A model for equipment replacement due to technological obsolescence *

Suresh K. Nair

Department of Operations and Information Management, School of Business, U41-IM, University of Connecticut, Storrs, CT 06269-0241, USA

Wallace J. Hopp

Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA

Received June 1990; revised January 1992

Abstract: We consider the problem of deciding whether to keep a piece of equipment or to replace it with a more advanced technology. This decision must take into account both the nature of the available replacement technology and the possibility of future technological advances. Existing models are restrictive in the way they model the appearance of future technologies and the costs and revenues associated with those technologies. In an earlier paper we allowed the probability of appearance of new technologies to be non-stationary in time but required the costs and revenues of technologies to be different but constant over time. In this paper, we allow the technology forecasts and revenue functions associated with technologies to be non-stationary in time and consider salvage values for technologies. We develop a simple and efficient algorithm for finding the optimal decision using a forecast horizon approach. This approach finds the optimal decision in any period with minimal reliance on forecast data.

Keywords: Equipment; innovation; investment; Markov decision; Dynamic Programming

1. Introduction

The replacement of productive equipment ranks among the most important strategic decisions faced by both manufacturing and service firms. This is because purchase of a new piece of equipment often involves a significant cost and can affect the productivity and competitiveness of the firm for several years into the future. In recent years, the difficulty of this problem has been compounded by the fact that technologies are rapidly changing, and what may appear to be a good equipment purchase can soon become obsolete. Under these circumstances, the driving motivation behind replacement decisions is likely to be technological obsolescence, rather than physical deterioration, of the existing equipment. This situation is typical of microcomputers, computerized numerically controlled machines, and other electronics technologies.

Much work has been done in the area of equipment replacement. For a detailed exposition, see the survey in Pierskalla and Voelker (1976).

The traditional approach to the equipment replacement problem emphasizes the physical deterioration of the existing equipment. The basic idea is to replace the equipment when the cost of operating and maintaining it become sufficiently high, in net expected present value terms, to justify a replacement.

Correspondence to: S.K. Nair, Department of Operations and Information Management, School of Business, U41-IM, University of Connecticut, Storrs, CT 06269-0241, USA.

^{*} This work was supported in part by the National Science Foundation under Grant No. ECS-8619732.

In most traditional models, technology is assumed to remain constant over time. These include Derman (1963) and Hatoyama (1984). This approach would lead to inappropriate decisions if technology does change, as is becoming increasingly common. For example, if technology is changing it may be better to keep the current equipment for a few more years and then take advantage of an improved technology that appears on the market than to replace the current equipment with what is available in the market at the present time. Typically, if the capital costs for purchasing the new equipment are not recouped by the time another advanced technology appears, it would be extremely difficult for decision makers to justify another replacement when the advanced technology appears.

There are a few models that do consider technological change. Traditionally improvements in technology are measured either in terms of increased revenue associated with the new technology or decreased costs of procurement and operation of the new technology. Most of these models assume that technology improves in a deterministic manner, that is, the extent (in terms of costs and revenues) and timing of the technological improvements is known beforehand (Sethi and Chand, 1979). Though assuming that technology improves in a deterministic manner is a good first approximation, these models do not accommodate the reality that technological improvements can rarely be forecasted with certainty.

Only recently have researchers started to model the equipment replacement problem due to technological change under uncertainty. Goldstein, Ladany and Mehrez (1986, 1988) introduce uncertainty using stationary forecasts (that is, forecasts that are constant over time). Hopp and Nair (1991) developed a model using non-stationary technology forecasts but where the revenue generated by various technologies is different but constant over time.

The models incorporating non-stationary technological forecasts use a forecast horizon procedure to find the optimal keep/replace decision. The *forecast horizon* is the minimum number of periods of forecasted information required to guarantee that the initial decision is optimal, regardless of forecasts in later periods. The reason forecast horizons are important is that it not only reduces the computation and forecasting necessary to arrive at an optimal decision, but also because it ensures that the optimal decision would be no different if data for additional periods were forecasted.

A forecast horizon procedure attempts to identify the forecast horizon as efficiently as possible. If a forecast horizon exists in a particular problem and can be found, then one needs to solve the problem only to that time horizon to compute the optimal decision in the initial period corresponding to any time horizon beyond it, and in particular the infinite horizon. Since the first decision is the only one that the decision maker is committed to make immediately, this approach is well-suited to the equipment replacement problem. The process can be repeated in the next period. We are interested in the infinite horizon here because it is usually not clear what time horizon a firm should consider, and what the status of the firm would be at the end of that time horizon. Using a forecast horizon obviates the need for using an arbitrary time horizon.

Forecast horizon techniques have been used in the context of a variety of applications (for a survey see Bhaskaran and Sethi, 1987) including equipment replacement (Chand and Sethi, 1982). More recently, conditions for the existence of forecast horizons for deterministic (Bean and Smith, 1984) and stochastic (Hopp, Bean and Smith, 1987; Bès and Sethi, 1988) problems have been developed.

If a forecast horizon does exist, it could be put to use only if it is identified. Algorithms for identifying forecast horizons in general deterministic (Chand and Morton, 1986) and stochastic optimization (Hopp, 1987; Bès and Sethi, 1988) problem are typically cumbersome or inefficient (i.e., they identify forecast horizons that are much longer than necessary). By exploiting the specific structure of the equipment replacement problem we can greatly improve the efficiency of identifying forecast horizons (see Nair, 1988). This paper is an extension of Hopp and Nair (1991) to the equipment replacement case where not only are the technological forecasts non-stationary in time, but also the revenue and cost functions are non-stationary in time. In addition, salvage values have been considered in the present paper. Structural results on the sensitivity of the optimal decision to changes in forecast are also discussed.

2. Problem formulation

We consider the problem of a firm trying to decide whether to keep a piece of equipment it currently owns or to replace it with a better technology that is currently available on the market. The problem is complicated by the fact that further improvements of technology could occur in the future and these improvements may have a bearing on the current decision. For example, it may be optimal to wait for a few years and take advantage of a better technology when it becomes available. However, whether or when the advanced technology will appear and what its revenue generating potential will be are uncertain. We assume that the firm's objective is to maximize expected discounted present revenue over the infinite horizon.

We model this problem as an infinite horizon non-stationary Markov decision process (MDP). We choose an infinite horizon as it would be better than assuming an arbitrary finite horizon and then making assumptions on the state of the firm at the end of that horizon. The model considered here is restricted to the case of three technologies, viz., the technology on hand, a 'better' technology currently available on the market, and an 'even better' technology that may become available in the future. Our notation and approach are described below.

2.1. Notation

We denote the technology currently in use by index 0, the better technology available on the market by index 1, and exactly one advanced technology that may appear in the future by index 2.

We define a 'state' at some point in time, t, by (i, l) where i represents the index of the technology in use by the firm and l the index of the latest technology available on the market. By our definition of the problem, at the beginning of the problem, the state is (0, 1) and the time period is t = 0. Let K_i represent the action 'keep technology i' and R_j $(j \le l)$ represent the action 'replace the current technology with technology j' in state (i, l).

To formulate a Markov decision process, we suppose the firm can make keep-replace decisions at the beginning of each period. We let r_{it} represent the (expected) one period revenue generated by technology i in period t, c_{it} represent the expected capital cost of purchasing technology i in period t, and s_{it} represent the salvage value received from selling technology i in period t. We assume r_{it} , c_{it} and s_{it} are bounded. Notice that the revenue, cost and salvage value functions are only time dependent and not age dependent. This would be reasonable for classes of equipment not subject to physical deterioration, like microcomputers and controls, etc.

Let p_t be the probability of appearance of technology 2 in period t given that it has not appeared in period t-1 or earlier. Note here that though the cost and revenue functions may vary over time, their values are essentially known (or can be estimated). However, the period of appearance of the new technology is uncertain. Again, this kind of situation is appropriate for advanced technology wherein the technical parameters and specifications of the new technology may be decided on beforehand, and the appearance of technology on the market would depend on breakthroughs achieved, market conditions, etc.

Let β be the one period discount factor (where $\beta < 1$) and $f_t^T(i, l)$ represent the maximal expected discounted return of being in state (i, l) in period t, if an optimal policy is followed from periods t through some time horizon T $(T \ge t)$. The 'boundary condition', i.e., when t = T, is represented as $L(i, l) \equiv f_T^T(i, l)$.

2.2. Approach

To find the optimal decision in any state, we need to calculate the value function $f_t^T(i, l)$. Further, we would like to solve the problem for an infinite horizon, i.e., when $T = \infty$. However, since that would entail making infinite calculations, we will first solve the problem for some finite T, and then extend

these results to the infinite horizon case using forecast horizon techniques. We start by first stating the following dynamic programming recursive equations describing the problem:

$$f_{t}^{T}(0,1) = \max \begin{cases} R_{1}: & -c_{1t} + s_{0t} + r_{1t} + \beta\{(1 - p_{t+1})f_{t+1}^{T}(1,1) + p_{t+1}f_{t+1}^{T}(1,2)\}, \\ K_{0}: & r_{0t} + \beta\{(1 - p_{t+1})f_{t+1}^{T}(0,1) + p_{t+1}f_{t+1}^{T}(0,2)\}. \end{cases}$$
(1)

The above expression states that in state (0, 1) in period t, there are only two actions possible, either 'replace with technology 1' (represented by R_1) or 'keep technology 0' (represented by K_0). In the former case, a cost of c_{1t} is incurred in purchase of technology 1, a salvage value of s_{0t} is collected by selling technology 0, and a one period revenue of r_{1t} accrues from the technology purchased. In addition, a discounted future stream of revenue accrues which depends on whether the new technology 2 does not appear (in which case the state in the next period will be (1, 1)), or appears (in which case the state in the next period will be (1, 2)). If K_0 is the action chosen, then a revenue of r_{0t} is collected in addition to the discounted future stream of revenue. This future stream would again depend on whether the new technology 2 does not appear (in which case the state will be (0, 1)) or it appears (the state would be (0, 2)).

Since in state (1, 1) the only decision available is to 'keep technology 1', (it can be proved using Lemma 1(a) presented later that it would never be optimal to go back to technology 0, once technology 1 is purchased) the functional equation in this case would be

$$f_t^T(1,1) = r_{1t} + \beta \{ (1 - p_{t+1}) f_{t+1}^T(1,1) + p_{t+1} f_{t+1}^T(1,2) \}.$$
 (2)

The functional equations for the other possible states are given by the following expressions using the same reasoning as explained above:

$$f_{t}^{T}(0, 2) = \max \begin{cases} R_{2}: & -c_{2t} + s_{0t} + r_{2t} + \beta f_{t+1}^{T}(2, 2), \\ R_{1}: & -c_{1t} + s_{0t} + r_{1t} + \beta f_{t+1}^{T}(1, 2), \\ K_{0}: & r_{0t} + \beta f_{t+1}^{T}(0, 2), \end{cases}$$
(3)

$$f_t^T(1,2) = \max \begin{cases} R_2: & -c_{2t} + s_{1t} + r_{2t} + \beta f_{t+1}^T(2,2), \\ K_1: & r_{1t} + \beta f_{t+1}^T(1,2), \end{cases}$$
(4)

$$f_t^T(2,2) = r_{2t} + \beta f_{t+1}^T(2,2). \tag{5}$$

The above expressions are valid for t = 0, ..., T - 1. For the backward recursion of dynamic programming to work, we need to define the functional value of each state when t = T. These are called the boundary conditions and we are representing it by L(i, l). As long as T is finite, these actions can be computed using backward recursion. In the usual T period problem, we take L(i, l) = 0 for all i, l. However, as we shall see later, using alternate values of L(i, l) can give useful results.

We let $\pi_t^T(i, l)$ represent the action, i.e., either K_i or R_j $(l \ge j > i)$, that achieves the above maximizations (i.e., $\pi_t^T(i, l)$ is the optimal action in state (i, l) in period t of the T period problem).

The results of Hopp, Bean and Smith (1987) show that, since $\beta < 1$ and r_{it} , c_{it} , and s_{it} are bounded, $f_t^T(i, l)$ converges to a limiting value function $f_t(i, l)$ as $T \to \infty$ (henceforth, we will drop the superscript T to represent the infinite horizon case). Furthermore, since the set of actions in state (i, l) is finite, their results imply that if the infinite horizon optimal decision, written $\pi_t(i, l)$, is unique (i.e., the revenue streams from alternative actions is not exactly equal), then there exists a finite horizon τ such that $\pi_t^T(i, l) = \pi_t(i, l)$ for all $T \ge \tau$, regardless of the boundary condition L(i, l), as long as these are bounded. Such a τ is called a forecast horizon.

Since we begin the problem in state (0, 1) and time t = 0, we are interested in determining the infinite horizon optimal decision $\pi_0(0, 1)$. In this paper we give a simple and efficient method for identifying the forecast horizon and using it to obtain this decision.

2.3. Assumptions

We make the following assumptions regarding the cost and revenue functions for our model:

- I. $r_{2t} \ge r_{1t} \ge r_{0t}$ for all $t \le T$.
- II. $c_{1t} \ge s_{1t} \ge s_{0t}$ for all $t \le T$.

III.
$$\beta(s_{1t+1} - s_{0t+1}) \ge (s_{1t} - s_{0t}) - (r_{1t} - r_{0t})$$
 for all $t < T$, and $r_{1t} - r_{0t} \ge s_{1t} - s_{0t}$ for $t = T$.

The first assumption implies that the per period revenues from new technologies are greater than those of old technologies, and this relation holds for all time periods. The first part of the statement is reasonable, the second part would be reasonable in cases where the revenue generated by an equipment is affected by market conditions (e.g., demand for goods or services produced by the piece of equipment under consideration) and changes in market conditions would affect all the technologies in a similar manner.

The second assumption is quite justifiable and merely states that the cost for purchasing technology 1 be greater than it's salvage value in all periods, and that the salvage value for technology 1 be greater than that of technology 0 in all periods.

The third assumption is sufficient to ensure that if the optimal decision in state (1, 2) in period t were to replace with technology 2, then the same decision would hold in state (0, 2) (using the result of Lemma 1, discussed later). This is a reasonable deduction to make, and the assumption just formalizes it to avoid unrealistic cases. If salvage values were not considered, then this result is intuitive and will directly follow from assumption A1 above (and Lemma 1(a)). However, when salvage values are considered, the above assumption is required. To see that this assumption is reasonable at least in certain cases, note that if $s_{it} = s_i$, $r_{it} = r_i$ for all t, then III reduces to $(r_1 - r_0)/(1 - \beta) \ge s_1 - s_0$, which means that the extra revenue earned by technology 1 over technology 0 over the infinite horizon is greater than the difference in their salvage values, which is very reasonable. Also note that if salvage values were not important (or very small or equal to some nominal value for all technologies) as in some of the advanced fast changing technologies, then III reduces to I. Clearly, there will be many other real life situations where III will hold and hence it should not be considered a major restriction on the use of the model.

Since no restrictions have been placed on the relationships between costs of purchase of technologies c_{1t} and c_{2t} , (other than II), the model allows these costs to move up or down in any manner with time. This comes in handy in modeling replacement of advanced technologies which may cost less in the future though its capabilities may increase.

3. Main results

In order to solve the infinite horizon problem using the finite horizon formulation in the dynamic programming recursions (1)–(5), we first show that under two sets of boundary conditions for a finite T period problem (as stated later in Lemmas 2 and 3), the difference between choosing 'keep technology 0' and 'replace with technology' in state (0, 1) changes monotonically (one non-decreasing and the other non-increasing) with T. We use these two sets of boundary conditions to develop a *stopping rule* to identify the forecast horizon for the problem. Once we identify the forecast horizon, τ , we are certain (by definition) that solving a τ time problem will give us an infinite horizon optimal decision.

3.1. A stopping rule

We define

$$\Delta_{t}^{T} = -c_{1t} + s_{0t} + (r_{1t} - r_{0t}) + \beta (1 - p_{t+1}) [f_{t+1}^{T}(1, 1) - f_{t+1}^{T}(0, 1)] + \beta p_{t+1} [f_{t+1}^{T}(1, 2) - f_{t+1}^{T}(0, 2)],$$
(6)

which from (1) is the difference between the value of using action R_1 or K_0 in state (0, 1). Hence, a necessary and sufficient condition for the optimal decision $\pi_t^T(0, 1) = R_1$ is $\Delta_t^T > 0$, and for $\pi_t^T(0, 1) = K_0$ it is $\Delta_t^T \leq 0$. Notice that we break ties in favor of K_0 .

If under certain conditions we can show that Δ_t^T is non-decreasing in T for all t, then this would mean that the optimal decision $\pi_t^{\tau}(0, 1) = R_1$ implies that $\pi_t^{\tau}(0, 1) = R_1$, for all $T \ge \tau$, and in particular the optimal infinite horizon decision $\pi_t(0, 1) = R_1$. Similarly, if under certain conditions we can show that Δ_t^T is non-increasing in T for all t, then this would mean that the optimal decision $\pi_t^{\tau}(0, 1) = K_0$ implies that $\pi_t^T(0, 1) = K_0$ for all $T \ge \tau$, and in particular the optimal infinite horizon decision $\pi_t(0, 1) = K_0$. After one technical lemma (which may be skipped by the casual reader), we present such conditions below as Lemmas 2 and 3.

The following technical lemma gives conditions for the boundary conditions that make certain decision combinations in different states to be clearly non-optimal. These conditions are simple to ensure and implement and make the identification of the optimal decision simpler.

Lemma 1. If

- (a) $L(i, j) \le L(i+1, j)$ for all i = 0, 1, 2, and $j \ge i+1$, then $f_t^T(i, j) \le f_t^T(i+1, j)$ for all i = 0, 1, 2; $j \ge i + 1$ and $t \le T$.
- (b) $L(1, 2) L(0, 2) \ge s_{1T} s_{0T}$, then $f_t^T(1, 2) f_t^T(0, 2) \ge s_{1t} s_{0t}$ for all $t \le T$. (c) $L(1, 1) L(0, 1) \ge L(1, 2)$, then $f_t^T(1, 1) f_t^T(0, 1) \ge f_t^T(1, 2) f_t^T(0, 2)$ for all $t \le T$.

Proof. The proof of (a) follows from straightforward backward induction on t. The proof of (b) and (c) follows from backward induction using assumption III.

It follows directly from (3) and (4), the result of Lemma 1(a) and 1(b) and by assumptions I-III that $\pi_t^T(1,2) = R_2$ implies $\pi_t^T(0,2) = R_2$ for all t < T. That is, if the optimal decision in state (1, 2) is to replace technology 1 with technology 2, then the optimal decision in state (0, 2) will also be to replace technology 0 with technology 2. This makes the problem more tractable by precluding cases where the optimal decision in state (1, 2) is to 'replace with technology 2' but the optimal decision in state (0, 2) is to 'replace with technology 1' or to 'keep technology 0'. Such cases would indeed be unusual, but one could think of unrealistic salvage values that could make it possible. It is for this reason that assumption III was made. Lemma 1(c) is needed for proving Lemma 4, stated later.

Using the result of Lemma 1 we can obtain expressions for $f_t^T(1, 1) - f_t^T(0, 1)$ and $f_t^T(1, 2) - f_t^T(0, 2)$ which help in evaluating (6) and hence arrive at the optimal decision. These expressions are given as part of the proof of the next lemma in the Appendix. The minimum and maximum values that these expressions take are used for specifying the boundary conditions in Lemmas 2 and 3 below.

We now present the two lemmas that state the boundary conditions that ensure monotonicity of Δ_t^T with respect to T in (6).

Lemma 2. If

$$L(1, 1) - L(0, 1) = \min[c_{1T} - s_{0T}, r_{1T} - r_{0T}],$$

$$L(1, 2) - L(0, 2) = s_{1T} - s_{0T},$$

$$L(2, 2) - L(1, 2) = c_{2T} - s_{1T}$$

then $\Delta_t^T(g)$ is non-decreasing in T, where $\Delta_t^T(g) \equiv \Delta_t^T$ under these particular boundary conditions.

Proof. See Appendix.

Lemma 3. If

$$L(1, 1) - L(0, 1) = c_{1T} - s_{0T},$$

$$L(1, 2) - L(0, 2) = c_{1T} - s_{0T},$$

$$L(2, 2) - L(1, 2) = \min[c_{2T} - s_{1T}, r_{2T} - r_{1T}],$$

then $\Delta_t^T(h)$ is non-increasing in T, where $\Delta_t^T(h) \equiv \Delta_t^T$ under these particular boundary conditions.

Proof. The proof is exactly analogous to that of Lemma 2 and is not repeated here. \Box

Note that the boundary conditions in Lemmas 2 and 3 satisfy the requirements of Lemma 1.

The conditions in Lemma 2 make the 'keep' decision as attractive as possible. The conditions in Lemma 3 make the 'replace' decision as attractive as possible. The intuition behind this is that the conditions in Lemma 2 state that if technology 2 has not appeared until T, the advantage of being in state (1, 1) over being in state (0, 1) is small (actually it is the minimum value the difference is being in these two states can take, as shown in the Appendix in expression (A.1)), hence being in state (0, 1), that is the 'keep' decision, is attractive. If in spite of this advantage the optimal decision for some τ is to 'replace', then it can be shown that this decision will be optimal for all $T \ge \tau$. The conditions of Lemma 3 are such that if technology 2 has not appeared until T the advantage of being in state (1, 1) over being in state (0, 1) is large (we ensure this difference is the maximum possible, as shown in the Appendix in expression (A.1)). This makes 'replace' and moving to state (1, 1) decision attractive. If under these conditions the optimal decision is to 'keep', then it can be shown that this decision will be optimal for all $T \ge \tau$. We present this important result as Theorem 1 below.

Theorem 1 below takes advantage of the monotonicity results in Lemmas 2 and 3 to identify the forecast horizon. Because minimum values are used in Lemma 2 it can be easily shown that $\Delta_t^T(g)$ is a lower bound on Δ_t^T in (6), and since the maximum values are used in Lemma 3, that $\Delta_t^T(h)$ is an upper bound on Δ_t^T in (6). From the discussion immediately following (6) we know that if $\Delta_t^T > 0$ then the optimal decision is to replace technology 0 with technology 1. Since $\Delta_t^T(g)$ is a lower bound on Δ_t^T , $\Delta_t^T(g) > 0$ should also give the same optimal decision. Similarly, $\Delta_t^T(h) \le 0$ should imply that the optimal decision is to keep technology 0. This observation is formalized in Theorem 1.

Theorem 1.

```
(a) If \Delta_t^{\tau}(g) > 0, then \pi_t^T(0, 1) = R_1 for all T \ge \tau.

(b) If \Delta_t^{\tau}(h) \le 0, then \pi_t^T(0, 1) = K_0 for all T \ge \tau.
```

Proof. The result follows directly from Lemmas 2 and 3 and the discussion immediately following expression (6). \Box

Theorem 1 (a) states that by using the conditions of Lemma 2, if the Δ_t^{τ} in (6) is positive for some horizon τ , then the optimal decision will be to replace technology 0 with technology 1 for all horizons beyond τ . Thus τ will be the forecast horizon. Theorem 1(b) states that if by using the conditions of Lemma 3, the value of Δ_t^{τ} from (6) is non-positive, then the optimal infinite horizon decision would be to keep technology 0 for all horizons beyond τ . τ in this case would be the forecast horizon. This discussion leads us immediately to an algorithm for finding the forecast horizon which will solve two dynamic programs, one with the boundary conditions in Lemma 2 and the other with the boundary conditions in Lemma 3. Since the $\Delta_t^T(g)$ and $\Delta_t^T(h)$ in these programs are lower and upper bounds, respectively, for Δ_t^T in (6), whenever these values are either both positive or both non-positive for some T, the forecast horizon is found. Since these values coverge as T increases, we iteratively increase the time horizon T in the algorithm until the forecast horizon is found.

We present Figure 1 to put all our results up to now in perspective, and also to facilitate the understanding of the results and algorithm stated later. Notice that the forecast horizon is identified

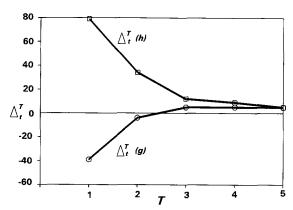


Figure 1. Monotonicity of Δ_t^T using boundary conditions of Lemmas 2 and 3. Data is from numerical example (a) in Section 3.6. Here the optimal infinite horizon decision is $\pi_t(0, 1) = R_1$, and the forecast horizon is $\tau = 3$

when both the curves are on the same side of $\Delta_t^T = 0$. When they are on the positive side, the optimal decision is $\pi_t(0, 1) = R_1$, and when they are on the negative side, the optimal decision is $\pi_t(0, 1) = K_0$.

3.2. The algorithm

Theorem 1 provides the basis for an algorithmic approach to computing the infinite horizon optimal 'keep' or 'replace' decision for any period, and specifically for t = 0. The algorithm can be stated as follows:

- Step 1. Choose a small time horizon T (say T=3), make forecasts p_t for periods $1, \ldots, T$.
- Step 2. Solve two T-period problems:
 - The first with boundary conditions of Lemma 2. One possible set could be:
 - $\triangleright L(0, 1) = 0,$
 - $\triangleright L(1, 1) = \min[c_{1T} s_{0T}, r_{1T} r_{0T}],$
 - $\triangleright L(0, 2) = 0,$
 - $\triangleright L(1, 2) = s_{1T} S_{0T}$, and
 - $\triangleright L(2, 2) = c_{2T} s_{0T}.$
 - The second problem should have boundary conditions of Lemma 3. One possible set could
 - $\triangleright L(0, 1) = 0,$
 - $\triangleright L(1, 1) = c_{1T} s_{0T},$
 - $\triangleright L(0, 2) = 0,$
 - $\triangleright L(1, 2) = c_{1T} s_{0T}$, and
- $> L(2, 2) = \min[c_{2T} s_{1T}, r_{2T} r_{1T}] + c_{1T} s_{0T}.$ Step 3. For the period of interest, in this case t = 0, check to see if either $\Delta_t^T(g) > 0$ in the first problem, or $\Delta_{\tau}^{T}(h) \leq 0$ in the second problem. If yes, stop. $\tau = T$ is the forecast horizon. The optimal decision in both the problems will be the same (from Theorem 1) and the optimal decision is $\pi_t(0, 1) = \pi_t^T(0, 1)$. If no, go to Step 4.
- Step 4. There are two options:
 - (a) Increment T by 1 and go to Step 1, or
 - (b) Compute the maximum error from using the T period problem decision (using result of Theorem 3, stated later). If acceptable, stop. If not acceptable, go to (a).

This approach has three advantages. First, it is extremely simple to implement, since testing for forecast horizons merely requires solving two finite horizon dynamic programs. An implementation of the above algorithm in Turbo Pascal takes a fraction of a second to run for each value of time horizon T on an IBM PS/2 Model 55 SX. This makes it roughly equivalent in complexity to the stopping rule of Bès and Sethi (1988) and much simpler than that of Bean, Hopp and Duenyas (1989). Second, it requires only incremental forecasting for the future technology, i.e., forecasts are made for only as many periods into the future as are necessary to identify the forecast horizon. And finally, as we will develop below, it is 'efficient' in the sense that it identifies short forecast horizons. This is important, since it enables the decision maker to base decisions on near term information, rather than requiring dependence on distant, unreliable forecasts.

3.3. Efficiency of the algorithm

In Bean, Hopp and Duenyas (1989) it is shown that the stopping rule of Bès and Sethi (1988) can lead to overly long forecast horizons (e.g., 55 periods when the minimum forecast horizon is 4 periods). The Hopp, Bean and Smith (1987) approach is more efficient in that it identifies shorter forecast horizons, but it is computationally very intensive and does not guarantee to the shortest possible forecast horizon. All these methods are for the general non-stationary problem and hence do not have the computational efficiency or the efficiency in finding short forecast horizons that our algorithm has. As we show in the following theorem, under certain conditions, the algorithm introduced above generates forecast horizons that are as short as they can be.

Theorem 2. If

$$\begin{split} &-c_{2\tau}+s_{0\tau}+\mathcal{R}_2\geq -c_{1\tau}+s_{0\tau}+\mathcal{R}_1\geq \mathcal{R}_0,\\ &-c_{2\tau}+s_{1\tau}+\mathcal{R}_2\geq \mathcal{R}_1, \end{split}$$

and

$$\tau = \min\{T : \Delta_t^T(g) > 0 \text{ or } \Delta_t^T(h) \le 0\}$$

where $\mathcal{R}_i = \sum_{t=\tau}^{\infty} (\beta^{t-\tau}) r_{it}$, i = 0, 1, 2, then τ is the shortest forecast horizon consistent with the partial forecast $\{p_t\}_{t=1}^{\tau}$.

Proof. In Appendix. \square

 \mathcal{R}_i is the total revenue for technology i from period τ through ∞ discounted back to period τ .

The above result shows that if the conditions of Theorem 2 hold, then τ is the shortest forecast horizon consistent with the partial forecast up to τ . We will show later that even if these conditions do not hold, our algorithm results in much shorter forecast horizons compared to other algorithms.

3.4. Epsilon optimality

The algorithm states that for any T that is short of the forecast horizon one could stop the procedure and calculate the maximum error entailed in choosing one of the decisions from the two T period problems. This is especially helpful if for some problems T becomes too large for forecasting probabilities and estimating costs and revenues. The following result gives a procedure for calculating this error:

Theorem 3. If $\Delta_t^T(g) \le 0$ and $\Delta_t^T(h) > 0$, then the maximum error from using (a) $\pi_t(0, 1) = R_1$ is $-\Delta_t^T(g)$; (b) $\pi_t(0, 1) = K_0$ is $\Delta_t^T(h)$.

Proof. The proof is straightforward from Lemmas 2 and 3. \Box

The intuition behind this result will be clear from Figure 1. Since $\Delta_t^T(g)$ and $\Delta_t^T(h)$ are bounds for Δ_t^T , if we choose the decision based on one of these, the maximum error will be given by the (absolute value of the) other.

3.5. Sensitivity to changes in forecasts

Since the probabilities in this model are non-stationary in time, it would be useful to have results which could predict the optimal decision with changes in probabilities. For example, if a decision maker forecasts probabilities upto T=5, and runs the algorithm to find that the optimal infinite horizon decision is to keep technology 0. It would be useful to know if this decision would change if the probability estimates were increased for some or all of the periods, without having to run the dynamic programs once again. We present one such result here after proving a technical lemma (which may be skipped by the casual reader).

Lemma 4. If $\mathbf{p} = (p_{t+1}, p_{t+2}, \dots, p_T)$ and $\hat{\mathbf{p}} = (\hat{p}_{t+1}, \hat{p}_{t+2}, \dots, \hat{p}_T)$ where $\hat{p}_i \ge p_i$ for all $t < i \le T$, then $\Delta_t^T(\hat{\mathbf{p}}) \le \Delta_t^T(\mathbf{p})$, where $\Delta_t^T(\cdot) \equiv \Delta_t^T$ when probability vector (\cdot) is used.

Proof. We know from Lemma 1(c) that $f_t^T(1, 1) - f_t^T(0, 1) \ge f_t^T(1, 2) - f_t^T(0, 2)$ for all $t \le T$. Hence from (6) we know that Δ_t^T is non-increasing in p_{t+1} . From (1)-(3) we can see that $f_t^T(1, 1) - f_t^T(0, 1)$ is non-increasing in p_{t+1} and from (3)-(4) that $f_t^T(1, 2) - f_t^T(0, 2)$ is independent of p_t for all t. Hence it follows that $f_t^T(1, 1) - f_t^T(0, 1)$ is non-increasing in p_m for all $m \ge t + 1$. Thus $\Delta_t^T(\hat{p}) \le \Delta_t^T(p)$. \square

The above result states that when the probability forecasts are increased, the value of Δ_t^T decreases (or more accurately, does not increase). This result helps in predicting the stability of the optimal decision, which we formally state in the result below.

Theorem 4. (a) If
$$\pi_t^T(p) = K_0$$
, then $\pi_t^T(\hat{p}) = K_0$;
(b) If $\pi_t^T(\hat{p}) = R_1$, then $\pi_t^T(p) = R_1$.

Proof. The result follows directly from Theorem 1 and Lemma 4. \square

Theorem 4 implies that if the optimal decision is to 'keep' using a particular technological forecast vector, then the optimal decision will be 'keep' also for a technological forecast vector that stochastically dominates the original vector. This is entirely intuitive. The opposite holds if the optimal decision were to 'replace'.

3.6. Numerical examples

We now discuss a few numerical examples to illustrate our algorithm for finding the optimal infinite horizon decision using the forecast horizon approach. Specifically, we are interested in $\pi_0(0, 1)$, the optimal infinite horizon decision in the initial period in state (0, 1).

Suppose
$$r_{2t} = 175$$
, $c_{2t} = 200$, $s_{1t} = 75$ and $s_{0t} = 35$ for all t ; $T = 4$ and $\beta = 0.9$. Let

$$egin{aligned} r_{00} = 50, & r_{01} = 60, & r_{02} = 45, & r_{03} = 50, & r_{04} = 65, \\ r_{10} = 100, & r_{12} = 100, & r_{12} = 95, & r_{13} = 90, & r_{14} = 75, \\ c_{10} = 125, & c_{11} = 175, & c_{12} = 100, & c_{13} = 100, & c_{14} = 200. \end{aligned}$$

Note that assumptions I–III are satisfied by the data. We use the boundary conditions stated in our algorithm in Section 3.2 to solve two dynamic programs.

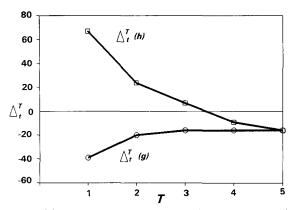


Figure 2. Results from numerical example (b). Here the optimal infinite horizon decision is $\pi_i(0, 1) = K_0$ and the forecast horizon is $\tau = 4$

- (a) Suppose the technological forecast were $p_1 = 0.1$, $p_2 = 0.2$, $p_3 = 0.3$, $p_4 = 0.6$. Using the boundary conditions in Lemmas 2 and 3 we have $\Delta_0^4(g) = 5$ and $\Delta_0^4(h) = 9$, respectively. It follows from Theorem 1 that $\pi_0(0, 1)$ will be R_1 , and so, the decision to replace the current equipment with that presently available in the market is optimal for any technological forecast in period 5 and beyond. The values of $\Delta_0^T(g)$ and $\Delta_0^T(h)$ for different values of T are shown in Figure 1. It can be seen from Figure 1 that in this case the forecast horizon $\tau = 3$. When the time horizon is T = 4, as in this example, the same decision holds by definition of forecast horizons as we have above.
- (b) If $p_1 = 0.5$, $p_2 = 0.3$, $p_3 = 0.3$, $p_4 = 0.6$, then using the boundary conditions in Lemmas 2 and 3 we have $\Delta_0^4(g) = -16$ and $\Delta_0^4(h) = -9$, respectively. It follows from Theorem 1 that $\pi_0(0, 1) = K_0$ regardless of the forecast for periods 5 and beyond. Hence, one should keep the existing piece of equipment so as to replace it later when technology 2 appears. The forecast horizon is $\tau = 4$. The values of $\Delta_0^T(g)$ and $\Delta_0^T(h)$ for different values of T are shown in Figure 2.
- (c) If $p_1 = 0.25$, $p_2 = 0.2$, $p_3 = 0.3$, $p_4 = 0.6$, then using the boundary conditions in Lemmas 2 and 3 we get $\Delta_0^4(g) = -3$ and $\Delta_0^4(h) = 2$, respectively, as shown in Figure 3. In such a case, the optimal decision depends on the forecast in periods 5 and beyond. The decision maker could then either increase the horizon by one period and repeat the exercise or use the epsilon-optimality result discussed in Theorem 3 which states that the maximum error in choosing K_0 is 2, and the maximum error in choosing K_0 is 3. To minimize the maximum error, we would choose K_0 , (i.e., keep the current equipment) and incur a maximum error of 2.

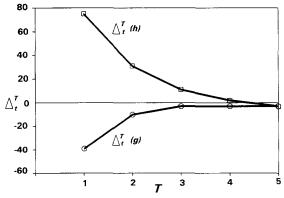


Figure 3. Results from numerical examples (c) and (d). Here the optimal infinite horizon decision is $\pi_t(0, 1) = K_0$ and the forecast horizon is $\tau = 5$

- (d) If, however, the decision maker in the previous case decides to increase T to 5 and forecasts the probability, costs and revenues to be $p_5 = 0.6$, $r_{05} = 85$, $r_{15} = 85$ and $c_{15} = 290$, then we get $\Delta_0^5(g) = -3$ and $\Delta_0^5(h) = -3$ as shown in Figure 3. Hence $\tau = 5$ is the forecast horizon and using Theorem 1, $\pi_0(0, 1) = K_0$.
- (e) Using Theorem 4 we can say that all forecast vectors that are dominated by the one in case (a) also will result in $\pi_0(0, 1) = R_1$ and all forecast vectors that dominate the one in case (b) and (d) will result in $\pi_0(0, 1) = K_0$.

If in the above examples we let $r_{it} = r_{i4}$ and $c_{it} = c_{i4}$ for all t > 4 and i = 0, 1, 2, then the first condition of Theorem 2 is not satisfied (as $-c_{2T} + s_{0T} + \mathcal{R}_2 = 1585$, $-c_{1T} + s_{0T} + \mathcal{R}_1 = 585$, and $\mathcal{R}_0 = 650$) and hence the forecast horizon obtained is not the shortest possible. It is interesting to note that the forecast horizon of 3 and 4 for our examples (a) and (b), respectively, is still appreciably shorter than that obtained by the method of Bès and Sethi (1988). To see this, we observe that for our problem, the Bès and Sethi stopping rule can be stated as follows:

Stopping rule (Bès and Sethi). If $\Delta_0^T \ge (|-c_{10} + s_{00} + r_{10}|)\beta^T/(1-\beta)$, then $\pi_0^M(0, 1) = \pi_0^T(0, 1)$ for all $M \ge T$.

For example (a), $\Delta_0^4 = 7$, and hence the forecast horizon works out to T = 36 using the Bès and Sethi rule, whereas in our method the forecast horizon was T = 4. Other examples yield similar results. Thus, our simple method shows the same level of improvement over the Bès and Sethi method achieved by the much more complicated approach of Bean, Hopp and Duenyas (1989).

4. Conclusions

We have given a method for computing the optimal equipment replacement decision where the technological forecasts and cost and revenue functions are non-stationary. This method assumes that only one new technology may appear in the future in addition to the technology already available on the market at the current time. A few assumptions about the nature of revenue and cost functions have been made with a short discussion of situations where these assumptions would be justifiable.

The algorithm developed in this paper is extremely simple computationally as well as to code, and, under certain conditions, is efficient in the sense of computing the optimal initial 'keep' or 'replace' action using a minimum number of periods of forecast data. Even when such conditions do not hold, the algorithm identifies forecast horizons that are much shorter and in lesser computational time than other existing methods.

Further work is needed to characterize the problem of equipment replacement when more than one alternative is available on the market at the outset, and the situation where more than one improvement can occur in the future. These situations do not lend themselves to choice of appropriate boundary conditions as was done in this paper. General-purpose methods can be used for these cases but such methods are cumbersome, as discussed earlier.

Appendix

Lemma 2. If

$$L(1, 1) - L(0, 1) = \min[c_{1T} - s_{0T}, r_{1T} - r_{0T}],$$

$$L(1, 2) - L(0, 2) = s_{1T} - s_{0T},$$

$$L(2, 2) - L(1, 2) = c_{2T} - s_{1T},$$

then $\Delta_t^T(g)$ is non-decreasing in T, where $\Delta_t^T(g) \equiv \Delta_t^T$ under these particular boundary conditions.

Proof. Let $g_t^T(\cdot, \cdot) = f_t^T(\cdot, \cdot)$ under these boundary conditions. Using (1)–(5) and Lemma 1, we can write

$$g_{t}^{T}(1,1) - g_{t}^{T}(0,1) = \min \begin{cases} c_{1t} - s_{0t}, \\ (r_{1t} - r_{0t}) + \beta(1 - p_{t+1}) \left[g_{t+1}^{T}(1,1) - g_{t+1}^{T}(0,1) \right] \\ + \beta p_{t+1} \left[g_{t+1}^{T}(1,2) - g_{t+1}^{T}(0,2) \right], \end{cases}$$
(A.1)

$$g_{t}^{T}(1,2) - g_{t}^{T}(0,2) = \max \begin{cases} s_{1t} - s_{0t}, \\ c_{2t} - s_{0t} - (r_{2t} - r_{1t}) - \beta \left[g_{t+1}^{T}(2,2) - g_{t+1}^{T}(1,2) \right], \\ c_{1t} - s_{0t}, \\ r_{1t} - r_{0t} + \beta \left[g_{t+1}^{T}(1,2) - g_{t+1}^{T}(0,2) \right], \end{cases}$$
(A.2)

$$g_t^T(2,2) - g_t^T(1,2) = \min \begin{cases} c_{2t} - s_{1t}, \\ r_{2t} - r_{1t} + \beta [g_{t+1}^T(2,2) - g_{t+1}^T(1,2)]. \end{cases}$$
(A.3)

By assumptions I-III and the boundary conditions used in this lemma, the conditions of Lemma 1 are satisfied, so

$$g_{t+1}^T(1,1) - g_{t+1}^T(0,1) \ge 0$$
 and $g_{t+1}^T(1,2) - g_{t+1}^T(0,2) \ge 0$.

Hence, (A.1) implies that

$$g_t^T(1, 1) - g_t^T(0, 1) \ge \min[c_{1t} - s_{0t}, r_{1t} - r_{0t}]$$
 for all $t < T$.

Similarly, (A.2) implies

$$g_t^T(1, 2) - g_t^T(0, 2) \ge s_{1t} - s_{0t}$$

and (A.3) implies

$$g_t^T(2, 2) - g_t^T(1, 2) \le c_{2t} - s_{1t}$$
 for all $t < T$.

These observations imply

$$\begin{split} g^{T+1}(2,2) - g_T^{T-1}(1,2) &\leq c_{2T} - s_{1T} = L(2,2) - L(1,2) = g_T^T(2,2) - g_T^T(1,2), \\ g_T^{T+1}(1,2) - g_T^{T+1}(0,2) &\geq s_{1T} - s_{0T} = L(1,2) - L(0,2) = g_T^T(1,2) - g_T^T(0,2), \\ g_T^{T+1}(1,1) - g_T^{T+1}(0,1) &\geq \min[c_{1T} - s_{0T}, c_{1T} - c_{0T}] = L(1,1) - L(0,1) \\ &= g_T^T(1,1) - g_T^T(0,1). \end{split}$$

Thus the Lemma holds for t = T.

Now suppose

$$\begin{split} g_{t+1}^{T+1}(2,2) - g_{t+1}^{T+1}(1,2) &\leq g_{t+1}^{T}(2,2) - g_{t+1}^{T}(1,2), \\ g_{t+1}^{T+1}(1,2) - g_{t+1}^{T+1}(0,2) &\geq g_{t+1}^{T}(1,2) - g_{t+1}^{T}(0,2), \\ g_{t+1}^{T+1}(1,1) - g_{t+1}^{T+1}(0,1) &\geq g_{t+1}^{T}(1,1) - g_{t+1}^{T}(0,1). \end{split}$$

Then from (7)–(9).

$$\begin{split} g_t^{T+1}(2,2) - g_t^{T+1}(1,2) &\leq g_t^T(2,2) - g_t^T(1,2), \\ g_t^{T+1}(1,2) - g_t^{T+1}(0,2) &\geq g_t^T(1,2) - g_t^T(0,2), \\ g_t^{T+1}(1,1) - g_t^{T+1}(0,1) &\geq g_t^T(1,1) - g_t^T(0,1). \end{split}$$

It follows by induction that $g_t^T(1, 1) - g_t^T(0, 1)$ and $g_t^T(1, 2) - g_t^T(0, 2)$ are non-decreasing in T. Thus using (6) $\Delta_t^T(g)$ is non-decreasing in T. \square

Theorem 2. If

$$\begin{split} &-c_{2\tau}+s_{0\tau}+\mathcal{R}_2 \geq -c_{1\tau}+s_{0\tau}+\mathcal{R}_1 \geq \mathcal{R}_0, \\ &-c_{2\tau}+s_{1\tau}+\mathcal{R}_2 \geq \mathcal{R}_1, \end{split}$$

and

$$\tau = \min\{T : \Delta_t^T(g) > 0 \text{ or } \Delta_t^T(h) \le 0\}$$

where $\mathcal{R}_i = \sum_{t=\tau}^{\infty} (\beta^{t-\tau}) r_{it}$, i = 0, 1, 2, then τ is the shortest forecast horizon consistent with the partial forecast $\{p_i\}_{t=1}^{\tau}$.

Proof. We prove the result by showing that the boundary conditions of Lemmas 2 and 3 are consistent with actual forecasts for periods $\tau + 1$, $\tau + 2$, Hence, if $\Delta_t^{\tau}(g) \le 0$ or $\Delta_t^{\tau}(h) > 0$, then there are forecasts beyond τ that lead to different decisions, so the initial decision depends on forecast data beyond τ and is not a forecast horizon.

First we show that the boundary conditions of Lemma 2 are consistent with any forecast that has $p_{\tau+1} = 1$. Notice that if $p_{\tau+1} = 1$, then the only states that are accessible beyond period τ are (0, 2), (1, 2), and (2, 2). Thus, to show that this forecast is consistent with the boundary conditions of Lemma 2, it is sufficient to show that

$$f_{\tau}(1,2) - f_{\tau}(0,2) = s_{1\tau} - s_{0\tau}$$
 and $f_{\tau}(2,2) - f_{\tau}(1,2) = c_{2\tau} - s_{1\tau}$

where $f_{\tau}(\cdot, \cdot)$ represents the infinite horizon value function in period τ . The condition of the theorem guarantees that for this forecast (from expressions (3)–(5)),

$$f_{\tau}(0,2) = -c_{2\tau} + s_{0\tau} + \mathcal{R}_2, \qquad f_{\tau}(1,2) = -c_{2\tau} + s_{1\tau} + \mathcal{R}_2, \qquad f_{\tau}(2,2) = \mathcal{R}_2,$$

so the result follows.

Next we show that the boundary conditions of Lemma 3 are consistent with the forecast $p_{\tau+1} = p_{\tau+2} = \cdots = 0$. Notice that in this case the only states that are accessible beyond period τ are (0, 1) and (1, 1). To show that this forecast is consistent with the conditions in Lemma 3, it is sufficient to show that $f_{\tau}(1, 1) - f_{\tau}(0, 1) = c_{1\tau} - s_{0\tau}$. From the condition of the theorem we have (from expressions (1)-(2))

$$f_{\tau}(0, 1) = -c_{1\tau} + s_{0\tau} + \mathcal{R}_{1}, \qquad f_{\tau}(1, 1) = \mathcal{R}_{1},$$

which completes the result. \Box

References

Bhaskaran, S., and Sethi, S. (1987), "Decision and forecast horizons in a stochastic environment: A survey", *Optimal Control Applications and Methods* 8, 201-207.

Bean, J., and Smith, R. (1984), "Conditions for the existence of planning horizons", *Mathematics of Operations Research* 9, 391-401.

Bean, J., Hopp, W., and Duenyas, I. (1989), "A stopping rule for forecast horizons in non-homogeneous Markov decision processes", TR 89-03, Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.

Bès, C., and Sethi, S. (1988), "Concepts of forecast and decision horizons: Applications to dynamic stochastic optimization problems", *Mathematics of Operations Research* 13, 295-310.

Chand, S., and Morton, T. (1986), "Minimal forecast horizon procedures for dynamic lot size models", *Naval Research Logistics* 33, 111-122.

Chand, S., and Sethi, S. (1982), "Planning horizon procedures for machine replacement models with several possible replacement alternatives", *Naval Research Logistics* 29, 483-493.

- Derman, C. (1963), "Optimal replacement and maintenance under Markovian deterioration with probability bounds on failure", *Management Science* 9, 478-481.
- Goldstein, T., Ladany, S., and Mehrez, A. (1986), "A dual machine replacement model: A note on planning horizon procedures for machine replacement", *Operations Research* 34, 938-941.
- Goldstein, T., Ladany, S., and Mehrez, A. (1988), "A discounted machine replacement model with an expected future technological breakthrough", *Naval Research Logistics* 35, 209-220.
- Hatoyama, Y. (1984), "On Markov maintenance problems", IEEE Transactions on Reliability 33, 280-283.
- Hopp, W. (1987), "Identifying forecast horizons in nonhomogeneous Markov decision processes", Operations Research 35, 339-343.
 Hopp, W., Bean, J., and Smith, R. (1987), "A new optimality criteria for nonhomogeneous Markov decision processes", Operations Research 35, 875-883.
- Hopp, W., and Nair, S. (1991), "Timing equipment replacement decisions due to discontinuous technological change", *Naval Research Logistics* 38, 203-220.
- Nair, S. (1988), "Replacement decisions due to technological obsolescence under uncertainty", Unpublished Ph.D. Thesis, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.
- Pierskalla, W., and Voelker, J. (1976), "A survey of maintenance models: The control and surveillance of deteriorating systems", Naval Research Logistics 23, 353-388.
- Sethi, S., and Chand, S. (1979), "Planning horizon procedures for machine replacement models", Management Science 25, 140-151.