Diyin TANG
Jinsong YU

# OPTIMAL REPLACEMENT POLICY FOR A PERIODICALLY INSPECTED SYSTEM SUBJECT TO THE COMPETING SOFT AND SUDDEN FAILURES

# OPTYMALNA POLITYKA WYMIANY DO ZASTOSOWANIA W SYSTEMACH PODDAWANYCH PRZEGLĄDOM OKRESOWYMNARAŻONYCH NA KONKURUJĄCE USZKODZENIA PARAMETRYCZNE I NAGŁE

*This paper analyzes a replacement problem for a continuously degrading system which is periodically inspected and subject to the competing risk of soft and sudden failures. The system should be correctively replaced by a new one upon failure, or it could be preventively replaced before failure due to safety and economic considerations. Dependent soft and sudden failures are considered. The degradation process of the system observed by inspections exhibits a monotone increasing pattern and is described by a gamma process. The failure rate of the sudden failure is characterized by its dependency on the system age and the degradation state. By formulating the optimization problem in a semi-Markov decision process framework, the specific form of the optimal replacement policy which minimizes the long-run expected average cost per unit time is found, considering a cost structure that includes the cost for inspections, the cost for preventive replacement, and the costs for different failure modes. The corresponding computational algorithm to obtain the optimal replacement policy is also developed. A real data set is utilized to illustrate the application of the optimal replacement policy.*

*Keywords*: *condition-based maintenance, periodic inspections, competing risk, soft failure, sudden failure.*

*W artykule przeanalizowano problem wymiany dotyczący poddawanego przeglądom okresowym systemu ulegającego ciągłej degradacji i narażonego na konkurujące zagrożenie uszkodzeniami parametrycznymi i nagłymi. System taki powinien zostać wymieniony na nowy w ramach konserwacji korygującej w przypadku wystąpienia uszkodzenia lub też, ze względów bezpieczeństwa i względów ekonomicznych, można dokonać wymiany profilaktycznej jeszcze przed wystąpieniem uszkodzenia. W artykule rozważono przypadek zależnych od siebie uszkodzeń parametrycznych i nagłych. Proces degradacji systemu obserwowany podczas przeglądów ma charakter monotonicznie rosnący i można go opisać za pomocą procesu gamma. Intensywność uszkodzeń nagłych zależy od wieku systemu i jego stanu degradacji. Formułując problem optymalizacyjny w ramach semi-markowskiego procesu decyzyjnego, można określić formę optymalnej polityki wymiany, która minimalizowałaby długookresowy średni koszt na jednostkę czasu, z uwzględnieniem struktury kosztów, która obejmuje koszty przeglądów, koszty wymiany profilaktycznej oraz koszty różnych przyczyn uszkodzeń. Opracowano odpowiedni algorytm obliczeniowy umożliwiający ustalenie optymalnej polityki wymiany. Zastosowanie proponowanej optymalnej polityki wymiany zilustrowano na przykładzie zbioru rzeczywistych danych.*

*Słowa kluczowe*: *utrzymanie ruchu zależne od stanu technicznego urządzeń, przegląd okresowy, zagrożenie konkurujące, uszkodzenie parametryczne, uszkodzenie nagłe.*

## 1. Introduction

Pressure from the market spurs the advancement of maintenance strategies. Traditional maintenance strategies such as corrective maintenance, which only takes place when the failure is observed, and age-based or length-of-usage-based policy, which are merely performed at scheduled intervals, are not always able to meet the ever-growing demand for high level of system reliability while balancing the operating costs. In recent years, the advance of instrumentation & measurement technology has brought about the idea of condition-based maintenance (CBM). In contrast with any pre-determined maintenance strategy, CBM enables maintenance actions based on the condition monitoring information collected at real-time. Some successful examples of implementing CBM in real systems have demonstrated its efficiency and effectiveness in preventing catastrophic failures and improving maintenance performance (e.g. [1, 4, 6, 11, 12, 14]).

The interest of most CBM policies has chiefly focused on a single failure mode, and the failure mode being mostly considered is the failure due to degradation, which is usually referred to as the soft failure.

Soft failure is identified when the system state defined by the degradation level exceeds a predetermined threshold, and when it happens, the system is no longer assumed to be able to function satisfactorily or safely and it should be stopped and replaced, even if no physical failure is observed. The most frequently used CBM strategy for a continuous-state degrading system subject to the soft failure is the control-limit policy. Wang [19] jointly optimized the inspection interval and a control limit of preventive maintenance for a degradation process described by a linear growth model with random coefficients. Liao *et al.* [13] investigated a condition-based availability limit policy for a gamma-process-based degrading system, considering imperfect maintenance and an availability constraint. Policies with multiple control limits have also been proposed with the aim to improve CBM effectiveness and to satisfy different requirements. In [7], a CBM policy using multi-level control limits was proposed and optimized. The degrading system was modeled by a gamma process, and multiple control limits were used to determine the current maintenance actions and future inspection times. Elwanly *et al.* [6] analyzed a replacement problem for the exponentially increasing degradation system. They

demonstrated that the optimal replacement policy was a multi-level control-limit policy with monotonically increasing control limits.

However, considering only the soft failure seems to be inadequate for the degrading systems that are also subject to the sudden failures. In many practical situations, sudden failures are very likely to interrupt the graceful degradation and then result in more serious consequences. Therefore, in this paper, we consider a competing risk maintenance situation. The system is regarded as failed when the degradation process reaches a critical threshold or when the sudden failure occurs although the degradation process has not reached the threshold. Most of the present papers that deal with such failure scenario assume that the degradation process and the sudden failure are independent with each other. Nevertheless, even if independence is demonstrated to be appropriate for certain types of competing risks, e.g. [2, 8, 20, 21], in many situations the dependent structure between the two failure modes is of importance and should not be neglected. We consider the dependence is described by the failure rate of the sudden failure, which is influenced by both the age and the degradation state of the system. Similar assumptions were also found in Huynh *et al.* [9], Liu *et al.* [15], Castro *et al.* [3] and Huynh *et al.* [10] to deal with competing risk maintenance situation. A preventive threshold in terms of the degradation state was optimized in Huynh *et al.* [9] and Liu *et al.* [15] with different monitoring strategies. The former focused on periodically inspected systems and the latter dealt with continuously monitored ones. Castro *et al.* [13] also considered a degradation state - based control limit as the alarm of preventive maintenance, but multiple degradation processes were involved in the failure mechanism. In Huynh *et al.* [10], a novel CBM strategy was proposed using the mean residual life as the control limit. The effectiveness and potential of condition indices that are not strictly limited to failure mechanism in CBM decision-making problems was firstly investigated.

In this paper, we will focus on analyzing the optimal replacement policy for a periodically inspected system subject to the competing soft and sudden failures. This policy performs the preventive replacement only at inspection instants, and correctively replaces the system at the time of failure. We describe the degrading system using a gamma degradation process, which implies that the soft failure of the system results from a gradual and irreversible accumulation of deterioration. The failure rate corresponding to the sudden failure is described by a proportional hazards model, which means that it is influenced by both the age of the system and the state degradation of the system. Using the above models, we will show that for a certain group of maintenance situations the optimal replacement policy minimizing the average cost has a specific form, and it is also in fact a monotonically non-increasing multi-level control-limit policy. Computational algorithm to calculate the optimal multi-level control limits is developed as well, using a semi-Markov decision process framework. Finally, we will present a case study using a real laser data set from Meeker and Escorbar [16] to illustrate the proposed policy.

The paper is organized as follows. Section 2 describes the model for system degradation and sudden failures. In Section 3, we present the replacement problem. In Section 4, we examine the structure of the optimal replacement policy. The computational algorithm to calculate the optimal replacement policy is developed in Section 5. Section 6 gives an example based on a real data set. Conclusions are in Section 7.

## 2. Model of system degradation and sudden failure

### 2.1. Model of system degradation

We consider a continuously deteriorating system subject to periodic inspections. Due to limitations such as difficulty in placing sensors, the costs by condition monitoring, the internal structure of the system,

and etc., not all systems can be continuously monitored. Therefore, the periodic monitoring strategy is a typical approach applied in many real applications. The degradation state of the system is hidden and can only be known by inspections. Generally, due to the physical nature of most degradation processes, the degradation state usually presents a monotonically increasing (or decreasing) trend. Even some fluctuations may occur due to measurement errors, self-recovering mechanism, etc., when the inspection interval is long enough, the observed increment between two inspections is still very likely to be non-negative. For degradation process involving s-independent and non-negative increments, gamma process is an appropriate stochastic model to describe it (see e.g. [18]). We assume that the system degradation process $\{Y(t), t \geq 0\}$ starts from a known initial state $y_0$ and follows a gamma process $Ga(\eta(t), \gamma)$. Let $Y_n$ denote the degradation state observed at inspection time $t_n = nh$, $n = 1, 2, 3, \ldots$, where $h$ is the inspection interval. Then, the probability density function (pdf) of the increment $z_n = Y_n - Y_{n-1}$ in the inspection interval $[(n-1)h, nh]$ has the form:

$$f_{z_n}(z) = \frac{\gamma (\gamma z)^{\eta_n - \eta_{n-1} - 1} \exp(-\gamma z)}{\Gamma(\eta_n - \eta_{n-1})} \quad (1)$$

where $\eta(t)$ is a given, monotone increasing function and $\eta_n = \eta(nh)$.

If the soft failure threshold $D_f$ is given, the probability of soft failure for the next inspection interval $(nh, (n+1)nh]$ conditioning on the current observation $Y_n$ ($Y_n < D_f$) can be written as:

$$\Pr(T < t_{n+1} \mid Y_n, Y_n < D_f) = \int_{D_f - Y_n}^{+\infty} f_{z_{n+1}}(z) dz \quad (2)$$

In the following, we use homogeneous gamma process, i.e. $\eta(t) = \eta t$, at first for rigorous mathematical derivation of the optimal replacement policy. Extensions of the policy to cover a more general form of the gamma degradation model will be discussed as well.

### 2.2. Model of sudden failure

Sudden failure is a common failure mode which may interrupt a graceful degradation. In many practical situations, the sudden failure rate is very likely to be influenced by the degradation process. For example, the higher the degradation, the more the system is prone to sudden failures. Thus, it is reasonable to assume that the failure rate of sudden failures depends on the degradation process. In this paper, we assume that the failure rate of sudden failures is described by the proportional hazards (PH) model (see e.g. [5]), which explicitly includes both the effect of the age and the degradation state. It can be expressed by the following relation:

$$\lambda(t, Y_t) = \lambda_0(t; \boldsymbol{\alpha}) \theta[Y(t); \boldsymbol{\beta}]. \quad (3)$$

where $\lambda_0(t; \boldsymbol{\alpha})$ denotes baseline failure rate at time $t$ with unknown vector of parameters $\boldsymbol{\alpha}$, and $\theta[Y(t); \boldsymbol{\beta}]$ with the unknown vector of parameters $\boldsymbol{\beta}$ is a positive function dependent only on the values of the degradation state $Y(t)$.

Due to the restraint of the periodic inspection policy, the values of $Y(t)$ are only known at some discrete points of time. Thus, we approximate the failure rate at time $t$ as:

$$\lambda(t, Y_t) = \lambda_0(t; \boldsymbol{\alpha})\theta(Y_{n-1}; \boldsymbol{\beta}), \quad if \ t \in [(n-1)h, nh). \qquad (4)$$

For the system which has not failed by time $nh$, if the latest observation is $Y_n$, the conditional reliability function of surviving beyond $nh + t$ for any $0 < t < h$ can be approximated as:

$$R(Y_n, n, t) = \Pr(T > nh + t, t < h \mid T > nh, Y_n)$$
$$= \exp\left\{-\int_{nh}^{nh+t} \lambda_0(s; \boldsymbol{\alpha})\theta(Y_n; \boldsymbol{\beta})ds\right\} = \exp\left\{-\theta(Y_n; \boldsymbol{\beta})\int_{nh}^{nh+t} \lambda_0(s; \boldsymbol{\alpha})ds\right\}. \qquad (5)$$

Then, the expected sojourn time for system to keep operating during interval $[nh, nh + a)$ ($0 < a < h$) is given by

$$\tau(Y_n, n, a) = \int_0^a R(Y_n, n, t)dt = \int_0^a \exp\left\{-\theta(Y_n; \boldsymbol{\beta})\int_{nh}^{nh+t} \lambda_0(s; \boldsymbol{\alpha})ds\right\} dt. \qquad (6)$$

## 3. The optimal replacement problem

Consider a non-repairable single-unit system described as in Section 2. The degradation state of the system is hidden and the soft failure is non-self-announcing. No indicator can exhibit the degradation state except to do an inspection. The system starts working at time $t = 0$ and is inspected every $h$ time units. The inspections are assumed to be perfect and incurred a cost $C_0$. Three kinds of replacement actions are available on the system:

1. If the system's degradation state identified by inspection exceeds its soft failure threshold $D_f$, a corrective replacement is performed with the expected cost $C_2 + C$.

2. If sudden failure happens before the degradation state reaches the threshold $D_f$, the system is also correctively replaced, but with a possibly more expensive expected cost $C_1 + C$ ($C_1 > C_2$).

3. At the time of inspection, if the system still operates and its degradation state observed by inspection is below the soft failure threshold $D_f$, a preventive replacement may exert on the system instantaneously at the expected cost $C$.

After the replacement, the system is back to as-good-as-new state. Even though both the preventive and the corrective maintenance actions bring the system back to the as-good-as-new state, they are generally different in practice because the unplanned maintenance actions (i.e. corrective replacements) may have to include a larger economic loss. Moreover, the corrective replacement for sudden failure is quite possible to be more expensive than that for soft failure because of its unexpected nature and the damage resulting from physical breakdown.

In addition, we introduce the following assumptions:
1. Any replacement, whether corrective or preventive, takes negligible time.
2. The baseline function $\lambda_0(t; \boldsymbol{\alpha})$ in the PH model is non-decreasing for $t \geq 0$.
3. The link function $\theta[Y(t); \boldsymbol{\beta}]$ is a non-decreasing function of the degradation state $Y(t)$.

4. The inspection cost $C_0$ is very small compared to any replacement cost. It is reasonable because condition-monitoring loses its significance if the monitoring cost is too high.

The objective of our replacement policy is to minimize the long-run expected average cost per unit time $g$ by performing appropriate preventive replacement. Denote the state space by $\Omega = (\mathbf{H}, \mathbf{K})$, where $\mathbf{H} = \{nh; n = 0, 1, 2, ...\}$ represents the inspection times and $\mathbf{K}$ represents the possible states of observed $Y_n$. We note that discretization of the continuous degradation state is applied in the following proof procedure in order to guarantee a finite state space for the policy optimization model. We will explain the details of the discretization scheme in Section 5. Let $V(n, Y_n)$ be the relative value function (see e.g. [17]) that formulates the relative cost in the infinite-horizon decision process when the system is currently in state $(n, Y_n) \in \Omega$, and $F_{Y_{n+1}}(y \mid n, Y_n)$ be the conditional cumulative distribution function (cdf) of $Y_{n+1}$, then the optimality equations can be expressed as follows:

$$V(n, Y_n) = \begin{cases} C_2 + C + V(0, Y_0), & if \ Y_n \geq D_f \\ \min\{C + V(0, Y_0), W(n, Y_n)\}, & if \ Y_n < D_f \end{cases} \qquad (7)$$

where

$$W(n, Y_n) = \left[C_1 + C + V(0, Y_0)\right]\left[1 - R(Y_n, n, h)\right] + R(Y_n, n, h)\int_{Y_n}^{D_f} V(n+1, y)F_{Y_{n+1}}(y \mid n, Y_n)$$
$$+ R(Y_n, n, h)\Pr\left(T < t_{n+1} \mid Y_n, Y_n < D_f\right)\left[C_2 + C + V(0, Y_0)\right] + C_0 R(Y_n, n, h) - g\tau(Y_n, n, h) \qquad (8)$$

The first line of Eq. (7) describes the situation that if the observed degradation state by current inspection exceeds the soft failure threshold $D_f$, we perform an immediate corrective replacement at the expected cost $C_2 + C$ and put the system back to as-good-as-new state. The second line of Eq. (7) proposes the maintenance rule that if the observed degradation state is below the soft failure threshold $D_f$, we can either choose to preventively replace the system, or do nothing and continue operation, depending on the relative costs of the two different maintenance actions.

## 4. Structure of the optimal replacement policy

In this section, we will examine the structure of the optimal replacement policy for the replacement problem defined in Section 3.

**Theorem 1.** If $C_1 > C_2 + C_0$, the relative value function $V(n, Y_n)$ defined by the second line of Eq. (7) is non-decreasing for $(n, Y_n) \in \Omega$ and for any positive constant $g$.

**Proof.** Following the second line of Eq. (7), since $V(n, Y_n) = \min\{C + V(0, Y_0), W(n, Y_n)\}$, $V(n, Y_n) \leq C + V(0, Y_0)$. Let $V^k(n, Y_n)$ denote the relative value function at the $k$ th iteration of the value iteration algorithm. We start by defining the initial value:

$$V^0(n, Y_n) = 0 \ for \ any \ n \qquad (9)$$

$V^0(n, Y_n)$ is non-decreasing for $(n, Y_n) \in \Omega$. Next, assume that $V^k(n, Y_n)$ is non-decreasing for $Y_n < D_f$ and $V^k(n, Y_n) \leq C + V^k(0, Y_0)$ are true at the $k$ th iteration, then from Eq. (7),

$$V^{k+1}(n, Y_n) = \min\{C + V^k(0, Y_0), W^k(n, Y_n)\} \qquad (10)$$

where:

$$W^k(n,Y_n) = \left[C_1 + C + V^k(0,Y_0)\right]\left[1 - R(Y_n,n,h)\right] + R(Y_n,n,h)\int_{Y_n}^{D_f} V^k(n+1,y)F_{Y_{n+1}}(y\,|\,n,Y_n)$$
$$+ C_0 R(Y_n,n,h) + R(Y_n,n,h)\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)\left[C_2 + C + V^k(0,Y_0)\right] - g\tau(Y_n,n,h)$$
$$= \left[C_1 + C + V^k(0,Y_0)\right] - R(Y_n,n,h)\int_{Y_n}^{D_f}\left[C + V^k(0,Y_0) - V^k(n+1,y)\right]F_{Y_{n+1}}(y\,|\,n,Y_n)$$
$$- R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] - g\tau(Y_n,n,h)$$

(11)

By induction hypothesis, $V^k(n,Y_n)$ is non-decreasing for $(n,Y_n) \in \Omega$ if $Y_n < D_f$. $R(Y_n,n,h)$ and $\tau(Y_n,n,h)$ in Eq. (11) are also non-increasing functions for $(n,Y_n) \in \Omega$. Moreover, since we consider a homogeneous gamma process, $F_{Y_{n+1}}(y\,|\,n,Y_n)$ is non-increasing and $\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)$ defined by Eq. (2) is also non-decreasing for $(n,Y_n) \in \Omega$. Therefore, if $C_1 > C_2 + C_0$, $W^k(n,Y_n)$ is non-decreasing for any positive constant $g$. Since both $C + V^k(0,Y_0)$ and $W^k(n,Y_n)$ in Eq. (10) are non-decreasing, $V^{k+1}(n,Y_n)$ is non-decreasing as well, which completes the proof.

Based on Theorem 1, we are able to find the form of optimal replacement policy by analyzing the optimality equations Eq. (7) and Eq. (8).

**Theorem 2.** Let 0 represent the immediate preventive replacement upon observation of the system state, 1 represent the immediate corrective replacement for the soft failure, and $+\infty$ corresponds to no replacements. If $C_1 > C_2 + C_0$, the optimal replacement policy has the following form:

$$\delta(n,Y_n) = \begin{cases} 0, \text{ if } C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] \geq g\tau(Y_n,n,h) \\ 1, \text{ if } Y_n \geq D_f \\ +\infty, \text{ if } C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] < g\tau(Y_n,n,h) \end{cases}$$

(12)

**Proof.**
Consider the case if:

$$C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] < g\tau(Y_n,n,h),$$

then:

$$W(n,z) - C - V(0,Y_0)$$
$$= C_0 R(Y_n,n,h) + C_1\left[1 - R(Y_n,n,h)\right] + R(Y_n,n,h)\int_{Y_n}^{D_f}\left[V(n+1,y) - C - V(0,Y_0)\right]F_{Y_{n+1}}(y\,|\,n,Y_n)$$
$$+ R(Y_n,n,h)\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)\left[C_2 + C + V(0,Y_0)\right]$$
$$- R(Y_n,n,h)\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)\left[C + V(0,Y_0)\right] - g\tau(Y_n,n,h)$$
$$= C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right]$$
$$+ R(Y_n,n,h)\int_{Y_n}^{D_f}\left[V(n+1,y) - C - V(0,Y_0)\right]F_{Y_{n+1}}(y\,|\,n,Y_n) - g\tau(Y_n,n,h)$$
$$< 0$$

So that in this case the optimal decision is no replacement.
Consider the case if:

$$C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] \geq g\tau(Y_n,n,h)$$

and assume that $V(n,Y_n) = W(n,Y_n) < C + V(0,Y_0)$. For $Y_{n+1} < D_f$, we have:

$$V(n+1,Y_{n+1}) - V(n,Y_n)$$
$$= V(n+1,Y_{n+1}) - \left[C_1 + C + V(0,Y_0)\right]\left[1 - R(Y_n,n,h)\right] + R(Y_n,n,h)\int_{Y_n}^{D_f} V(n+1,y)F_{Y_{n+1}}(y\,|\,n,Y_n)$$
$$+ R(Y_n,n,h)\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)\left[C_2 + C + V(0,Y_0)\right] - C_0 R(Y_n,n,h) - g\tau(Y_n,n,h)$$
$$= V(n+1,Y_{n+1}) - \left[C + V(0,Y_0)\right] + R(Y_n,n,h)\int_{Y_n}^{D_f}\left[V(n+1,y) - C - V(0,Y_0)\right]F_{Y_{n+1}}(y\,|\,n,Y_n)$$
$$+ C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n\right) - C_0\right] - g\tau(Y_n,n,h)$$
$$< C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n\right) - C_0\right] - g\tau(Y_n,n,h) \leq 0$$

which is a contradiction to Theorem 1.
So that if:

$$C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right] \geq g\tau(Y_n,n,h),$$

the optimal decision is an immediate preventive replacement. This establishes the result.

Next, we will show in Theorem 3 that for the replacement problem defined in Section 3, the optimal replacement policy is a monotonically non-increasing multi-level control-limit policy in terms of the degradation state.

**Theorem 3.** If $C_1 > C_2 + C_0$, for all inspection times $t_n = nh$, if the observed system state $Y_n$ is below the soft failure threshold $D_f$, the optimal decision is to preventively replace the system if and only if $Y_n \geq w_n^*$, where $w_n^*$ is the optimal control limit at $t_n = nh$. The control limit $w_n^*$ is monotonically non-increasing in $n$.

**Proof.** By Theorem 2, consider the condition for the preventive replacement:

$$\frac{C_1 - R(Y_n,n,h)\left[C_1 - C_2\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right) - C_0\right]}{\tau(Y_n,n,h)} \geq g \quad (13)$$

For any $t_n = nh$, the left-hand side of Eq. (13) is non-decreasing in $Y_n$. Therefore, if $\exists w_n^* < D_f$ makes Eq. (13) holds for $Y_n = w_n^*$, then Eq. (13) still holds for any $Y_n \geq w_n^*$. In other words, given that the optimal decision for degradation state at $Y_n = w_n^*$ is to preventively replace, the optimal decision for degradation state $Y_n \geq w_n^*$ is also to preventively replace. Thus, the optimal replacement policy at time $t_n = nh$ is a control-limit policy with control limit $w_n^*$.

On the other hand, since the left-hand side of Eq. (13) is non-decreasing in $n$, so that for any $Y_n$, there exists a starting time $n^*h$ such that for any $nh \geq n^*h$ the optimal decision is to preventively replace the system. By the existence of a control limit $w_n^*$ for each inspection time and a starting time $n^*h$ for each degradation state, the control limit is $w_n^*$ monotonically non-increasing in $n$.

**Remark.** We only use the homogeneous gamma process to derive the above optimal replacement policy. However, the optimality still establishes using other degradation models, if that $\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)$ is non-decreasing for $(n,Y_n) \in \Omega$ can be validated. The non-decreasing $\Pr\left(T < t_{n+1}\,|\,Y_n, Y_n < D_f\right)$ for $(n,Y_n) \in \Omega$ indicates two characteristics of a degradation process. Firstly, at the same inspection time, the degradation process which has a more severe deterioration would have a larger probability of soft failure during next inspection interval; secondly, for the same degradation state, the system which is "older" would be more likely to confront soft failure for the next inspection

interval. These two phenomena can be found in many real situations. Even if $\Pr\left(T < t_{n+1} \mid Y_n, Y_n < D_f\right)$ is not strictly non-decreasing for $(n, Y_n) \in \Omega$, the above replacement policy still might be the optimal since $R(Y_n, n, h)$, the conditional reliability function of sudden failure, may dominate the trend of $W(n, Y_n)$. Similarly, $C_1 > C_2 + C_0$ is also a quite strong assumption in order to ensure the monotonicity; in fact, if $R(Y_n, n, h)$ decreases quickly enough as the system ages, the theorems still hold.

## 5. Computation of the Optimal Replacement Policy

We demonstrate in Section 4 that the optimal replacement policy has a specific form, and it is also a multi-level control-limit policy with non-increasing control limits $w_n^*$. To apply this policy, it is necessary to compute the minimum long-run expected average cost per unit time $g$. We thus develop a computational procedure using semi-Markov decision process (SMDP) to obtain $g$. The computation of SMDP requires discretizing the possible range of values of $Y_n \in [y_0, +\infty)$ into a finite set of states. Define $[D_f, +\infty)$ as the failure state $\mathbf{F}$. We can then divide the continuous state space of $[y_0, D_f]$ into $L$ equidistant intervals with constant length $\Delta = (D_f - y_0) / L$. If for some fixed integer $0 < \overline{k}_n < L$, $\overline{k}_n\Delta$ is defined as the control limit $\overline{w}_n$ at inspection time $t_n = nh$, then the warning state $\mathbf{W}_n$ at inspection time $t_n = nh$ will be $[\overline{k}_n\Delta + y_0, L\Delta + y_0)$, and the healthy state $\mathbf{S}_n$ will be $[y_0, \overline{k}_n\Delta + y_0)$. Thus, we approximate the state space $\Omega = (\mathbf{H}, \mathbf{K})$ by a non-decreasing homogeneous Markov chain with countable state space considering states $(k, n)$, $(\mathbf{W}_n, n)$, $(\mathbf{F}, n)$.

Based on the above discretization scheme, for the system which has not failed by time $t_n$, if $Y_n$ is known, the conditional reliability function defined in Eq. (5) is calculated as follows:

$$R(Y_n, n, t) = R(k, n, t) = \begin{cases} \Pr(T > t, t < h \mid Y_0 = y_0), & \text{for } k = 0, n = 0 \\ \Pr(T > nh + t, t < h \mid T > nh, Y_n = (k + 0.5)\Delta + y_0), & \text{for } 0 \le k < \overline{k}_n, n = 1, 2, \ldots \end{cases}$$
$$= \begin{cases} \exp\left\{-\theta(y_0; \boldsymbol{\beta}) \int_0^t \gamma_0(s; \boldsymbol{\alpha}) ds\right\}, & \text{for } k = 0, n = 0 \\ \exp\left\{-\theta[(k + 0.5)\Delta + y_0; \boldsymbol{\beta}] \int_{nh}^{nh+t} \gamma_0(s; \boldsymbol{\alpha}) ds\right\}, & \text{for } 0 \le k < \overline{k}_n, n = 1, 2, \ldots \end{cases}$$
(14)

Then, the expected sojourn time defined in Eq. (6) is calculated by:

$$\tau(Y_n, n, a) = \tau(k, n, a) = \int_0^a R(k, n, t) dt$$
$$= \begin{cases} \int_0^a \exp\left\{-\theta(y_0; \boldsymbol{\beta}) \int_0^t \lambda_0(s; \boldsymbol{\alpha}) ds\right\} dt, & \text{for } k = 0, n = 0 \\ \int_0^a \exp\left\{-\theta[(k + 0.5)\Delta + y_0; \boldsymbol{\beta}] \int_{nh}^{nh+t} \lambda_0(s; \boldsymbol{\alpha}) ds\right\} dt, & \text{for } 0 \le k < \overline{k}_n, n = 1, 2, \ldots \end{cases}$$
(15)

Next, we will derive the one-step transition probabilities in the SMDP. Using homogeneous gamma process, the probability that the value of $Y_{n+1}$ will be in state $[l\Delta + y_0, (l+1)\Delta + y_0]$ given the current value of $Y_n$ is in the interval $[k\Delta + y_0, (k+1)\Delta + y_0]$, $0 \le k < \overline{k}_n, 0 \le l < L$, is calculated by:

$$p_{(k,n),(l,n+1)} =$$
$$\begin{cases} \Pr\left(Y_{n+1} \in [l\Delta + y_0, (l+1)\Delta + y_0] \mid Y_0 = y_0\right), & \text{for } k = 0, n = 0 \\ \Pr\left(Y_{n+1} \in [l\Delta + y_0, (l+1)\Delta + y_0] \mid Y_n = (k + 0.5)\Delta + y_0\right), & \text{for } 0 \le k < \overline{k}_n, k \le l < L, n = 0, 1, \ldots \end{cases}$$
$$= R(k, n, h) \int_{(l-k-0.5)\Delta}^{(l-k+0.5)\Delta} \frac{\gamma(\gamma z)^{h-1} \exp(-\gamma z)}{\Gamma(h)} dz.$$
(16)

Similarly, the probability that the value of $Y_{n+1}$ will be in the warning state $\mathbf{W}_{n+1}$ given the current value of $Y_n$ is in the interval $[k\Delta + y_0, (k+1)\Delta + y_0]$, $0 \le k < \overline{k}_n$, is calculated by:

$$p_{(k,n),(\mathbf{W}_{n+1}, n+1)} = \begin{cases} \Pr\left(Y_{n+1} \in \mathbf{W}_{n+1} \mid Y_0 = y_0\right), & \text{for } k = 0, n = 0 \\ \Pr\left(Y_{n+1} \in \mathbf{W}_{n+1} \mid Y_n = (k + 0.5)\Delta + y_0\right), & \text{for } 0 \le k < k_n, n = 0, 1, \ldots \end{cases}$$
$$= R(k, n, h) \int_{k_{n+1}-k-0.5}^{(L-k+0.5)\Delta} \frac{\gamma(\gamma z)^{h-1} \exp(-\gamma z)}{\Gamma(h)} dz.$$
(17)

Moreover, the probability that the value of $Y_{n+1}$ will be in the failure region determined by state $\mathbf{F}$ given the current value of $Y_n$ is in the interval $[k\Delta + y_0, (k+1)\Delta + y_0]$, $0 \le k < \overline{k}_n$, is calculated by:

$$p_{(k,n),(\mathbf{F}, n+1)} = \begin{cases} \Pr\left(Y_{n+1} \in \mathbf{F} \mid Y_0 = y_0\right), & \text{for } k = 0, n = 0 \\ \Pr\left(Y_{n+1} \in \mathbf{F} \mid Y_n = (k + 0.5)\Delta + y_0\right), & \text{for } 0 \le k < k_n, n = 0, 1, \ldots \end{cases}$$
$$= R(k, n, h) \left(1 - \int_0^{(L-k+0.5)\Delta} \frac{\gamma(\gamma z)^{h-1} \exp(-\gamma z)}{\Gamma(h)} dz\right).$$
(18)

With the definition of the state space, for a fixed control limit $\chi$, the warning state $\mathbf{W}_n$ at inspection time $t_n = nh$ will be $[\overline{k}_n\Delta + y_0, L\Delta + y_0)$, where $\overline{k}_n$ can be approximated by

$$\overline{k}_n = \min\left\{k; C_1 - R(k, n, h)\left(C_1 - C_2 \Pr(T < t_{n+1} \mid Y_n = k\Delta + y_0, k < L) - C_0\right) \ge \chi\tau(k, n, h)\right\}$$
(19)

If $\overline{k}_n \ge L$, only corrective replacement is allowed at inspection time $t_n = nh$; and if $\overline{k}_n \le 0$, the system should stop operating at inspection time $t_n = nh$ with an immediate preventive or corrective replacement. The smallest time to stop operating and enforce replacement is defined as $\tilde{t}_\chi = \tilde{n}_\chi h$. Thus, the determination of a control limit $\chi$ can be transformed into determining the $\tilde{n}_\chi$. For a fixed integer $\tilde{n}_\chi$, the corresponding control limit $\chi$ is given by:

$$\chi = \frac{C_1 - R(0, \tilde{n}_\chi, h)\left[C_1 - C_2 \Pr\left(T < t_{\tilde{n}_\chi+1} \mid Y_n = y_0\right) - C_0\right]}{\tau(0, \tilde{n}_\chi, h)}$$
(20)

Once all of the quantities above are defined, for a fixed integer $\tilde{n}_\chi$, the long-run expected average cost per unit time $g(\chi)$ for the competing risk of soft and sudden failure can be obtained by solving the following systems of linear equations (see e.g. [17]):

$$V(0,0) = C_0 + (C_1 + C)[1 - R(0,0,h)] - g(\chi)\tau(0,0,h) + \sum_{l=0}^{\bar{k}_1} p_{(0,0),(l,1)} V(1,l) + p_{(0,0),(\mathbf{W}_1,1)} V(1,\mathbf{W}_1)$$
$$+ p_{(0,0),(\mathbf{F},1)} V(1,\mathbf{F}) + [1 - R(0,0,h)] V(0,0)$$

$$V(n,k) = C_0 R(k,n,h) + (C_1 + C)[1 - R(k,n,h)] - g(\chi)\tau(k,n,h) + \sum_{l=0}^{\bar{k}_{n+1}} p_{(k,n),(l,n+1)} V(n+1,l)$$
$$+ p_{(k,n),(\mathbf{W}_{n+1},n+1)} V(n+1,\mathbf{W}_{n+1}) + p_{(k,n),(\mathbf{F},n+1)} V(n+1,\mathbf{F}) + [1 - R(k,n,h)] V(0,0),$$
$$\text{for } k = 0,1,...,\bar{k}_{\tilde{n}_\chi} \text{ and } n = 1,2,...,\tilde{n}_\chi - 1$$

$$V(n,\mathbf{W}_n) = C + V(0,0), \text{ for } n = 1,2,...,\tilde{n}_\chi$$
$$V(n,\mathbf{F}) = C_2 + C + V(0,0), \text{ for } n = 1,2,...,\tilde{n}_\chi \qquad (21)$$
$$V(\tilde{n}_\chi,k) = V(\tilde{n}_\chi,\mathbf{W}_{\tilde{n}_\chi}), \text{ for } k = 0,1,...,\bar{k}_{\tilde{n}_\chi}$$
$$V(p,q) = 0, \text{ for some } p = 0,1,...,L \text{ and } q = 0,1,...,\tilde{n}_\chi.$$

where $V(\bullet,\bullet)$ is the same relative value function as defined in Section 3, considering discretized degradation states.

So that the optimal control limit $\chi^*$ and corresponding minimum long-run expected average cost per unit time $g(\chi^*)$ can be found by:

$$g(\chi^*) = \inf_{\tilde{n}_\chi} \{g(\chi)\}. \qquad (22)$$

A stopping rule for determining an appropriate $L$ to partition the continuous state space $[y_0, D_f]$ is needed in this algorithm, which could be:

$$g(\chi^* \mid L = 2^{U+1}) - g(\chi^* \mid L = 2^U) \le \delta. \qquad (23)$$

where $U$ is positive integer and $\delta$ is the selected small number.

## 6. Case study

In this section, we use a real degradation data presented by Meeker and Escobar [16] (Chapter 13, Example 13.5) to illustrate the application of our proposed optimal replacement policy. This data set consists of 20 degradation histories, describing the degradation process of some GaAs lasers subject to the competing soft and sudden failures. During the life of GaAs lasers, the degradation causes an increase in the laser's operating current. The laser is considered to be failed if the operating current increases to $D_f$ percent of its original value. On the other hand, physical breakdown due to the sudden failures may also occur and consequently interrupts the graceful degradation. For illustration purpose, we assume $D_f = 5$ as the soft failure threshold, and the operating currents were inspected every $h = 100$ hours up to $\tau = 2000$ hours or until the sudden failure happens. The degradation paths are plotted in Fig.1, in which 13 out of 20 degrade gradually till the censored time, while the other 7 samples show a sharp increase in operating current when the sudden failure occurs.

First, we fit the data using the model described in Section 2. We assume the baseline hazard function is Weibull hazard function denoted as $\lambda_0(t) = \rho t^{\rho-1} / \sigma^\rho$ and the function to quantify the effect of system degradation state on the failure rate is exponential having the form $\theta(Y_t) = \exp(cY_t)$, where $c$ is a real coefficient. Using joint likelihood function and the interior-reflective Newton method, we obtain the ML estimates for the model parameters, as shown in Table 1.

In order to validate the fitted model, we use the probability plot to assess whether the increments of laser degradation follow the gamma process $Ga(\eta h, \gamma)$ and compare the ML estimates with non-paramet-
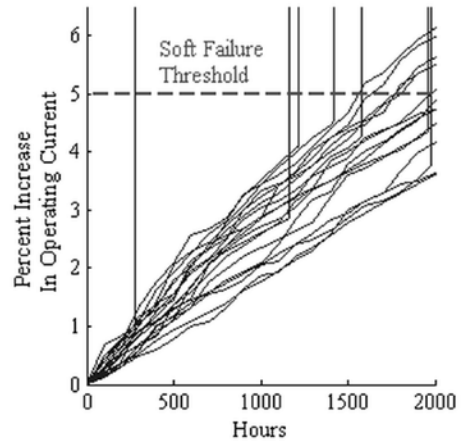


*Fig. 1. Degradation paths for the laser data*

ric Kaplan-Meier estimates for the PH model. Fig. 2 and Fig. 3 show the results, demonstrating that the fitted model is well suited for the laser data. Note that in Fig. 3, we approximate the cumulative distribution function by:

$$F(nh) = 1 - \exp\left(-\int_0^{nh} \lambda_0(t;\rho,\sigma)\theta[Y(t);c]dt\right) = 1 - \exp\left(-\sum_{k=0}^{n-1} \exp[cE(Y_k)]\int_{kh}^{(k+1)h} \rho t^{\rho-1} / \alpha^\rho dt\right). \qquad (24)$$
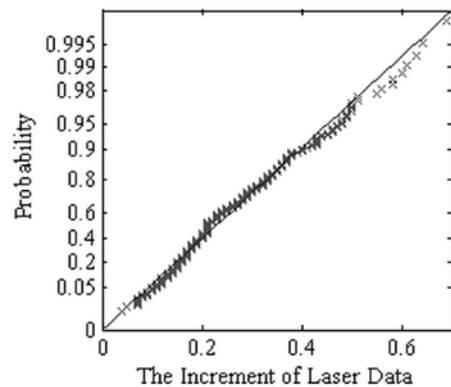


*Fig.2. Probability plot of the increments for the laser data*

Next, we consider the following replacement costs to illustrate our replacement policy: $C_0 = 100, C = 1000, C_1 = 4000, C_2 = 3000$. Since $C_1 > C_2 + C_0$, the optimal replacement policy proposed in Section 4 is applicable to this case. We partition the continuous state space of $[y_0 = 0, D_f = 5]$ into $L = 128$ intervals. The long-run expected average cost per thousand hours $g(\chi)$ for different control limits $\chi$ are plotted in Fig.4, in which we can find the minimum long-run expected average cost per thousand hours is $g(\chi^*) = 1586.052$ with the optimal control limit $\chi^* = 1587.573$. Using the control limit $\chi^* = 1587.573$, the optimal control limits $w_n^*$ in terms of degradation level at each inspection time $t_n = nh$ presents a decreasing trend in time, as shown in Fig. 5. When the observed degradation level $Y_n$ exceeds $w_n^*$, the optimal maintenance policy is to preventively replace the system. We demonstrate the appropriateness of $L = 128$ using the stopping rule of Eq. (23). The results are shown in Table 2.
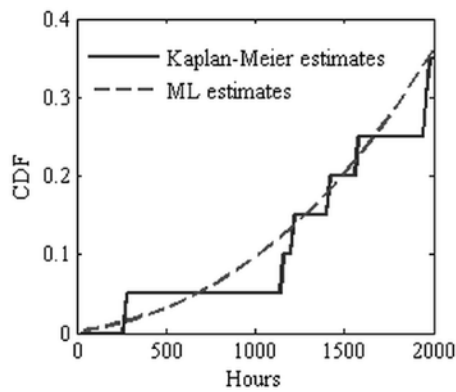
*Table 1. ML estimates for the model parameters*

| Parameters | $\eta$ | $\gamma$ | $\rho$ | $\sigma$ | $c$ |
|---|---|---|---|---|---|
| ML estimates | 4.7676 | 19.5353 | 1.3932 | 8.3859 | 0.3540 |

Fig. 3. *ML estimates and the Kaplan-Meier estimates for the marginal CDF of sudden failure times*
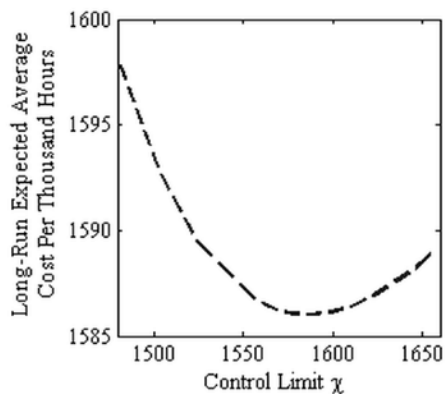


Fig. 4. *The long-run expected average costs per thousand hours for different control limits*

degradation process can be described by a gamma process and sudden failure rate by a PH model. If the preventive replacement only performs at inspection times and the sudden failure dominates in the failure mechanism, it is demonstrated that the optimal replacement policy has a specific form and it is actually a multi-level control-limit policy with control limits in terms of the degradation level. A computational
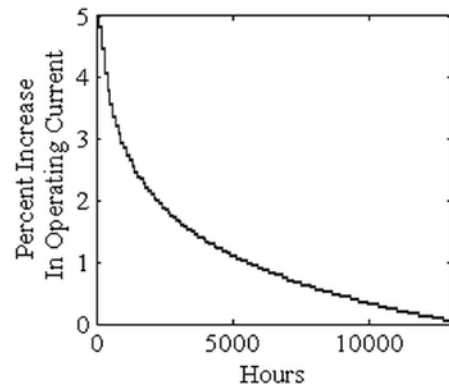


Fig. 5. *The optimal control limits* $w_n^*$ *at each inspection time*

Table 2. $\chi^*$ *and* $g(\chi^*)$ *under different combinations of* $L$

| $L$ | 16 | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|
| $\chi^*$ | 1521.154 | 1538.410 | 1570.134 | 1587.573 | 1590.965 |
| $g(\chi^*)$ | 1520.362 | 1538.742 | 1569.721 | 1586.052 | 1590.750 |

algorithm based on a SMDP framework has also been developed to obtain the optimal replacement policy. The entire procedure of applying this policy is illustrated by a real laser example.

## 7. Conclusion

In this paper, we have investigated the optimal replacement policy for a periodically inspected system subject to the competing risk of soft and sudden failures. This policy focuses on the system whose

## References

1. Besnard F, Bertling L. An approach for condition-based maintenance optimization applied to wind turbine blades. IEEE Transactions on Sustainable Energy 2010; 1(2): 77-83, http://dx.doi.org/10.1109/TSTE.2010.2049452.
2. Bocchetti D, Giorgio M, Guida M, Pulcini G. A competing risk model for the reliability of cylinder liners in marine diesel engines. Reliability Engineering & System Safety 2009; 94(8): 1299-1307, http://dx.doi.org/10.1016/j.ress.2009.01.010.
3. Castro IT, Caballé NC, Pérez CJ. A condition-based maintenance for a system subject to multiple degradation processes and external shocks. International Journal of Systems Science 2013; ahead-of-print: 1-13, http://dx.doi.org/10.1080/00207721.2013.828796.
4. Coats D, Cho K, Shin YJ, Goodman N, Blechertas V, Bayoumi AE. Advanced time-frequency mutual information measures for condition based maintenance of helicopter drive trains. IEEE Transactions on Instrumentation and Measurement 2004; 60(8): 2984-2994, http://dx.doi.org/10.1109/TIM.2011.2122370.
5. Cox DR, Oakes D. Analysis of survival data. Chapman and Hall; 1984.
6. Elwany AH, Gebraeel NZ, Maillart LM. Structured replacement policies for components with complex degradation processes and dedicated sensors. Operations research 2011; 59(3): 684-695, http://dx.doi.org/10.1287/opre.1110.0912.
7. Grall A, Bérenguer C, and Dieulle L. A condition-based maintenance policy for stochastically deteriorating systems. Reliability Engineering & System Safety 2002; 76(2): 167-180, http://dx.doi.org/10.1016/S0951-8320(01)00148-X.
8. Huang W, Askin RG. Reliability analysis of electronic devices with multiple competing failure modes involving performance aging degradation. Quality and Reliability Engineering International 2003; 19(3): 241-254, http://dx.doi.org/10.1002/qre.524.
9. Huynh KT, Barros A, Bérenguer C, Castro IT. A periodic inspection and replacement policy for systems subject to competing failure modes due to degradation and traumatic events. Reliability Engineering & System Safety 2011; 96(4): 497-508, http://dx.doi.org/10.1016/j.ress.2010.12.018.
10. Huynh KT, Castro IT, Barros A, Bérenguer C. On the use of mean residual life as a condition index for condition-based maintenance decision making. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2014; 44(7): 877-893, http://dx.doi.org/10.1109/TSMC.2013.2290772.
11. Jiang R, Yu J, Makis V. Optimal Bayesian estimation and control scheme for gear shaft fault detection. Computers & Industrial Engineering

2012; 63(4): 754-762, http://dx.doi.org/10.1016/j.cie.2012.04.015.

12. Kim MJ, Jiang R, Makis V, Lee CG. Optimal Bayesian fault prediction scheme for a partially observable system subject to random failure. European Journal of Operational Research 2011; 214(2): 331-339, http://dx.doi.org/10.1016/j.ejor.2011.04.023.

13. Liao H, Elsayed EA, Chan LY. Maintenance of continuously monitored degrading systems. European Journal of Operational Research 2006; 175(2): 821-835, http://dx.doi.org/10.1016/j.ejor.2005.05.017.

14. Lin D, Wiseman M, Banjevic D, Jardine, AK. An approach to signal processing and condition-based maintenance for gearboxes subject to tooth failure. Mechanical Systems and Signal Processing 2004; 18(5): 993-1007, http://dx.doi.org/10.1016/j.ymssp.2003.10.005.

15. Liu X, Li J, Al-Khalifa KN, Hamouda AS, Coit DW, and Elsayed E A. Condition-based maintenance for continuously monitored degrading systems with multiple failure modes. IIE Transactions 2013; 45(4): 422-435, http://dx.doi.org/10.1080/0740817X.2012.690930.

16. Meeker WQ, Escobar LA. Statistical methods for reliability data. John Wiley & Sons; 1998.

17. Tijms HC. Stochastic models: an algorithmic approach. John Wiley & Sons; 1994.

18. van Noortwijk JM. A survey of the application of gamma processes in maintenance. Reliability Engineering & System Safety 2009; 94(1): 2-21, http://dx.doi.org/10.1016/j.ress.2007.03.019.

19. Wang WB. A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance. International Journal of Production Research 2000; 38(6): 1425-1436, http://dx.doi.org/10.1080/002075400188933.

20. Zhu Y, Elsayed EA, Liao H, Chan LY. Availability optimization of systems subject to competing risk. European Journal of Operational Research 2010; 202(3): 781-788, http://dx.doi.org/10.1016/j.ejor.2009.06.008.

21. Ye ZS, Xie M, Tang LC, Shen Y. Degradation-based burn-in planning under competing risks. Technometrics 2012; 54(2): 159-168, http://dx.doi.org/10.1080/00401706.2012.676946.

**Diyin TANG**
**Jinsong YU**
School of Automation Science and Electrical Engineering
Beihang University
No. 37 Xueyuan Road, Haidian District
Beijing, China, 100191

Collaborative Innovation Center of Advanced Aero-Engine
Beijing, China, 100191
E-mails: amytdy@asee.buaa.edu.cn, yujs@buaa.edu.cn