

## Research Article

# Generalized Accelerated Failure Time Frailty Model for Systems Subject to Imperfect Preventive Maintenance

Huilin Yin,<sup>1</sup> Xiaohan Yang,<sup>2</sup> and Rui Peng<sup>3</sup>

<sup>1</sup> Department of Electrical & Information Engineering, CDHK, Tongji University, Shanghai 200092, China

<sup>2</sup> Department of Mathematics, Tongji University, Siping Road 1239, Shanghai 200092, China

<sup>3</sup> School of Electrical & Information Engineering, Tongji University, Shanghai 200092, China

Correspondence should be addressed to Xiaohan Yang; xiaohyang@tongji.edu.cn

Received 14 August 2014; Revised 13 October 2014; Accepted 13 October 2014

Academic Editor: Gang Li

Copyright © 2015 Huilin Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Imperfect preventive maintenance (PM) activities are very common in industrial systems. For condition-based maintenance (CBM), it is necessary to model the failure likelihood of systems subject to imperfect PM activities. In this paper, the models in the field of survival analysis are introduced into CBM. Namely, the generalized accelerated failure time (AFT) frailty model is investigated to model the failure likelihood of industrial systems. Further, on the basis of the traditional maximum likelihood (ML) estimation and expectation maximization (EM) algorithm, the hybrid ML-EM algorithm is investigated for the estimation of parameters. The hybrid iterative estimation procedure is analyzed in detail. In the evaluation experiment, the generated data of a typical degradation model are verified to be appropriate for the real industrial processes with imperfect PM activities. The estimates of the model parameters are calculated using the training data. Then, the performance of the model is analyzed through the prediction of remaining useful life (RUL) using the testing data. Finally, comparison between the results of the proposed model and the existing model verifies the effectiveness of the generalized AFT frailty model.

## 1. Introduction

Condition-based maintenance (CBM) has continuously been an important issue in the area of maintenance strategy. As degradation processes before failure of many systems can be measured, CBM is more effective as corrective maintenance and time-based preventive maintenance in some aspects, such as catastrophic failure reduction and availability maximization [1, 2]. CBM program includes three steps: data acquisition, signal processing, and maintenance decision support [1]. Maintenance decision support is categorized into diagnostics and prognostics. Prognostics are often characterized by estimating the remaining useful life (RUL) of systems using available condition monitoring information. The RUL estimation ensures enough time to perform the necessary maintenance actions prior to failure [3, 4].

The proportional hazards model (PHM) [5, 6], which is a popular model in survival analysis, can be applied in reliability evaluation and maintenance optimization. It

has shown its effectiveness for RUL prediction and CBM scheduling of industrial systems [7–9]. PHM is an effective CBM method due to its strength in handling the influence of variable operational conditions. The accelerated failure time (AFT) model [10, 11] is an important alternative to the PHM in survival analysis and has the advantage of being more intuitively interpretable than the PHM. However, the AFT method has been rarely applied in reliability-related fields. It is shown by reviewing and comparing the major mathematical models of survival analysis and reliability theory that both fields address the same mathematical problems [6]. Because of the unified mathematical models of both fields, the methods for survival analysis might be used for reliability analysis and related fields such as CBM. In this paper, we introduce the AFT model in reliability fields and investigate the AFT-based model for CBM.

Imperfect preventive maintenance (PM) restores a system to a better state but not “as good as new.” Imperfect PM activities are considerably common in industrial systems and

there are many studies on the model of imperfect PM [12–17]. The extended PHM (EPHM) [16, 17] has been proved to be a superiorly effective method to predict RUL of systems subject to imperfect PM. Inspired by the EPHM, this paper investigates a new alternative model, the generalized AFT frailty model.

The rest of the paper is organized as follows. Section 2 introduces the generalized AFT frailty model and compares it with the existing model. Section 3 proposes the hybrid maximum likelihood- (ML-) expectation maximization (EM) algorithm for the calculation of the parameters in the model. Section 4 proves the effectiveness of the proposed model by the simulation experiment of RUL prediction for systems subject to imperfect PM. Finally, Section 5 concludes the paper.

## 2. Generalized AFT Frailty Model

*2.1. Models.* First, we introduce the Cox PHM, the AFT model, the generalized AFT model, and the PHM frailty model in sequence. Then, we propose the generalized AFT frailty model.

The Cox PHM [5] is a popular model in survival analysis, and it is an effective method for RUL prediction and CBM scheduling of industrial systems. Let  $\lambda(t | \mathbf{X})$  be the hazard/failure function at time  $t$  given the covariates  $\mathbf{X}$ ; the PHM is expressed as

$$\lambda(t | \mathbf{X}) = \lambda_0(t) \exp(\boldsymbol{\beta}\mathbf{X}), \quad (1)$$

where  $\lambda_0(t)$  is the baseline hazard/failure function, which considers the age of the system at the time of inspection.  $\mathbf{X}$  is the vector of the covariates and  $\boldsymbol{\beta}$  is the regression coefficient vector. In a CBM program, the systems' degradation indicators are usually chosen to be the covariates.

The AFT model [10, 11, 18] is an important alternative to the PHM. Suppose  $T_{ij}$  is the failure time of system  $j$  in cluster (or group)  $i$ . There is a censoring random variable, which we denote as  $C_{ij}$ . We only observe  $T_{ij}^* = \min(T_{ij}, C_{ij})$  and the linear regression (i.e., AFT) model is

$$\log(T_{ij}) = \boldsymbol{\beta}\mathbf{X}_{ij} + \zeta_{ij}, \quad (2)$$

where  $\mathbf{X}_{ij}$  is the vector of observed covariates,  $\boldsymbol{\beta}$  is the regression coefficient vector, and  $\zeta_{ij}$  is the random error. This linear form of the AFT model deals with the regression relationship of the covariates and the failure times logarithms, which has the similar form to the generic linear regression model. In the PHM, the baseline failure function and the covariates in PHM are independent, which limits the modeling of some types of the survival data in medical research. Collett [19] has introduced the influence of the covariates in the baseline failure function and proposed the following generalized AFT model:

$$\lambda(t | \mathbf{X}_{ij}) = \exp(-\boldsymbol{\beta}\mathbf{X}_{ij}) \lambda_0(t \exp(-\boldsymbol{\beta}\mathbf{X}_{ij})). \quad (3)$$

In this model, the whole part of the covariates  $\exp(-\boldsymbol{\beta}\mathbf{X}_{ij})$  has been introduced into the baseline failure function.

The effectiveness of this model has been verified by the modeling of the survival data in medical field [19].

PHM and AFT are suitable for the mutually independent failure time data. In reality, correlated or clustered failure time data are very common in the fields of survival analysis and reliability. For example, systems operate in the same environment with the same temperature and humidity. The shared environment of the subjects leads to the dependence among the observed failure time. Frailty is a good tool to represent the random effect shared by subjects in the same cluster (or group) and it induces dependence among the correlated or clustered failure time data. The PHM frailty model [20] is

$$\lambda(t | \mathbf{X}_{ij}, \omega_i) = \omega_i \lambda_0(t) \exp(\boldsymbol{\beta}\mathbf{X}_{ij}), \quad (4)$$

where  $\omega_i$  is the frailty term and models the random effect that is shared by the systems in the  $i$ th cluster (group). Frailty has been also introduced in the AFT model to represent the possible correlation among failure times [12]. Considering both the generalized AFT model (3) and the PHM frailty model (4), the following generalized AFT frailty model [21] is proposed:

$$\lambda(t | \mathbf{X}_{ij}, \omega_i) = \omega_i \exp(-\boldsymbol{\beta}\mathbf{X}_{ij}) \lambda_0(t \exp(-\boldsymbol{\beta}\mathbf{X}_{ij})). \quad (5)$$

*2.2. Comparison between the Generalized AFT Frailty Model and the Existing Model.* The EPHM has been proposed by You et al. [16, 17] to model the failure likelihood for systems subject to imperfect PM activities. The EPHM has the following form:

$$\begin{aligned} h_n \left( t' + \sum_{k=1}^n P_k, \mathbf{X} \left( t' + \sum_{k=1}^n P_k \right) \right) \\ = A_n h_0(t' + b_n y_n) \exp \left( \boldsymbol{\beta}\mathbf{X} \left( t' + \sum_{k=1}^n P_k \right) \right), \end{aligned} \quad (6)$$

where  $h_n$  is the hazard/failure function of the system after the  $n$ th PM activity and before the  $(n+1)$ th PM activity,  $P_k$  is the  $k$ th PM interval, and  $t'$  is the random local time between the  $k$ th and the  $(k+1)$ th PM activity.  $A_n = a_0 \cdot a_1 \cdots a_n$ , where  $a_k$  is the hazard rate increase factor (HRIF) due to the  $k$ th PM activity, and  $b_n$  is the age reduction factor (ARF) due to the  $n$ th PM activity,  $y_n = P_n + b_{n-1}P_{n-1} + b_{n-1}b_{n-2}P_{n-2} + \cdots + (\prod_{j=1}^{n-1} b_j)P_1$ . The HRIF modifies the increase rate of the hazard/failure rate of the system after an imperfect PM activity, and the ARF measures the extent to which the PM activity brings the system to a younger age.

Comparing (5) and (6), we observe that the HRIF  $A_n$  and the frailty term  $\omega_i$  act multiplicatively on the hazard/failure rate function and play the same role in the models. When the system starts operation after imperfect PM activities, the system has already aged somewhat. The ARF measures the extent to which the imperfect PM affects the effective age. The generalized AFT frailty model describes the hybrid influence of imperfect PM on both the hazard/failure rate and the effective age, which is shown in Figure 1.

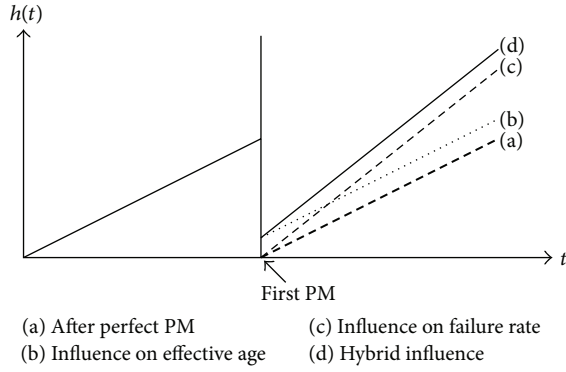


FIGURE 1: Hybrid influence of imperfect PM on both failure rate and effective age.

To some extent, the generalized AFT frailty model is unified with the EPHM. So we introduce the generalized AFT frailty model (5) into the CBM field to model the failure likelihood of the systems subject to imperfect PM.

### 3. Parameter Estimation Algorithm

After the introduction of the generalized AFT frailty model, we investigate its parameter estimation algorithm. The EM algorithm has been a popular method for computing ML estimates from incomplete data. Due to the unknown frailty terms, the EM algorithm has been used to estimate parameters in the PHM frailty model [22] and AFT frailty model [12]. Vu and Kuniman [23] have proposed a hybrid ML-EM method for the calculation of ML estimates in PHM frailty models and have verified that the hybrid method is more computationally efficient than the standard EM method and faster than the standard direct ML method. In this paper, we estimate the parameters in the generalized AFT frailty model with a hybrid ML-EM algorithm.

If the failure time is smaller than the censoring random variable, the indicator function has a value of one; otherwise, it has a value of zero:

$$I_{ij} = \begin{cases} 1 & \text{if } T_{ij} \leq C_{ij} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$(i = 1, 2, \dots, n; j = 1, 2, \dots, n_i).$$

The likelihood function of right censoring data is

$$L = \prod_{i=1}^n \prod_{j=1}^{n_i} \lambda(t_{ij})^{I_{ij}} R(t_{ij}), \quad (8)$$

where  $R(\cdot)$  is the survival/reliability function.

The frailty terms are random variables, and we suppose  $\omega_i$  obeys the distribution  $g(\omega_i, \theta)$ , with the parameter  $\theta$ , which can be a vector. Thus there are two unknown vectors  $\theta$  and  $\beta$  in model (5).

The parameter  $\theta$  is computed with the ML estimation method. The ML estimation is based on the marginal likelihood function, which is written as

$$\begin{aligned} L_{\text{marginal}} &= \prod_{i=1}^n \int_0^{\infty} \prod_{j=1}^{n_i} \lambda_{ij}(t_{ij})^{I_{ij}} R_{ij}(t_{ij}) g(\omega_i, \theta) d\omega_i \\ &= \prod_{i=1}^n \int_0^{\infty} \prod_{j=1}^{n_i} \omega_i^{I_{ij}} \exp(-\beta I_{ij} \mathbf{X}_{ij}) \lambda_0^{I_{ij}}(t_{ij} e^{-\beta \mathbf{X}_{ij}}) \\ &\quad \times \exp(-\omega_i \Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})) \\ &\quad \times g(\omega_i, \theta) d\omega_i, \end{aligned} \quad (9)$$

where  $\Lambda_0$  is the cumulative failure function. For the sake of convenient calculation, we take the logarithm of the above equation so that the multiplications can be converted into additions. The logarithmic form is

$$\begin{aligned} L_M &= \sum_{i=1}^n \sum_{j=1}^{n_i} \left[ -\beta I_{ij} \mathbf{X}_{ij} + I_{ij} \log(\lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})) \right. \\ &\quad \left. + \log\left(\int_0^{\infty} \omega_i^{D_i} \exp(-\omega_i \Lambda_i) g(\omega_i, \theta) d\omega_i\right) \right], \end{aligned} \quad (10)$$

where  $D_i = \sum_{j=1}^{n_i} I_{ij}$ ,  $\Lambda_i = \sum_{j=1}^{n_i} \Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})$ .

Given the initial values of  $\beta$  and  $\Lambda_0$ , we can calculate the ML estimate of  $\theta$ . Given the estimate of  $\theta$ , we can estimate  $\beta$  by an EM algorithm. Namely, the estimations of  $\theta$  and  $\beta$  interact with each other as both preconditions and results. So we update both parameters with an iterative process until they converge.

The parameter  $\beta$  is estimated with the EM algorithm. We start the EM algorithm with the full likelihood function, which is written as

$$\begin{aligned} L_{\text{full}}(\theta, \beta, \lambda_0, \omega) &= \prod_{i=1}^n \prod_{j=1}^{n_i} \lambda_{ij}(t_{ij})^{I_{ij}} R_{ij}(t_{ij}) g(\omega_i, \theta) \\ &= \prod_{i=1}^n \prod_{j=1}^{n_i} \omega_i^{I_{ij}} \exp(-\beta I_{ij} \mathbf{X}_{ij}) \lambda_0^{I_{ij}}(t_{ij} e^{-\beta \mathbf{X}_{ij}})^{I_{ij}} \\ &\quad \times \exp(-\omega_i \Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})) \\ &\quad \times g(\omega_i, \theta). \end{aligned} \quad (11)$$

The logarithmic form of  $L_{\text{full}}$  is denoted as  $L_F$  and the corresponding full log-likelihood is

$$L_F(\theta, \beta, \lambda_0, \omega) = L_1(\theta, \omega) + L_2(\beta, \lambda_0, \omega), \quad (12)$$

where

$$L_1(\boldsymbol{\theta}, \boldsymbol{\omega}) = \sum_{i=1}^n \left[ \log(g(\omega_i, \boldsymbol{\theta})) + \sum_{j=1}^{n_i} I_{ij} \log \omega_i \right], \quad (13)$$

$$L_2(\boldsymbol{\beta}, \lambda_0, \boldsymbol{\omega}) = \sum_{i=1}^n \sum_{j=1}^{n_i} \left[ -\beta I_{ij} \mathbf{X}_{ij} + I_{ij} \log(\lambda_0 (t_{ij} e^{-\beta \mathbf{X}_{ij}})) - \omega_i \Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}}) \right]. \quad (14)$$

The  $E$ -step calculates the expectation of the full likelihood function with respect to  $\boldsymbol{\omega}$ , in order to eliminate its effect on the likelihood function. The corresponding equations for the expectation calculation are

$$E(L_1) = \sum_{i=1}^n \left[ E(\log(g(\omega_i, \boldsymbol{\theta}))) + \sum_{j=1}^{n_i} I_{ij} E(\log \omega_i) \right], \quad (15)$$

$$E(L_2) = \sum_{i=1}^n \sum_{j=1}^{n_i} \left[ -\beta I_{ij} \mathbf{X}_{ij} + I_{ij} \log(\lambda_0 (t_{ij} e^{-\beta \mathbf{X}_{ij}})) - E(\omega_i) \Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}}) \right]. \quad (16)$$

From (15) and (16), we observe that the calculation of  $E(L_1)$  depends only on the parameter  $\boldsymbol{\theta}$  and is independent of  $\boldsymbol{\beta}$ . The calculation of  $E(L_2)$  contains  $\boldsymbol{\beta}$ . So we estimate  $\boldsymbol{\beta}$  only by analyzing (16).

The  $M$ -step maximizes  $E(L_2)$  and calculates the ML estimate of  $\boldsymbol{\beta}$ . Due to both the unknown parts  $\lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})$  and  $\Lambda_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})$  in  $E(L_2)$ , we cannot handle  $E(L_2)$  directly. With kernel smoothing processes [24], we convert the semiparametric form of (16) into the parametric form. The likelihood function is written as

$$E(L_2^s(\boldsymbol{\beta})) = \sum_{i=1}^n \sum_{j=1}^{n_i} \left[ I_{ij} \log \left( \frac{1}{m_i a_n} \sum_{k=1}^n \sum_{l=1}^{n_i} I_{kl} K \left( \frac{r_{kl} - r_{ij}}{a_n} \right) \right) - I_{ij} \log \left( \frac{1}{m_i} \sum_{k=1}^n \sum_{l=1}^{n_i} \int_{-\infty}^{(r_{kl} - r_{ij})/a_n} E(\omega_i) K(s) ds \right) \right], \quad (17)$$

where  $r_{ij} = \log t_{ij} - \beta \mathbf{X}_{ij}$ ,  $K(s)$  is the kernel function, and  $a_n$  is the window width of the kernel function. By maximizing the likelihood function  $E(L_2^s(\boldsymbol{\beta}))$ , we obtain the ML estimate of  $\boldsymbol{\beta}$ .

Based on the estimate  $\hat{\boldsymbol{\beta}}$ , we can calculate the failure density function, the reliability function, and the corresponding cumulative failure function:

$$\hat{f}(t) = (n a_n t)^{-1} \sum_{i=1}^n \sum_{j=1}^{n_i} I_{ij} K \left( \frac{r_{ij}(\hat{\boldsymbol{\beta}}) - \log t}{a_n} \right), \quad (18)$$

$$\hat{R}(t) = 1 - \int_0^t \hat{f}(x) dx, \quad (19)$$

$$\begin{aligned} \hat{\Lambda}_0(t) &= -\log \hat{R}(t) \\ &= \int_{-\infty}^{\log t} \left[ \frac{(n a_n)^{-1} \sum_{i=1}^n \sum_{j=1}^{n_i} I_{ij} K \left( (r_{ij}(\hat{\boldsymbol{\beta}}) - s) / a_n \right)}{n^{-1} \sum_{i=1}^n \sum_{j=1}^{n_i} \int_{-\infty}^{(r_{ij}(\hat{\boldsymbol{\beta}}) - s) / a_n} K(u) du} \right] ds. \end{aligned} \quad (20)$$

With the calculated estimates  $\hat{\boldsymbol{\beta}}$  and  $\hat{\Lambda}_0$ , we calculate the ML estimate  $\boldsymbol{\theta}$  by maximizing  $L_M$  in (10). Then, with the estimated value  $\hat{\boldsymbol{\theta}}$ , we follow the EM process to calculate  $\hat{\boldsymbol{\beta}}$  and  $\hat{\Lambda}_0$  again. Thus, we repeat this iterative process until the parameters convergence.

We summarize the above-analyzed process as the following estimation procedure:

- Step 1: obtain initial estimate  $\hat{\boldsymbol{\beta}}$  from the independent AFT model and estimate the cumulative failure function  $\hat{\Lambda}_0(t_{ij} e^{-\beta \mathbf{X}_{ij}})$ ;
- Step 2 (ML estimation): given the current estimates  $\hat{\boldsymbol{\beta}}$  and  $\hat{\Lambda}_0$ , update the estimate  $\hat{\boldsymbol{\theta}}$  by maximizing the marginal log-likelihood  $L_M$  in (10);
- Step 3: prepare the full log-likelihood function for the EM algorithm;
- Step 4 ( $E$ -step): calculate the expectation of the full log-likelihood function with respect to  $\boldsymbol{\omega}$  in order to eliminate its effect on the likelihood function;
- Step 5 ( $M$ -step): given the current estimate  $\hat{\boldsymbol{\theta}}$  (from Step 2), update the estimate  $\hat{\boldsymbol{\beta}}$  by maximizing the partial log-likelihood function  $E(L_2^s(\boldsymbol{\beta}))$  in (17) and calculate the corresponding cumulative failure function  $\hat{\Lambda}_0$  with (20);
- Step 6: repeat Steps 2, 3, and 4 until both parameters  $\hat{\boldsymbol{\beta}}$  and  $\hat{\boldsymbol{\theta}}$  converge.

This iterative ML-EM procedure is illustrated in Figure 2.

## 4. Evaluation Experiment

We generated the data on the basis of a typical degradation model to simulate the industrial processes of the systems subject to imperfect PM activities. With the simulated data for training, we calculated the parameter estimates of the generalized AFT frailty model. In order to test the effectiveness of the model, we calculated the RUL prediction with the testing

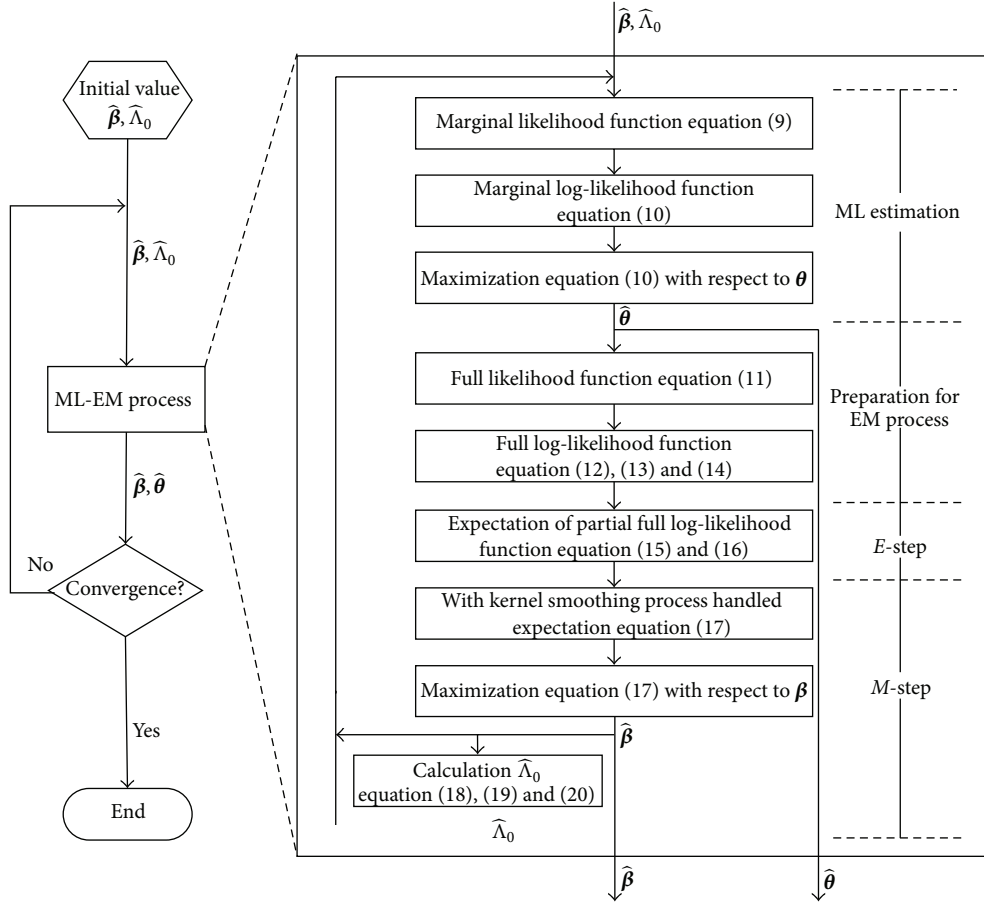


FIGURE 2: Parameters estimation procedure.

data and compare the results with those of the EPHM. By the comparison, it was proved that the generalized AFT frailty model is effective for the systems subject to imperfect PM activities.

**4.1. Process-Data Generation.** The functional form of the typical degradation process is as follows:

$$M(t) = \delta + \eta t + \varepsilon(t), \quad (21)$$

where  $M(t)$  denotes the degradation indicator, and the stochastic parameters are assumed to follow the normal distribution; namely,  $\delta \sim N(\mu_0, \sigma_0^2)$  and  $\eta \sim N(\mu_1, \sigma_1^2)$ . The error term  $\varepsilon(t)$  is included in the model to capture system and environmental noise, signal transients, measurement errors, and variations due to monitoring system.  $\varepsilon(t)$  is assumed to follow a Brownian motion process  $\varepsilon(t) \sim N(0, \sigma^2 t)$ . Equation (21) is used to generate the degradation indicator over time, and a higher magnitude of degradation indicator corresponds to a worse system state.

On the basis of (21), the degradation process with imperfect PM activities was simulated. The simulation mechanism and experiment design are introduced by You and Meng [17]. It was demonstrated that the simulated degradation processes

were close to real data in practice. In our simulation experiment, we also considered the single degradation indicator  $M(t)$  as a covariate for consistency with the simulation work of You and Meng [17]. It was assumed that the PM duration was ignored as zero, and the  $k$ th PM activities occur at time  $T_k$ . After PM, the degradation process was simulated as a new process plus a residual degradation due to the PM activity with a random imperfect effect. For example, after the first PM, sample  $j$  has the following degradation process:

$$M_j(t) = \rho_{j1} M_j(T_1) + (\theta_{j1} + \eta_{j1}(t - T_1) + \varepsilon_{j1}(t - T_1)). \quad (22)$$

After the  $k$ th PM the corresponding degradation signal is

$$\begin{aligned} M_j(t) &= \rho_{jk} M_j(T_1 + \dots + T_k) + M_j(t - T_1 - \dots - T_k) \\ &= \rho_{jk} M_j(T_1 + \dots + T_k) \\ &\quad + (\theta_{jk} + \eta_{jk}(t - T_1 - \dots - T_k) \\ &\quad + \varepsilon_{jk}(t - T_1 - \dots - T_k)), \end{aligned} \quad (23)$$

TABLE 1: Updated degradation model parameters.

Number of PM activities	$\mu_{0k}$	$\mu_{1k}$	$\sigma_{0k}$	$\sigma_{1k}$	$\sigma_k$
$k = 0$	1.00	4.00	1.00	1.00	0.15
$k = 1$	1.53	4.77	0.99	0.98	0.15
$k = 2$	1.99	5.48	1.04	1.02	0.16
$k = 3$	2.45	5.89	1.56	1.54	0.16

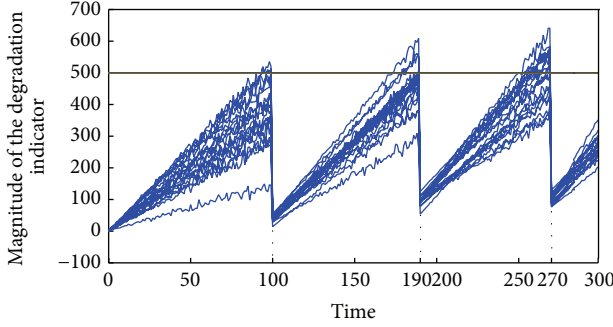


FIGURE 3: Degradation processes with a maximum of three imperfect PM activities.

where  $\rho$  is the coefficient quantifying the amount of residual degradation, which is defined as

$$\rho_{jk} \sim N(1 - \exp(-0.5k), 0.005); \quad (24)$$

$\delta_{jk}$ ,  $\eta_{jk}$ , and  $\varepsilon_{jk}$  are the model parameters of sample  $j$  after it receives its  $k$ th PM activity. The typical method for the estimate of the degradation model parameters after each PM performs a Bayesian update, which combines the prior distribution parameters and the real-time information to obtain the posterior distribution parameters [25].

We simulate the degradation processes of 400 systems, which are subject to a maximum of three imperfect PM activities. Some systems fail before they receive any PM. Some systems fail after the first PM, second PM, or third PM. Some systems operate in a normal state up to the end of the simulation time. The updated degradation model parameters of one sample are shown in Table 1.

Twenty degradation processes are illustrated in Figure 3. The horizontal axis represents time. The vertical axis represents the magnitude of the degradation indicator.

The three moments of PM activities are  $T_1 = 100$ ,  $T_2 = 190$ , and  $T_3 = 270$ , respectively. Each section represents the degradation process before the corresponding PM. When the degradation indicator exceeds the threshold (defined as  $D = 500$ ), the process is considered to have failed. Among the twenty illustrated degradation processes, four processes fail before receiving any PM activity, three processes fail after receiving the first PM activity, and eleven processes fail after receiving the second PM activity.

Next, we demonstrate the effectiveness of the simulated data for imperfect PM. We take the local lifetime as the time interval between the PM moment and the time of failure in each section. In Figure 4,  $t_0$ ,  $t_1$ , and  $t_2$  are the times of failure in the three sections, respectively.

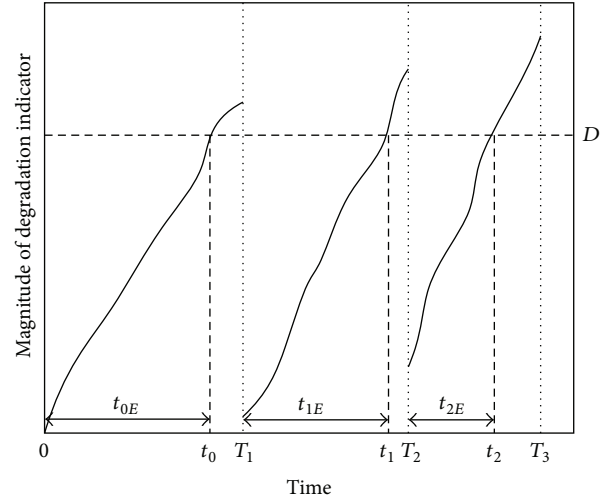


FIGURE 4: Schematic drawing of local lifetimes.

TABLE 2: Average local lifetime in each section.

Section	Number of failed signals	Average local lifetime
Before 1st PM	33	96.5
Between 1st and 2nd PM	49	85.3
Between 2nd and 3rd PM	80	72.9

The corresponding local lifetimes are  $t_{0E} = t_0 - 0$ ,  $t_{1E} = t_1 - T_1$ , and  $t_{2E} = t_2 - T_2$ . Concerning all of the 400 simulated processes, the number of failed processes and the average local lifetime in each section are shown in Table 2.

From Table 2, we observe that the average local lifetime decreases with the increases in the times of imperfect PM activities. These simulated data are appropriate for processes with imperfect PM activities, which restore the system to better but not “as good as new” states.

**4.2. RUL Prediction and Results Comparison.** We used the RUL prediction, an example to test the effectiveness of the generalized AFT frailty model. Among the 400 simulated degradation processes, we used 300 samples for training and 100 samples for testing. Parameters of the model were estimated with training samples, and RUL predictions were calculated with testing samples. During the parameters estimation procedure, we used the typical Weibull distribution as the baseline failure function. We chose the gamma distribution with mean 1 and variance  $\theta$  as the frailty distribution for the following two reasons. Firstly, the gamma distribution is the most popular choice for the frailty distribution [12, 18, 23]. The posterior distribution of the gamma frailty is still the gamma distribution. Secondly, the gamma distribution with mean 1 and variance  $\theta$  is appropriate for the characterization of the frailty terms in our model. Frailty acts multiplicatively on the failure rate function and represents the change in the failure rate after PM relative to that before PM. This multiplication factor is theoretically greater than or equal to 1 but sometimes smaller than 1 in reality [17].

TABLE 3: Statistical analysis of the predicted RUL using G-AFT-F and EPHM for the testing samples that fail before receiving any PM activity.

Actual RUL	10	9	8	7	6	5	4	3	2	1
G-AFT-F										
Mean	9.93	8.96	7.90	7.06	6.05	4.86	3.80	3.07	2.09	1.24
SD	0.65	0.76	0.43	0.86	0.75	0.75	0.86	0.43	0.35	0.27
EPHM										
Mean	9.91	9.03	7.95	7.04	6.01	5.03	3.98	3.06	2.12	1.32
SD	0.57	0.69	0.97	0.48	0.59	0.76	0.57	0.58	0.63	0.59

\*SD = standard deviation.

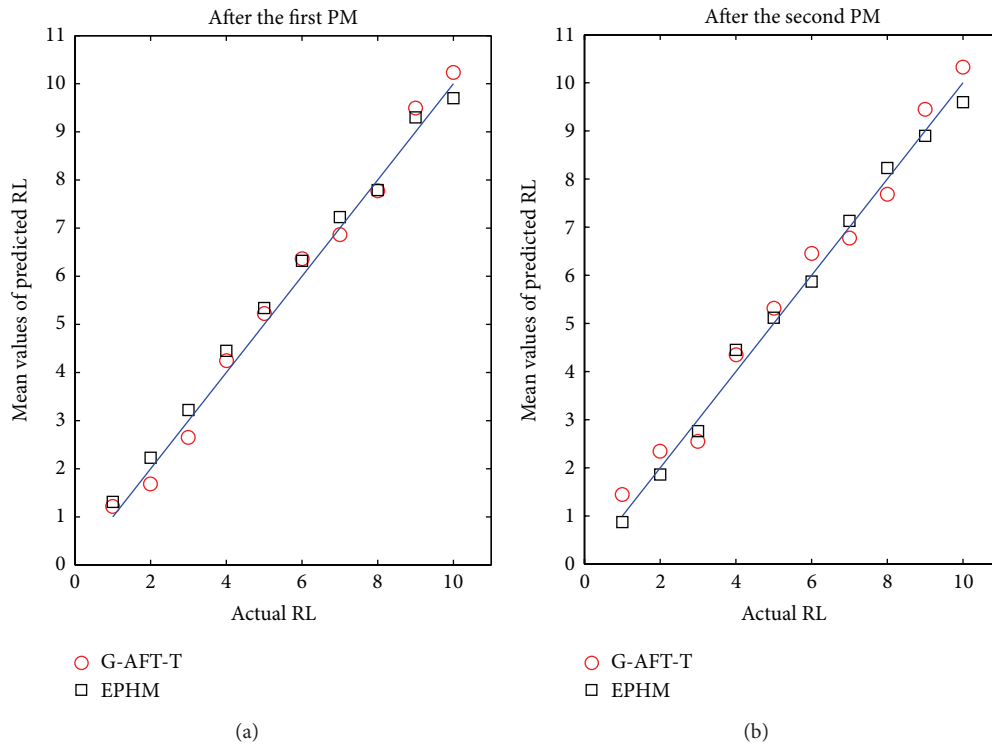


FIGURE 5: Statistical analysis of the predicted RUL using G-AFT-F and EPHM for the testing samples that fail after the first and second PM.

With the estimated parameters  $\hat{\theta}$  and  $\hat{\beta}$ , we calculated the RUL prediction of the 100 testing samples following the typical procedure [17]. The minimum of the actual local lifetimes in three sections were 43.5, 39.6, and 37.3, respectively. Theoretically, all the samples whose RUL is less than the minimum of the actual local lifetime could be used for further analysis. However, because the prediction of being close to failure is the most essential information in a CBM program, we only collect the samples whose actual RUL is less than 10. One part of the statistical analysis of the predicted RUL based on the generalized AFT frailty model (G-AFT-F for short) and the EPHM with respect to the actual RUL are summarized in Table 3.

The statistical analysis for the samples that fail after receiving the first PM activity and the second PM activity is shown in Figure 5. Mean values of the statistical analysis results are regarded as predicted RUL values, so only mean values are illustrated. The horizontal axis represents the actual

RUL, and the vertical axis represents the mean values of the predicted RUL.

From Table 3 and Figure 5 we can see that there is no obvious performance difference between the results using both models. The superior performance of the EPHM over other classical models has been proved [17]. Similar to the EPHM, the generalized AFT frailty model can achieve reasonably accurate and reliable RUL prediction. Thus, it is demonstrated that the generalized AFT frailty model is appropriate to model the failure likelihood of the systems subject to imperfect PM activities.

### 5. Conclusions and Future Work

Imperfect PM activities are considerably common in industrial systems. This paper has introduced the methods of survival analysis into the CBM field and used the statistical tools to solve practical industrial problems. The generalized

AFT frailty model has been investigated to model the failure likelihood of systems subject to imperfect PM activities. On the basis of the ML estimation and the EM method, the hybrid ML-EM algorithm for parameter estimation has been analyzed. In the evaluation experiment, the data of the typical degradation model has been generated. The RUL prediction has been taken as a calculation case and the results of the proposed model have been compared with those of the existing model. The comparison demonstrates the effectiveness of the generalized AFT frailty model.

The data in the experiment are generated from the typical degradation processes and are consistent to practical industrial cases. That is why we performed the evaluation only with simulation data. However, in future research, it will be necessary to evaluate the performance of the proposed model using real data. In addition, it would be interesting to introduce more methods of survival analysis into reliability-related fields and investigate their applicability.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgment

This research work was supported by the National Natural Science Foundation of China under the Contract nos. 60903033 and 11271289.

### References

- [1] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1519, 2006.
- [2] F. Camci, *Process monitoring, diagnostics and prognostics using support vector machines and hidden Markov models [Ph.D. thesis]*, Department of Industrial Engineering, Wayne State University, Detroit, Mich, USA, 2005.
- [3] C.-H. Hu, X.-S. Si, and D.-H. Zhou, "Nonlinear degradation process modeling and remaining useful life estimation subject to measurement error," *Acta Automatica Sinica*, vol. 39, no. 5, pp. 530–541, 2013.
- [4] X. S. Si, W. Wang, C. H. Hu, D. H. Zhou, and M. G. Pecht, "Remaining useful life estimation based on a nonlinear diffusion degradation process," *IEEE Transactions on Reliability*, vol. 61, no. 1, pp. 50–67, 2012.
- [5] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society Series B*, vol. 34, no. 2, pp. 187–220, 1972.
- [6] Z. Ma and A. W. Krings, "Survival analysis approach to reliability, survivability and prognostics and health management(PHM)," in *Proceedings of the IEEE Aerospace Conference*, pp. 1–20, 2008.
- [7] A. K. S. Jardine, D. Banjevic, M. Wiseman, S. Buck, and T. Joseph, "Optimizing a mine haul truck wheel motors' condition monitoring program: use of proportional hazards modeling," *Journal of Quality in Maintenance Engineering*, vol. 7, no. 4, pp. 286–301, 2001.
- [8] D. Lin, D. Banjevic, and A. K. S. Jardine, "Using principal components in a proportional hazards model with applications in condition-based maintenance," *Journal of the Operational Research Society*, vol. 57, no. 8, pp. 910–919, 2006.
- [9] H. Liao, W. Zhao, and H. Guo, "Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model," in *Proceedings of the Annual Reliability and Maintainability Symposium*, pp. 127–132, Newport Beach, Calif, USA, January 2006.
- [10] L. J. Wei, "The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis," *Statistics in Medicine*, vol. 11, no. 14-15, pp. 1871–1879, 1992.
- [11] D. R. Cox, "Some remarks on the analysis of survival data," in *Proceedings of the 1st Seattle Symposium in Biostatistics: Survival Analysis*, Springer, New York, NY, USA, 1997.
- [12] W. Pan, "Using frailties in the accelerated failure time model," *Lifetime Data Analysis*, vol. 7, no. 1, pp. 55–64, 2001.
- [13] H. Pham and H. Wang, "Imperfect maintenance," *European Journal of Operational Research*, vol. 94, no. 3, pp. 425–438, 1996.
- [14] M. J. Kallen, "Modelling imperfect maintenance and the reliability of complex systems using superposed renewal processes," *Reliability Engineering & System Safety*, vol. 96, no. 6, pp. 636–641, 2011.
- [15] S. El-Ferik and M. Ben-Daya, "Age-based hybrid model for imperfect preventive maintenance," *IIE Transactions*, vol. 38, no. 4, pp. 365–375, 2006.
- [16] M.-Y. You, H. Li, and G. Meng, "Control-limit preventive maintenance policies for components subject to imperfect preventive maintenance and variable operational conditions," *Reliability Engineering & System Safety*, vol. 96, no. 5, pp. 590–598, 2011.
- [17] M.-Y. You and G. Meng, "Residual life prediction of repairable systems subject to imperfect preventive maintenance using extended proportional hazards model," *Proceedings of the Institution of Mechanical Engineers E: Journal of Process Mechanical Engineering*, vol. 226, no. 1, pp. 50–63, 2012.
- [18] J. Zhang and Y. Peng, "An alternative estimation method for the accelerated failure time frailty model," *Computational Statistics & Data Analysis*, vol. 51, no. 9, pp. 4413–4423, 2007.
- [19] D. Collett, *Modelling Survival Data in Medical Research*, Chapman & Hall, CRC, Boca Raton, Fla, USA, 2003.
- [20] D. Clayton and J. Cuzick, "Multivariate generalizations of the proportional hazards model," *Journal of the Royal Statistical Society Series A*, vol. 148, no. 2, pp. 82–117, 1985.
- [21] P. Chen, *Generalized accelerated failure time frailty model [M.S. dissertation]*, East China Normal University, 2010.
- [22] J. P. Klein, "Semiparametric estimation of random effects using the Cox model based on the EM algorithm," *Biometrics*, vol. 48, no. 3, pp. 795–806, 1992.
- [23] H. T. V. Vu and M. W. Knuiman, "A hybrid ML-EM algorithm for calculation of maximum likelihood estimates in semiparametric shared frailty models," *Computational Statistics & Data Analysis*, vol. 40, no. 1, pp. 173–187, 2002.
- [24] D. Zeng and D. Y. Lin, "Efficient estimation for the accelerated failure time model," *Journal of the American Statistical Association*, vol. 102, no. 480, pp. 1387–1396, 2007.
- [25] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: a Bayesian approach," *IIE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.





# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

