Simultaneous Determination of Production and Maintenance Schedules Using In-Line Equipment Condition and Yield Information

Thomas Sloan

Department of Manufacturing and Management Information Systems, University of Massachusetts Lowell, College of Management, Lowell, Massachusetts 01854

Abstract: In many manufacturing environments, equipment condition has a significant impact on product quality, or yield. This paper presents a semi-Markov decision process model of a single-stage production system with multiple products and multiple maintenance actions. The model simultaneously determines maintenance and production schedules, accounting for the fact that equipment condition affects the yield of each product differently. It extends earlier work by allowing the expected time between decision epochs to vary by both action and machine state, by allowing multiple maintenance actions, and by treating the outcome of maintenance as less than certain. Sufficient conditions are developed that ensure the monotonicity of both the optimal production and maintenance actions. While the maintenance conditions closely resemble previously studied conditions for this type of problem, the production conditions represent a significant departure from earlier results. The simultaneous solution method is compared to an approach commonly used in industry, where the maintenance and production problems are treated independently. Solving more than one thousand test problems confirms that the combination of both features of the model—accounting for product differences and solving the problems simultaneously—has a significant impact on performance. © 2007 Wiley Periodicals, Inc. Naval Research Logistics 55: 116–129, 2008

Keywords: maintenance; semi-Markov decision process; variable yield

1. INTRODUCTION

Equipment condition is an important determinant of product quality in a variety of manufacturing environments: in machine shops tool bits wear down and cause defects; in pharmaceutical production facilities, the ultra-clean equipment becomes contaminated and causes product spoilage. This article explores how information about equipment condition can be used to improve maintenance and production scheduling decisions and increase yield.

Consider, for example, the etch operation in semiconductor wafer fabrication in which chemicals are used to strip materials from the surface of silicon wafers. The inner chambers of the etch equipment become more and more contaminated as wafers are processed, and eventually one must stop production and clean the machine, returning it to an improved state. In recent years, the development of *in situ* particle monitors

and other devices has made it possible to assess the level of contamination while the equipment is in use. Experts have promoted the use of such devices to improve maintenance procedures and process control (see Hunter and Nguyen [16], Takahashi and Daugherty [41], and Lee and Tien [27], for example). However, little effort has been made to use such information in production scheduling decisions, despite the fact that the level of contamination may have a bigger impact on some products than on others. For example, leading-edge computer chips typically have smaller circuit sizes than more mature chips and thus are more susceptible to damage by particles in the equipment.

Only recently have researchers begun to explore the interaction between equipment condition and yield, and this research has focused almost exclusively on single-product systems, ignoring how equipment condition may have a different impact on the yield of different product types. In this article we develop a semi-Markov decision process model that simultaneously determines the production and maintenance schedules for a single-stage, multiple-product system, accounting for product differences. The equipment condition, which deteriorates over time, is characterized by a

discrete set of states. At each decision epoch, the decision maker can choose to produce one of several products or to perform one of several maintenance actions. There is uncertainty associated with each action, and the expected rewards or costs depend on the machine state. In addition, the time between decision epochs—marked by the completion of production or the completion of maintenance—is a random variable that depends on the machine condition and the action taken.

One goal of the article is to find conditions which are sufficient to ensure that a monotone policy is optimal, i.e., the optimal action increases as the machine condition gets worse. Products are indexed in descending order of their expected profit and their sensitivity to the equipment condition; maintenance actions are indexed in ascending order of their cost and their effectiveness. After establishing that the usual conditions sufficient for monotonicity of Markov decision process (MDP) problems are not adequate in this case, new monotonicity conditions are developed. While the results for the maintenance actions are closely related to the original MDP conditions, the conditions for the production actions, which incorporate the conditional probability of the machine deteriorating, are significantly different.

After exploring these structural properties, numerical test problems are used to study how much improvement can be achieved using the simultaneous solution method which uses product yield and equipment condition information. Traditionally, the maintenance and production schedules are determined sequentially. First, a maintenance problem is solved to determine the equipment deterioration level at which maintenance should be undertaken. Next, the production schedule is determined, without regard to the effect of the equipment condition on different products. We compare this traditional approach to three others: simultaneous solution with non-yield-based production scheduling, sequential solution with yield-based production scheduling, and simultaneous solution of the two problems with yield-based production scheduling. In the more than one thousand problems solved, the simultaneous, yield-based approach was an average of roughly 42 percent better than the traditional, sequential approach, roughly 27 percent better than the simultaneous, non-yield-based approach, and nearly 18 percent better than the sequential approach with yield-based dispatching. These results reinforce some conclusions of Sloan and Shanthikumar [38] but also demonstrate how the model extensions reveal additional insights.

The rest of the article is organized as follows. Section 2 presents an overview of the literature related to the problem described above. The details of the model and the structural results are presented in Section 3. Section 4 discusses different solution approaches, and Section 5 reports results of the numerical test problems. Conclusions and possible extensions are discussed in Section 6.

2. LITERATURE REVIEW

There is a vast body of research addressing problems of equipment maintenance, i.e., when to repair or replace equipment; see Valdez-Flores and Feldman [43] and Wang [45] for extensive reviews. The work most closely related to our problem allows for multiple repair actions, beginning with Kamien and Schwartz [20] and continuing with works such as Hopp and Wu [14], Su et al. [40], Ivy and Pollack [18], and Gilardoni and Colosimo [9]. However, these models ignore the effect that equipment condition may have on the yield or quality of the output.

Research on production management with equipment failures includes continuous flow control models (Lou et al. [28], Song [39]), models of discrete part production systems (Hong et al. [13], Taylor et al. [42]), lot-sizing models (Groenevelt et al. [10, 11]), and models of two-stage flowlines (Van der Duyn Schouten and Vanneste [44] Meller and Kim [30], Filliger and Hongler [7]). In all of these models, the interaction between production and maintenance is reduced to one dimension: Is the equipment working or not? Production models with intermediate machine states—i.e., between "up" and "down"—include single-product systems (Boukas et al. [2]) and multi-product systems (Dedopoulos and Shah [4]). Again, while the equipment condition may affect costs and throughput in these models, it does not affect yield.

Yano and Lee [47] present a comprehensive review of production models with variable yield. Most of the work in this area has focused on single-product systems and does not treat yield as a function of process condition. Rosenblatt and Lee [34] and Porteus [32] made the first connections in this area by extending the classical economic manufacturing quantity (EMQ) framework to incorporate imperfect process condition in the determination optimal production lot sizes. Various extensions to these seminal works have been pursued: allowing process inspections during the production cycle (Kim et al. [23], Wang [46]), studying different cost structures (Lee and Rosenblatt [25], Lee and Park [26]), considering the effects of machine failures (Makis and Fung [29], Ben-Daya [1], Sheu and Chen [36]), and allowing process improvements to be made (Freimer et al. [8]). Other attempts to examine the interaction between production, maintenance, and quality have been pursued using other, non-EMQ approaches (Iravani and Duenyas [17], Sloan [37], Kenné et al. [22]). All of these models examine *how much* to produce, but do not address the question of dispatching, or *which* product to process next.

In a similar vein, Nurani et al. [31] study a single-stage system in which the process can drift out of control, resulting in defective items. Although product quality depends on the equipment condition, the purpose of the model is to determine issuing policies, i.e., in what order finished goods are passed to the customer, rather than production dispatching policies.

Cunningham [3] does explicitly connect product yield to dispatching by showing that using in-line yield estimates to guide scheduling decisions can increase throughput, reduce flow time, and improve yield predictability. However, the process yield is not directly connected to the equipment condition.

The model most closely related to the current paper is that of Sloan and Shanthikumar [38], a Markov decision process model that simultaneously determines maintenance and production schedules for a single-stage, multiple-product system. While the model of [38] makes important steps towards linking the areas of equipment maintenance, product yield, and production scheduling, the model presented in the next section extends it in three important ways: allowing production and maintenance times to vary, including multiple maintenance actions, and incorporating uncertainty with respect to the maintenance action outcome. These extensions have a profound effect on the model's structural properties and the insights revealed by the model.

3. MODEL

3.1. Formulation

Using a semi-Markov decision process (SMDP) framework, we consider the problem of determining production and maintenance schedules for a single machine that produces multiple products. The condition of the equipment deteriorates over time, and this condition affects the yield of different products differently. The set of machine states is defined as $\mathcal{S} = \{0, 1, \ldots, J\}$, where state 0 is the best possible condition, and J is the worst condition. (Note that we are implicitly assuming that machine states can be ordered from better to worse. Referring to the earlier example, a higher state corresponds to an increased number of particles in the equipment.) At each decision epoch, the state of the machine is observed, and the decision maker has the option to produce one of K products or to stop production and perform one of M maintenance actions. The set of production actions is defined as $\mathcal{A}_{\mathbb{K}} = \{1, 2, \ldots, K\}$ and is indexed by k. The set of maintenance actions is defined as $\mathcal{A}_{\mathbb{M}} = \{K+1, K+2, \ldots, K+M\}$ and is indexed by m. The set of all possible actions is defined as $\mathcal{A} = \mathcal{A}_{\mathbb{K}} \cup \mathcal{A}_{\mathbb{M}}$, and the index a refers to any action in this set.

A new decision epoch occurs whenever the processing of a unit is completed or a maintenance action is completed. The time between decision epochs is a random variable that depends on the action taken at the current epoch. Define $F_{ia}(t)$ as the probability that the next decision epoch occurs within t time units of the current decision epoch given that action a is taken when the equipment is in state i, and define $\tau(i, a)$ as the mean of the distribution F_{ia}. In keeping with

the usual SMDP framework, the mean of F_{ia} is assumed to be non-zero and finite (i.e., $0 < \tau(i, a) < \infty$ for all $i \in \mathcal{S}$ and all $a \in \mathcal{A}$). This condition eliminates the possibility of an infinite number of decision epochs occurring in a finite period of time.

Define $P_{ij}(a)$ as the probability that the equipment will be in state j at the next decision epoch given that the decision maker chooses action a when the equipment is in state i at the current decision epoch. In general, producing causes the equipment state to increase (move to a higher level of deterioration) and performing maintenance causes the equipment state to decrease (move to a lower level of deterioration). However, performing maintenance does *not necessarily* return the equipment to state 0. Note that the equipment condition may change between decision epochs; however, the decision maker must wait until an action is completed to observe the state and choose the next action. Connections between the machine state transition probabilities and the time between decision epochs, while not explicit, can easily be incorporated. For example, longer expected production times may result in a higher likelihood of equipment deterioration, and the transition probabilities can reflect this type of relationship.

Let $r(i, a)$ denote the expected reward earned when the machine state is i and action a is taken at the current decision epoch. The structure of the reward will depend on whether a production action or a maintenance action is taken. For production actions, define R_k as the profit earned for a finished, working unit of product k. It is assumed that a new setup must be performed each time an item is produced, regardless of the product type, and any cost associated with this setup is accounted for in the R_k values. Let β_{ik} be the probability that product k works when production begins while the machine is in state i. Put differently, β_{ik} is the expected yield of product k, so the expected reward for choosing to produce product k in when the machine is in state i is $r(i, k) = R_k \beta_{ik}$.

For maintenance actions, the reward has a different structure. Most notably, the reward will be negative, so we will often refer to it as a cost when discussing maintenance actions. When the equipment is maintained, a fixed cost of $c(i, m)$ is incurred immediately, and a variable cost of $v(i, m)$ per unit time is incurred. The expected cost of performing maintenance action m when the equipment is in state i is expressed as $r(i, m) = c(i, m) + v(i, m)\tau(i, m)$.

Regardless of whether one chooses production or maintenance, the expected rewards are real-valued and bounded, i.e., $|r(i, a)| < \infty$ for each $i \in \mathcal{S}$ and each $a \in \mathcal{A}$. This property, which follows from the fact that the state space, action space, and the time between decision epochs are all finite, ensures that we do not have an infinitely profitable product nor an infinitely costly maintenance action, independent of the time.

Since the time between decision epochs is a random variable and changes in the machine state depend only on the current state and the action taken, a semi-Markov decision process framework is used to model the problem. Let X_n denote the machine state at the nth decision epoch ($n \geq 0$), and let a_n be the corresponding action taken. The associated expected reward earned is $r(X_n, a_n)$, and the expected time until the next decision epoch is $\tau(X_n, a_n)$.

Our objective is to determine a policy, or decision rule, that maximizes the long-run expected average reward. Let π denote a particular policy, i.e., a prescription for which action to take in each state. Define $\bar{g}_\pi(i)$ as the long-run expected average reward under policy π when the process begins in state i. The average reward $\bar{g}_\pi(i)$ is also referred to as the *gain* and is expressed as

$$\bar{g}_\pi(i) = \liminf_{N \to \infty} \frac{E_\pi \left[\sum_{n=0}^N r(X_n, a_n) \mid X_0 = i \right]}{E_\pi \left[\sum_{n=0}^N \tau(X_n, a_n) \mid X_0 = i \right]}, \quad (1)$$

where E_π denotes the expected value under policy π. An optimal policy, denoted as π^*, is a policy (from the set of all possible policies) for which $\bar{g}^*(i) = \sup_\pi \{\bar{g}_\pi(i)\}$ for each $i \in \mathcal{S}$. A *stationary* policy is a decision rule that is nonrandomized and does not depend on time, i.e., depends only on the current machine state.

The problem specified by these objects—$\mathcal{S}$, $\mathcal{A}$, $P_{ij}(a)$, $r(i, a)$, $\tau(i, a)$, and $\bar{g}$—is referred to as "the SMDP" throughout the rest of the article. Some preliminary results are needed before the existence of an optimal policy can be established and the structure of this policy characterized.

3.2. Chain Structure of the SMDP

The chain structure of the SMDP is determined by the machine state transition probabilities. Equipment deterioration is caused by production, so the machine state will not decrease when production is performed. Conversely, the equipment deterioration level only decreases when maintenance is performed, so the machine state will not increase when a maintenance action is taken. To avoid any ambiguity, we explicitly assume that once the equipment reaches the worst state, it cannot leave this state unless maintenance is performed. In addition, the machine cannot leave the best state unless production is performed. These conditions are summarized in the following assumption:

(A1)   a. For $i = 0, 1, \ldots, J-1$, $P_{ij}(k) > 0$, for at least one $j > i$ and $P_{ij}(k) = 0$ for all $j < i$, for each $k \in \mathcal{A}_{\mathbb{K}}$; in addition, $P_{i,i+1}(k) > 0$ for at least one $k \in \mathcal{A}_{\mathbb{K}}$;
   b. For $i = 1, 2, \ldots, J$, $P_{ij}(m) > 0$ for at least one $j < i$ and $P_{ij}(m) = 0$ for all $j > i$, for each $m \in \mathcal{A}_{\mathbb{M}}$.

By definition, $\sum_{j=0}^J P_{ij}(a) = 1$ for all $i \in \mathcal{S}$ and all $a \in \mathcal{A}$, so the above assumption implies that $P_{JJ}(k) = 1$ for all $k \in \mathcal{A}_{\mathbb{K}}$, and $P_{00}(m) = 1$ for all $m \in \mathcal{A}_{\mathbb{M}}$. Aside from these two special cases, assumption (A1) eliminates the possibility of an absorbing state. In addition, the requirement that $P_{i,i+1}(k)$ for at least one production action creates a potential connection between all states. Assumption (A1) is quite mild, especially when compared to the restrictions imposed in other models. For example, the model of Kao [21] assumes that $P_{ij} = 0$ for $j \leq i$ when maintenance is not performed and the models of Derman [6], Kao [21], and Sloan and Shanthikumar [38] all assume $P_{i0}(m) = 1$ for all i when maintenance is performed.

The fact that $\mathcal{S}$ is finite, combined with assumption (A1), ensures that the underlying Markov chain is communicating, which means that every state can eventually reach every other state in a finite number of transitions under *some* policy, even if this is not true for *every* policy. The following lemma summarizes this point:

LEMMA 3.1: If assumption (A1) is met, then the SMDP is communicating.

PROOF: The proof relies on Proposition 8.3.1a from [33, p. 350], which states that the (embedded) Markov decision process "is communicating if and only if there exists a randomized stationary policy which induces a recurrent Markov chain," i.e., a chain with a single closed irreducible class of states. Let m be a fixed maintenance action, and choose this action with probability one in state J and with probability 0.5 in all states $i < J$. Choose a production action k_i with $P_{i,i+1}(k_i) > 0$ with probability 0.5 in all states $i < J$. This randomized stationary policy results in a recurrent Markov chain, and therefore, the SMDP is communicating. □

3.3. Existence of a Stationary Optimal Policy

In this section, the existence of a policy that optimizes the long-run average reward (1) is established. Furthermore, we show that the optimal policy is stationary and deterministic. Following the approach of Sennott [35] and others, the average reward results are built on properties of the associated discounted reward problem. For a discount factor $\alpha > 0$, the discounted reward optimality equation is

$$V_\alpha(i) = \max_{a \in \mathcal{A}} \left\{ r(i, a) + \sum_{j=0}^J P_{ij}(a) \int_0^\infty e^{-\alpha t} dF_{ia}(t) V_\alpha(j) \right\}, \quad i \in \mathcal{S}. \quad (2)$$

The following lemma shows that the discounted reward problem has an optimal solution.

LEMMA 3.2: The optimality equation (2) has a unique solution. Furthermore, a stationary deterministic optimal policy exists.

PROOF: Since the rewards are bounded, the distribution of time between decision epochs is proper, and the state space is finite, the result follows immediately from Theorem 11.3.2 of Puterman [33].   □

Note that the above lemma makes no assumptions regarding the recurrence properties of the embedded Markov chain. Using this result, we may now show that a solution exists for the average reward model.

PROPOSITION 3.1: If assumption (A1) is met, then there exists a stationary deterministic optimal policy that satisfies the optimality equation

$$h(i) = \max_{a \in \mathcal{A}} \left\{ r(i,a) - g^* \tau(i,a) + \sum_{j=0}^{J} P_{ij}(a) h(j) \right\},$$
$$i \in \mathcal{S}, \quad (3)$$

where g^* is a constant such that $g^* = \lim_{\alpha \downarrow 0} \alpha V_{\alpha}(i)$ for all $i \in \mathcal{S}$. Furthermore, $\bar{g}^*(i) = g^*$ for all $i \in \mathcal{S}$.

The details of the proof appear in the Appendix. With this fundamental result in hand, we can begin exploration of the structural properties of the optimal solution.

3.4. Structural Properties

Insight into the structural properties can have important implications for both the solution of the problem and implementation of the results. Under what conditions will the optimal policy be monotone? That is, under what conditions will the optimal action be nondecreasing with respect to the equipment condition?

First, define $a(i)$ as the optimal action; to avoid ambiguity, let it denote the *smallest* action that optimizes a given problem. The following proposition identifies sufficient conditions for monotonicity.

PROPOSITION 3.2: Suppose that assumption (A1) holds and that the following conditions are met:

(C1) $r(i,a)$ is nonincreasing in i for all $a \in \mathcal{A}$;

(C2) $\sum_{j=l}^{J} P_{ij}(a)$ is nondecreasing in i for all $l \in \mathcal{S}$ and for all $a \in \mathcal{A}$;

(C3) $r(i,a)$ is superadditive; that is, $r(i,\hat{a}) - r(i,\bar{a})$ is nondecreasing in i for $\hat{a}, \bar{a} \in \mathcal{A}$ such that $\hat{a} > \bar{a}$;

(C4) For all $l \in \mathcal{S}$, $\sum_{j=l}^{J} P_{ij}(a)$ is subadditive; that is, $\sum_{j=l}^{J} [P_{ij}(\hat{a}) - P_{ij}(\bar{a})]$ is nonincreasing in i for all $l \in \mathcal{S}$ and all $\hat{a}, \bar{a} \in \mathcal{A}$ such that $\hat{a} > \bar{a}$;

(C5) $\tau(i,a)$ is subadditive; that is, $\tau(i,\hat{a}) - \tau(i,\bar{a})$ is nonincreasing in i for $\hat{a}, \bar{a} \in \mathcal{A}$ such that $\hat{a} > \bar{a}$.

Then there exists a deterministic stationary optimal policy such that the optimal action $a(i)$ is nondecreasing in i.

The proof appears in the Appendix. Conditions (C1) through (C4) are analogous to the well-known monotonicity conditions for MDPs (see, for example, Theorem 8.11.4 of [33] or Corollary 8.5a of Heyman and Sobel [12]), and (C5) is new. The interpretation of the conditions will be discussed briefly with respect to a maintenance control limit, the monotonicity of maintenance actions, and the monotonicity of production actions. Throughout the remainder of the article, the terms "increasing" and "nondecreasing" are used interchangeably, as are the terms "decreasing" and "nonincreasing."

3.4.1. Maintenance Control Limit

Building on the discounted-reward MDP results of Hopp and Wu [14], we first we establish the existence of a maintenance *control limit*, i.e., a threshold machine state at (or above) which it is optimal to perform maintenance:

COROLLARY 3.1: If conditions (C1) through (C5) are met for $\hat{a} \in \mathcal{A}_{\mathbb{M}}$ and $\bar{a} \in \mathcal{A}_{\mathbb{K}}$, and assumption (A1) holds, then there exists a threshold state, $\hat{\imath}$, such that the optimal action $a(i) > K$ for $i \geq \hat{\imath}$, and $a(i) \leq K$ for $i < \hat{\imath}$.

The result follows immediately from Proposition 3.2. For production actions, Condition (C1) means that the expected yield for each product, β_{ik}, decreases as the machine state gets worse. Similarly, the expected maintenance costs increase as the equipment condition deteriorates. Condition (C2) means that for production actions, as the equipment condition gets closer to state J, it becomes increasingly likely that state J will be reached. Condition (C2) also means that performing maintenance is more challenging in a worse state than in a better state, so for a given maintenance action, transitions to "good" states become less likely as the equipment condition gets worse.

Condition (C3) means that for a given increase in equipment deterioration, the decrease in expected profit for production is greater than the increase in expected cost for maintenance. Condition (C4) requires that the difference between failure rates for maintenance actions and production actions

decreases as the machine state gets worse. If performing maintenance does not provide an increasing improvement compared to non-maintenance, then (C4) will not be met, and there is possibility that it will be optimal to switch from production to maintenance and then back to production. Condition (C5) means that for a given increase in machine state, the difference between the expected time to perform maintenance and the expected time to produce a product decreases.

3.4.2.   *Monotonicity of Maintenance Actions*

Now we identify conditions sufficient to ensure that the optimal maintenance action is nondecreasing in the machine state. First, the maintenance actions are arranged in increasing order of effectiveness and cost; that is, a maintenance action with a higher index is more likely to move the equipment condition to a better state but also costs more to apply. For maintenance actions $m \in \mathcal{A}_{\mathbb{M}}$, we have

- (A2) $r(i, m)$ is nonincreasing in m for all $i \in \mathcal{S}$.
- (A3) $\sum_{j=0}^{l} P_{ij}(m)$ is nondecreasing in m for all i, $l \in \mathcal{S}$.

We can now state the following corollary:

COROLLARY 3.2: If conditions (C1) through (C5) are met for $\hat{a}, \bar{a} \in \mathcal{A}_{\mathbb{M}}$, and assumptions (A1) through (A3) hold, then the optimal action, $a(i)$, is nondecreasing in i for $a(i) > K$.

This result, which follows immediately from Proposition 3.2, means that once the maintenance threshold is reached, progressively more effective (and more costly) maintenance actions will be applied as the equipment condition gets worse.

Conditions (C1) and (C2) are discussed above. In this context, Condition (C3) means that the difference between expected rewards for more- and less-effective maintenance actions increases as the equipment condition gets worse—for a given increase in equipment deterioration, the expected cost of a higher-index maintenance action increases *less* than the cost of a lower-index maintenance action. Condition (C4) can be reframed as: $\sum_{j=0}^{l}[P_{ij}(\hat{m}) - P_{ij}(\bar{m})]$ is nondecreasing in i, where $\hat{m}, \bar{m} \in \mathcal{A}_{\mathbb{M}}$ such that $\hat{m} > \bar{m}$. This means that for a given increase in the machine state, the likelihood of a high-index maintenance action moving the equipment to an improved state decreases *less* than the likelihood of a low-index maintenance action moving the equipment to an improved state. Condition (C5) means that as the equipment deterioration level increases, the difference in the expected time required for high-index and low-index maintenance actions decreases.

3.4.3.   *Monotonicity of Production Actions*

Now we address the issue of production policies for which the optimal production action is nondecreasing in the machine state. To do so, production actions are arranged in descending order of profit. In addition, the yield for each product decreases as the machine condition gets worse, and for a given machine state, the yield is higher for lower-profit products than for higher-profit products. These assumptions, which are in keeping with the semiconductor manufacturing scenario discussed earlier, are expressed as

- (A4) R_k is nonincreasing in k for all $k \in \mathcal{A}_{\mathbb{K}}$, and
- (A5) β_{ik} is nonincreasing in i for any fixed $k \in \mathcal{A}_{\mathbb{K}}$ and nondecreasing in k for any fixed $i \in \mathcal{S}$.

These assumptions require that products with lower indices are more sensitive to the machine condition but earn higher profits; however, they do *not* imply that $r(i, k)$ decreases in k.

Given that Sloan and Shanthikumar [38] showed that (C1) through (C4) are sufficient to ensure a monotone production policy in an MDP model when all products induce the same machine state transition probabilities and have the same mean production times, one might expect that conditions (C1) through (C5) would be sufficient for the SMDP case. However, this is not so. Indeed, condition (C4) is unlikely to be met by any transition probability matrices that also meet assumption (A1). The following proposition, which represents the main analytical result of the article, identifies new conditions sufficient to ensure the monotonicity of production actions:

PROPOSITION 3.3: Suppose that assumptions (A1), (A4), and (A5) hold, conditions (C1) and (C2) are met, and the following conditions are met:

- (C3$'$) $\dfrac{r(i, \hat{k})}{[1 - P_{ii}(\hat{k})]} - \dfrac{r(i, \bar{k})}{[1 - P_{ii}(\bar{k})]}$ is nondecreasing in i,
- (C4$'$) $\dfrac{\sum_{j=i+1}^{l} P_{ij}(\hat{k})}{[1 - P_{ii}(\hat{k})]} - \dfrac{\sum_{j=i+1}^{l} P_{ij}(\bar{k})}{[1 - P_{ii}(\bar{k})]} \leq 0$ and is nondecreasing in i, where $i < l \leq J$, and
- (C5$'$) $\dfrac{\tau(i, \hat{k})}{[1 - P_{ii}(\hat{k})]} - \dfrac{\tau(i, \bar{k})}{[1 - P_{ii}(\bar{k})]}$ is nonincreasing in i,

where $i = 0, 1, \ldots, J - 1$ and $\hat{k}, \bar{k} \in \mathcal{A}_{\mathbb{K}}$ such that $\hat{k} > \bar{k}$. Then the optimal action, $a(i)$, is nondecreasing in i for $a(i) \leq K$.

The proof appears in the Appendix. The new conditions clearly bear a strong resemblance to (C3) through (C5). The

expression $[1 - P_{ii}(k)]$ is the probability that the machine condition leaves state i when production action k is taken, and it is a key feature of the new conditions. For a particular state i, we can think of the decision process as either being in state i or outside of state i. The underlying machine state may make several "virtual" transitions from state i back to itself before it makes a "real" transition to another state, and this pattern may be repeated several times between decision epochs. The expression $1/[1 - P_{ii}(k)]$ is equal to the number of virtual transitions before the process makes a real transition to a state outside of i when production action k is taken (see Howard [15, p. 632]). So $r(i, k)/[1 - P_{ii}(k)]$ can be thought of as the reward *rate* earned for choosing to produce product k in state i. Condition (C3′) therefore means the difference between more-sensitive and less-sensitive products with respect to the reward rate increases as the equipment condition gets worse.

Similarly, $P_{ij}(k)/[1 - P_{ii}(k)]$ is the *conditional* machine state transition probability, and condition (C4′) requires that the difference between high-index and low-index products with respect to the conditional failure rate increases as the machine condition gets worse. In addition, the conditional failure rate is higher for more-sensitive products, and therefore the difference between high- and low-index products will be less than or equal to zero.

The expression $\tau(i, k)/[1 - P_{ii}(k)]$ refers to the mean duration of time that the process spends in state i before making a real transition to another state when product k is produced and is often called the *holding time* [15, p. 632]. Condition (C5′) can therefore be interpreted as follows: the mean holding time difference between higher-index (less-sensitive) and lower-index (more-sensitive) products decreases as the machine state gets higher.

The fact that the optimal action is nondecreasing with respect to production simply means that we will produce more-sensitive products when the machine condition is good, and as the machine condition deteriorates, we will produce less-sensitive products. Eventually, the machine condition will reach the maintenance threshold, and we will stop production and perform maintenance.

4. SOLUTION METHOD

4.1. Linear Programming Solution

Many solution procedures are available for unichain SMDPs; however, such procedures may not work when there is a possibility of multiple recurrent classes, as is the case here (Puterman [33]). One procedure that does work for multichain SMDPs such as this one is the linear programming method of Denardo and Fox [5]. Denoting $r(i, a)$ as the expected reward for taking action a when the equipment is in state i and maximizing $\sum_{i \in \mathcal{S}} \sum_{a \in \mathcal{A}} r(i, a) x_{ia}$ subject to

constraints on the state frequencies, yields an average-reward optimal policy with objective function value equal to g^* from Eq. (3). In this framework, the decision variable x_{ia} can be thought of as the fraction of decision epochs that action a is being taken while the system is in state $i \in \mathcal{S}^x$, where $\mathcal{S}^x$ is the set of recurrent states.

It may be of interest to add production targets to the model presented in the last section, and doing so requires the addition of constraints to the linear program (LP) described above. The addition of constraints to a multichain SMDP linear program presents a non-trivial challenge due to the possibility of multiple closed sets of states. The linear programming approach proposed by Krass and Vrieze [24] is adapted here to address this issue. The LP is as follows:

$$\textbf{maximize} \quad \sum_{i^1 \in \mathcal{S}^1} \sum_{a \in \mathcal{A}} r(i^1, a) x_{i^1 a} \tag{4}$$

$$\textbf{subject to} \quad \sum_{a \in \mathcal{A}} x_{j^1 a} - \sum_{i^1 \in \mathcal{S}^1} \sum_{a \in \mathcal{A}} P_{i^1 j^1}(a) x_{i^1 a} = 0$$
$$\text{for all } j^1 \in \mathcal{S}^1, \tag{5}$$

$$y_{j^0 \delta} + \sum_{a \in \mathcal{A}} y_{j^0 a} - \sum_{i^0 \in \mathcal{S}^0} \sum_{a \in \mathcal{A}} P_{i^0 j^0}(a) y_{i^0 a} = \alpha_{j^0}$$
$$\text{for all } j^0 \in \mathcal{S}^0, \tag{6}$$

$$y_{j^0 \delta} - \sum_{a \in \mathcal{A}} \tau(j, a) x_{j^1 a} = 0 \quad \text{for all } j^0 \in \mathcal{S}^0, \tag{7}$$

$$x_{i^1 a} \geq 0 \quad \text{for all } i^1 \in \mathcal{S}^1, \text{ and all } a \in \mathcal{A},$$
$$y_{i^0 a} \geq 0 \quad \text{for all } i^0 \in \mathcal{S}^0, \text{ and all } a \in \mathcal{A}, \tag{8}$$
$$\text{and } y_{i^0 \delta} \geq 0 \quad \text{for all } i^0 \in \mathcal{S}^0,$$

where the superscript 1 indicates the real machine states, and the superscript 0 signifies a "virtual" machine state. The process begins in the virtual state j^0 with probability α_{j^0}. In addition to the previously defined actions $a \in \mathcal{A}$, the action δ indicates the one-way (irreversible) switch from the virtual to the real states. A solution to this LP yields the same objective function value as the LP described previously but does so in a way that ensures a separation between the transient and recurrent states, which allows constraints on the long-term (recurrent) behavior of the process to be enforced. The linear program specified by (4) through (8) is referred to as LP-1.

Now consider the addition of a production target, meaning that over the long run a certain proportion of total production must consist of a particular product type. Let γ_k be the long-run proportion of product k required, where $k \in \mathcal{A}_{\mathbb{K}}$, and $\sum_{k \in \mathcal{A}_{\mathbb{K}}} \gamma_k = 1$. The fraction of decision epochs where the embedded Markov chain is in state i implementing action a is equal to x_{ia}, where $i \in \mathcal{S}^x$ and $a \in \mathcal{A}$. The expected yield for product k when production begins in state i is β_{ik}. Weighting the fraction of decision epochs where production is performed by the expected yield of the different

products tells us what fraction of good output is made up of product k:

$$\frac{\sum_{i \in \mathcal{S}} x_{ik} \beta_{ik}}{\sum_{k \in \mathcal{A}_{\mathbb{K}}} \sum_{i \in \mathcal{S}} x_{ik} \beta_{ik}} = \gamma_k \quad \text{for each } k \in \mathcal{A}_{\mathbb{K}}. \quad (9)$$

These K constraints, rewritten in linear form and added to LP-1, constitute was it referred to as LP-2. Also note that no claims about the structural properties of optimal policy for the production-constrained problem are being made here.

4.2. Solution Approaches: Different Timing and Yield Information

In the next section, test problems are presented that explore the impact of solving the maintenance and production aspects of this problem sequentially versus simultaneously and solving the problems with yield information versus without yield information. The three approaches employed by Sloan and Shanthikumar [38] with the more basic model are used here, as well as a new approach which highlights the influence of yield information. We seek to reveal how much better the simultaneous, yield-based solution method is and whether the improvement comes more from the simultaneity of the solution or the use of yield information. The four approaches used are discussed below.

4.2.1. Approach 1 (sequential, non-yield-based dispatch)

Two problems are solved. First, a maintenance schedule is determined, ignoring product differences. The problem is reduced to choosing between "produce" and "maintain" for each state. Next, a production schedule is determined. When producing, the next item to process is picked without regard to the equipment condition. In the long run, this is equivalent to choosing product k with probability γ_k since that is the proportion of finished output desired. For example, if there are two products produced, and they are required in equal proportions, then there is a 50–50 chance of choosing each product at the next decision epoch. The expected reward earned for the "produce" action is a function of the expected reward for each product weighted by the production ratios. Denote this weighted expected reward as $\bar{r}(i) = \sum_{k \in \mathcal{A}_{\mathbb{K}}} \gamma_k R_k \beta_{ik}$. Under Approach 1, $\bar{r}(i)$ will be earned for each $k \in \mathcal{A}_{\mathbb{K}}$. This procedure will produce equivalent results—in terms of yield—as any other dispatching policy that ignores the equipment condition, such as first-come, first-serve (FCFS).

The maintenance threshold is determined by solving LP-1, using the $\bar{r}(i)$ values in the objective function (4). Once the maintenance threshold has been established, the production schedule is determined by solving LP-2, requiring the production ratios be attained and also that the maintenance threshold

be enforced. This procedure is referred to as the "traditional approach."

4.2.2. Approach 2 (simultaneous, non-yield-based dispatch)

The maintenance and production problems are determined simultaneously by solving LP-2 using the $\bar{r}(i)$ values in the objective function for each $k \in \mathcal{A}_{\mathbb{K}}$. Like Approach 1, Approach 2 does not account for yield differences in deciding which product to produce.

4.2.3. Approach 3 (sequential, yield-based dispatch)

The maintenance threshold is taken to be the one determined using Approach 1. A production schedule is determined by solving LP-2 with the original objective function coefficients (i.e., accounting for the product differences) and with the constraint that the maintenance threshold be enforced. Approaches 1 and 3 therefore have the same maintenance policy, but the production policy of Approach 3 will take yield differences into account.

4.2.4. Approach 4 (simultaneous, yield-based dispatch)

The maintenance and production schedules are determined simultaneously by solving LP-2 using the original objective function coefficients.

The results of the test problems are presented in the next section.

5. NUMERICAL RESULTS

The monotonicity conditions for this model are significantly different from those reported in previous research, and this section examines the practical effects of incorporating more in-line equipment condition and yield information in production and maintenance scheduling decisions by reporting the results of some numerical test problems.

5.1. Detailed Examples

Some simplified examples are presented to illustrate how all of the pieces of the model fit together. Consider a problem with two products (actions 1 and 2), two maintenance actions (actions 3 and 4), and four machine states (in increasing order of deterioration, denoted as states $\mathcal{S} = \{0, 1, 2, 3\}$). For each $i \in \mathcal{S}$, the mean production times are $\tau(i, 1) = 2$ and $\tau(i, 2) = 1$, and the mean maintenance times are $\tau(i, 3) = 1$ and $\tau(i, 4) = 2$, respectively. The machine condition deteriorates while production is being performed, and since $\tau(i, 1) > \tau(i, 1)$, the likelihood of going to a worse state

increases, as shown in the following machine state transition probability matrices:

$$[P_{ij}(1)] = \begin{bmatrix} 0.35 & 0.35 & 0.2 & 0.1 \\ 0 & 0.35 & 0.35 & 0.3 \\ 0 & 0 & 0.35 & 0.65 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$[P_{ij}(2)] = \begin{bmatrix} 0.6 & 0.215 & 0.075 & 0.11 \\ 0 & 0.6 & 0.215 & 0.185 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Action 3 can be thought of as "minor" maintenance and action 4 as "major" maintenance, and the corresponding machine state transition probabilities are

$$[P_{ij}(3)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.7 & 0.30 & 0 & 0 \\ 0.7 & 0.25 & 0.05 & 0 \\ 0.7 & 0.25 & 0.025 & 0.025 \end{bmatrix}$$

$$[P_{ij}(4)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The profit for a finished, working unit of product 1 is $R_1 = 600$ and for product 2 is $R_2 = 300$. The yield values for product 1 are $\beta_{01} = 0.5$, $\beta_{11} = 0.15$, $\beta_{21} = 0.05$, and $\beta_{31} = 0$. For product 2 the yields are $\beta_{02} = 1$, $\beta_{12} = 0.8$, $\beta_{22} = 0.7$, and $\beta_{32} = 0$. These values illustrate the situation in which one product has a higher profit but lower yield and longer processing time than the other. For both products, the yield decreases as the machine condition deteriorates, but the drop-off is much steeper for product 1 than for product 2.

The variable maintenance cost is 25 for both maintenance actions: $v(i, m) = -25$ for each i and both m. The fixed cost for minor maintenance starts at 225 and increases by 25 per state as the machine condition deteriorates: $c(0, 3) = -225$ and $c(i, 3) = c(0, 3) - 25i$ for $i = 1, 2, 3$. The fixed cost for major maintenance is $c(i, 4) = -500$ for each i.

All of the conditions for Proposition 3.3 are met, and optimal policy is monotone: $\pi^* = [2, 2, 2, 3]$, meaning produce product 2 in states 0, 1, and 2, and perform minor maintenance in state 3. Incorporating production targets will result in a different optimal policy. For the remainder of the example, suppose that equal proportions of each product are needed: $\gamma_1 = \gamma_2 = 0.5$.

The results of the this and two other example problems are reported in Table 1. In Example 1, the Approach 1 policy calls for production in the first three states and minor maintenance in the last state. (The notation "1/2" in the table indicates that both products 1 and 2 are being produced.) As the results in the table show, solving the problems simultaneously without yield information (Approach 2), adjusting the dispatching policy (Approach 3), and solving the problems simultaneously with yield information (Approach 4), produce increasingly greater average rewards. Treating the maintenance and production dimensions in isolation—as is traditionally the case—can produce very poor results. This insight is not just of theoretical interest: The author is familiar with several semiconductor manufacturers making use of particle monitors and the like to improve maintenance decisions but ignoring the potential to use this information for production scheduling.

Two other examples are considered. For Example 2, suppose that mean minor maintenance time is $\tau(i, 3) = 1.5$ for each i. The results in Table 1 demonstrate that this small change has a big impact—particularly with respect to Approach 3. For Example 3, suppose that the mean minor maintenance time is $\tau(i, 3) = 0.25$ for each i. Again, this seemingly minor change greatly influences the results. In these examples, solving the problems simultaneously yields some improvement, but using yield information in the production decisions has a bigger impact.

5.2. Overview of Test Problems

For the larger numerical study, we examine a single-stage system that produces four products (labeled $k = 1, 2, 3, 4$) and has five machine states (labeled $i = 0, 1, 2, 3, 4$). Two maintenance actions are available: "minor" maintenance

Table 1. Summary of detailed example results.

Machine state	Example 1				Example 2				Example 3			
	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4
0	1/2	1/2	1	1	1/2	1/2	1	1	1/2	1/2	1/2	1
1	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	2/3
2	1/2	1/2	1/2	4	1/2	1/2	1/2	4	3	3	3	3
3	3	4	3	4	4	4	4	4	3	3	3	3
Reward	32.1	40.8	42.4	57.0	38.8	40.8	52.5	57.0	48.2	48.2	48.2	67.2
% Imp. over A1	—	27.2	32.1	77.9	—	5.0	35.2	46.9	—	0	0	40.8

Table 2. Summary of test problem results.

| | | Improvement (%) | | |
	All problems	A2/A1	A3/A1	A4/A1
	Average	9.5	19.8	42.5
	Maximum	474.3	612.6	774.0
	Minimum	0.0	0.0	0.0
Factor	Level			
Deterioration rate	low	1.7	15.9	23.0
	med.	3.5	20.9	41.8
	high	23.3	22.6	62.8
Production times	low	7.3	35.9	36.5
(increase per state)	med.	14.2	11.2	42.1
	high	7.1	12.2	49.0
Minor maint. base mean	1.0	10.2	16.7	39.3
time, $\tau(0, K+1)$	1.5	8.8	22.9	45.8
Major maint. base mean	1.5	9.7	19.6	42.4
time, $\tau(0, K+2)$	2.0	9.3	20.0	42.7
Unit profits	100	23.6	37.9	93.8
(% difference)	150	2.7	14.2	21.8
	200	2.2	7.3	12.0
Minor maint. fixed cost	5	10.0	19.0	41.0
($ increase per state)	10	9.0	20.6	44.1
Major maint. fixed cost	0	9.5	19.8	42.5
($ increase per state)	5	9.5	19.8	42.5
Expected yields	0	0.4	0.0	0.6
	1	8.8	28.1	62.6
	2	19.4	31.3	64.5

Each line of the section labeled "Factor" reports the average improvement attained when the specified factor is held constant at the level indicated. Details of the parameter values are reported in Appendix B (available as an on-line supplement).

and "major" maintenance. Thus, we can choose one of six actions at each decision epoch. The required output is the same for all products, i.e., that $\gamma_k = 0.25$ for all k. Other parameters values are chosen to reflect the semiconductor manufacturing environment which motivated the study. Each model parameter (factor) is tested at several different values (levels), while holding the other parameters fixed. The factors and levels are summarized in Table 2, which also reports the test results. The details of the factors and levels, including a discussion of how they were generated, are reported in Appendix B, which is available as an on-line supplement. All problems meet the conditions of Proposition 3.3.

5.3. Summary of Results

Table 2 summarizes the results of the test problems; it reports the percentage improvement provided by Approach 2 over Approach 1 (A2/A1), Approach 3 over Approach 1 (A3/A1), and Approach 4 over Approach 1 (A4/A1). The section labeled "All Problems" at the top of the table lists summary results for the entire set of 1,296 test problems. Each line

of the section labeled "Factor" reports the average improvement attained when the specified factor is held constant at the level indicated. Over all, the results mirror the pattern observed in the example problems above. Specifically, solving the maintenance and production problems simultaneously is beneficial, but using yield information for dispatching is more important. In combination, i.e., solving the problems simultaneously and using yield information, substantially improves performance. While not reported explicitly, the A4/A2 improvement is roughly equal to the difference between A4/A1 and A2/A1, and the A4/A3 improvement is roughly equal to the difference between A4/A1 and A3/A1.

The effect of individual parameters can be seen in the results reported in Table 2. In summary, differences in the equipment deterioration rate, spread of the mean production times, unit profit variance, and expected yields all have a significant impact on how much better Approach 4 performs. These results seems logical given that these parameters are at the heart of the differences between products. Approaches 1 and 2 can effectively adjust to changes in maintenance-related parameters, as one might expect. And when the Approach 1 maintenance policy is reasonably good, then Approach 3 can make up for some—but not all—of the reward forgone by Approach 1 by adjusting the dispatching policy. When the differences between products are pronounced, then Approaches 1 and 2 typically do not fare well. When Approach 1 generates a maintenance policy that is far from optimal, then there is little opportunity for Approach 3 to make improvements based on production, and thus Approach 4 is substantially better.

While the model differences make direct comparison difficult, these results generally echo the findings of Sloan and Shanthikumar [38]. Specifically, the equipment deterioration rate and spread of the profits were found to have a significant impact on the A4/A1 differences in both studies. By including Approach 2 (simultaneous, non-yield-based), we can see that using equipment condition and yield information has a bigger effect than solving the problems simultaneously. However, the current study reveals much bigger Approach 4 improvements over Approach 3 than reported in [38]. This finding suggests that the model extensions—i.e., accounting for differences in expected production times, allowing multiple maintenance actions, and incorporating uncertainty with respect to maintenance action transition probabilities— amplify the effects of solving the production and maintenance problems simultaneously.

While it may seem obvious at first glance that using more information about the machine state is better than ignoring it, this study shows that the current practice of gathering equipment condition information solely for maintenance scheduling or yield prediction is a potentially severe mistake. The magnitude of improvement suggested by these test problems is extremely significant, especially for a

highly competitive industry like semiconductor manufacturing.

6. CONCLUSIONS

In many manufacturing systems, the process or equipment condition has an important effect on product quality, and this article explores the effects of incorporating equipment condition information into decisions about production and maintenance scheduling. The article makes three contributions. First, it extends the MDP model of Sloan and Shanthikumar [38] to a more realistic environment by using an SMDP model. Specifically, the decision maker can choose among several maintenance actions as well as several production actions, the time to complete an action can vary by both action and by state, and the effects of performing maintenance are uncertain, i.e., the equipment does not necessarily return to the best state.

Second, the article reaffirms and extends the lessons learned in [38]: simultaneously solving the production and maintenance scheduling problems using available equipment condition and yield information substantially improves performance. In the 1,296 test problems studied, the simultaneous, yield-based approach generated rewards that were an average of more than 42% greater than the sequential, non-yield-based approach; roughly 27% greater than the simultaneous, non-yield-based approach; and approximately 18% greater than the sequential approach with yield-based dispatching. These results indicate that the use of equipment condition and yield information has a bigger impact than solving the problems simultaneously. The fact that the improvements are even greater in a richer and more realistic setting than that of [38] is noteworthy.

Third, the article develops new conditions sufficient to ensure that a monotone policy is optimal, i.e., the optimal actions are nondecreasing with respect to the equipment condition. The "usual" monotonicity conditions for MDP models are extended to SMDP models by adding a condition to account for time differences between actions. After observing that these conditions generally do not hold for the production decisions in this model, new conditions are developed. These new conditions, which are based on conditional transition probabilities, exploit the structure of the machine state transition probabilities associated with the production actions, and hold for a much broader range of problems.

APPENDIX: PROOFS OF MAIN RESULTS

A.1   Proof of Proposition 3.1

Corollary 5.2 of Jianyong and Xiaobo [19] shows that for (weakly) communicating SMDPs, a stationary deterministic optimal policy exists and

$\bar{g}^*(i)$ is constant for all $i \in \mathcal{S}$. Since the state and action spaces are finite, and $\tau(i, a)$ and $r(i, a)$ are bounded for all $i \in \mathcal{S}$ and all $a \in \mathcal{A}$, the conditions of Theorem 11.4.8 of [33] are met, and we may conclude that optimality Eq. (3) holds and g^* is a constant such that $g^* = \lim_{\alpha \downarrow 0} \alpha V_\alpha(i)$ for all $i \in \mathcal{S}$.

Lastly, Theorem 9.1.6 of [33] shows that $\bar{g}^*(i) = g^*$ for all $i \in \mathcal{S}$ for multichain MDPs (including communicating ones) with finite state and action spaces. Proposition 4.1 of [19] shows that this MDP result carries over to (weakly) communicating SMDPs, and thus the result holds. This completes the proof.   □

A.2   Proof of Proposition 3.2

Preliminary Results for Proposition 3.2

Some preliminary results are needed for the proof of Proposition 3.2.

LEMMA A.1: If assumption (A1) is met and conditions (C1) and (C2) hold, then $h(i)$ is bounded and nonincreasing in i.

PROOF: Define $V_n(i)$ as the finite-horizon version of the value function in the discounted optimality equation (2), where i is the initial machine state, and n refers to the number of periods remaining. (Note that $V_n(i)$ also depends on α; however, for notational convenience, the α is omitted.) By condition (C1), $r(i, a)$ is nonincreasing in i, so $V_0(i)$ is nonincreasing in i. Suppose that $V_n(i)$ is nonincreasing in i for $n = 1, 2, \ldots, N-1$. The optimality equation is

$$V_N(i) = \max_{a \in \mathcal{A}} \left\{ r(i, a) + \sum_{j=0}^{J} P_{ij}(a) \int_0^\infty e^{-\alpha t} dF_{ia}(t) V_{N-1}(j) \right\}.$$

By condition (C2), $\sum_{j=l}^{J} P_{ij}(a)$ is nondecreasing in i. Since $V_n(i)$ is nonincreasing in i, $V_N(i)$ is also nonincreasing in i (see Lemma 1 of Derman [6]). Since the expected rewards are bounded and the time between decision epochs is bounded, Theorem 6.3.1 of Puterman [33] can be applied to show that $V(i) = \lim_{N \to \infty} V_N(i)$. Therefore, $V(i)$ is nonincreasing in i. Since all expected costs and times are bounded, the number of machine states is finite, and the number of actions is finite, $V(i)$ is bounded. Note that $h(i) \equiv V(i) - V(z)$, where $V(i)$ is the optimal discounted reward given that the process begins in state i, and z is any distinguished state [33]. Since $V(z)$ is constant with respect to i, and $V(i)$ is bounded and nonincreasing, then $h(i)$ is bounded and nonincreasing in i.   □

PROOF OF PROPOSITION 3.2: The proof builds on the MDP results of Theorem 8.11.4 of Puterman [33]. The objective is to show that if the optimal action $a(i) > a^*$ for some $i \in \mathcal{S}$, then $a(l) > a^*$ for all $l > i$. Suppose that $a(i) = \hat{a}$ but $a(l) = \bar{a}$ for some $l > i$, where $\hat{a} > \bar{a}$. Since $a(i)$ is defined as the smallest action that maximizes (3), for state i we have

$$r(i, \hat{a}) - g^*\tau(i, \hat{a}) + \sum_{j=0}^{J} P_{ij}(\hat{a})h(j) > r(i, \bar{a}) - g^*\tau(i, \bar{a}) + \sum_{j=0}^{J} P_{ij}(\bar{a})h(j),$$

and for state l

$$r(l, \bar{a}) - g^*\tau(l, \bar{a}) + \sum_{j=0}^{J} P_{lj}(\bar{a})h(j) \geq r(l, \hat{a}) - g^*\tau(l, \hat{a}) + \sum_{j=0}^{J} P_{lj}(\hat{a})h(j).$$

Together, the above inequalities imply that

$$r(i,\hat{a}) - r(i,\bar{a}) - g^*[\tau(i,\hat{a}) - \tau(i,\bar{a})] + \sum_{j=0}^{J}[P_{ij}(\hat{a}) - P_{ij}(\bar{a})]h(j) >$$

$$r(l,\hat{a}) - r(l,\bar{a}) - g^*[\tau(l,\hat{a}) - \tau(l,\bar{a})] + \sum_{j=0}^{J}[P_{lj}(\hat{a}) - P_{lj}(\bar{a})]h(j).$$

Note that g^* cannot be less than the smallest of the largest earning rates over all of the states (Howard [15]). Since performing maintenance incurs a negative expected reward, and the yield for any product must be greater than or equal to zero, $r(i,k) \geq 0 \geq r(i,m)$ for all $i \in \mathcal{S}$, where $k \in \mathcal{A}_{\mathbb{K}}$ and $m \in \mathcal{A}_{\mathbb{M}}$; therefore, $g^* \geq 0$.

Since $r(i,\hat{a}) - r(i,\bar{a})$ is nondecreasing in i by condition (C3) and $\tau(i,\hat{a}) - \tau(i,\bar{a})$ is nonincreasing in i by condition (C5), we must have

$$\sum_{j=0}^{J}[P_{ij}(\hat{a}) - P_{ij}(\bar{a})]h(j) > \sum_{j=0}^{J}[P_{lj}(\hat{a}) - P_{lj}(\bar{a})]h(j).$$

However, condition (C4), together with Lemma A.1 and Lemma 3 of Hopp and Wu [14], imply that

$$\sum_{j=0}^{J}[P_{ij}(\hat{a}) - P_{ij}(\bar{a})]h(j) \leq \sum_{j=0}^{J}[P_{lj}(\hat{a}) - P_{lj}(\bar{a})]h(j).$$

Thus, we have a contradiction, and we can conclude that $a(i)$ is nondecreasing in i for $a(i) > K$. $\square$

A.3 Proof of Proposition 3.3

Preliminary Result for Proposition 3.3

The next lemma is needed for the proof of Proposition 3.3.

LEMMA A.2: If condition (C4′) holds, then

$$\sum_{j=i+1}^{J}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)f(j) \text{ is nondecreasing in } i$$

for any bounded, nonincreasing function f, for production actions $\hat{k}, \bar{k} \in \mathcal{A}_{\mathbb{K}}$ such that $\hat{k} > \bar{k}$.

PROOF: The proof builds on the proofs of the analogous lemmas of Derman [6] and Hopp and Wu [14]. First, note that any bounded, nonincreasing function $f(j)$ can be expressed as $f(j) = \sum_{l=0}^{J}\lambda_l f_l(j)$, where $\lambda_l \geq 0$ for $l = 0, 1, \ldots, J$, and $f_l(j) = 1$ for $j \leq l$ and 0 otherwise. Then

$$\sum_{j=i+1}^{J}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)f(j)$$

$$= \sum_{j=i+1}^{J}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)\sum_{l=0}^{J}\lambda_l f_l(j)$$

$$= \sum_{l=0}^{J}\lambda_l \sum_{j=i+1}^{l}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right).$$

Since $\lambda_l \geq 0$ and, by condition (C4′), $\sum_{j=i+1}^{l}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)$ is nondecreasing in i for each l, then the result follows. $\square$

PROOF OF PROPOSITION 3.3: The objective is to show that the optimal action $a(i)$ is nondecreasing in i for $a(i) \leq K$. Examining equation (3), we see that an $h(\cdot)$ term appears on both the right- and left-hand sides. Furthermore, $P_{ij}(k) = 0$ for $j < i$ for all production actions $k \in \mathcal{A}_{\mathbb{K}}$ by assumption (A1), so the optimality equation (3) can be rewritten as

$$h(i) = \frac{r(i,k) - g^*\tau(i,k)}{[1-P_{ii}(k)]} + \sum_{j=i+1}^{J}\frac{P_{ij}(k)}{[1-P_{ii}(k)]}h(j), \qquad (10)$$

for production actions $k \in \mathcal{A}_{\mathbb{K}}$.

Let us assume that a non-monotone policy is optimal, i.e., that $a(i) = \hat{k}$ for some $i < J - 1$, but $a(l) = \bar{k}$ for some $l > i$, where $\bar{k}, \hat{k} \in \mathcal{A}_{\mathbb{K}}$, and $\hat{k} > \bar{k}$. Since $a(i)$ is the smallest action that maximizes (10), for state i we have

$$\frac{r(i,\hat{k}) - g^*\tau(i,\hat{k})}{[1-P_{ii}(\hat{k})]} + \sum_{j=i+1}^{J}\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]}h(j)$$

$$> \frac{r(i,\bar{k}) - g^*\tau(i,\bar{k})}{[1-P_{ii}(\bar{k})]} + \sum_{j=i+1}^{J}\frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}h(j),$$

and for state l

$$\frac{r(l,\hat{k}) - g^*\tau(l,\hat{k})}{[1-P_{ll}(\hat{k})]} + \sum_{j=l+1}^{J}\frac{P_{lj}(\hat{k})}{[1-P_{ll}(\hat{k})]}h(j)$$

$$\leq \frac{r(l,\bar{k}) - g^*\tau(l,\bar{k})}{[1-P_{ll}(\bar{k})]} + \sum_{j=l+1}^{J}\frac{P_{lj}(\bar{k})}{[1-P_{ll}(\bar{k})]}h(j).$$

Together, these equations imply that

$$\left(\frac{r(i,\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{r(i,\bar{k})}{[1-P_{ii}(\bar{k})]}\right) - g^*\left(\frac{\tau(i,\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{\tau(i,\bar{k})}{[1-P_{ii}(\bar{k})]}\right)$$

$$+ \sum_{j=i+1}^{J}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)h(j) >$$

$$\left(\frac{r(l,\hat{k})}{[1-P_{ll}(\hat{k})]} - \frac{r(l,\bar{k})}{[1-P_{ll}(\bar{k})]}\right) - g^*\left(\frac{\tau(l,\hat{k})}{[1-P_{ll}(\hat{k})]} - \frac{\tau(l,\bar{k})}{[1-P_{ll}(\bar{k})]}\right)$$

$$+ \sum_{j=l+1}^{J}\left(\frac{P_{lj}(\hat{k})}{[1-P_{ll}(\hat{k})]} - \frac{P_{lj}(\bar{k})}{[1-P_{ll}(\bar{k})]}\right)h(j). \qquad (11)$$

However, condition (C3′) implies that the inequality goes the other way for the first term in parentheses on either side of (11). In addition, (C5′) along with the fact that $g^* \geq 0$ (as shown above in the proof of Proposition 3.2) implies that the inequality also goes the other way for the second term in parentheses on either side of (11). Together, these imply that

$$\sum_{j=i+1}^{J}\left(\frac{P_{ij}(\hat{k})}{[1-P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1-P_{ii}(\bar{k})]}\right)h(j)$$

$$> \sum_{j=l+1}^{J}\left(\frac{P_{lj}(\hat{k})}{[1-P_{ll}(\hat{k})]} - \frac{P_{lj}(\bar{k})}{[1-P_{ll}(\bar{k})]}\right)h(j).$$

However, condition (C4′), along with Lemmas A.1 and A.2 imply

$$\sum_{j=i+1}^{J} \left(\frac{P_{ij}(\hat{k})}{[1 - P_{ii}(\hat{k})]} - \frac{P_{ij}(\bar{k})}{[1 - P_{ii}(\bar{k})]} \right) h(j)$$

$$\leq \sum_{j=l+1}^{J} \left(\frac{P_{lj}(\hat{k})}{[1 - P_{ll}(\hat{k})]} - \frac{P_{lj}(\bar{k})}{[1 - P_{ll}(\bar{k})]} \right) h(j).$$

Thus, we have a contradiction, and we can conclude that $a(i)$ is nondecreasing in i for $a(i) \leq K$. $\square$

ACKNOWLEDGMENTS

REFERENCES

[1] M. Ben-Daya, The economic production lot-sizing problem with imperfect production processes and imperfect maintenance, Int J Prod Econ 76 (2002), 257–264.

[2] E.K. Boukas, J. Yang, Q. Zhang, and G. Yin, Periodic maintenance and repair rate control in stochastic manufacturing systems, J Optim Theory Appl 91 (1996), 347–361.

[3] S.P. Cunningham, Applications of Spatial Statistics to Semiconductor Wafer Defects, Ph.D. thesis, University of California, Berkeley, Department of Industrial Engineering and Operations Research (1995).

[4] I.T. Dedopoulos and N. Shah, Long-term maintenance policy optimization in multipurpose process plants, Chem Eng Res Des 74 (1996), 307–320.

[5] E.V. Denardo and B.L. Fox, Multichain Markov renewal programs, SIAM J Appl Math 16 (1968), 468–487.

[6] C. Derman, On optimal replacement rules when changes of state are Markovian, Mathematical Optimization Techniques, in: R. Bellman, editor, University of California Press, Berkeley, California, 1963, pp. 201–210.

[7] R. Filliger and M.-O. Hongler, Optimal threshold control for failure-prone tandem production systems, IIE Trans 37 (2005), 861–875.

[8] M. Freimer, D. Thomas, and J. Tyworth, The value of setup cost reduction and process improvement for the economic production quantity model with defects, Eur J Oper Res 173 (2006), 241–251.

[9] G.L. Gilardoni and E.A. Colosimo, Optimal maintenance time for repairable systems, J Qual Technol 39 (2007), 48–53.

[10] H. Groenevelt, L. Pintelon, and A. Seidmann, Production batching with machine breakdowns and safety stocks, Oper Research 40 (1992), 959–971.

[11] H. Groenevelt, L. Pintelon, and A. Seidmann, Production lot sizing with machine breakdowns, Manage Sci 38 (1992), 104–123.

[12] D.P. Heyman and M.J. Sobel, Stochastic models in operations research, Vol. II: Stochastic optimization, McGraw-Hill, New York, 1984.

[13] Y. Hong, C.R. Glassey, and D. Seong, The analysis of a production line with unreliable machines and random processing times, IIE Trans 24 (1992), 77–83.

[14] W.J. Hopp and S.-C. Wu, Machine maintenance with multiple maintenance actions, IIE Trans 22 (1990), 226–232.

[15] R.A. Howard, Dynamic probabilistic systems, Vol. II: Semi-Markov and decision processes, Wiley, New York, New York, 1971.

[16] J. Hunter and H.K. Nguyen, In-situ particle monitoring reduces wafer defects, Semicond Int 16 (1993), 80–84.

[17] S.M.R. Iravani and I. Duenyas, Integrated maintenance and production control of a deteriorating production system, IIE Trans 34 (2002), 423–435.

[18] J.S. Ivy and S.M. Pollock, Marginally monotonic maintenance policies for a multi-state deteriorating machine with probabilistic monitoring, and silent failures, IEEE Trans on Reliabil 54 (2005), 489–497.

[19] L. Jianyong and Z. Xiaobo, On average reward semi-Markov decision processes with a general multichain structure, Math Oper Res 29 (2004), 339–352.

[20] M.I. Kamien and N.L. Schwartz, Optimal maintenance and sale age for a machine subject to failure, Manage Sci 17 (1971), B495–B504.

[21] E.P.C. Kao, Optimal replacement rules when changes of state are semi-Markovian, Oper Res 21 (1973), 1231–1249.

[22] J.P. Kenné, A. Gharbi, and M. Beit, Age-dependent production planning and maintenance strategies in unreliable manufacturing systems with lost sale, Eur J Oper Res 178 (2007), 408–420.

[23] C.H. Kim, Y. Hong, and S.Y. Chang, Optimal production run length and inspection schedules in a deteriorating production process, IIE Trans 33 (2001), 421–426.

[24] D. Krass and O. Vrieze, Achieving target state-action frequencies in multichain average-reward Markov decision processes, Math Oper Res 27 (2002), 545–566.

[25] H.L. Lee and M.J. Rosenblatt, A production and maintenance planning model with restoration cost dependent on detection delay, IIE Trans 21 (1989), 368–375.

[26] J.S. Lee and K.S. Park, Joint determination of production cycle and inspection intervals in a deteriorating production system, J Oper Res Soc 42 (1991), 775–783.

[27] S. Lee and Y.-C. Tien, In situ and real-time monitoring of plasma process chamber component qualities and predictive controlling of wafer yields, IEEE Trans Semicond Manufact 19 (2006), 432–436.

[28] S. Lou, S. Sethi, and G. Sorger, Stability of real-time lot scheduling policies for an unreliable machine, IEEE Trans Automatic Control 37 (1992), 1966–1970.

[29] V. Makis and J. Fung, An EMQ model with inspections and random machine failures, J Oper Res Soc 49 (1998), 66–76.

[30] R.D. Meller and D.S. Kim, The impact of preventive maintenance on system cost and buffer size, Eur J Oper Res 95 (1996), 577–591.

[31] R.K. Nurani, S. Seshadri, and J.G. Shanthikumar, Optimal control of a single stage production system subject to random process shifts, Oper Res 45 (1997), 713–724.

[32] E.L. Porteus, Optimal lot sizing, process quality improvement and setup cost reduction, Oper Res 34 (1986), 137–144.

[33] M.L. Puterman, Markov decision processes: Discrete stochastic dynamic programming, Wiley, New York, 1994. [Corrections available at http://coe.ubc.ca/users/marty/].

[34] M.J. Rosenblatt and H.L. Lee, Economic production cycles with imperfect production processes, IIE Trans 18 (1986), 48–54.

[35] L.I. Sennott, Average cost semi-Markov decision processes and the control of queueing systems, Probabil Eng Inform Sci 3 (1989), 247–272.

[36] S.-H. Sheu and G.J.-A. Chen, Optimal lot-sizing problem with imperfect maintenance and imperfect production, Int J Syst Sci 35 (2004), 69–77.

[37] T.W. Sloan, A periodic review production and maintenance model with random demand, deteriorating equipment, and binomial yield, J Oper Res Soc 55 (2004), 647–656.

[38] T.W. Sloan and J.G. Shanthikumar, Combined production and maintenance scheduling for a multiple-product, single-machine production system, Production and Operations Management 9 (2000), 379–399.

[39] D.-P. Song, Optimal production and backordering policy in failure-prone manufacturing systems, IEEE Trans Automatic Control 51 (2006), 906–911.

[40] C.-T. Su, S.-C. Wu, and C.-C. Chang, Multiaction maintenance subject to action-dependent risk and stochastic failure, Eur J Oper Res 125 (2000), 133–148.

[41] K.M. Takahashi and J.E. Daugherty, Current capabilities and limitations of in situ particle monitors in silicon processing equipment, J Vacuum Sci Technol, A: Vacuum, Surfaces, Films 14 (1996), 2983–2993.

[42] B.W. Taylor, E.R. Clayton, and E.T. Grasso, Simulation of a production system with machine breakdowns using network modeling, Comput Oper Res 9 (1982), 255–264.

[43] C. Valdez-Flores and R.M. Feldman, A survey of preventive maintenance models for stocastically deteriorating single-unit systems, Naval Res Logist 36 (1989), 419–446.

[44] F.A. Van der Duyn Schouten and S.G. Vanneste, Maintenance optimization of a production system with buffer capacity, European J Oper Res 82 (1995), 323–338.

[45] H. Wang, A survey of maintenance policies of deteriorating systems, Eur J Oper Res 139 (2002), 469–489.

[46] W.-Y. Wang, The development of an integrated production maintenance and quality model, with positive inspection time lengths, Int J Manage 21 (2004), 481–491.

[47] C.A. Yano and H.L. Lee, Lot sizing with random yields: A review, Oper Res 43 (1995), 311–334.