

Technical Memorandum Number 796

**Optimal Maintenance Policies for Serial, Multi-Machine
Systems with Non-Instantaneous Repairs**

by

**Lisa Maillart
Xiang Fang**

December 2004

**Department of Operations
Weatherhead School of Management
Case Western Reserve University
330 Peter B Lewis Building
Cleveland, Ohio 44106**

Optimal Maintenance Policies for Serial, Multi-Machine Systems with Non-Instantaneous Repairs

Lisa M. Maillart
Xiang Fang

*Department of Operations, Weatherhead School of Management,
Case Western Reserve University, Cleveland, Ohio 44106*

Abstract

We study optimal age-replacement policies for machines in series with non-instantaneous repair times by formulating two nonlinear programs: One that minimizes total cost-rate subject to a steady-state throughput requirement, and another that maximizes steady-state throughput subject to a cost-rate budget constraint. Under reasonable assumptions, the single-machine cost-optimal and throughput-optimal solutions are unique and orderable, and the multi-machine optimal solutions have appealing structure. Furthermore, we establish equivalence between the two formulations and provide an illustrative numerical example.

Keywords: age-replacement, nonlinear programming, efficient frontier

1 Introduction

Maintenance optimization models find the optimum balance between the costs and benefits of performing maintenance activities on stochastically degrading equipment, subject to various constraints. In this paper, we consider such a model for a system comprised of n failure-independent, but non-identical machines in series. The objective is to find an age-replacement policy for each machine that either minimizes the throughput of the entire system, subject to a maintenance-budget constraint, or alternatively, one that minimizes the total maintenance cost-rate, subject to a throughput constraint.

In contrast, the majority of the existing maintenance optimization literature considers a single unit in isolation, and seeks a maintenance policy that optimizes some criterion for that one unit, typically total expected discounted cost, expected cost-rate or stationary availability (see surveys by McCall [9], Pierskalla and Voelker [11], Sherif and Smith [13], Valdez-Flores [15]). Moreover, few of these models incorporate production considerations, as stated in Iravani and Duenyas [7]: “Most of the existing literature on maintenance policies does not consider the interactions between production/inventory and repair/maintenance decisions.”

The maintenance optimization surveys by Thomas [14], Cho and Parlar [3], Dekker et al [5], and Wang [16] review the growing body of literature on multi-unit systems. Not surprisingly, even fewer of these papers incorporate production considerations. Among the exceptions are Meller [10], Schouten and Vanneste [12] and Marquez et. al. [8], who attempt to do by making preventive maintenance decisions as a function of a unit’s age and/or the

amount of inventory in the buffer. However, these models seek a preventive maintenance policy for the first unit only, not all units.

This paper is most closely related to Hsu [6], who considers a serial production system with infinite-capacity buffers between every pair of machines. In Hsu [6], each machine has its own lifetime distribution, processing time distribution, repair time distribution, and preventive maintenance time distribution. The objective is to find an age-replacement policy for each machine i that minimizes the total long-run average cost of the system, which is comprised of each machine's maintenance costs, waiting costs and lost capacity costs. Due to the complexity of this model, Hsu provides numerical results only, for the case of exponential lifetimes and processing times. Furthermore, Hsu's cost function includes the cost of lost capacity for every machine, regardless of whether or not it is the bottleneck machine.

Like Hsu, we seek an age-replacement policy for each machine (i.e. restore machine i to a status "as-good-as-new" when it fails or reaches age T_i , whichever comes first) in a serial production system with infinite-capacity buffers. We assume that demand is sufficiently large to ensure that it is not the bottleneck and that raw material is always available at the first machine. Unlike Hsu, we formulate constrained models, rather than taking an unconstrained cost-minimization approach. Furthermore, we analyze steady-state system performance, which allows us to parameterize our model using average production rates and mean replacement times. This simplification enables us to gain analytical insights on model behavior and policy structure.

The remainder of the paper is organized as follows. In §2 we formulate nonlinear pro-

gramming models under the two criteria of interest: throughput maximization and cost-rate minimization. In §3 we review existing structural results for the single-machine problem under the throughput criterion and establish new, analogous results under the cost-rate criterion. Furthermore, we derive relationships between the single-machine cost-rate minimizing and throughput-maximizing solutions. These results act as building blocks in §4, in which we establish structural results for the multi-machine problems, including a useful equivalence result. We illustrate these results using an insightful numerical example in §5 and summarize our contributions in §6.

2 Model Formulation

An age replacement policy for a particular machine i consists of a single parameter, T_i . Under such a policy, the machine is restored to a status “as-good-as-new” at age T_i (preventive maintenance) or failure (reactive maintenance), whichever comes first. For each machine i , $i = 1, \dots, n$, let

u_i = the average throughput of machine i when it is working

X_i = the lifetime of machine i

$f_i(x)$ = the pdf of X_i

$F_i(x)$ = the cdf of X_i

$\overline{F}_i(x)$ = $1 - F_i(x)$

$h_i(x)$ = the hazard function of X_i , i.e. $f_i(x)/\overline{F}_i(x)$

- v_i = the mean duration of preventive maintenance on machine i
 r_i = the mean duration of reactive maintenance on machine i
 c_{vi} = the mean cost of preventive maintenance on machine i
 c_{ri} = the mean cost of reactive maintenance on machine i
 $A_i(T_i)$ = the stationary availability of machine i under policy T_i .

First, consider a single-machine system comprised of machine i alone. The throughput of this single machine under age replacement policy T_i is given by

$$\begin{aligned}
\epsilon_i(T_i) &\equiv u_i A_i(T_i) \\
&= u_i \frac{T_i \bar{F}_i(T_i) + E[X_i | X_i < T_i] F_i(T_i)}{(T_i + v_i) \bar{F}_i(T_i) + (E[X_i | X_i < T_i] + r_i) F_i(T_i)} \\
&= u_i \frac{T_i \bar{F}_i(T_i) + \int_0^{T_i} t f_i(t) dt}{(T_i + v_i) \bar{F}_i(T_i) + \int_0^{T_i} t f_i(t) dt + r_i F_i(T_i)}, \tag{1}
\end{aligned}$$

and the corresponding cost per unit time is given by

$$\begin{aligned}
C_i(T_i) &\equiv \frac{c_{vi} \bar{F}_i(T_i) + c_{ri} F_i(T_i)}{(T_i + v_i) \bar{F}_i(T_i) + (E[X_i | X_i < T_i] + r_i) F_i(T_i)} \\
&= \frac{c_{vi} \bar{F}_i(T_i) + c_{ri} F_i(T_i)}{(T_i + v_i) \bar{F}_i(T_i) + \int_0^{T_i} t f_i(t) dt + r_i F_i(T_i)}. \tag{2}
\end{aligned}$$

Let

$$\begin{aligned}
\tau_i^* &\equiv \arg \max \{ \epsilon_i(T_i) \} \\
\xi_i^* &\equiv \arg \min \{ C_i(T_i) \}.
\end{aligned}$$

Next, consider the multi-machine problem under the criterion of throughput maximization. Because the throughput of the system is determined by the bottleneck, the objective

is to find a set of age-replacement times, $\underline{T}^* = \{T_1^*, T_2^*, \dots, T_n^*\}$, that maximizes

$$\max_{\underline{T}} \min\{\epsilon_1(T_1), \epsilon_2(T_2), \dots, \epsilon_n(T_n)\}. \quad (3)$$

If the total average cost-rate is constrained by some fixed budget B , then we have the following mathematical program

$$\max_{\underline{T}} \min\{\epsilon_1(T_1), \epsilon_2(T_2), \dots, \epsilon_n(T_n)\} \quad (4)$$

$$\text{s.t.} \quad \sum_{i=1}^n C_i(T_i) \leq B. \quad (5)$$

Now consider the multi-machine problem under the criterion of cost-rate minimization.

The objective is to find the set of age-replacement times that minimizes total cost, or

$$\min_{\underline{T}} \sum_{i=1}^n C_i(T_i). \quad (6)$$

If the system throughput is required to be greater than or equal to some minimum rate $\underline{\epsilon}$, then we have the following mathematical program

$$\min_{\underline{T}} \sum_{i=1}^n C_i(T_i) \quad (7)$$

$$\text{s.t.} \quad \min\{\epsilon_1(T_1), \epsilon_2(T_2), \dots, \epsilon_n(T_n)\} \geq \underline{\epsilon}. \quad (8)$$

3 Analytical Results for the Single-Machine Problems

In this section, we focus on properties of the unconstrained single machine throughput-optimal and cost-optimal solutions, τ_i^* and ξ_i^* , respectively. We review some existing results, and build upon them to offer new insights. The results in this section are inherently inter-

esting on their own, but also provide the foundation for the multi-machine structural results discussed in §4.

We make the following assumptions throughout the remainder of the paper:

- (A1) For every machine i , the hazard function, $h_i(x)$, is strictly increasing to infinity.
- (A2) The expected duration of reactive maintenance is longer than that of preventive maintenance for all machines, i.e. $r_i > v_i$ for all i .
- (A3) The expected cost of reactive maintenance is greater than that of preventive maintenance for all machines, i.e. $c_{ri} > c_{vi}$ for all i .

Note that the Weibull distribution with shape parameter greater than one satisfies assumption (A1). Moreover, this type of distribution is arguably the most widely used distribution in reliability theory; as stated in Barlow and Proschan [1]: “Experience has shown that, in many cases, [the Weibull distribution] fits the observations better than other known distribution functions (p. 2)....It is perhaps the most popular parametric family of failure distributions at the present time (p. 16).”

For the remainder of §3, we drop the subscript i 's for notational convenience.

Proposition 1 establishes that the single-machine throughput-optimal solution is unique and finite under assumptions (A1) and (A2). Proposition 2 establishes analogous results for the single-machine cost-optimal solution as a function of the cost ratio $\frac{c_r}{c_v}$.

Proposition 1. (i) *The throughput function $\epsilon(T)$ is strongly quasiconcave.*

(ii) *τ^* is unique and satisfies $0 < \tau^* < \infty$.*

Proposition 2. (i) If $\frac{r+E[X]}{v+E[X]} \geq \frac{c_r}{c_v}$, then $C(T)$ is strictly decreasing and $\xi^* = \infty$.

(ii) If $\frac{c_r}{c_v} \geq \frac{r}{v} + \frac{1}{vf(0)}$, then $C(T)$ is strictly increasing and $\xi^* = 0$.

(iii) If $\frac{r}{v} + \frac{1}{vf(0)} > \frac{c_r}{c_v} > \frac{r+E[X]}{v+E[X]}$, then $C(T)$ is strongly quasiconvex and there exists a unique ξ^* that satisfies $0 < \xi^* < \infty$.

Propositions 1 and 2 are fundamental to the multi-machine results derived in §4. They also imply the results in Proposition 3, which establish conditions under which we can order the single-machine throughput-optimal and cost-optimal solutions (Figure 1). The intuition behind Proposition 3 is as follows. If the relative reactive maintenance “cost” is greater in terms of dollars than in terms of time ($\frac{c_r}{c_v} \geq \frac{r}{v}$), then the optimal age-replacement time for the cost-minimization problem is less than that for the throughput-maximization problem, or $\xi^* < \tau^*$. In other words, under this condition, it is more critical to avoid reactive maintenance under the cost-minimization criterion than under the throughput-maximization criterion. The reverse argument applies for $\xi^* > \tau^*$.

Proposition 3. (i) If $\frac{r+E[X]}{v+E[X]} \geq \frac{c_r}{c_v}$, then $0 < \tau^* < \xi^* = \infty$.

(ii) If $\frac{c_r}{c_v} \geq \frac{r}{v} + \frac{1}{vf(0)}$, then $0 = \xi^* < \tau^* < \infty$.

(iii) If $\frac{c_r}{c_v} = \frac{r}{v}$, then $0 < \tau^* = \xi^* < \infty$.

(iv) If $\frac{r}{v} > \frac{c_r}{c_v} > \frac{r+E[X]}{v+E[X]}$, then $0 < \tau^* < \xi^* < \infty$.

(v) If $\frac{r}{v} + \frac{1}{vf(0)} > \frac{c_r}{c_v} > \frac{r}{v}$, then $0 < \xi^* < \tau^* < \infty$.

————— insert Figure 1 here —————

Figure 1: Proposition 2 establishes properties of the cost-optimal solution, ξ^* , for ranges of c_r/c_v as depicted. Similarly, Proposition 3 establishes the relative magnitudes of the cost-optimal and throughput-optimal solutions (ξ^* and τ^* , respectively) in each of the intervals as depicted.

4 Analytical Results for the Multi-Machine Problems

In this section, we establish structural properties of the optimal policies for the multi-machine constrained problems, (4)-(5) and (7)-(8). Proposition 4 establishes conditions under which, for each machine, the optimal age-replacement time for the constrained throughput-maximization problem (T_i^*) is between the cost-optimal age-replacement time (ξ_i^*) and the throughput-optimal age-replacement time (τ_i^*). For example, if the cost-optimal age-replacement time is less than the throughput-optimal age replacement time ($\xi_i^* < \tau_i^*$), then the optimal policy

for the constrained throughput-maximization problem delays preventive maintenance to a time later than the cost-optimal age-replacement time, but no later than the throughput-optimal age-replacement time, i.e. $\xi_i^* \leq T_i^* \leq \tau_i^*$. This result is rather intuitive since to solve the constrained model is, in a sense, to strike an optimal balance between the maximum throughput and the minimum cost.

Proposition 4. *If problem (4)-(5) is feasible, then there exists an optimal solution*

$\{T_1^, T_2^*, \dots, T_n^*\}$ such that $\min\{\tau_i^*, \xi_i^*\} \leq T_i^* \leq \max\{\tau_i^*, \xi_i^*\}, i = 1, \dots, n$.*

Now consider the multi-machine constrained cost-rate minimization problem given by (7)-(8), and notice that unlike problem (4)-(5), it is separable into n single-machine problems of the form

$$\min \quad C_i(T_i) \tag{9}$$

$$\text{s.t.} \quad \epsilon_i(T_i) \geq \underline{\epsilon}, \tag{10}$$

$i = 1, \dots, n$. It is not surprising, therefore, that we are able to establish stronger results for problem (7)-(8) than those established in Proposition 4 for problem (4)-(5). In fact, Proposition 5 proves that the optimal solution to problem (7)-(8) is unique, and that each machine's optimal age-replacement time is one of three easily computable values. The intuition behind this result is that under the optimal solution each machine is either a bottleneck or operates under its cost-rate minimizing solution. In other words, the optimal solution "balances the line" at the throughput requirement $\underline{\epsilon}$, with the exception of the machines that can operate at a throughput larger than $\underline{\epsilon}$ for a lower cost. The throughput of these machines can be lowered to $\underline{\epsilon}$ to balance the entire line by adjusting their u_i values.

Proposition 5. *If problem (7)-(8) is feasible, then its optimal solution $\underline{T}^* \equiv \{T_1^*, T_2^*, \dots, T_n^*\}$*

is unique and satisfies

$$T_i^* = \begin{cases} \xi_i^*, & \xi_i^* \in [T_i^-(\underline{\epsilon}), T_i^+(\underline{\epsilon})] \\ T_i^-(\underline{\epsilon}), & \xi_i^* < T_i^-(\underline{\epsilon}) \\ T_i^+(\underline{\epsilon}), & \xi_i^* > T_i^+(\underline{\epsilon}), \end{cases} \quad (11)$$

where, for any given throughput value $\epsilon \in [0, \epsilon_i(\tau_i^*)]$,

$$T_i^-(\epsilon) \equiv \min\{T : \epsilon_i(T) = \epsilon\} \quad (12)$$

$$T_i^+(\epsilon) \equiv \begin{cases} \max\{T : \epsilon_i(T) = \epsilon\}, & \text{if } \epsilon_i(\infty) = u_i \frac{r_i}{E[X_i] + r_i} < \epsilon \\ \infty, & \text{if } \epsilon_i(\infty) = u_i \frac{r_i}{E[X_i] + r_i} \geq \epsilon. \end{cases} \quad (13)$$

Note that because $T_i^-(\epsilon) \leq \tau_i^* \leq T_i^+(\epsilon)$, Proposition 5 implies that Proposition 4 also holds for problem (7)-(8).

As shown in Theorem 1, in fact, problems (4)-(5) and (7)-(8) are actually equivalent. That is, for every problem of the form (4)-(5) there exists a unique problem of the form (7)-(8) with identical machine parameters, that has the the same optimal solution. To establish this result, we need the following two lemmas. Lemma 1 establishes behavior of (7)-(8) for various ranges of $\underline{\epsilon}$. Analogously, Lemma 2 establishes behavior of (4)-(5) for various ranges of B . Let ϵ^* (respectively, C^*) be the optimal objective function value for problem (4)-(5) (respectively, problem (7)-(8)). Furthermore, let b denote the unconstrained bottleneck machine, that is the machine that satisfies

$$\epsilon_b(\tau_b^*) = \min\{\epsilon_1(\tau_1^*), \epsilon_2(\tau_2^*), \dots, \epsilon_n(\tau_n^*)\}, \quad (14)$$

and let $\epsilon_{max} \equiv \epsilon_b(\tau_b^*)$. Obviously, ϵ_{max} represents the greatest achievable system throughput.

Let $\epsilon_{min} \equiv \min\{\epsilon_1(\xi_1^*), \dots, \epsilon_n(\xi_n^*)\}$, so that ϵ_{min} is the system throughput when each machine operates at its cost-optimal solution. Similarly, define $B_{min} \equiv \sum_{i=1}^n C_i(\xi_i^*)$, so that B_{min} is the lowest achievable maintenance cost-rate for the system. And lastly,

$$B_{max} \equiv \min_{T_i \in [T_i^-(\epsilon_{max}), T_i^+(\epsilon_{max})], i=1, \dots, n, i \neq b} \left\{ C_b(\tau_b^*) + \sum_{i=1, i \neq b}^n C_i(T_i) \right\},$$

so that B_{max} is the smallest budget under which ϵ_{max} is attainable. Consequently, we have Lemma 1 and Lemma 2 as follows.

Lemma 1. *For the cost-rate minimization problem (7)-(8), if*

(i) $\underline{\epsilon} > \epsilon_{max}$, then (7)-(8) is infeasible.

(ii) $\underline{\epsilon} < \epsilon_{min}$, then the throughput requirement (8) is nonbinding at optimality and $C^* = B_{min}$.

(iii) $\epsilon_{min} \leq \underline{\epsilon} \leq \epsilon_{max}$, then the throughput requirement (8) is binding at optimality and C^* is strictly increasing in $\underline{\epsilon}$.

Lemma 2. *For the throughput-maximizing problem (4)-(5), if*

(i) $B < B_{min}$, then (4)-(5) is infeasible.

(ii) $B > B_{max}$, then the budget constraint (5) is nonbinding at optimality and $\epsilon^* = \epsilon_{max}$.

(iii) $B_{min} \leq B \leq B_{max}$, then the budget constraint (5) is binding at optimality and ϵ^* is strictly increasing in B .

Theorem 1. *For every $\underline{\epsilon} \in [\epsilon_{min}, \epsilon_{max}]$, problem (7)-(8) has a unique optimal solution, \underline{T}^* described by equation (11), with optimal objective function value $C^* \in [B_{min}, B_{max}]$. For $B = C^*$, \underline{T}^* is also the unique optimal solution to problem (4)-(5) with optimal objective function value $\epsilon^* = \underline{\epsilon}$.*

The usefulness of Theorem 1 stems from the fact that, because of the decomposability of problem (7)-(8) and Proposition 5, solving this cost-minimization problem is much less computationally intensive than solving the throughput-maximization problem. A production manager who wishes to explore the throughput-cost tradeoff using a numerically generated efficient frontier, need only consider instances of the cost-rate minimization problem, because Theorem 1 implies that problems (4)-(5) and (7)-(8) produce the same efficient frontier. We illustrate this idea in §5.

5 A Numerical Example

Table 1 contains the parameter values for the five-machine example that we use to illustrate the results in §3 and §4. The lifetime distribution of machine i , $i = 1, \dots, 5$, is Weibull with scale parameter $\lambda_i = 0.00893$ and shape parameter β_i , which varies between $[1.5, 2.5]$. We do not vary λ_i from machine to machine because it is simply the scale parameter. Similarly, we fix all of the working throughput values, u_i $i = 1, \dots, 5$, to the same value, 100, to eliminate trivial cases in which machines with relatively high u_i values have sufficiently large throughput such that they are never the bottleneck. Lastly, we set $v_i = c_{vi} = 1$ for $i = 1, \dots, 5$ for simplicity (although this choice is partially justified by the fact that τ_i^* only

depends on the ratio r_i/v_i). The other parameters vary from machine to machine as shown in the table.

Table 1: Example Parameter Values for Machine i , $i = 1, \dots, 5$

i	λ_i	β_i	$E[X_i]$	u_i	r_i	v_i	c_{ri}	c_{vi}
1	0.000893	1.5	1010.91	100	5	1	400	1
2	0.000893	1.8	995.842	100	50	1	500	1
3	0.000893	2	992.415	100	40	1	5	1
4	0.000893	2.2	991.741	100	60	1	10	1
5	0.000893	2.5	993.577	100	70	1	50	1

Because assumptions (A1)-(A3) and the condition of Proposition 2 part (iii) are satisfied, we are able to obtain finite throughput-optimal and cost-optimal solutions for each single machine in isolation as shown in the first two columns of Table 2. Figure 2 is generated by varying the value of $\underline{\epsilon}$ within the interval $[\epsilon_{min}, \epsilon_{max}]$ and solving the cost-minimization problem (7)-(8). By Theorem 1, we know that repeating the analogous steps for the throughput-maximization problem, i.e. varying the value of B within the interval $[B_{min}, B_{max}]$ and solving (4)-(5), would yield the same curve. As a result, this efficient frontier can be used as in the following example: If we are interested in the solution to the throughput-maximization problem under a budget of \$0.19 per unit time, we can use the curve to identify that the equivalent cost-rate minimization problem is problem (7)-(8) with a throughput requirement of $\underline{\epsilon} = 98.4$. We can then solve this instance of problem (7)-(8) and implement the corresponding cost-rate minimizing solution, $\underline{T}^* = \{65.31, 102.32, 462.27, 363.34, 201.17\}$, rather

Table 2: Example Results

i	τ_i^*	ξ_i^*	$\epsilon_i(\xi_i^*)$	$\epsilon_i(\tau_i^*)$	$C_i(\xi_i^*)$	$C_i(\tau_i^*)$
1	756.43	30.8675	96.8009= ϵ_{min}	99.5616	0.0886551	0.1148240
2	146.248	39.0775	97.2197	98.478= ϵ_{max}	0.0544651	0.0726157
3	179.699	581.431	98.1199	98.8946	0.00357995	0.0103168
4	161.681	386.991	98.3168	98.8763	0.00478559	0.0104297
5	175.151	201.17	99.0424	99.0561	0.00823273	0.0111222
				Sum of cost =	0.159718 = B_{min}	0.2193084 > B_{max}
						$B_{max} = 0.210725$

than solve the computationally intensive problem (4)-(5).

————— insert Figure 2 here —————

Figure 2: Efficient frontier for the numerical example.

Because we number the machines such that $\epsilon_1(\xi_1^*) = \epsilon_{min} < \epsilon_2(\xi_2^*) < \epsilon_3(\xi_3^*) < \epsilon_4(\xi_4^*) < \epsilon_{max}$, as $\underline{\epsilon}$ increases from ϵ_{min} to ϵ_{max} these machines join the set of bottlenecks in numerical order. That is, for $\underline{\epsilon} \in [\epsilon_{min}, 0.9722]$ only machine one is a bottleneck; for $\underline{\epsilon} \in [0.9722, 98.12]$ both machine one and machine two are bottlenecks; etc... Note that machine five is never the bottleneck since $\epsilon_5(\xi_5^*) = 99.0424 > \epsilon_{max} = 98.478$. Consistent with Proposition 5, for the bottleneck machines $T_i^* = T_i^-(\underline{\epsilon})$ if $\xi < T_i^-(\underline{\epsilon})$ and $T_i^* = T_i^+(\underline{\epsilon})$ if $\xi < T_i^+(\underline{\epsilon})$. For example, for $\underline{\epsilon} \geq 0.9812$, $T_3^* = T_3^+(\underline{\epsilon})$. The efficient frontier is also strictly increasing, as established in Lemmas 1 and 2. Moreover, the efficient frontier is concave, which is intuitive because of

economies of scale. That is, it becomes more and more expensive to meet the throughput requirement $\underline{\epsilon}$ as it approaches the maximum achievable throughput, ϵ_{max} .

6 Summary and Future Work

We formulate two constrained nonlinear maintenance optimization models, one which maximizes throughput subject to a budget constraint and one that minimizes cost-rate subject to a throughput requirement, for multi-machine production systems in series. Unlike most of the existing maintenance optimization literature, these models capture one form of interaction between production and maintenance decisions – namely variable capacity – by examining throughput under non-instantaneous replacement times. That is, each machine’s capacity alternates between zero (during replacements) and its working-capacity.

Under some reasonable assumptions, we prove the uniqueness and finiteness of the unconstrained, cost-optimal age-replacement time for a single machine. Additionally, we establish conditions under which this cost-optimal age-replacement time is larger (smaller) than the throughput-optimal age-replacement time for a single machine system. For multi-machine systems, we show that the solution to the cost minimization model has an appealing, simple structure. Analogous results are difficult to obtain for the throughput maximization model because it is not decomposable and therefore difficult to solve compared to the cost minimization model, especially for large systems. Fortunately, however, we establish that the models have one-to-one correspondence, and thus produce a unique efficient frontier. This result enables us to solve throughput maximization problems by converting them into

equivalent cost minimization problems.

All of the results presented here are derived based on assumption (A1), namely that each machine's hazard rate function is strictly increasing to infinity. A natural extension would be to relax this assumption to simply nondecreasing hazard, although doing so would not likely lead to different insights. Another natural extension would be to consider maintenance policies other than age replacement, such as opportunistic replacement schemes.

Appendix

Proof. Proposition 1

Both results are well known (Barlow and Proschan [1], pg. 87) and follow directly from (A1) and (A2). \square

Proof. Proposition 2

Let $k_1 = \frac{c_r}{c_v}$ and $k_2 = \frac{r}{v}$, and note that (A2) and (A3) imply that $k_2 > 1$ and $k_1 > 1$, respectively. Taking the first derivative of equation (2) yields

$$C'(T) = \frac{(c_r - c_v)f(T)(\int_0^T \bar{F}(t) dt + rF(T) + v\bar{F}(T)) - (c_v\bar{F}(T) + c_rF(T))((r - v)f(T) + \bar{F}(T))}{(\int_0^T \bar{F}(t) dt + rF(T) + v\bar{F}(T))^2} \quad (15)$$

and

$$C'(0) = \frac{c_v f(0)}{v} \left(k_1 - k_2 - \frac{1}{v f(0)} \right) \quad (16)$$

$$C'(\infty) = \frac{c_v f(\infty)(v + E[X])}{(r + E[X])^2} \left(k_1 - \frac{r + E[X]}{v + E[X]} \right). \quad (17)$$

Setting equation (15) equal to zero and simplifying yields

$$C'(T) = 0 \Leftrightarrow P(T) = \frac{1}{k_1 - 1}, \quad (18)$$

where

$$P(T) \equiv h(T) \int_0^T \bar{F}(t) dt - F(T) + \frac{v(k_1 - k_2)h(T)}{k_1 - 1}. \quad (19)$$

Furthermore,

$$P(0) = \frac{vf(0)(k_1 - k_2)}{k_1 - 1} = \frac{vf(0) \left(k_1 - k_2 - \frac{1}{vf(0)} \right) + 1}{k_1 - 1} \quad (20)$$

$$P(\infty) = \frac{h(\infty)(v + E[X])}{k_1 - 1} \left(k_1 - \frac{r + E[X]}{v + E[X]} \right) - 1 \quad (21)$$

$$P'(T) = h'(T) \left(\int_0^T \bar{F}(t) dt + \frac{v(k_1 - k_2)}{k_1 - 1} \right). \quad (22)$$

Lastly, note that assumption (A2) implies

$$k_2 > \frac{r + E[X]}{v + E[X]} > 1. \quad (23)$$

(i) If $\frac{r+E[X]}{v+E[X]} \geq k_1$, then

$$P(0) \leq 0 \quad \text{by (23)}$$

$$P'(T) < 0 \quad \text{by (23) and (A1).}$$

Hence, $P(T)$ is strictly decreasing and equation (18) has no solution, which implies that $C'(T)$ is either strictly increasing or strictly decreasing. However, since $C'(0) < 0$ by inequality (23), we then know that $C(T)$ is strictly decreasing and ξ^* is uniquely ∞ . (Dagpunar [4] establishes the same result, namely that $\xi^* < \infty$ if and only if $\frac{r+E[X]}{v+E[X]} < \frac{c_r}{c_v}$.)

(ii) If $k_1 \geq k_2 + \frac{1}{vf(0)}$, then

$$P'(T) > 0 \quad \text{by (A1)}$$

$$P(0) \geq \frac{1}{k_1 - 1}.$$

Hence, $P(T)$ is strictly increasing and equation (18) has either no solution or a unique solution, 0, which implies that $C'(T)$ is either strictly increasing or strictly decreasing. However, since $C'(\infty) \geq 0$ by inequality (23), we then know that $C(T)$ is strictly increasing and ξ^* is uniquely 0.

(iii) If $k_2 + \frac{1}{vf(0)} > k_1 > \frac{r+E[X]}{v+E[X]}$, then we consider the following two mutually exclusive and jointly exhaustive cases. We show that $C(T)$ is strongly quasiconvex and $0 < \xi^* < \infty$ for both cases.

Case 1: If $k_2 + \frac{1}{vf(0)} > k_1 \geq k_2 > \frac{r+E[X]}{v+E[X]}$, then

$$P'(T) > 0, \forall T > 0 \quad \text{by (A1)}$$

$$P(0) < \frac{1}{k_1-1}$$

$$P(\infty) = \infty \quad \text{by (A1)}.$$

Hence, equation (18) has a unique, finite, positive solution. Moreover, since in this case $C'(0) \leq 0$ and $C'(\infty) \geq 0$, we then know that $C(T)$ is strongly quasiconvex with a unique global minimum and $0 < \xi^* < \infty$ (see Bazaraa et. al. [2], pg. 123).

Case 2: If $k_2 + \frac{1}{vf(0)} > k_2 > k_1 > \frac{r+E[X]}{v+E[X]}$, then (A1) implies that

$$P(0) \leq 0, P(\infty) = \infty$$

$$P'(0) < 0, P'(\infty) > 0$$

$$P'(T) = 0 \Leftrightarrow \int_0^T \bar{F}(t) dt = \frac{v(k_2-k_1)}{k_1-1}.$$

Since $P'(0) < 0$ and $P'(\infty) > 0$, $P'(T) = 0$ has solution(s). Furthermore, because $\int_0^T \bar{F}(t) dt$ is strictly increasing and $\frac{v(k_2-k_1)}{k_1-1}$ is a positive constant, $P'(T) = 0$ has a unique solution. Consequently, $P(T)$ is strongly quasiconvex with a unique global minimum. As a result, since $P(0) \leq 0$ and $P(\infty) = \infty$, equation (18) has a unique solution. As we have argued in Case 1, $C(T)$ is therefore strongly quasiconvex and $0 < \xi^* < \infty$.

□

Proof. Proposition 3

Let $k_1 = \frac{c_x}{c_v}$ and $k_2 = \frac{r}{v}$. Setting the first derivative of $\epsilon(T)$ equal to zero and simplifying yields

$$\epsilon'(T) = 0 \Leftrightarrow Q(T) = \frac{1}{k_2 - 1}, \quad (24)$$

where

$$Q(T) \equiv h(T) \int_0^T \bar{F}(T) dt - F(T). \quad (25)$$

Note that (A1) implies that

$$Q'(T) = h'(T) \int_0^T \bar{F}(t) dt > 0. \quad (26)$$

Substituting equation (25) into equation (19) yields

$$P(T) \equiv Q(T) + \frac{v(k_1 - k_2)h(T)}{k_1 - 1}. \quad (27)$$

- (i) This result follows directly from Proposition 1 and part (i) of Proposition 2.
- (ii) This result follows directly from Proposition 1 and part (ii) of Proposition 2.
- (iii) Proposition 1 implies that there exists a unique $\tau^* \in (0, \infty)$ that satisfies equation (24). If $k_1 = k_2$, then $P(T) = Q(T)$ and condition (18) is equivalent to condition (24). Hence, we have $0 < \xi^* = \tau^* < \infty$.
- (iv) There exists a unique $\tau^* \in (0, \infty)$ and $\xi^* \in (0, \infty)$ by Proposition 1 and part (iii) of Proposition 2, respectively. By equations (18), (24), and (27),

$$Q(\tau^*) = \frac{1}{k_2 - 1} \quad \text{and} \quad P(\xi^*) = Q(\xi^*) + \frac{v(k_1 - k_2)h(\xi^*)}{k_1 - 1} = \frac{1}{k_1 - 1}$$

or

$$Q(\tau^*) - Q(\xi^*) = \frac{1}{k_2 - 1} - \frac{1}{k_1 - 1} + \frac{v(k_1 - k_2)h(\xi^*)}{k_1 - 1} < 0$$

where the last inequality follows from the condition of this part (iv). Inequality (26) implies that $Q(T)$ is strictly increasing, hence $0 < \tau^* < \xi^* < \infty$.

(v) The proof of part (v) is analogous to that of part (iv) and is omitted.

□

Proof. Proposition 4

Propositions 1 and 2 imply that if $T_i \in [0, \min\{\tau_i^*, \xi_i^*\})$, then $\epsilon_i(\min\{\tau_i^*, \xi_i^*\}) > \epsilon(T_i)$ and $C_i(\min\{\tau_i^*, \xi_i^*\}) < C_i(T_i)$; if $T_i \in (\max\{\tau_i^*, \xi_i^*\}, \infty)$, then $\epsilon_i(\max\{\tau_i^*, \xi_i^*\}) > \epsilon(T_i)$ and $C_i(\max\{\tau_i^*, \xi_i^*\}) < C_i(T_i)$. Hence, for any optimal solution to problem (4)-(5), say $\{T_1^*, T_2^*, \dots, T_n^*\}$, with at least one machine i such that $T_i^* \notin [\min\{\tau_i^*, \xi_i^*\}, \max\{\tau_i^*, \xi_i^*\}]$, there exists another optimal solution $\{T_1^{**}, T_2^{**}, \dots, T_n^{**}\}$ with lower cost where

$$T_i^{**} = \begin{cases} \min\{\tau_i^*, \xi_i^*\}, & T_i^* < \min\{\tau_i^*, \xi_i^*\} \\ \max\{\tau_i^*, \xi_i^*\}, & T_i^* > \max\{\tau_i^*, \xi_i^*\} \\ T_i^*, & \text{otherwise.} \end{cases}$$

□

Proof. Proposition 5

Proposition 1 and equations (12)-(13) imply that

$$\epsilon_i(T_i) \geq \underline{\epsilon} \Leftrightarrow T_i \in [T_i^-(\underline{\epsilon}), T_i^+(\underline{\epsilon})].$$

Proposition 2 implies that $C_i(T_i)$ is strictly decreasing for all $0 \leq T_i \leq \xi_i^*$ and strictly increasing for all $T_i > \xi_i^*$. Hence, T_i^* as defined in equation (11) satisfies

$$T_i^* = \arg \min \{C_i(T_i) : \epsilon_i(T_i) \geq \underline{\epsilon}\}$$

which implies that T_i^* is a unique optimal solution to problem (9)-(10) for machine i . Because this result holds for all $i = 1, \dots, n$, $\{T_1^*, T_2^*, \dots, T_n^*\}$ is the unique optimal solution to problem (7)-(8). \square

Proof. Lemma 1

(i) Because $\epsilon_{max} \equiv \epsilon_b(\tau_b^*)$, if $\underline{\epsilon} > \epsilon_{max}$, then problem (9)-(10) is infeasible for machine b .

Because problem (7)-(8) is equivalent to the collection of n separate problems of the form of (9)-(10), problem (7)-(8) is also infeasible.

(ii) If $\underline{\epsilon} < \epsilon_{min} \equiv \min\{\epsilon(\xi_1^*), \epsilon(\xi_2^*), \dots, \epsilon(\xi_n^*)\}$, then $\epsilon_i(\xi_i^*) > \underline{\epsilon}$ for all $i = 1, 2, \dots, n$. Hence, the optimal solution to problem (7)-(8) is $\{\xi_1^*, \xi_2^*, \dots, \xi_n^*\}$, and the throughput requirement (8) is not binding at optimality.

(iii) If $\underline{\epsilon} \leq \epsilon_{max}$, then problem (7)-(8) is feasible. If $\underline{\epsilon} \geq \epsilon_{min}$, then there exists at least one machine $j \in \{1, 2, \dots, n\}$ such that $\epsilon_j(\xi_j^*) \leq \underline{\epsilon}$. Proposition 5 then implies that $\epsilon_j(T_j^*) = \underline{\epsilon}$, hence (8) is binding at optimality. To establish the strictly increasing property of C^* , we proceed by contradiction. Suppose C^* is not strictly increasing in $\underline{\epsilon}$ when $\epsilon_{max} \geq \underline{\epsilon} \geq \epsilon_{min}$; then there exist two throughput values $\underline{\epsilon}_1$ and $\underline{\epsilon}_2$ such that $\epsilon_{max} \geq \underline{\epsilon}_1 > \underline{\epsilon}_2 \geq \epsilon_{min}$ and $C^*(\underline{\epsilon}_2) = C^*(\underline{\epsilon}_1)$. As a result, (8) is not necessarily binding at optimality for problem (7)-(8) with $\underline{\epsilon} = \underline{\epsilon}_2$, and we reach a contradiction.

□

Proof. Lemma 2

(i) Because $C_i(\xi_i^*)$ is the minimum value of $C_i(T_i)$ for each $i = 1, \dots, n$, if $B < \sum_i^n C_i(\xi_i^*)$, then the budget constraint (5) is infeasible.

(ii) Proposition 1 implies that if $T_i \in [T_i^-(\epsilon_{max}), T_i^+(\epsilon_{max})]$, then $\epsilon_i(T_i) \geq \epsilon_{max}$. Let

$$\hat{T}_i^* \equiv \arg \min_{T_i \in [T_i^-(\epsilon_{max}), T_i^+(\epsilon_{max})]} \{C_i(T_i)\}, \forall i = 1, \dots, n, i \neq b$$

and $\hat{T}_b^* \equiv \tau_b^*$. Consequently, we have $\sum_{i=1}^n C_i(\hat{T}_i^*) = B_{max}$. It is straightforward to verify that for $B > B_{max}$, $\{\hat{T}_1^*, \dots, \hat{T}_n^*\}$ is an optimal solution to problem (4)-(5) such that the budget constraint (5) is not binding.

(iii) The definition of B_{max} implies that if $B = B_{max}$, then $\epsilon^* = \epsilon_{max}$ and (5) is binding at optimality. Consequently, we only need to prove that the results hold when $B_{min} \leq B < B_{max}$. Clearly, parts (i) and (ii) imply that if $B_{min} \leq B < B_{max}$, then problem (4)-(5) is feasible and $\epsilon^* < \epsilon_{max}$. By Proposition 4, there exists an optimal solution, say $\underline{T}^* = \{T_1^*, T_2^*, \dots, T_n^*\}$, that satisfies $T_i^* \in [\min\{\tau_i^*, \xi_i^*\}, \max\{\tau_i^*, \xi_i^*\}]$, $i = 1, \dots, n$. Without loss of generality, we assume that under this solution there is only one bottleneck machine, say β . Thus, $\epsilon^* = \epsilon_\beta(T_\beta^*) < \epsilon_{max} \leq \epsilon_\beta(\tau_\beta^*)$ implies that $T_\beta^* \neq \tau_\beta^*$. To prove that the budget constraint (5) must be binding at optimality, we proceed by contradiction. We let B^* be the left-hand side value of constraint (5) at optimality, assume that $B^* < B$, and construct a feasible solution to problem (4)-(5), denoted by $\hat{\underline{T}} \equiv \{\hat{T}_1, \hat{T}_2, \dots, \hat{T}_n\}$, which is strictly better than \underline{T}^* . For all $i = 1, \dots, n, i \neq \beta$, let

$\hat{T}_i = T_i^*$. To construct \hat{T}_β , we consider the following two mutually exclusive and jointly exhaustive cases.

- (a) If $B - B^* \geq C_\beta(\tau_\beta^*) - C_\beta(T_\beta^*)$, let $\hat{T}_\beta = \tau_\beta^*$. Since $T_\beta^* \neq \tau_\beta^*$, we have $\epsilon_\beta(\hat{T}_\beta) > \epsilon_\beta(T_\beta^*)$. It is straightforward to show that $\sum_{i=1}^n C_i(\hat{T}_i) \leq B$ and $\min_{i=1, \dots, n} \epsilon_i(\hat{T}_i) > \min_{i=1, \dots, n} \epsilon_i(T_i^*)$.
- (b) If $B - B^* < C_\beta(\tau_\beta^*) - C_\beta(T_\beta^*)$, let \hat{T}_β be the value such that $C_\beta(\hat{T}_\beta) = C_\beta(T_\beta^*) + (B - B^*)$ and $\epsilon_\beta(\hat{T}_\beta) > \epsilon_\beta(T_\beta^*)$. Propositions 1 and 2 imply that for all $i = 1, \dots, n$, if $T_i \in [\min\{\tau_i^*, \xi_i^*\}, \max\{\tau_i^*, \xi_i^*\}]$, then $C_i(T_i)$ and $\epsilon_i(T_i)$ are either both strictly decreasing (if $\tau_i^* < \xi_i^*$) or both strictly increasing (if $\tau_i^* > \xi_i^*$). And due to the fact that $T_\beta^* \neq \tau_\beta^*$, there exists a \hat{T}_β which satisfies $C_\beta(\hat{T}_\beta) = C_\beta(T_\beta^*) + (B - B^*)$ and $\epsilon_\beta(\hat{T}_\beta) > \epsilon_\beta(T_\beta^*)$. It is straight forward to show that $\sum_{i=1}^n C_i(\hat{T}_i) = B$ and $\min_{i=1, \dots, n} \epsilon_i(\hat{T}_i) > \min_{i=1, \dots, n} \epsilon_i(T_i^*)$.

In summary, if $B^* < B$, there exists a solution \underline{T} , that is strictly better than the optimal solution, \underline{T}^* . Thus, we reach a contradiction. The strictly increasing property of B^* follows by the same argument as that of part (iii) of Lemma 1.

□

Proof. Theorem 1

If $\epsilon_{min} \leq \underline{\epsilon} \leq \epsilon_{max}$, then $B_{min} \leq C^* \leq B_{max}$ by the definitions of ϵ_{min} , B_{min} , ϵ_{max} , and B_{max} . Moreover, part (iii) of Lemma 1 implies that $\min\{\epsilon_1(T_1^*), \epsilon_2(T_2^*), \dots, \epsilon_n(T_n^*)\} = \underline{\epsilon}$. Now suppose that $\underline{T}^* \equiv \{T_1^*, T_2^*, \dots, T_n^*\}$ is not an optimal solution to problem (4)-(5) with $B =$

C^* , and let $\hat{\underline{T}}^* \equiv \{\hat{T}_1^*, \dots, \hat{T}_n^*\} \neq \underline{T}^*$ be an optimal solution. Consequently, $\sum_{i=1}^n C_i(\hat{T}_i^*) \leq B = C^*$ and $\min\{\epsilon_1(\hat{T}_1^*), \epsilon_2(\hat{T}_2^*), \dots, \epsilon_n(\hat{T}_n^*)\} > \min\{\epsilon_1(T_1^*), \epsilon_2(T_2^*), \dots, \epsilon_n(T_n^*)\} = \underline{\epsilon}$, which implies that $\hat{\underline{T}}^*$ is also an optimal solution to problem (7)-(8), but that constraint (8) is not binding when $\underline{T} = \hat{\underline{T}}^*$. The result follows by contradiction with part (iii) of Lemma 1. To establish the uniqueness of this optimal solution to problem (4)-(5) with $B = C^*$, assume that this problem has another optimal solution \underline{T}^{**} other than \underline{T}^* . Clearly, \underline{T}^{**} is also an optimal solution to problem (7)-(8) with $\underline{\epsilon} = \epsilon^*$ which contradicts Proposition 5. \square

References

- [1] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. Society for Industrial and Applied Mathematics, New York, 1996.
- [2] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming theory and algorithms*. John Wiley & Sons, Inc., New York, 1993.
- [3] D. I. Cho and M. Parlar. A survey of maintenance models for multi-unit systems. *European Journal of Operational Research*, 51:1–23, (1991).
- [4] J. S. Dagpunar. Some necessary and sufficient conditions for age replacement with non-zero downtimes. *The Journal of the Operations Research Society*, 45(2):2251–229, 1994.
- [5] F. S. Dekker, R. and R. Wildeman. A review of multi-component maintenance models with economic dependence. *Mathematical Methods of Operations Research*, 45(3):411–435, 1997.
- [6] L. F. Hsu. Optimal preventive maintenance policies in a serial production system. *International Journal of Production Research*, 29:2534–2555, 1991.
- [7] S. M. R. Iravani and I. Duenyas. Integrated maintenance and production control of a deteriorating production system. *IIE Transactions*, 34:423–435, 2002.

- [8] A. C. Marquez, J. N. D. Gupta, and A. S. Heguedas. Maintenance policies for a production system with constrained production rate and buffer capacity. *International Journal of Production Research*, 41(9):1909–1926, 2003.
- [9] J. J. McCall. Maintenance policies for stochastically failing equipment: A survey. *Management Science*, 11(5):493–524, 1965.
- [10] R. D. Meller and D. S. Kim. The impact of preventive maintenance on system cost and buffer size. *European Journal of Operational Research*, 95:577–591, 1996.
- [11] W. P. Pierskalla and J. A. Voelker. A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23:353–388, 1976.
- [12] F. A. Schouten and S. G. Vanneste. Maintenance optimization of a production system with buffer capacity. *European Journal of Operational Research*, 82:323–338, 1995.
- [13] Y. S. Sherif and M. L. Smith. Optimal maintenance modles for systems subject to failure – a review. *Naval Research Logistics Quarterly*, 28:47–74, 1981.
- [14] L. C. Thomas. A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering*, 16:297–308, 1986.
- [15] C. Valdez-Flores and R. M. Feldman. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, 36:419–446, 1989.

- [16] H. Z. Wang. A survey of maintenance policies of deterioration systems. *European Journal of Operational Research*, 139:469–489, 2002.

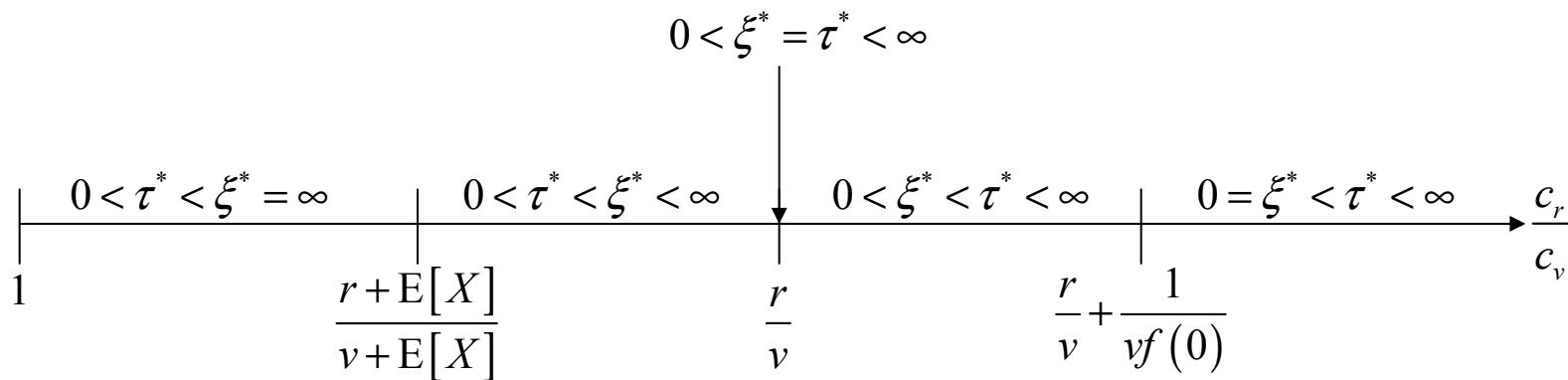


Figure 1

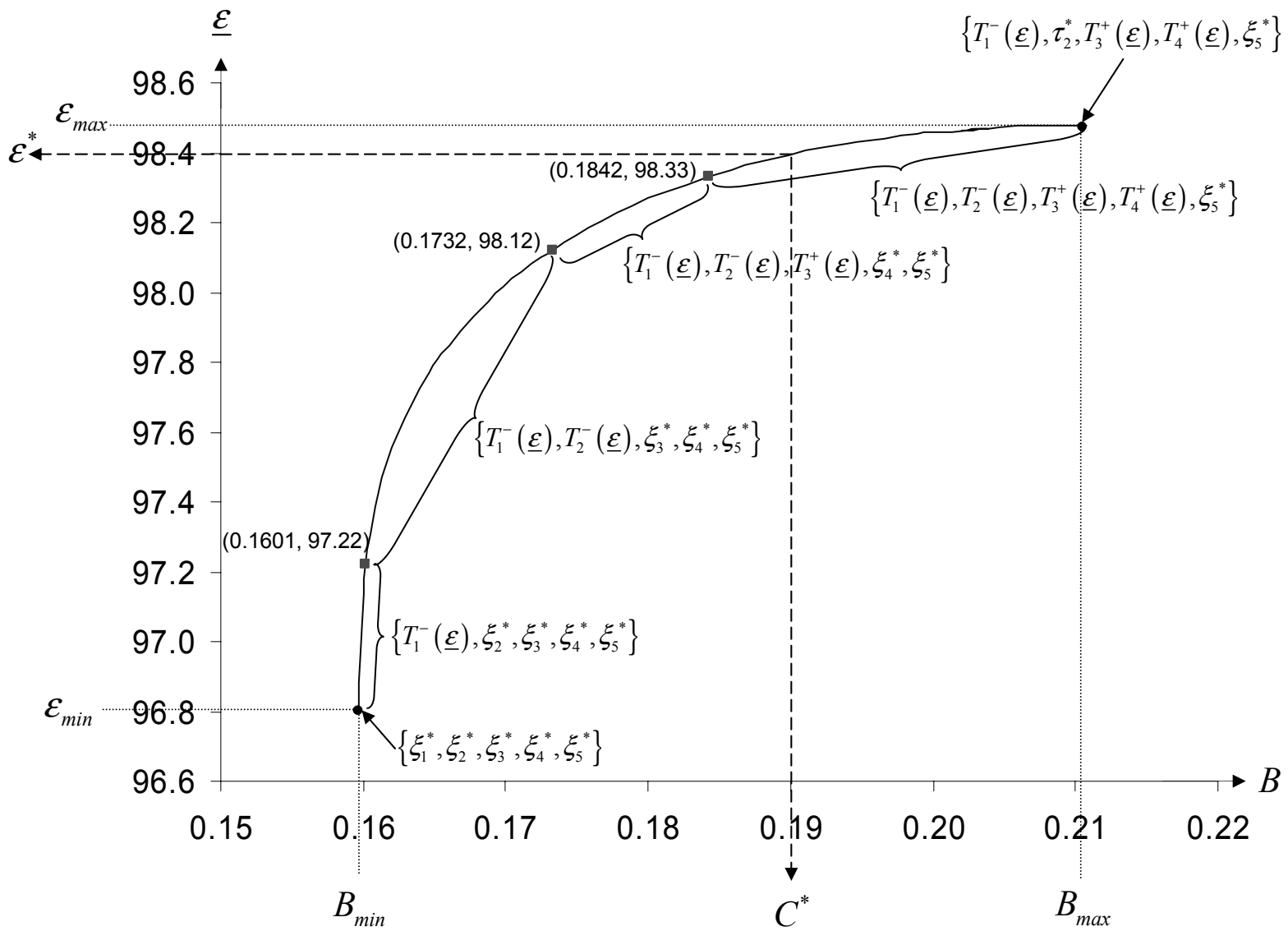


Figure 2