23
27	$(8, \mu_I, \phi_{II}, r_{II})$	(1,50, 20)	14.362	(2,8,14,10)	14.403	0.28	(3,6,14,9)	14.411	0.34
28	$(8, \mu_I, \phi_{II}, r_{II})$	(1,50, 200)	15.274	(3,8,14,6)	15.312	0.25	(3,6,14,12)	15.296	0.14
29	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,10, 20)	8.678	(1,3,8,8)	8.678	0	(1,3,8,8)	8.678	0
30	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,10, 200)	9.721	(2,3,8,–)	9.721	0	(2,3,8,–)	9.721	0
31	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,50, 20)	14.476	(1,3,13,12)	14.476	0	(1,3,13,12)	14.476	0
32	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,50, 200)	15.626	(2,3,13,–)	15.626	0	(2,3,13,–)	15.626	0

Table 2. A comparison of the double-threshold and the heuristic with the optimal policy for variable repair costs

Number	*(λ,μ,φ,r)*	*m*	*TC(opt.)*	*Double-threshold* (K^*, L^*, M^*, N^*)	*Cost*	ϵ_D	*Heuristic* (K^*, L^*, M^*, N^*)	*Cost*	ϵ_H
1	$(8, \mu_I, \phi_I, r_I)$	m_1	10.644	(6,10,8,8)	10.646	0.02	(8,10,8,7)	10.722	0.73
2	$(8, \mu_I, \phi_I, r_I)$	m_2	11.450	(2,10,8,6)	11.543	0.81	(4,6,8,5)	11.884	3.78
3	$(8, \mu_I, \phi_I, r_{II})$	m_1	10.874	(3,6,8,4)	10.915	0.37	(3,6,8,8)	11.064	1.74
4	$(8, \mu_I, \phi_I, r_{II})$	m_2	10.847	(1,6,8,6)	10.987	1.29	(2,4,8,8)	11.041	1.79
5	$(8, \mu_I, \phi_{II}, r_I)$	m_1	8.815	(2,10,8,8)	8.816	0.02	(4,7,8,7)	8.880	0.74
6	$(8, \mu_I, \phi_{II}, r_I)$	m_2	8.801	(2,10,8,7)	8.807	0.06	(2,4,8,7)	8.850	0.55
7	$(8, \mu_{II}, \phi_I, r_I)$	m_1	13.330	(2,5,9,2)	13.373	0.32	(3,5,10,2)	13.613	2.12
8	$(8, \mu_{II}, \phi_I, r_I)$	m_2	12.989	(2,5,9,1)	13.037	0.37	(2,4,9,1)	13.040	0.40
9	$(8, \mu_{II}, \phi_{II}, r_I)$	m_1	8.983	(1,3,8,8)	8.983	0	(1,3,8,8)	8.983	0
10	$(8, \mu_{II}, \phi_{II}, r_I)$	m_2	8.916	(1,3,8,8)	8.916	0	(1,3,8,8)	8.916	0
11	$(8, \mu_{II}, \phi_I, r_{II})$	m_1	11.756	(2,4,9,1)	11.888	1.12	(2,4,9,1)	11.888	1.12
12	$(8, \mu_{II}, \phi_I, r_{II})$	m_2	11.432	(1,3,9,6)	11.432	0	(1,3,9,6)	11.432	0
13	$(8, \mu_I, \phi_{II}, r_{II})$	m_1	8.730	(1,6,8,8)	8.732	0.02	(2,4,8,7)	8.760	0.35
14	$(8, \mu_I, \phi_{II}, r_{II})$	m_2	8.697	(1,6,8,7)	8.701	0.04	(1,3,8,7)	8.725	0.32
15	$(8, \mu_{II}, \phi_{II}, r_{II})$	m_1	8.761	(1,3,8,7)	8.761	0	(1,3,8,7)	8.761	0
16	$(8, \mu_{II}, \phi_{II}, r_{II})$	m_2	8.725	(1,3,8,7)	8.725	0	(1,3,8,7)	8.725	0

6.2. *Independent inventory and repair policies*

In almost all practical situations we are aware of, most maintenance/repair and inventory level decisions are made in an *ad-hoc* manner. In particular, when managers typically make these decisions, they ignore the interactions between them. We explored the magnitude of the cost difference between considering these decisions in an integrated manner and ignoring these interactions.

Consider a maintenance/repair department which devises a maintenance/repair program considering only the repair/maintenance-related costs. The production control department then takes this maintenance program as given and establishes their optimal production/inventory levels based on this information.

As is also typical in most maintenance literature, the maintenance departments usually find a threshold state (L_r) such that once the machine enters this state, it has to be maintained. If one aims to minimize the total maintenance/repair-related costs per unit time, then the total average repair cost per unit time for a repair policy which starts repair as soon as the machine enters state L_r can be obtained by solving a simple Markov chain with L_r states and transition rate matrix Q with elements q_{ij}, where

$$q_{ij} = \begin{cases} \phi_i, & j = i+1,\ i = 1, 2, \ldots, L_r - 1, \\ r_{L_r}, & j = 1,\ i = L, \\ 0, & \text{otherwise.} \end{cases}$$

After solving the above Markov chain, the steady-state probabilities of the system being in state i, π_i, are obtained and the total average repair cost per unit time will be $m\pi_{L_r}$. Computing $m\pi_{L_r}$ for different values of $L_r = 2, 3, \ldots, I$ for an I-state machine problem by solving $I - 1$ Markov chains, L_r^* will be obtained. L_r^* is the value of L_r which has the least $m\pi_{L_r}$.

Now considering this repair policy, the best base-stock level M_r^* in order to minimize the total average inventory cost per unit time can be obtained. We call this policy a single-threshold policy since it is a special case of the double-threshold policy for a machine with L_r^* state where $K = L_r^* - 1$ (threshold N does not exist). The optimal threshold M_r^* can be obtained using our exact model or heuristic. Note that in this case, the optimal inventory levels are being set by taking into account the maintenance policy to be used. However, the maintenance policy is not taking into account the costs of carrying inventory and lost sales.

As Table 3 displays, ignoring the full interactions between maintenance and inventory decisions can be rather costly. Table 3 shows the cost of the above-described single-threshold policy as compared to the optimal dynamic policy for the same examples in Table 1. The average error of this policy is 39.7% and errors can go as high as 330%. We believe that these results clearly make the case for using analytical models such as the ones developed in this paper that help users make these decisions in an integrated fashion.

6.3. *The effect of machine deterioration*

The effects of machine deterioration on maintenance decisions are sometimes even larger than the effects of machine failure. Brek and Moinzadeh (2000) present examples of the effects of machine deterioration in semiconductor fabrication. Their model utilizes the information about the output yield (which deteriorates with machine age) in order to establish a cost-effective maintenance policy.

Our model also allows the production rate to decrease as the machine deteriorates. This is often the case in practice due to the machine producing a larger number of defective items as it ages (therefore, the "effective" production rate of good parts decreases), or requiring larger adjustment times. However, most of the models in the literature, especially those considering maintenance and production (Van der Duyn Schouten and Vanneste, 1995) assume that the production rate of a machine is constant as long as the machine is operational. Thus a user who is using one of those models has to devise a single aggregated production rate that best reflects the "average" capability of the machine. We demonstrate the importance of collecting and using information on how the production rate deteriorates as a function of machine state with the following example.

Consider case 18 in Table 1 for a 10-state machine with decreasing production rate $\mu = [10\ 10\ 8\ 8\ 6\ 6\ 4\ 4\ 2\ 2]$. The optimal double-threshold policy for this case has parameters (2, 5, 8, 1) and a total average cost of 10.233. However, if we could only use an aggregate production rate for all operational states, we would have to select a reasonable aggregate production rate. If we replace the production rate $\mu = [10\ 10\ 8\ 8\ 6\ 6\ 4\ 4\ 2\ 2]$ with the average production rate $\mu g = [6\ 6\ 6\ 6\ 6\ 6\ 6\ 6\ 6\ 6]$ in case 18 of Table 1, the optimal double-threshold policy obtained will have parameters $(7, 10, 23, 24)$. This change in parameters of the double-threshold policy from $(2, 5, 8, 1)$ to $(7, 10, 23, 24)$ is the result of neglecting the decrease in production rate by using a model which ignores the machine deterioration process. To show how far this new solution is from the optimal, we compute the total average cost of applying double-threshold policy with parameters $(7, 10, 23, 24)$ in the original problem where the production rates are $\mu = [10\ 10\ 8\ 8\ 6\ 6\ 4\ 4\ 2\ 2]$. It is found that the total average cost of this policy is 19.798 which is a 93% increase in the total cost. In fact, regardless of which value of μ_g is used, as long as a single value is used for all operational states, the result will be similar. For example, if we use $\mu_g = 10$, then the cost is 47% higher, while for $\mu_g = 8$, the cost is 50% higher. This additional cost can be viewed as the cost that the firm will incur if it does not use data on how its production rate deteriorates as a function of machine state appropriately in a model that uses this information.

We have run the other examples in Table 1 and obtain similar results. These results indicate that the information

Table 3. Performance of the single-threshold policy

Number	*(λ,μ,ϕ,r)*	*(h,C, m)*	*TC(opt.)*	*Single-threshold policy* $(L\hat{}*,M\hat{}*)$	Cost	ϵ_D
1	$(8, \mu_I, \phi_I, r_I)$	(1,10, 20)	9.111	(10,8)	9.293	2.00
2	$(8, \mu_I, \phi_I, r_I)$	(1,10, 200)	12.306	(10,8)	12.346	0.32
3	$(8, \mu_I, \phi_I, r_I)$	(1,50, 20)	15.296	(10,15)	15.601	1.99
4	$(8, \mu_I, \phi_I, r_I)$	(1,50, 200)	18.692	(10,15)	18.768	0.41
5	$(8, \mu_I, \phi_I, r_{II})$	(1,10, 20)	9.630	(6,9)	9.893	0.28
6	$(8, \mu_I, \phi_I, r_{II})$	(1,10, 200)	15.387	(6,8)	15.387	0
7	$(8, \mu_I, \phi_I, r_{II})$	(1,50, 20)	16.367	(6,16)	16.707	0.21
8	$(8, \mu_I, \phi_I, r_{II})$	(1,50, 200)	22.320	(6,15)	22.392	0.48
9	$(8, \mu_I, \phi_{II}, r_I)$	(1,10, 20)	8.608	(10,8)	8.692	0.97
10	$(8, \mu_I, \phi_{II}, r_I)$	(1,10, 200)	9.411	(10,8)	9.444	0.35
11	$(8, \mu_I, \phi_{II}, r_I)$	(1,50, 20)	14.348	(10,14)	14.551	1.41
12	$(8, \mu_I, \phi_{II}, r_I)$	(1,50, 200)	15.268	(10,14)	15.342	0.48
13	$(8, \mu_{II}, \phi_I, r_I)$	(1,10, 20)	11.794	(10,21)	23.388	98.30
14	$(8, \mu_{II}, \phi_I, r_I)$	(1,10, 200)	18.966	(10,20)	27.257	43.71
15	$(8, \mu_{II}, \phi_I, r_I)$	(1,50, 20)	20.791	(10,40)	91.310	339.18
16	$(8, \mu_{II}, \phi_I, r_I)$	(1,50, 200)	30.353	(10,40)	95.221	213.71
17	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,10, 20)	8.829	(10,9)	14.097	59.67
18	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,10, 200)	10.233	(10,9)	14.946	46.06
19	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,50, 20)	14.698	(10,32)	37.449	154.79
20	$(8, \mu_{II}, \phi_{II}, r_I)$	(1,50, 200)	16.275	(10,32)	38.364	135.72
21	$(8, \mu_{II}, \phi_I, r_{II})$	(1,10, 20)	10.776	(6,12)	13.529	25.55
22	$(8, \mu_{II}, \phi_I, r_{II})$	(1,10, 200)	18.092	(6,11)	19.797	9.42
23	$(8, \mu_{II}, \phi_I, r_{II})$	(1,50, 20)	18.444	(6,28)	29.584	60.40
24	$(8, \mu_{II}, \phi_I, r_{II})$	(1,50, 200)	26.959	(6,28)	36.193	34.25
25	$(8, \mu_I, \phi_{II}, r_{II})$	(1,10, 20)	8.611	(4,8)	8.694	0.96
26	$(8, \mu_I, \phi_{II}, r_{II})$	(1,10, 200)	9.464	(4,8)	9.571	0.56
27	$(8, \mu_I, \phi_{II}, r_{II})$	(1,50, 20)	14.362	(4,14)	14.537	1.22
28	$(8, \mu_I, \phi_{II}, r_{II})$	(1,50, 200)	15.274	(4,14)	15.384	0.72
29	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,10, 20)	8.678	(4,8)	9.340	7.62
30	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,10, 200)	9.721	(4,8)	10.182	4.74
31	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,50, 20)	14.476	(4,16)	16.662	15.10
32	$(8, \mu_{II}, \phi_{II}, r_{II})$	(1,50, 200)	15.626	(4,15)	17.536	12.22

on how machine deterioration affects production rates is critical and firms should collect this information and make use of models that take this information into account in their calculations.

7. Conclusions

We have presented an integrated maintenance/repair and production/inventory model. The optimal policy for this problem is found to have a very complex structure. However, we have described an easily implementable double-threshold policy which performs very well and have also derived exact and heuristic performance evaluation methods for systems using this policy. Our results indicate that the common practice of making maintenance and production decisions separately (or even consecutively, where inventory decisions take maintenance decisions into account but not *vice versa*) can be rather costly and that there are significant benefits to making these decisions in an integrated fashion. Finally, we have shown that collecting good data on how machine deterioration affects production rates and making use of models that can process this information is also critical in reducing total maintenance and inventory costs.

Acknowledgement

We would like to thank Professor Van der Duyn Schouten for drawing our attention to the paper of Van der Duyn Schouten and Vanneste (1995).

References

Berg, M., Posner, M.J.M. and Zhao, H. (1994) Production-inventory systems with unreliable machines. *Operations Research*, **42**, 111–118.

Berk, E. and Moinzadeh, K. (2000) Analysis of maintenance policies for M machines with deteriorating performance. *IIE Transactions*, **32**, 433–444.

Das, T.K. and Sarkar, S. (1999) Optimal preventive maintenance in a production inventory system. *IIE Transactions*, **31**, 537–551.

Dekker, R. (1996) Application of maintenance optimization models: a review and analysis. *Reliability Engineering and System Safety*, **51**, 229–240.

Gallego, G. (1990) Scheduling the production of several items with random demands in a single facility. *Management Science*, **36**, 1579–1592.

Gallego, G. (1994) When is a base stock policy optimal in recovering cyclic schedules? *Naval Research Logistics*, **41**, 317–333.

Groenevelt, H., Pintelon, L. and Seidmann, A. (1992a) Production lot sizing with machine breakdowns. *Management Science*, **38**, 104–123.

Groenevelt, H., Pintelon, L. and Seidmann, A. (1992b) Production batching with breakdowns and safety stocks. *Operations Research*, **40**, 959–971.

Iravani, S. and Duenyas, I. (1999) Integrated maintenance and production control of a deteriorating production system. (Full version of current paper). Technical report, Industrial Engineering and Management Sciences Department, Northwestern University, Evanston, IL, 60201.

Ivy, J.S. and Pollock, S.M. (1999) Maintenance of deteriorating machines with probabilistic monitoring and silent failures. Technical report, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI 48109.

Lee, H.L. and Rosenblatt, M.J. (1987) Simultaneous determination of production cycle and inspection schedules in a production system. *Management Science*, **33**, 1125–1136.

Lippman, S. (1975) Applying a new device in the optimization of exponential queueing systems. *Operations Research*, **23**, 687–710.

McCall, J.J. (1985) Maintenance policies for stochastically failing equipment: a survey. *Management Science*, **11**, 493–525.

Meyer, R.R., Rothkopf, M.H. and Smith, S.A. (1979) Reliability and inventory in a production-storage system. *Management Science*, **25**, 799–807.

Pierskalla, W.P. and Voelker, J.A. (1976) A survey of maintenance models the control and survilance of deteriorating systems. *Naval Research Logistics Quarterly*, **23**, 353–388.

Posner, M.J.M. and Berg, M. (1989) Analysis of a production–inventory system with unreliable production facility. *Operations Research Letters*, **8**, 339–345.

Sharafali, M. (1984) On a continuous review production–inventory problem. *Operations Research Letters*, **3**, 199–204.

Sharifnia, A. (1988) Production control of a manufacturing system with multiple machine states. *IEEE Transactions on Automatic Control*, **33**, 620–625.

Sherif, Y.S. and Smith, M.L. (1981) Optimal maintenance models for systems subject to failure: a review. *Naval Research Logistics Quarterly*, **28**, 47–74.

Srinivasan, M.M. and Lee, H.S. (1996) Production/inventory system with preventive maintenance. *IIE Transactions*, **28**, 879–890.

Valdez–Flores, C. and Feldman, R.M. (1989) A survey of preventive maintenance models for stochastically deteriorating single–unit systems. *Naval Research Logistics*, **36**, 419–446.

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

Biographies

Seyed Iravani is an Assistant Professor in the Department of Industrial Engineering and Management Sciences at Northwestern University. Dr. Iravani's publications have appeared in *Queueing Systems* and *Operations Research*. His research interests are in the stochastic modeling of manufacturing systems and the analysis and control of queueing systems.

Izak Duenyas is John Psarouthakis Professor of Manufacturing Management and Professor of Operations Management at the University of Michigan. He also holds a joint appointment in the Industrial Engineering Department. His research interests are in the modelling of complex production/inventory systems, supply chain management and investments in new technology. He currently serves on the Editorial Boards of *IIE Transactions, Operations Research* and *MSOM*.

Contributed by the Inventory Department