Linear and nonlinear preventive maintenance models

Shaomin Wu, and Ming J. Zuo *Senior member, IEEE*,

Abstract

Preventive maintenance (PM) is a maintenance program with activities initiated at predetermined intervals, or according to prescribed criteria, and intended to reduce the probability of failure, or the degradation of the functioning of an item. In the literature, a number of PM models have been introduced to depict the effectiveness of PM. Based on these models, approaches to scheduling PM policies have been considerably studied. This paper attempts to review existing PM models, and investigate their inter-relationships. We then categorize these models into three classes: linear, nonlinear, and a hybrid of both. These three PM model classes depict the relationships of the hazard functions before, and after a PM. Possible extensions to these three PM models are discussed. The statistical properties for models are derived, and approaches to optimizing the PM policy are given.

Index Terms

Preventive maintenance, corrective maintenance, linear preventive maintenance, nonlinear preventive maintenance, hazard function, maintenance effectiveness.

ACRONYM[1]

CM	Corrective Maintenance
PM	Preventive Maintenance
IFR	Increasing Failure Rate
DFR	Decreasing Failure Rate
IFRA	Increasing Failure Rate on Average
DFRA	Decreasing Failure Rate on Average
NBU	New Better than Used
NWU	New Worse than Used

Corresponding author: Shaomin Wu is with School of Applied Sciences, Cranfield University, Bedfordshire MK43 0AL, UK. E-mail: shaomin.wu@cranfield.ac.uk.

Ming J. Zuo is with Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta T6G 2G8, Canada. E-mail: ming.zuo@ualberta.ca

[1]The singular and plural of an acronym are always spelled the same .

NOTATION

t	A non-negative number, it is reset to zero at each PM.
$h_0(t)$	Hazard function of the system when no PM is conducted on it, or naked hazard function.
$h_k(t)$	Hazard function of the system at time t after the kth PM, where $k = 1, 2, 3, \ldots..$
t_0	Scheduled interval before the first PM.
t_k	Scheduled interval between the kth PM, and the $(k + 1)$th PM, where $k = 1, 2, 3\ldots.$
a, b, α, β	Non-negative parameters.

I. INTRODUCTION

Maintenance activities (e.g., CM, and PM) can change a maintained system in one of two ways. Accordingly, the maintenance can be described as better, or worse. A better maintenance decreases the hazard rate and/or virtual age of a system, whereas a worse one increases them, or even brings the system to fail or break down.

The effectiveness of a better maintenance can be further classified into one of the three situations: *perfect*, *minimal*, and *imperfect*. A perfect maintenance is assumed to restore a system to be as good as new (AGAN). A minimal maintenance restores a system to a state the same as just before the maintenance, or as bad as old (ABAO). An imperfect maintenance may bring a system to any condition between AGAN and ABAO. In reality, however, both CM, and PM are usually imperfect. Pham & Wang [1], and more recently Doyen & Gaudoin [2] have given useful surveys on imperfect maintenance models.

A worse maintenance deteriorates the health of a system compared to what it was prior to maintenance. In the case of technology changes, a system can be brought to a state better than its AGAN state, but these two situations are outside the scope of this article.

Modelling the effectiveness of PM is an active research topic; see [3]–[21] for examples. It is an essential requirement in various scenarios, for example, when people plan maintenance strategies, select maintenance contractors, or estimate the residual lifetime for some important industrial systems (for example, nuclear power plants, planes, trains) put up for re-sale at the end of their planned life.

There are many papers modelling the effectiveness of PM. Most of them model the hazard rates of maintained systems after PM interventions [3]– [4]. According to a taxonomy given by [5], existing PM models are categorized into three groups: age reduction models, hazard rate reduction models, and a hybrid of both.

- *Age reduction models* These models are developed by considering age reduction in the hazard function ([3], [6]– [7]). Using the concept of age reduction, we might say that a certain PM has reduced the virtual age of a maintained system to a younger age, for example, from an age of t years old to "as good as $t - \tau$ years old" where $\tau > 0$. That is, the hazard function changes from $h(t)$ before a PM to $h(t - \tau)$ after a PM.
- *Hazard rate reduction models* These models are developed considering the reduction of the hazard rate of a system [5]– [8]. This group of models assumes that the hazard rate of a system changes, for example, from $h(t)$ before a PM to $ah(t)$ after a PM.

- *Hybrid models* These models are combinations of the above two groups [4], [9]; the hazard rate changes from $h(t)$ to $ah(t-t_0)$, for example.

However, if one reviews existing PM models, the following weaknesses can be found in those models.

- Although many PM models have been developed, little record on comparing these models has been found.
- Parameters in lifetime distributions (such as the Weibull distribution) usually have their physical meanings (for example, the location parameter, and the scale parameter, in the Weibull distribution). However, parameters in the existing PM models are not given any physical explanation, which can limit the applications of these models because of a lack of proper interpretation of the parameters.
- Assumptions in some existing PM models may not be appropriate, which are elaborated upon in Section III.

The main theme of this paper is to briefly review existing PM models, explore their interrelationships, and extend them to three new ones: linear, nonlinear and their hybrid. A PM model is *linear* if the maintained system has hazard rates $h_k(t)(k=1,2,...)$ immediately after the kth PM with $h_k(t)=ah_{k-1}(t)+b$, *nonlinear* if $h_k(t)=h_{k-1}(\alpha t+\beta)$, or *hybrid* if $h_k(t)=ah_{k-1}(\alpha t+\beta)+b$, where a,b,α, and β are non-negative parameters, and $t>0$. Physical interpretation of the parameters in these models will be given. More general extensions of these models will also be provided.

Although we discuss PM models and PM policy development in this paper, our primary focus is on the comparison of commonly studied PM models. The paper does not pretend to give a comprehensive view of the topic of existing PM models, and it emphasizes on modelling rather than on statistical inference. We have tried to make Section II reasonably complete; however, those papers which are not included were either considered not to bear directly on the topic of this paper, or were inadvertently overlooked. Our apologies are extended to both the researchers and readers if any relevant papers are omitted.

The paper is structured as follows. Section II briefly reviews the existing PM models, and explores their interrelationships. Section III introduces two new PM models, accompanied by their statistical properties. Section IV studies the periodic PM policy for the two PM models, and provides conditions on the existence of solutions. Section V offers discussion on possible extensions of the newly proposed PM models, and their corresponding properties. Finally, Section VI arrives at conclusions, and further work that might be important.

II. Existing PM models, and their properties

Assume that PM actions are carried out at times $t_0, t_0+t_1, t_0+t_1+t_2, ...$ CM on failure between adjacent PM actions is assumed to be minimal. Time on either PM or CM is negligible. $h_k(t)$ is the hazard rate of the system after the kth PM intervention (the initial hazard function is $h(t)(=h_0(t))$). If the kth PM is performed, the induced hazard function changes from $h_{k-1}(t)$ to $h_k(t)$ after the PM. $h_k(t)$ is assumed to be increasing in t. The time t is reset to zero at each PM. Denote $F_k(t)=1-\exp[-\int_0^t h_k(s)ds]$, and $F(t)=1-\exp[-\int_0^t h_0(s)ds]$.

Definition 1: Better PM: The kth PM is a better PM if $h_{k-1}(t+t_{k-1})\geq h_k(t)$.

A. Existing PM models

The effectiveness of a PM may be AGAN or ABAO. The criteria are given as follows.

Definition 2: The kth PM is AGAN if

$$h_k(t) = h_0(t). \tag{1}$$

The kth PM is ABAO if

$$h_k(t) = h_{k-1}(t + t_{k-1}). \tag{2}$$

However, we should note that, in the case of ABAO, a PM is useless.

In the Malik model [3], the improvement of the kth PM is that the t years old system is no longer that old, and its post-maintenance age is reduced from t to $\mu_k t$ in terms of its reliability, where μ_k varies between zero and one. If we simply concern the age reduction and hazard function, then the effect of the maintenance can also be expressed by hazard functions as follows.

Definition 3: (Malik [3]) We say that the kth PM is MAL if

$$h_k(t) = h_0(t + \sum_{i=0}^{k-1} \mu_i t_i), \tag{3}$$

where $0 < \mu_i < 1$.

The values of parameter μ_i in the MAL model (3) are restricted to (0,1). However, if the range of μ_i can be extended to include the two end points, then the MAL model is an extension of the ABAO model, or the AGAN model. That is, the MAL model reduces to the ABAO model if $\mu_i = 1$, and to the AGAN model if $\mu_i = 0$.

Nakagawa [5] proposes two PM models; one is an age reduction model, and the other is a hazard rate reduction model. In what follows, these two models are referred to as NAK1, and NAK2, respectively.

Definition 4: (Nakagawa [5]) We say that the kth PM is NAK1 if

$$h_k(t) = h_0(t + \sum_{i=0}^{k-1}(t_i \prod_{j=i}^{k-1} \nu_{1j})), \tag{4}$$

where $0 < \nu_{10} < \nu_{11} < ... < \nu_{1k-1} < 1$.

If we compare (3) with (4), and let $\sum_{i=0}^{k-1}(t_i \prod_{j=i}^{k-1} \nu_{1j}) = \sum_{i=0}^{k-1} \mu_i t_i$ for $\forall k$, then model NAK1 is equivalent to model MAL. Obviously, $\forall j$, model NAK1 reduces to model ABAO if $\nu_{1j} = 1$, and to model AGAN if $\nu_{1j} = 0$.

Definition 5: (Nakagawa [5]) We say that the kth PM is NAK2 if

$$h_k(t) = h_0(t) \prod_{i=0}^{k-1} \nu_{2i}, \tag{5}$$

where $1 \leq \nu_{20} \leq \nu_{21} \leq \ldots \leq \nu_{2k-1}$.

Obviously, if ν_{2k} is set to 1 for all k, then the NAK2 model promotes to the AGAN model.

Model NAK1 assumes that PM activities can bring the maintained system younger, and Model NAK2 regards that a PM can first bring the hazard rate to zero, and then increase more quickly than it did in the previous PM interval.

Lin et al. [4] combine models NAK1 and NAK2, and propose the following PM model to link the hazard functions between two adjacent working periods.

Definition 6: (Lin et at [4]) We say that the kth PM is LIN if

$$h_k(t) = \lambda_{1k} h_{k-1}(\lambda_{2k} t_{k-1} + t). \tag{6}$$

Obviously, if $\lambda_{1k} = 1$ (or $\lambda_{2k} = 0$) for all k, then the LIN model is equivalent to the NAK1 (or NAK2) model.

All of the above PM models can be applied to both periodic, and sequential PM modelling. Canfield [6] only considers the periodic PM case. He distinguishes between the level of the hazard, and the shape of the hazard function as they are related to system degradation with time. The hazard level reflects the extent of the system degradation. The shape of the hazard function at a given time reflects the rate at which the hazard is changing. He regards the effective age after PM reduces to $t - \tau$ if the item's effective age was t just prior to this PM, while the hazard level remains unchanged, where $\tau(\geq 0)$ is the restoration interval at the effective age of the item due to the kth PM. The restoration interval τ in this model is an index for measuring the quality of PM.

Definition 7: We say that the kth PM is CAN if

$$h_k(t) = h_0(t + k(T - \tau)) + \sum_{i=1}^{k} \left\{ (h_0((i-1)(T-\tau) + T) - h_0(i(T-\tau)) \right\}, \tag{7}$$

where T is a fixed constant time length between two adjacent PM actions.

When $\tau = T$, and suppose $h_0(0) = 0$, the CAN model reduces to

$$h_k(t) = h_0(t) + k h_0(T). \tag{8}$$

Parameter τ in the CAN model is assumed to be a fixed constant. Reference [10] considers τ as a random variable and develops PM policies.

Kijima et al. [11], and Kijima [12] introduce two types of CM models, type I and type II, using the concept of virtual age. The idea is to distinguish between the system's age, which is the time elapsed since the system was new, usually at time $t = 0$; and the virtual age of the system, which describes its present health condition when compared to a new system. The two models are $V_k = V_{k-1} + \kappa_k X_k$, and $V_k = \kappa_k(V_{k-1} + X_k)$, where V_k is the virtual age of the system immediately after the kth repair, and κ_k is a parameter. Interesting extensions on the virtual age concept have been made by other authors. For example, Dagpunar [14] considers the case in which the virtual age after the kth CM can be expressed as $V_k = \phi(V_{k-1} + X_k)$ (where $\phi(.)$ is an arbitrary scaling function that models the effects of CM); Dorado [15] studied nonparametric statistical inference in a model slightly more general than Kijima's models. More references can be found in [16], [17]. Kijma's virtual age concept was originally introduced to model the effectiveness of CM activities. It has been applied to the PM case recently by some authors (for example, [18], [19]).

The type I, and type II models [11], [12] share similar expressions of hazard functions. Therefore, we only discuss type I as an example.

Definition 8: We say that the kth PM is KIJ if

$$h_k(t) = h_0(\kappa_k t + \sum_{i=1}^{k-1} \kappa_i t_i). \tag{9}$$

The KIJ model can be seen as an extension of the NAK1, and MAL models.

Seo & Bai [13] introduced a periodic PM model. They define $h_k(\omega_{k-1}(T)) = h_{k-1}(\Omega(\omega_{k-2}(T), T))$, where $\Omega(.)$ and $\omega(.)$ are specified functions, and T is a fixed constant time length between two adjacent PM actions. This model can be regarded as an extension of the KIJ model for periodic PM.

Another interesting model is the geometric process for CM. Lam [22] defines the geometric process as an alternative to the NHPP: a sequence of random variables $\{X_k, k = 1, 2, ...\}$ is a geometric process if the distribution function of X_k is given by $F(\alpha^{k-1}t)$ for $k = 1, 2, ...$, and α is a positive constant. The hazard rate changes from $h(t)$ before a CM activity to $\alpha h_{k-1}(\alpha t)$ after the CM. The change is similar to the hybrid PM models. Wang & Pham [23] later refer a process similar to the geometric process as a quasi-renewal process. Finkelstein [24] develops a very similar model where he defines a general deteriorating renewal process such that $F_{k+1}(t) \leq F_k(t)$. Wu & Clements-Croome [25] extend the geometric process by replacing its parameter α^{k-1} with $\alpha_1\alpha^{k-1} + \beta_1\beta^{k-1}$, where $\alpha > 1$, and $0 < \beta < 1$. The geometric process has been studied by many authors (for example, see [7], [26]–[29]). However, we have found very few works in the application of the geometric process to modelling PM. We hence will not discuss this model in detail.

B. Interrelationships

On the basis of the above discussion, if we use $Y \Longrightarrow Z$ to denote that Z can be derived from Y, the chain of implications in Fig. 1 exists among the existing sequential PM models.

$$\begin{tikzcd}
\mathrm{LIN} \arrow[r, Rightarrow] \arrow[d, Rightarrow] \arrow[dr, Rightarrow] & \mathrm{MAL} \arrow[r, Rightarrow] \arrow[d, Leftrightarrow] & \mathrm{ABAO} \\
\mathrm{NAK2} \arrow[d, Rightarrow] & \mathrm{NAK1} \arrow[dl, Rightarrow] & \mathrm{KIJ} \arrow[l, Rightarrow] \arrow[ul, Rightarrow] \\
\mathrm{AGAN} & &
\end{tikzcd}$$

Fig. 1. Interrelationships of existing PM models.

From the relationships shown in Fig. 1, we conclude that all of the existing PM models can be categorized to be special cases of the LIN model, and the KIJ model. The CAN model, and the model by Seo & Bai [13] are not included in Fig. 1 as these two models are periodic ones.

C. A new categorization

Unlike the classification of the PM models used by [5], the following category is created from the perspective of the aging property. All of the PM models in the preceding subsection can be categorized into one of the following classes, or a combination of the two classes.

- **Age reduction PM models:** PM modelled by AGAN, MAL, NAK1, or CAN assumes that the PM restores the maintained system to a younger age. Apparently, the model introduced by Seo & Bai [13] also falls into this category. After a PM, the system will follow the deteriorating speed from its younger age point. The parameters in the age reduction PM models indicate how much a PM has reduced the functional age of a maintained system. These parameters are μ_i, ν_{1j}, and τ in the MAL, NAK1, and CAN models, respectively. They are called *age reduction parameters* in what follows.
- **Age defying PM models:** PM modelled by NAK2 defies the age of the system after it has been maintained. The ageing of the system after a PM will slow down (or speed up in some cases). The effect of the PM mainly influences the future system degradation. The parameters in this category measure the speed of deterioration of the maintained system after PM has been conducted. These parameters are ν_{2i} in the NAK2 model, and they are called *age defying parameters*.

PM in the LIN model can function as both age reduction, and age defying. The parameters λ_{2k}, and λ_{1k} in the LIN model are the age reduction parameter, and the age defying parameter, respectively.

III. Linear, and nonlinear PM

Aside from the LIN model, the PM models reviewed in Section II consider the PM improvement simply from one aspect: either age reduction, or age defying. The PM models introduced in this section depict the PM effectiveness from both aspects. These models are called linear PM models, and nonlinear PM models, respectively.

A. Linear PM

The NAK2 model assumes that the hazard rate right after PM reduces to 0, and then increases more quickly than it did in the previous PM interval. These assumptions might be too rigorous, and even unrealistic in some scenarios, which can limit the models applications in practice. For example, in some scenarios, a PM, such as cleaning, adjustment, alignment, and lubrication work, may not always reduce the system's age or hazard rate to zero [30]. Instead, it may only reduce the degradation rate of the system to a certain level. Therefore, a reasonable extension is to relax the assumptions that the hazard rate reduces to a certain level after better maintenance, and then increases more quickly than it did in the previous PM interval. This relaxation of the assumption leads to the following model.

Definition 9: The kth PM is called *linear PM* if

$$h_k(t) = ah_{k-1}(t) + b, \tag{10}$$

where a, and b are parameters; $t \in (0, t_k)$ for $k = 1, 2, ...$, and $a, b > 0$.

The reason we call the kth PM a *linear PM* is that the relationship between the two adjacent hazard functions before and after the kth PM is linear. We also call (10) a *linear PM model*.

Parameters in (10) can have their physical meanings. Parameter a indicates a degree of deterioration after PM. The system deteriorates faster than before if $a > 1$, deteriorates slower than before if $a < 1$, or keeps the same shape but different locations of the hazard rate as before if $a = 1$ and $b \neq 0$. a is called an *age defying parameter* of the linear PM model, although the PM does not defy the age of the maintained system in the case of $a > 1$. Parameter b indicates the starting value of the hazard rate immediately after a PM. The PM is a worse maintenance if $b > h_{k-1}(t_{k-1})$, and it is a better maintenance if $b < h_{k-1}(t_{k-1})$. Therefore, we call b a *location parameter*.

Equation (10) reduces to the NAK2 model if $b = 0$; and to the AGAN model if $b = 0$, and $a = 1$. Parameters a, and b reflect the performance of the linear PM. a can be limited within $(0, 1)$ for better PM.

If all of the first k PM are linear PM, (10) can also be written as

$$h_k(t) = A_k h_0(t) + B_k, \tag{11}$$

where $A_0 = 1$, $B_0 = 0$, $B_1 = b$, $A_k = a^k$, and $B_k = \dfrac{a^k + a - 2}{a - 1} b$.

It is easy to derive the following Lemma.

Lemma 1: If all of the first k PM are linear PM, we have

$$F_k(t) = 1 - e^{-B_k t}(1 - F(t))^{A_k}. \tag{12}$$

Theorem 1: If all of the first k PM are linear PM, and $1 - F(t)$ belongs to IFR, DFR, IFRA, DFRA, NBU, or NWU, then $1 - F_k(t)$ is in the same category for $k = 2, 3, \ldots$.

The proof of Theorem 1 is given in Appendix.

B. Nonlinear PM

Models MAL, NAK1, and CAN consider age-reduction phenomenon after PM. If a PM can defy and also reduce the age of a maintained system, then the following model is more appropriate to describe such scenarios.

Definition 10: Assume $h_0(t)$ is a nonlinear function with respect to t. The kth PM is called *nonlinear PM* if

$$h_k(t) = h_{k-1}(\alpha t + \beta), \tag{13}$$

where $\alpha(> 0)$, and $\beta(\geq 0)$ are parameters; and $t \in (0, t_k)$.

In this model, α plays a role in defying or accelerating degradation of a maintained system due to the effectiveness of PM. A PM defies the age of a maintained system for $\alpha \in (0, 1)$, and it accelerates the deterioration of the system for $\alpha \in (1, \infty)$. β is the value that a PM brings the system's age to in terms of the immediate proceeding PM interval. Hence, we call α an *age defying parameter*, and β a *location parameter*.

The linearity between $h_k(t)$ and $h_{k-1}(t)$ in (13) depends on $h_{k-1}(t)$: if $h_{k-1}(t)$ is linear with respect to t, then the relationship is linear; otherwise, it is nonlinear. In the case that $h_{k-1}(t)$ is linear, (13) is equivalent to (10).

This equivalence is the reason we assume $h_0(t)$ is nonlinear in the definition. We also call (13) a *nonlinear PM model*.

The two parameters α, and β in (13) estimate the effectiveness of the PM: a better maintenance if $0 < \alpha < 1$, and $\beta < t_{k-1}$.

If all of the first k PM are nonlinear PM, another expression of (13) is given as

$$h_k(t) = h_0(\Phi_k t + \Psi_k), \tag{14}$$

where $\Phi_0 = 1$, $\Psi_0 = 0$, $\Psi_1 = \beta$, $\Phi_k = \alpha^k$, and $\Psi_k = \dfrac{\alpha^k + \alpha - 2}{a-1}\beta$.

Lemma 2: If all of the first k PM are nonlinear PM, for nonlinear PM, we have

$$F_k(t) = 1 - (1 - F(\Phi_k t + \Psi_k))^{\frac{1}{\Phi_k}} (1 - F(\Psi_k))^{-\frac{1}{\Phi_k}}. \tag{15}$$

Theorem 2: If all of the first k PM are nonlinear PM, and $1 - F(t)$ is IFR (or DFR), then $1 - F_k(t)$ of the nonlinear PM model (13) is IFR (or DFR) for $k = 1, 2, 3, \ldots$.

IV. PM POLICY OPTIMIZATION

A PM policy specifies how PM activities should be scheduled. In the reliability and maintenance literature, two PM policies are commonly discussed: periodic PM, and sequential PM. The periodic policy schedules PM activities at fixed time periods, for example, $T, 2T, 3T, \ldots$, whereas the sequential policy schedules PM activities at a sequence of time intervals, $t_1, t_2, \ldots$, that can be unequal. Obviously, if we let $t_1 = t_2 = \ldots = T$, then the sequential PM is equivalent to the periodic PM. Hence, we focus on the sequential PM in what follows, and searching maintenance intervals aiming at optimizing overall cost.

We make the following assumptions for the maintenance policy optimization.

- The planning horizon is infinite.
- The hazard functions, $h_0(t)$, is continuous, and strictly increasing if there are no PM interventions.
- The times for PM, minimal repair, and replacement are negligible.
- PM is performed at $t_1, t_1 + t_2, \ldots, \displaystyle\sum_{i=1}^{N-1} t_i$, and the system is replaced at $\displaystyle\sum_{i=1}^{N} t_i$.
- Minimal repairs are used for failures between PM.
- The system is restored to as good as new state at replacement.

To derive the expected cost expression, we assume that the planning horizon is infinite, the system is replaced after $N-1$ PM, and the system is brought to an AGAN state at replacement.

In what follows, we consider both the linear, and nonlinear cases.

A. Linear PM model case

Assume all PM are linear, the total number of minimal repairs is given by $\sum_{k=0}^{N-1} \int_0^{t_k} (A_k h_0(x) + B_k) dx$, the total number of PM is $N-1$, and there is one replacement. Then the expected cost per unit time between two

Inputs:	
$h_0(t)$:	hazard function;
c_m:	cost of minimal repair;
c_p:	unit cost of PM;
c_r:	replacement cost;
Outputs:	
N^*:	optimal number of PM before a replacement;
t_k^*:	optimal time interval for the kth PM;
Steps:	
1:	for $N_L = 1, 2, ..., N$ do;
2:	obtain t_k by solving equation (17);
3:	if inequalities (18) and (19) are satisfied, then;
4:	calculate $C_L(t_1, t_2, ..., t_{N-1}, N)$;
5:	set $T^* \leftarrow T$; $N^* \leftarrow N$;
6:	break;
7:	end;
8:	end;

TABLE I

SEARCHING THE OPTIMAL t_k^*, AND N^* FOR THE LINEAR PM MODEL.

adjacent replacements is given by

$$C_L(t_0, t_1, ..., t_{N-1}, N) = \frac{c_m \sum_{k=0}^{N-1} \int_0^{t_k} (A_k h_0(x) + B_k) dx + c_p(N-1) + c_r}{\sum_{k=0}^{N-1} t_k}, \tag{16}$$

where, c_m is the cost per CM, c_p is the cost per PM, c_r is replacement cost, and $N-1$ is the number of PM between two adjacent replacements.

The optimal t_k^* should satisfy the following conditions. $\frac{\partial C_L(t_0, t_1, ..., t_{N-1}, N)}{\partial t_k}|_{t_k = t_k^*} = 0$, for $k = 0, 1, 2, ..., N-1$. This implies that the optimum $t_0^*, t_1^*, ..., t_{N-1}^*$ should satisfy

$$A_k h_0(t_k^*) + B_k = \frac{C_L(t_0, t_1, ..., t_{N-1}, N)}{c_m}, \tag{17}$$

for $k = 0, 1, 2, ..., N-1$. The optimal N^* should satisfy the following conditions.

$$C_L(t_0^*, t_1^*, ..., t_{N^*-2}^*, (N^*-1)) \geq C_L(t_0^*, t_1^*, ..., t_{N^*-1}^*, N^*), \tag{18}$$

and

$$C_L(t_0^*, t_1^*, ..., t_{N^*}^*, (N^*+1)) \geq C_L(t_0^*, t_1^*, ..., t_{N^*-1}^*, N^*). \tag{19}$$

Table I presents a method to search the optimal values of t_k^*, and N^* for the linear PM model case.

B. *Nonlinear PM model case*

Assume all PM are nonlinear, and the system is aged from Ψ_k just before the kth PM to $\Phi_k t_k + \Psi_k$ just before the next PM. Then the total cost is given by

$$C_N(t_0, t_1, ..., t_{N-1}, N) = \frac{c_m \sum_{k=0}^{N-1} \int_0^{t_k} h_0(\Phi_k x + \Psi_k)dx + c_p(N-1) + c_r}{\sum_{k=0}^{N-1} t_k}. \quad (20)$$

The optimal t_k^* should satisfy the conditions $\frac{\partial C_N(t_0,t_1,...,t_{N-1},N)}{\partial t_k}|_{t_k=t_k^*} = 0$, for $k = 0, 1, 2, ..., N-1$. This implies that the optimum $t_0^*, t_1^*, ..., t_{N-1}^*$ should satisfy $h_0(t_0^*) = h_0(\Phi_1 t_1^* + \Psi_1) = h_0(\Phi_2 t_2^* + \Psi_2) = \ldots = h_0(\Phi_{N-1} t_{N-1}^* + \Psi_{N-1})$, and $c_m h_0(\Phi_k t_k^* + \Psi_k) = C_N(t_0^*, t_1^*, t_2^*, ..., t_{N-1}^*, N)$.

Similarly, the optimum value N^* should satisfy the conditions $C_N(t_1^*, t_2^*, ..., t_{N^*-2}^*, (N^*-1)) \geq C_N(t_1^*, t_2^*, ..., t_{N^*-1}^*, N^*)$, and $C_N(t_1^*, t_2^*, ..., t_{N^*}^*, (N^*+1)) \geq C_N(t_1^*, t_2^*, ..., t_{N^*-1}^*, N^*)$.

V. Discussion

Comparing to the existing PM models reviewed in Section II, we can see that the linear PM model relaxes the assumption of the NAK2 model, while the nonlinear PM model relaxes the assumption of models MAL, NAK1, and CAN. Here, by relaxation, we mean that the proposed models either relax the assumption of parameters $0 < \nu_{10} < \nu_{11} < ... < \nu_{1k-1} < 1$ in the NAK1 model, and $1 \leq \nu_{20} \leq \nu_{21} \leq ... \leq \nu_{2k-1}$ in the NAK2 model; or add one more parameter, b_k in the linear PM model, and β_k in the nonlinear PM model. The following hybrid model can be seen as extensions of all of the existing PM models.

If we combine the linear and nonlinear PM models, a hybrid PM model can be derived. For all PM models, parameter estimation is of interest in both practical, and theoretical perspectives. This section addresses both the hybrid PM model, and its parameter estimation.

A. Hybrid PM

Combining both the linear and nonlinear PM models, we can extend them to a hybrid PM model as follows.

Definition 11: The kth PM is called *hybrid PM* if the hazard functions before, and after the kth maintenance have the relationship

$$h_k(t) = a h_{k-1}(\alpha t + \beta) + b, \quad (21)$$

for $k = 1, 2, \ldots$.

B. More generic extensions

All PM models discussed in Section II assume different parameters after each PM. Similarly, we have the following extensions for the linear, nonlinear, and hybrid PM models.

1) Extended linear PM model:

Definition 12: The kth PM is called an *extended linear PM* if

$$h_k(t) = a_k h_{k-1}(t) + b_k, \quad (22)$$

where a_k, and b_k are parameters; $t \in (0, t_k)$, and $a_k, b_k > 0$. A typical, reasonable choice for the age reduction parameter b_k is to assume that it depends on t_{k-1}. For example, set $b_k = \rho_k h_{k-1}(t_{k-1})$ with $\rho_k \in (0, 1)$ for better maintenance.

If all of the first k PM are extended linear PM, then (22) can also be written as

$$h_k(t) = A'_k h_0(t) + B'_k, \tag{23}$$

where

$$A'_0 = 1, B'_0 = 0, B'_1 = b_1, A'_k = \prod_{i=1}^{k} a_i, \text{ and } B'_k = \sum_{i=1}^{k-1} \left(b_i \prod_{j=i+1}^{k} a_j \right) + b_k.$$

It is easy to derive the following Lemma.

Lemma 3: If all of the first k PM are extended linear PM, for (23), we have

$$F_k(t) = 1 - e^{-B'_k t}(1 - F(t))^{A'_k}. \tag{24}$$

If $1 - F(t)$ belongs to IFR, DFR, IFRA, DFRA, NBU, or NWU, then $1 - F_k(t)$ is in the same category for $k = 2, 3, \ldots$.

2) Extended nonlinear PM model:

Definition 13: Assume $h_0(t)$ is a nonlinear function with respect to t. The kth PM is called an *extended nonlinear PM* if

$$h_k(t) = h_{k-1}(\alpha_k t + \beta_k), \tag{25}$$

where $t \in (0, t_k)$, α_k is an age defying parameter, and β_k is an age reduction parameter.

Similar to the linear PM model, the expressions of parameters, α_k, and β_k, in the extended nonlinear PM model are important. If we recall the existing PM models, parameters μ_k in the MAL model, ν_{1k} in the NAK1 model, λ_{2k} in the LIN model, and $kT - k\tau$ in the CAN model, are related to the time intervals t_k, and playing a similar role as the parameter β_k in the extended nonlinear PM model. Therefore, β_k can be set to $\gamma_k t_{k-1}$, which will be used in the PM policy optimization section, where $\gamma_k \in (0, 1)$.

If all of the first k PM are extended nonlinear PM, another expression of e (25) is

$$h_k(t) = h_0(\Phi'_k t + \Psi'_k), \tag{26}$$

where

$$\Phi'_0 = 1, \Psi'_0 = 0, \Psi'_1 = \beta_1, \Phi'_k = \prod_{i=1}^{k} \alpha_i, \text{ and } \Psi'_k = \sum_{i=1}^{k-1} (\beta_i \prod_{j=i+1}^{k} \alpha_j) + \beta_k.$$

The proof for the following Lemma is simple, so we do not show the proof in this paper.

Lemma 4: If all of the first k PM are extended nonlinear PM, for (25), we have

$$F_k(t) = 1 - (1 - F(\Phi'_k t + \Psi'_k))^{\frac{1}{\Phi'_k}} (1 - F(\Psi'_k))^{-\frac{1}{\Phi'_k}}. \tag{27}$$

If $1 - F(t)$ is IFR (or DFR), then $1 - F_k(t)$ is IFR (or DFR) for $k = 1, 2, 3, \ldots$.

3) Extended hybrid PM model:

Definition 14: The kth PM is called an *extended hybrid* PM if the hazard functions before, and after the kth maintenance have the relationship

$$h_k(t) = a_k h_{k-1}(\alpha_k t + \beta_k) + b_k. \tag{28}$$

All PM models in Section II are special cases of the extended hybrid model.

If all of the first k PM are extended hybrid linear PM, then another expression of model (28) is given as

$$h_k(t) = A_k h_0(\Phi_k t + \Psi_k) + B_k, \tag{29}$$

and we have the following lemma.

Lemma 5: For Model (29), we have

$$F_k(t) = 1 - e^{-B_k t}(1 - F_0(\Phi_k t + \Psi_k))^{\frac{A_k}{\Phi_k}}(1 - F_0(\Psi_k))^{-\frac{A_k}{\Phi_k}}. \tag{30}$$

If $h_k(t)$ is IFR (DFR), then $h_{k+1}(t)$ is IFR (DFR).

C. More complex situations

Obviously, the introduced PM models do not consider more complex situations that can exhibit more complex non-linear relationship between $h_k(t)$ and $h_{k-1}(t)$. For example, $h_k(t)$ can be a $G(h_{k-1}(t))$, or $h_k(t) = h_{k-1}(g(t))$ where $G(.)$ and $g(.)$ are nonlinear functions. In practice, however, estimating parameters for a nonlinear relationship can be problematic as there might be limited data available.

D. Parameter estimation

In practice, there are two approaches to estimating the parameters in PM models. The first approach estimates the parameters on the basis of reliability data sets. For example, one can utilize maximum likelihood estimation to estimate the parameters of the linear, and non-linear PM models. The second one uses domain experience to estimate the parameters. This approach is used only if few failure data are collected (see [31]). A combination of these two approaches can also be used. For example, [7] considers the scenarios where the maintenance effect is a random variable. It assumes that both parameter τ in the CAN model, and parameter ν_{1i} in the NAK1 model are random variables with certain probability distributions. Under such assumptions, they show that more cost-effective PM policies can be obtained.

Note that a PM model with many parameters might not be applicable in practice. This is due to a lack of sufficient data for parameter estimation. However, (10) of the linear PM model, (13) of the nonlinear PM model, and (21) of the hybrid PM model include fewer parameters which should be easier to estimate, and more realistic for applications in practice.

VI. CONCLUSIONS

PM models are important in both designing maintenance policies, and assessing the residual lifetime of systems being preventively maintained. Many PM models are proposed in the reliability literature. As discussed in Section III, however, existing PM models present weaknesses in the sense that either their parameters might not have physical meanings, or their model assumptions are too restrictive. The linear, nonlinear, and hybrid PM models proposed in this paper overcome such weaknesses.

The main contributions of this paper are as follows.

- We have reviewed the existing PM models, investigated their interrelationships, and proposed a new classification of the PM models.
- Three PM models are introduced, and their relationships are investigated.
- The properties of the PM models are derived.
- The expected costs for the three PM models for sequential PM are formulated, and the necessary conditions of obtaining the optimal PM policies for both the general, and special cases are derived.

Our future research will include

- estimating the parameters within the three models, and comparing the three models with those reviewed in Section II with respect to their performance on the basis of field test data; and
- investigating the application of the proposed PM models to various scenarios, including optimizing warranty policies for products with linear or nonlinear preventive maintenance.

ACKNOWLEDGMENT

The authors would like to thank the referees for their valuable comments and suggestions that appreciably improved the quality of this paper.

This research was supported by Engineering and Physical Sciences Research Council (EPSRC) of the United Kingdom (EPSRC Grant reference: EP/G039674/1), and the Natural Sciences and Engineering Research Council of Canada (NSERC)

REFERENCES

[1] H. Pham and H. Wang, "Imperfect maintenance," *European Journal of Operational Research*, vol. 94, no. 3, pp. 425–438, 1996.

[2] L. Doyen and O. Gaudoin, "Classes of imperfect repair models based on reduction of failure intensity or virtual age," *Reliability Engineering and System Safety*, vol. 84, no. 1, pp. 45–56, 2004.

[3] M. A. K. Malik, "Reliable preventive maintenance scheduling." *AIIE Transactions (American Institute of Industrial Engineers)*, vol. 11, no. 3, pp. 221–228, 1979.

[4] D. Lin, M. J. Zuo, and R. Yam, "General sequential imperfect preventive maintenance models," *International Journal of Reliability, Quality and Safety Engineering*, vol. 7, no. 3, pp. 253–266, 2000.

[5] T. Nakagawa, "Sequential imperfect preventive maintenance policies." *IEEE Transactions on Reliability*, vol. 37, no. 3, pp. 295–298, 1988.

[6] R. V. Canfield, "Cost optimization of periodic preventive maintenance." *IEEE Transactions on Reliability*, vol. R-35, no. 1, pp. 78–81, 1986.

[7] S. Wu and D. Clements-Croome, "Optimal maintenance policies under different operational schedules," *IEEE Transactions on Reliability*, vol. 54, no. 2, pp. 338–346, 2005.

[8] J.-K. Chan and L. Shaw, "Modeling repairable systems with failure rates that depend on age & maintenance," *IEEE Transactions on Reliability*, vol. 42, no. 4, pp. 566–571, 1993.

[9] F. Zhang and A. K. S. Jardine, "Optimal maintenance models with minimal repair, periodic overhaul and complete renewal," *IIE Transactions (Institute of Industrial Engineers)*, vol. 30, no. 12, pp. 1109–1119, 1998.

[10] S. Wu and D. Clements-Croome, "Preventive maintenance models with random maintenance quality," *Reliability Engineering and System Safety*, vol. 90, no. 1, pp. 99–105, 2005.

[11] M. Kijima, H. Morimura, and Y. Suzuki, "Periodical replacement problem without assuming minimal repair," *European Journal of Operational Research*, vol. 37, no. 2, pp. 194–203, 1988.

[12] M. Kijima, "Some results for repairable systems with general repair," *Journal of Applied Probability*, vol. 26, no. 1, pp. 89–102, 1989.

[13] J. H. Seo and D. S. Bai, "An optimal maintenance policy for a system under periodic overhaul," *Mathematical and Computer Modelling*, vol. 39, no. 4-5, pp. 373–380, 2004.

[14] J. S. Dagpunar, "Renewal-type equations for a general repair process," *Quality and Reliability Engineering International*, vol. 13, no. 4, pp. 235–245, 1997.

[15] C. Dorado, M. Hollander, and J. Sethuraman, "Nonparametric estimation for a general repair model," *Annals of Statistics*, vol. 25, no. 3, pp. 1140–1160, 1997.

[16] B. H. Lindqvist, "On the statistical modeling and analysis of repairable systems," *Statistical Science*, vol. 21, no. 4, pp. 532–551, 2006.

[17] L. A. Baxter, M. Kijima, and M. Tortorella, "A point process model for the reliability of a maintained system subject to general repair," *Communications in Statistics. Part C: Stochastic Models*, vol. 12, no. 1, pp. 37–65, 1996.

[18] W. Kahle, "Optimal maintenance policies in incomplete repair models," *Reliability Engineering and System Safety*, vol. 92, no. 5, pp. 563–565, 2007.

[19] M. Bartholomew-Biggs, M. J. Zuo, and X. Li, "Modelling and optimizing sequential imperfect preventive maintenance," *Reliability Engineering and System Safety*, vol. 94, pp. 53–62, 2009.

[20] S. H. Sheu and C. C. Chang, "An extended periodic imperfect preventive maintenance model with age-dependent failure type," *IEEE Transactions on Reliability*, vol. 58, no. 2, pp. 397–405, 2009.

[21] L. Cui, W. Kuo, H. T. Loh, and M. Xie, "Optimal allocation of minimal and perfect repairs under resource constraints," *IEEE Transactions on Reliability*, vol. 53, no. 2, pp. 193–199, 2004.

[22] Y. Lam, "Geometric processes and replacement problem," *Acta Mathematicae Applicatae Sinica*, vol. 4, no. 4, pp. 366–377, 1988.

[23] H. Wang and H. Pham, "A quasi renewal process and its applications in imperfect maintenance," *International Journal of Systems Science*, vol. 27, no. 10, pp. 1055–1062, 1996.

[24] M. S. Finkelstein, "A scale model of general repair," *Microelectronics Reliability*, vol. 33, no. 1, pp. 41–44, 1993.

[25] S. Wu and D. Clements-Croome, "A novel repair model for imperfect maintenance," *IMA Journal Management Mathematics*, vol. 17, no. 3, pp. 235–243, 2006.

[26] H. Wang and H. Pham, *Reliability and optimal maintenance*. Springer, 2006.

[27] Y. L. Zhang and S. Wu, "Reliability analysis for a k / n (f) system with repairable repair-equipment," *Applied Mathematical Modelling*, vol. 33, no. 7, pp. 3052–3067, 2009.

[28] J. Jia and S. Wu, "A replacement policy for a repairable system with its repairman having multiple vacations," *Computers and Industrial Engineering*, 2009, article in Press.

[29] ——, "Optimizing replacement policy for a cold-standby system with waiting repair times," *Applied Mathematics and Computation*, 2009, article in Press.

[30] C. Y. Cheng, M. Chen, and R. Guo, "The optimal periodic preventive maintenance policy with degradation rate reduction under reliability limit," in *Proceedings of IEEE International Conference on Industrial Engineering and Engineering Management (IEEM2007)*, 2007, pp. 649–653.

[31] A. O'Hagan, C. Buck, A. Daneshkhah, J. Eiser, P. Garthwaite, D. Jenkinson, J. Oakley, and T. Rakow, *Uncertain Judgements: Eliciting Experts' Probabilities*. Wiley, 2006.

APPENDIX

The proof of Theorem 1 is as follows.

Proof If $h_0(t)$ is increasing (or decreasing), then (11) $h_k(t)$ is increasing (or decreasing).

$1-F(t)$ is IFRA (or DFRA) if $[1-F(t)]^{1/t}$ is decreasing (or increasing).

Because

$$(1-F_k(t))^{1/t} = (e^{-B_k t}(1-F(t))^{A_k})^{1/t} = e^{-B_k}\left((1-F(t))^{1/t}\right)^{A_k} \tag{31}$$

from (11), assuming that $(1-F(t))^{1/t}$ is increasing (decreasing) with respect to t, then $(1-F_k(t))^{1/t}$ is increasing (decreasing) in t.

$1-F(t)$ is NBU (or NWU) if $1-F(t_1+t_2) \leq (1-F(t_1))(1-F(t_2))$ (or $1-F(t_1+t_2) \geq (1-F(t_1))(1-F(t_2))$).

Assume that $(1-F(t_1))(1-F(t_2)) \geq 1-F(t_1+t_2)$. According to (1), then it follows that

$$\begin{aligned} 1-F_k(t_1+t_2) &= e^{-B_k(t_1+t_2)}(1-F(t_1+t_2))^{A_k} \\ &\geq e^{-B_k(t_1+t_2)}(1-F(t_1))^{A_k}(1-F(t_2))^{A_k} \\ &= (1-F_k(t_1))(1-F_k(t_2)) \end{aligned} \tag{32}$$

A similar proof exists for the case $1-F(t_1+t_2) \geq (1-F(t_1))(1-F(t_2))$. This proves the theorem.