Degradation-Based Maintenance Using Stochastic Filtering for Systems under Imperfect Maintenance

ABSTRACT

The stationary Wiener process is widely used in modeling degradation processes, mainly due to the existence of an analytical expression of the first hitting time distribution. However, it is only appropriate for modelling linearly drifted stochastic processes. This paper investigates a maintenance policy in which a family of non-stationary Wiener processes is used to simulate degradation phenomena. An approximation to the first hitting time distribution is invoked. A novel technique to treat imperfect maintenance is initiated by extending the improvement-factor method to the degradation rate function. A recursive filter is utilized to dynamically update the estimate of the degradation rate function. An algorithm, using the Laplace approximation, is developed for estimating model parameters. The applications of the proposed methods are demonstrated via both real-world and numerical examples.

Keywords: Maintenance; Degradation rate; Kalman filter; Laplace approximation; Non-stationary Wiener process.

1 INTRODUCTION

Degradation modeling serves as an efficient approach to evaluating reliability and predicting failure events for some highly reliable products. An appealing dominance of the degradation-based approach is that massed degradation measurements can be recorded on each individual product. Early work on degradation modeling is referenced by Meeker and Escobar (1998) and Ray and Phoha (1999), while more recent results are mentioned by Elwany et al. (2011), Zhou et al. (2014) and Zhou et al. (2012). The stationary Wiener process $\{Z_t,\ t \geq 0\}$ with drift λ and infinitesimal variance σ^2 is defined by $Z_t = \lambda t + \sigma W_t$, in which $\lambda > 0$ and $\sigma > 0$. $\{W_t,\ t \geq 0\}$ is the standard Brownian motion. The stationary Wiener process is a widely used stochastic process to characterize degradation phenomena (Mercier and Castro, 2013; Nikulin et al., 2010). The stationary Wiener process has stationary, independent, and normally distributed increments. That is, for all $0 \leq s < t$, $Z_t - Z_s$ is independent of Z_s and has the normal distribution $N\left(\lambda(t-s),\ \sigma^2(t-s)\right)$. A failure occurs when the degradation exceeds a pre-determined failure threshold. The failure threshold corresponds to the limit of deterioration beyond which the mission of the product is no longer fulfilled. The stationary Wiener process possesses a mathematical advantage in that the distribution of the first hitting time can be formulated analytically. However, the stationary Wiener process is inappropriate for modeling non-linearly drifted degradation processes. Non-linearly drifted degradation processes exist pervasively in real-life practice (Bian and Gebraeel, 2012; Son et al., 2013; Zhou et al., 2011). The model in which the degradation rate increases with the level of degradation is, for example, used in modeling the growth of fatigue crack. Hudak et al. (1978) collected fatigue crack-size data from alloy to obtain information on crack-growth rates. See also Meeker and Escobar (1998) for modeling the size of fatigue crack as a function of the number of cycles. The convex curves in Figure 13.1 (Meeker and Escobar, 1998) verify that the crack-growth rate increases with the size of crack. Compared with linearly drifted degradation processes, a rather limited amount of research has been done on simulating non-linearly drifted degradation processes. The present paper investigates a maintenance policy in which a family of non-stationary Wiener processes is utilized to characterize degradation phenomena. A non-stationary Wiener process with drift function $v(t)$ and infinitesimal variance σ^2 is defined by $X_t = v(t) + \sigma W_t$. $v(t)$ is a continuous, non-linear, real-valued function of $t \geq 0$ with $v(0) = 0$. The process $\{X_t,\ t \geq 0\}$ has independent, non-stationary, and normally distributed random increments: For all $0 \leq s < t$, $X_t - X_s$ is independent of X_s and has the normal distribution $N(v(t) - v(s),\ \sigma^2(t-s))$. The non-stationary Wiener process reduces to a stationary Wiener process with $v(t)$ being a linear function of t.

A non-trivial problem in the degradation-based maintenance need be solved: the treatment to imperfect maintenance. Widely used assumptions on maintenance actions are good-as-new (or, perfect) and bad-as-old (or, minimal). It is more realistic in true experience that maintenance actions merely restore a product's condition to somewhere between good-as-new and bad-as-old. This type of maintenance is known as the imperfect maintenance. Extensive research on imperfect maintenance has been documented (Lindqvist, 2006; Nicolai et al., 2009; Zhang et al., 2013). One of the most popular treatments is the improvement-factor method. The improvement-factor method assumes that each maintenance action changes the time of the failure rate curve to some newer time but not all the way to zero; see Pham and Wang (1996). A general framework of the improvement-factor method is given as follows. Let $h(t)$ $(t \geq 0)$ denote the hazard rate function of the target product. Given that a maintenance action is performed at time $t_1 (> 0)$, right after the maintenance action the hazard rate function assumes a new expression: $\beta h(t - t_1 + \alpha t_1)$. Here, $t - t_1$ represents the time elapsed from the maintenance action. $0 < \alpha < 1$ is an age-reduction factor. $\beta > 1$ is a hazard-rate-increase factor. A representative sample of works on improvement-factor method includes Chen et al. (2013), Doyen and Gaudoin (2011), Kurt and Kharoufeh (2010) and Li et al. (2012). However, the issue of treating imperfect maintenance in the context of degradation-based maintenance has not received much attention and remains widely open. By assuming that the degradation conforms to the non-stationary Wiener process, $\{X_t,\ t \geq 0\}$, the expected degradation up to time t is $E[X_t] = v(t)$. Therefore, the first-order derivative of $v(t)$, denoted to be $v^{'}(t)$, can be treated as the degradation rate function of the underlying degradation process. The degradation rate defined herein is the rate of accumulation of degradation. The degradation rate has been addressed a lot (Meeker and Escobar, 1998), and recent research can be found in Nikulin et al. (2010) and Rafiee et al. (2013). The improvement-factor method can be extended to the degradation rate function, i.e. $v^{'}(t)$, to treat imperfect maintenance. Maintenance actions slowing down degradation rate can be easily visualized in the context of, e.g., coating operations. Steel structures such as bridges, tanks and pylons are exposed to outdoor weathering conditions. In order to slow down the deteriorating process, they are usually protected by certain organic coating systems (Nicolai et al., 2009; Perrin et al., 2009). The degradation level after each coating operation may or may not change. Another typical maintenance action is lubricating rotating gears, by which the wearing processes of the gears will slow down. The degradation level after lubrication remains unchanged.

Another distinguishing feature of the proposed maintenance strategy is the utilization of stochastic filtering and the Laplace approximation. The bulk of the documented research on the improvement-factor method assumes a constant age-reduction factor and a constant hazard-rate-increase factor during a product's whole operational life cycle. Apparently, the assumption of constant improving factors will be inappropriate and problematical in many circumstances. Since the operational condition varies wildly, each maintenance action should have a different degree of impact on the degradation rate. Therefore, the improving factors are indeed random variables, and the sequential maintenance actions result in a dynamical system. In the present paper, it is assumed that the improving factors are random variables following certain probability distributions. The realizations of the improving factors are estimated dynamically, by using stochastic filtering. Stochastic filtering is commonly used in the field of analyzing condition-monitoring data; see Ece and Basaran (2011), Lall et al. (2012) and Wang and Wang (2012). By assuming probability distributions for the improving factors, more unknown model parameters will be introduced. The problem of estimating model parameters need be solved. When a likelihood function involves a multiple integral, the parameter-estimation problem will be very difficult. Most of the documented research on proposing maintenance policies made tangential remarks on the parameter-estimation problem. The present paper uses Laplace's method (Pinheiro and Bates, 1995) to deal with the multiple integral in a likelihood function.

In what follows, we first propose a degradation-based maintenance policy in which maintenance actions are assumed to be imperfect. Then we dynamically estimate each realization of an improving factor by using a recursive filter. In the third step, we develop a procedure for estimating unknown model parameters. The remainder of the paper is organized as follows. Section 2 is devoted to giving the framework of the advanced techniques. The method of treating imperfect maintenance via the degradation rate function is mathematically detailed. The proposed maintenance scheme is presented and the expression of the cost rate function is derived. Section 3 develops a stochastic filter for on-line estimating the realizations of an improving factor. Section 4 develops an algorithm for estimating unknown model parameters. Section 5 gives numerical examples to illustrate the applicability and competence of the advanced techniques. Section 6 outlines concluding remarks and future extensions.

2 MODEL FORMULATION

Maintenance actions are all assumed to take negligible time, in comparison to the expected operational lifespan of the targeted system. The idea of extending the improvement-factor method to the degradation rate function is detailed as follows. Suppose that a maintenance action is taken at time $t > 0$. Right before the maintenance the degradation rate is $v^{'}(t)$. Right after the maintenance the degradation rate changes to $bv^{'}(t)$, with $0 < b < 1$. Here, b is a degradation-rate-reduction factor. From time t forward, the degradation rate function assumes a new expression: $bv^{'}(s+t)$, $s > 0$. Correspondingly, the drift function assumes a new expression: $bv(s+t)+c$, $s > 0$. Here, c is a shifting factor. The shifting factor is introduced due to the fact that the maintenance action may or may not change (or, shift) the degradation level. By definition, the degradation level right before the maintenance is $v(t)+\sigma w_t$. w_t is a realization of W_t. The degradation level right after the maintenance is $bv(t)+c+\sigma w_t$. By adjusting the shifting factor, $bv(t)+c+\sigma w_t$ can take any value, including $v(t)+\sigma w_t$. The case, in which $c = v(t) - bv(t)$, indicates that the degradation level remains unchanged. The case, in which $c \neq v(t) - bv(t)$, indicates that the maintenance action results in a shift in the degradation level. Since the operational condition varies wildly, each maintenance action should have a different degree of impact on the degradation process. Therefore, in what follows, we assume that the factors, b and c, are random variables. With $b = 1$, we arrive at the bad-as-old assumption. Note that the proposed method involves the degradation-rate-reduction factor and the shifting factor, yet not the age-reduction factor. The topic of involving the age-reduction factor is beyond the scope of this paper. In terms of degradation-based maintenance, treating imperfect maintenance via the degradation rate function instead of the hazard rate function has two main advantages. Firstly, deriving the hazard rate function via the first hitting time distribution function is mathematically intractable, especially when the degradation process is non-stationary. Secondly, unlike the hazard rate which is unobservable and unmeasurable, the impact of maintenance actions on the degradation rate can be exteriorized from consecutive degradation measurements.

It is noteworthy that each realization of the shifting factor can be valued till the end, when all the other unknown parameters are estimated. For example, suppose that a maintenance action is performed at time $t_1 > 0$. From time t_1 forward, the degradation process $\{X_t,\ t \geq 0\}$ is defined by $X_t = b_1 v(t + t_1) + c_1 + \sigma w_{t_1} + \sigma W_t$. b_1 and c_1 are respective realizations of b and c at time t_1. w_{t_1} is a realization of W_{t_1}. The degradation level right before the maintenance is $x_{t_1^-} = v(t_1) + \sigma w_{t_1}$. The degradation level right after the maintenance is $x_{t_1^+} = b_1 v(t_1) + c_1 + \sigma w_{t_1}$. Suppose that another maintenance action is taken at time $t_2 > t_1$. The degradation level at time t_2, right before the second maintenance, is $x_{t_2^-} = b_1 v(t_2 + t_1) + c_1 + \sigma w_{t_1} + \sigma w_{t_2}$. w_{t_2} is a realization of W_{t_2}. Because the drift function remains unchanged between two consecutive maintenance actions, c_1 is eliminated in the degradation increment: $x_{t_2^-} - x_{t_1^+} = b_1 v(t_2 + t_1) + \sigma w_{t_2} - b_1 v(t_1)$. Given data $\{x_{t_1^-},\ x_{t_1^+},\ x_{t_2^-}\}$, we first estimate the other model parameters based on the degradation increments $x_{t_1^-}$ and $x_{t_2^-} - x_{t_1^+}$. Then, the realization of the shifting factor, c_1, can be easily calculated based on the shift, $x_{t_1^+} - x_{t_1^-}$, and the estimated model parameters. Evidently, what plays a decisive role in estimating the other model parameters are the degradation increments. The shift of the degradation level is of no effect on the estimation of the other model parameters. Consequently, when estimating the other model parameters, there is no loss of generality by assuming that maintenance actions do not shift the degradation level. This assumption is made only for ease of exposition: without introducing too many subscripts or superscripts.

The maintenance scheme proposed in this paper is implemented as follows. Preventive maintenance (PM) is assumed to be imperfect. Replacement, either preventive or unexpected, is assumed to be perfect. Degradation-measuring (DM) points are equally spaced with $\triangle > 0$ being the constant time between two consecutive DM points. The DM duration is negligible. Upon the ith $(i = 1, 2, 3, \ldots)$ DM point, $i\,\triangle$, a degradation measurement is obtained instantly, denoted by x_i. Inspection is perfect in the sense that it reveals the true degradation level of the system and does not change the condition of the system. By using a recursive filter, the degradation rate function is immediately updated utilizing the history of degradation measurements and maintenance actions up to date. Having updated the degradation rate function, engineers have to decide whether or not to take PM at the ith DM point. Then, engineers have to calculate the optimal preventive-replacement time, denoted by t^*. If the optimal preventive-replacement time is larger than the inspection interval, i.e. $t^* > \triangle$, the maintained system will be inspected at the next DM point, and the process repeats. If $t^* \leq \triangle$, the maintained system will be preventively replaced at point $i \triangle + t^*$, and the process resumes. The DM points need not be equally spaced. Generalization of the proposed maintenance strategy to irregular DM points is a straightforward matter. The DM points could be included as decision variables driving the overall degradation-based maintenance policy. However, the complexity of inspections in practice might allow measurements to be taken only at selected epochs.

Figure 1 demonstrates a possible course of a replacement cycle, i.e. the time from the in-

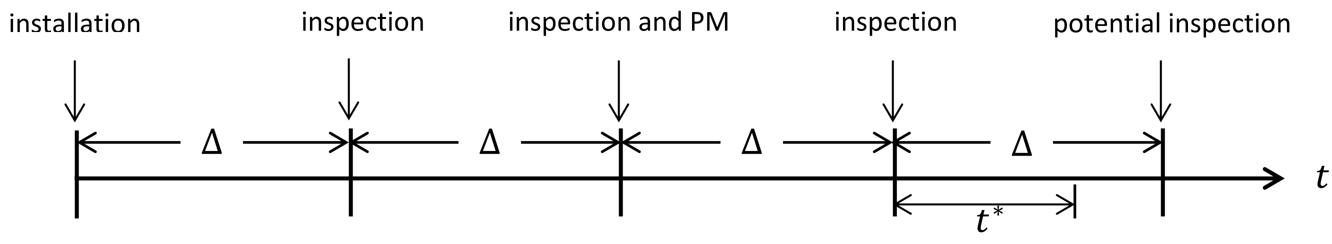

Figure 1: A possible course of a replacement cycle.

stallation of a system to its replacement. At the first DM point, $\triangle$, after obtaining a degradation measurement and updating the degradation rate function, the optimal maintenance decision is to subject the system to the next inspection with no PM action taken at time $\triangle$. At the second DM point, $2\triangle$, after obtaining another degradation measurement and updating the degradation rate function, the optimal maintenance decision is to subject the system to the next inspection with a PM action taken at time $2\triangle$. At the third DM point, the optimal maintenance decision is to preventively replace the system at time $3\triangle + t^*$ with no PM action taken at time $3\triangle$. If the system fails unexpectedly, it will be immediately replaced by a new physically and statistically identical one.

Unexpected failures can be observed instantaneously, each costing C_f. The average costs of a PM action and a preventive replacement are C_p and C_r, respectively. The cost of each inspection is C_I. Practical conditions define the constraints of the maintenance costs as follows: $C_I < C_p < C_r < C_f$.

Define a random variable, T_l, to be the first hitting time of the degradation process $\{X_t,\ t \geq 0\}$ to a pre-determined threshold, l. Given that the maintained system is still operating at the ith DM point, $i\triangle$, we introduce the following notation.

- $\ddot{X}_i$: history of degradation measurements up to and including the ith, $\ddot{X}_i = \{x_1, x_2, ..., x_i\}$.

- $\ddot{M}_i$: history of PM actions up to and including the ith, $\ddot{M}_i = \{m_1\ m_2, ..., m_i\}$. $m_i = 0$ indicates that no maintenance action is taken at the ith DM point. $m_i = 1$ indicates that an imperfect PM action is taken at the ith DM point.

- $EC^i_{pm}(t)$: expected cost per unit time with a PM action taken at the ith DM point, given that the system will be preventively replaced at time $i\triangle + t$.

- $EC^i_{na}(t)$: expected cost per unit time with no maintenance action at the ith DM point, given that the system will be preventively replaced at time $i\triangle + t$.

- $P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_i\right) = P\left(T_l \le t + i\bigtriangleup |\ddot{X}_i,\ \ddot{M}_i\right)$: remaining useful life distribution function at the ith DM point (after the PM, if any).

The optimization is based on the minimization of a cost rate function. If the system is replaced, a new cycle of operation starts. The failure process constructs a renewal process. According to the renewal theorem, the long-run cost rate is equal to the expected cost in one renewal cycle divided by the expected length of the renewal cycle. Suppose that a PM action is taken at time $i\bigtriangleup$ (instantaneously after the ith inspection) and that a preventive replacement is scheduled to be performed at time $i\bigtriangleup + t$. We have

$$\begin{aligned} EC_{pm}^{i}(t) &= \frac{ic_I + \sum_{j=1}^{i} m_j c_p + c_r\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)\right] + (c_f + c_r)P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)}{i\bigtriangleup + t\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)\right] + \int_0^t s\, dP_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)} \\ &= \frac{ic_I + \sum_{j=1}^{i} m_j c_p + c_r + c_f P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)}{i\bigtriangleup + t - \int_0^t P_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right) ds}. \end{aligned}$$

Here, ic_I is the overall inspection cost. $\sum_{j=1}^{i} m_j c_p$ is the overall PM cost. $c_r\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_i\right)\right]$ is the expected cost of a preventive replacement at time $i\bigtriangleup + t$. $(c_f + c_r)P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_i\right)$ is the expected cost of an unexpected failure within interval $(i\bigtriangleup,\ i\bigtriangleup + t)$. The denominator is the expected length of a replacement cycle:

$$\begin{aligned} E[T_l] &= i\bigtriangleup + E[T_l - i\bigtriangleup |T_l > i\bigtriangleup] = i\bigtriangleup + \int_0^\infty s\, dP_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right) \\ &= i\bigtriangleup + t\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right)\right] + \int_0^t s\, dP_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 1\right). \end{aligned}$$

Given no maintenance action at time $i\bigtriangleup$ and a scheduled preventive replacement at time $i\bigtriangleup + t$, we have

$$\begin{aligned} EC_{na}^{i}(t) &= \frac{ic_I + \sum_{j=1}^{i} m_j c_p + c_r\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right)\right] + (c_f + c_r)P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right)}{i\bigtriangleup + t\left[1 - P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right)\right] + \int_0^t s\, dP_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right)} \\ &= \frac{ic_I + \sum_{j=1}^{i} m_j c_p + c_r + c_f P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right)}{i\bigtriangleup + t - \int_0^t P_{rul}\left(s|\ddot{X}_i,\ \ddot{M}_{i-1}, m_i = 0\right) ds}. \end{aligned}$$

Denote t_p^* to be the optimal preventive-replacement time that minimizes the cost rate function $EC_{pm}^{i}(t)$. Denote t_n^* to be the optimal preventive-replacement time that minimizes the cost rate function $EC_{na}^{i}(t)$.

- If $EC_{pm}^{i}(t_p^*) < EC_{na}^{i}(t_n^*)$, a maintenance action is taken at time $i\bigtriangleup$, and the system will be preventively replaced at time $i\bigtriangleup + t_p^*$ $\left(if\ t_p^* < \bigtriangleup\right)$ or will be inspected at the next DM point $(i+1)\bigtriangleup$ $\left(if\ t_p^* > \bigtriangleup\right)$.

- If $EC_{pm}^{i}(t_p^*) > EC_{na}^{i}(t_n^*)$, no maintenance action is taken at time $i\bigtriangleup$, and the system will be preventively replaced at time $i\bigtriangleup + t_n^*$ $(if\ t_n^* < \bigtriangleup)$ or will be inspected at the next DM point $(i+1)\bigtriangleup$ $(if\ t_n^* > \bigtriangleup)$.

It is obvious that the remaining useful life distribution function, $P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_i\right)$, is conclusive in the cost rate function. $P_{rul}\left(t|\ddot{X}_i,\ \ddot{M}_i\right)$ can be analytically expressed in very few cases

(Buonocore et al., 2011). For a non-linearly drifted stochastic process, deriving the remaining useful life distribution function involves solving the Fokker-Planck-Kolmogorov equation with boundary constraints (Cox and Miller, 1965). Many works, such as Gebraeel et al. (2005), Park and Padgett (2005) and Wang (2010), make non-linear transformations on degradation data and fit the stationary Wiener process to the transformed data. An implicit assumption in the data-transformation approach is that the random part of the transformed process is still Brownian motion, which may not always be the case (Si et al., 2012). To reduce the heavy computational burden, we develop, in Section 3, a closed-form expression approximating the remaining useful life distribution function.

3 ON-LINE UPDATING

A commonly used formulation of the drift function in the non-stationary Wiener process is given by $v(t) = \lambda t^{\theta}$ (Son et al., 2013; Wang et al., 2014). $\lambda > 0$ is a scale parameter. $\theta > 0$ is a shape parameter. The degradation rate function is the derivative of the drift function: $v^{'}(t) = \lambda \theta t^{\theta-1}$. If a maintenance action is performed at time $\triangle$, immediately after the maintenance the degradation rate function assumes a new expression: $b_1 v^{'}(t) = b_1 \lambda \theta t^{\theta-1}$, $t > \triangle$. b_1 is a realization of b. We might define $\lambda_1 = b_1 \lambda$ and write $b_1 v^{'}(t) = \lambda_1 \theta t^{\theta-1}$. Instead of saying that the degradation rate function changes from $v^{'}(t)$ to $b_1 v^{'}(t)$, it is equivalent to saying that immediately after the maintenance the scale parameter changes from λ to λ_1. Therefore, to update the estimate of the degradation rate function, it is sufficient to update the estimate of the scale parameter. An updating procedure, i.e. the Kalman filter, is developed herein for recursively producing a statistically optimal estimate of the scale parameter. The Kalman filter (Chui and Chen, 2009) is an algorithm which iteratively involves two steps: the predict step and the update step. In the predict step, the Kalman filter produces an estimate of the state and an estimate of the covariance matrix. The estimate of the state is termed as the a prior state estimate. The estimate of the covariance matrix is termed as the a prior estimate error covariance. In the update step, the Kalman filter produces the a posterior state estimate and the a posterior estimate error covariance by using a weighted average. Theory, methods and applications of the Kalman filter have been extensively reported in operational research, such as Carr and Wang (2011), Humpherys et al. (2012), and Kluge et al. (2010); see Si et al. (2011) for a recent review.

Denote λ_i $(i = 1,2,3,...)$ to be the hidden true value of the scale parameter at time $i \triangle$ before any maintenance action. Denote x_i $(i = 0,1,2,...)$ to be the degradation measurement obtained at time $i \triangle$ with $x_0 = 0$. Denote y_i $(i = 1,2,3,...)$ to be the ith degradation increment: $y_i = x_i - x_{i-1}$. For $i = 1,2,3,...$, we have the process equation

$$\lambda_i = \mu_{i-1} \lambda_{i-1} + \delta_{i-1},$$

and the measurement equation

$$y_i = \rho_i \lambda_i + \omega_i.$$

The derivation of the process equation and the measurement equation is given in the Appendix. The process noise is additive, white and normally distributed with zero mean and variance Q_{i-1}. The measurement noise is additive, white and normally distributed with zero mean and variance $\sigma^2 \triangle$.

Denote $\hat{\lambda}_i^-$ to be the a prior estimate of λ_i, calculated by utilizing the knowledge of $\ddot{X}_{i-1}$ and $\ddot{M}_{i-1}$. Denote $\hat{\lambda}_i$ to be the a posterior estimate of λ_i, calculated by utilizing the degradation increment, y_i, and the a prior estimate, $\hat{\lambda}_i^-$. When making the a posterior estimate, $\hat{\lambda}_i$, be a linear combination of the a prior estimate, $\hat{\lambda}_i^-$, and the degradation increment, y_i, we arrive at the Kalman filter which is a minimum mean-square error estimator. Define P_i^- to be the a prior estimate error variance:

$$P_i^- = E[(\lambda_i - \hat{\lambda}_i^-)(\lambda_i - \hat{\lambda}_i^-)].$$

Define P_i to be the a posterior estimate error variance:

$$P_i = E[(\lambda_i - \hat{\lambda}_i)(\lambda_i - \hat{\lambda}_i)].$$

For i=1, 2, 3, ..., the recursive estimation of the scale parameter entails the following updating equations (Welch and Bishop, 2006):

- state estimate propagation
$$\hat{\lambda}_i^- = \mu_{i-1}\hat{\lambda}_{i-1},$$
- estimate error variance propagation
$$P_i^- = \mu_{i-1}^2 P_{i-1} + Q_{i-1},$$
- Kalman gain
$$G_i = \frac{P_i^- \rho_i}{P_i^- \rho_i^2 + \sigma^2 \vartriangle},$$
- state estimate update
$$\hat{\lambda}_i = \hat{\lambda}_i^- + G_i(y_i - \rho_i \hat{\lambda}_i^-),$$
- estimate error variance update
$$P_i = (1 - G_i\rho_i)P_i^-.$$

The starting values, $\hat{\lambda}_0$ and P_0, are chosen arbitrarily, reflecting our best assessment of the scale parameter prior to any DM information. Substituting $\hat{\lambda}_i^-$ and G_i into the state-estimate-update equation, we have

$$\hat{\lambda}_i = \mu_{i-1}\hat{\lambda}_{i-1} + \frac{(\mu_{i-1}^2 P_{i-1} + Q_{i-1})\rho_i}{(\mu_{i-1}^2 P_{i-1} + Q_{i-1})\rho_i^2 + \sigma^2 \vartriangle} \times \left(y_i - \rho_i \mu_{i-1}\hat{\lambda}_{i-1}\right).$$

Substituting P_i^- and G_i into the estimate-error-variance-update equation, we have

$$P_i = \frac{(\mu_{i-1}^2 P_{i-1} + Q_{i-1}) \times \sigma^2 \vartriangle}{(\mu_{i-1}^2 P_{i-1} + Q_{i-1})\rho_i^2 + \sigma^2 \vartriangle}.$$

Recall that the random variable T_l is defined as the first hitting time of the process $\{X_t,\ t \geq 0\}$ to the threshold l. For the non-stationary Wiener process with drift function $v(t) = \lambda t^\theta$, the analytical form of the density function, $P(T_l = t)$, cannot be constructed. Si et al. (2012) developed a closed-form expression[1] approximating the density function, which is given by

$$\tilde{f}(t;\, \lambda,\, \theta,\, \sigma,\, l) = \kappa \times \frac{l + (\theta - 1)\lambda t^{\theta}}{\sigma\sqrt{2\pi t^3}} \exp\left(-\frac{\left(l - \lambda t^{\theta}\right)^2}{2\sigma^2 t}\right), \qquad t \geq 0. \tag{1}$$

Here, κ is a normalizing constant:

$$\frac{1}{\kappa} = \int_0^{\infty} \frac{l + (\theta - 1)\lambda t^{\theta}}{\sigma\sqrt{2\pi t^3}} \exp\left(-\frac{\left(l - \lambda t^{\theta}\right)^2}{2\sigma^2 t}\right) dt.$$

Given that the maintained system is still operating at time $i \triangle$ $(i = 1, 2, 3, ...)$, once the degradation measurement x_i $(< l)$ is obtained, we immediately update the estimate of the scale parameter to be $\hat{\lambda}_i$. If there is no maintenance action at time $i \triangle$, the value of the scale parameter remains unchanged and is estimated by $\hat{\lambda}_i$. From time $i \triangle$ forward, the degradation process is defined by

$$X_{s+i\triangle} - X_{i\triangle} = \hat{\lambda}_i (s + i \triangle)^{\theta} - \hat{\lambda}_i (i \triangle)^{\theta} + \sigma W_s, \qquad s \geq 0.$$

We might denote $X_{s+i\triangle} - X_{i\triangle}$ by Y_s^i: $Y_s^i = X_{s+i\triangle} - X_{i\triangle}$. The remaining useful life of the system at time $i \triangle$ is indeed the first hitting time of the degradation process $\{Y_s^i,\, s \geq 0\}$ to the threshold $l_i = l - x_i$. The degradation process $\{Y_s^i,\, s \geq 0\}$ is a non-stationary Wiener process with drift function

$$\upsilon_i(s) = \hat{\lambda}_i (s + i \triangle)^{\theta} - \hat{\lambda}_i (i \triangle)^{\theta}.$$

We have $\upsilon_i(0) = 0$. Therefore, according to Equation (1), the remaining useful life distribution function, $P_{rul}\left(t | \ddot{X}_i,\, \ddot{M}_{i-1},\, m_i = 0\right)$, can be approximated by

$$\tilde{f}_i(t) = \int_0^t \kappa_i \times \frac{l_i - \upsilon_i(s) + \theta \hat{\lambda}_i s (s + i \triangle)^{\theta - 1}}{\sigma\sqrt{2\pi s^3}} \exp\left(-\frac{\left[l_i - \upsilon_i(s)\right]^2}{2\sigma^2 s}\right) ds, \quad t \geq 0, \tag{2}$$

with κ_i being the normalizing constant

$$\frac{1}{\kappa_i} = \int_0^{\infty} \frac{l_i - \upsilon_i(s) + \theta \hat{\lambda}_i s (s + i \triangle)^{\theta - 1}}{\sigma\sqrt{2\pi s^3}} \exp\left(-\frac{\left[l_i - \upsilon_i(s)\right]^2}{2\sigma^2 s}\right) ds. \tag{3}$$

If a maintenance action is performed at time $i \triangle$, then after the maintenance the scale parameter is estimated by $\hat{\lambda}_{i+1}^{-} = \mu \hat{\lambda}_i$. The remaining useful life distribution function, $P_{rul}\left(t | \ddot{X}_i,\, \ddot{M}_{i-1},\, m_i = 1\right)$, can be approximated by $\tilde{f}_i(s)$ in which $\hat{\lambda}_i$ is replaced by $\hat{\lambda}_{i+1}^{-}$. By using the closed-form expression, $\tilde{f}_i(s)$, approximating the remaining useful life distribution function, $P_{rul}\left(t | \ddot{X}_i,\, \ddot{M}_i\right)$, the expected cost rate functions, $EC_{pm}^i(t)$ and $EC_{na}^i(t)$, can be readily minimized.

[1] As can be evidenced, the integral of the approximating expression, i.e. Equation (11) in Si et al. (2012), over the interval $(0, +\infty)$ does not equal one. Therefore, in order to make it a probability density function, we multiply the approximating expression by a normalizing constant.

4 PARAMETER ESTIMATION

In the proposed maintenance policy we have five model parameters: three degradation parameters, $\{\lambda,\ \theta,\ \sigma\}$, and two maintenance parameters, $\{\mu,\ Q\}$. When estimating the scale parameter by the Kalman filter, the other model parameters, θ, σ, μ and Q, are assumed to be known. It is worth stressing that the performance of the recursive filter relies on whether or not the other four parameters are accurately estimated. We develop herein an approximating approach to estimating the model parameters: θ, σ, μ and Q. Up to time t, with $n\ \triangle \leq t < (n+1)\ \triangle$, we obtain degradation measurements $\ddot{X}_n = \{x_1, x_2, ..., x_n\}$ and maintenance indicators $\ddot{M}_{n-1} = \{m_1, m_2, ..., m_{n-1}\}$. Since the $(n+1)$th degradation measurement is not available, the nth maintenance indicator, m_n, is of no help in estimating the model parameters and hence omitted. Note that if a maintenance action is performed, the realization of the degradation-rate-reduction factor, on which the likelihood function is conditioned, is latent. Define a column vector $\boldsymbol{b}_{n-1} = (b_1, b_2, ..., b_{n-1})^T$ in which b_i $(1 \leq i \leq n-1)$ is the ith realization of b. By the superscript T, we mean the transpose of a vector/matrix. If there is no maintenance at time $i\ \triangle$, we simply set $b_i = 1$. For illustrative purpose, we might assume that $m_i = 1$ for all $1 \leq i \leq n-1$. Generalization to $m_i = 0$, for some $1 \leq i \leq n-1$, is straightforward. Given degradation measurements $\ddot{X}_n$, a general formulation of the likelihood function of the parameters $\Theta = \{\lambda,\ \theta,\ \sigma,\ \mu,\ Q\}$ is given by

$$\ell\left(\Theta|\ddot{X}_n\right) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \frac{1}{(\sigma\sqrt{2\pi\ \triangle})^n} \exp\left(-\frac{(x_1-\lambda\rho_1)^2}{2\sigma^2\ \triangle} - \frac{1}{2\sigma^2\ \triangle}\sum_{i=2}^{n}\left(x_i - x_{i-1} - \prod_{j=1}^{i-1} b_j\lambda\rho_i\right)^2\right)$$
$$\times \frac{1}{(\sqrt{2\pi Q})^{n-1}} \exp\left(-\sum_{j=1}^{n-1}\frac{(b_j-\mu)^2}{2Q}\right) d\boldsymbol{b}_{n-1}.$$

Each evaluation of $\ell\left(\Theta|\ddot{X}_n\right)$ will, in general, require numerical approximation of the integral of dimension $n-1$. When the number of maintenance actions is large, maximizing $\ell\left(\Theta|\ddot{X}_n\right)$ is extremely difficult and time-consuming. One possible solution to overcoming the computational inefficiency is to use approximations to the likelihood function. We use the Laplace approximation to serve the purpose. Once the estimates of the model parameters are obtained, the realizations of the shifting factor will be readily determined based on the estimates and the degradation measurements.

Define a new function $gr\left(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n\right)$:

$$gr\left(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n\right) = \frac{(x_1-\lambda\rho_1)^2}{2\sigma^2\ \triangle} + \frac{1}{2\sigma^2\ \triangle}\sum_{i=2}^{n}\left(x_i - x_{i-1} - \prod_{j=1}^{i-1} b_j\lambda\rho_i\right)^2 + \sum_{j=1}^{n-1}\frac{(b_j-\mu)^2}{2Q}.$$

The likelihood function, $\ell\left(\Theta|\ddot{X}_n\right)$, can be re-formulated to be

$$\ell\left(\Theta|\ddot{X}_n\right) = (2\pi\sigma^2\ \triangle)^{-n/2} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (2\pi Q)^{-(n-1)/2} \exp\left(-gr\left(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n\right)\right) d\boldsymbol{b}_{n-1}.$$

Let $gr^{'}(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n)$ denote the first order partial derivative of $gr(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n)$ with respect to $\boldsymbol{b}_{n-1}$:

$$gr^{'}(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n) = \frac{\partial gr(\Theta,\ \boldsymbol{b}_{n-1},\ \ddot{X}_n)}{\partial \boldsymbol{b}_{n-1}}.$$

Let $gr''(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ denote the second order partial derivative of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ with respect to $\boldsymbol{b}_{n-1}$:

$$gr''(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n) = \frac{\partial^2 gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)}{\partial \boldsymbol{b}_{n-1} \partial \boldsymbol{b}_{n-1}^T}.$$

The expressions of $gr'(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ and $gr''(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ are given in the Appendix. Let $\hat{\boldsymbol{b}}_{n-1}$ denote the minimizer of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$:

$$\hat{\boldsymbol{b}}_{n-1} = \arg\min_{\boldsymbol{b}} gr(\Theta,\, \boldsymbol{b},\, \ddot{X}_n).$$

Because $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ is differentiable, we have $gr'(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n) = 0$. The calculation of $\hat{\boldsymbol{b}}_{n-1}$ is straightforward, for which standard and efficient code is available. The second order Taylor expansion of $gr\left(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n\right)$ around $\hat{\boldsymbol{b}}_{n-1}$ is:

$$gr\left(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n\right) \approx gr\left(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n\right) + \frac{1}{2}[\boldsymbol{b}_{n-1} - \hat{\boldsymbol{b}}_{n-1}]^T gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)[\boldsymbol{b}_{n-1} - \hat{\boldsymbol{b}}_{n-1}],$$

in which the linear term of the Taylor series vanishes since $gr'(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n) = 0$. Treating $[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}$ as a covariance matrix, we have

$$\sqrt{|[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}|} = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^{n-1}}} \exp\left(-\tfrac{1}{2}[\boldsymbol{b}_{n-1} - \hat{\boldsymbol{b}}_{n-1}]^T gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)[\boldsymbol{b}_{n-1} - \hat{\boldsymbol{b}}_{n-1}]\right) d\boldsymbol{b}_{n-1},$$

in which $|[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}|$ is the determinant of $[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}$. Substituting the second order Taylor expansion into the likelihood function, $\ell\left(\Theta|\ddot{X}_n\right)$, we have

$$\ell\left(\Theta|\ddot{X}_n\right) \approx \frac{1}{\sqrt{(2\pi\sigma^2 \bigtriangleup)^n}} \exp\left(-gr\left(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n\right)\right) \frac{1}{\sqrt{Q^{n-1}}} \sqrt{|[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}|}.$$

The Laplace approximation to the log-likelihood function is

$$\ell_{LA}\left(\Theta|\ddot{X}_n,\, \hat{\boldsymbol{b}}_{n-1}\right) \cong -n\log(\sigma) - gr\left(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n\right) - \frac{n-1}{2}\log(Q) + \frac{1}{2}\log(|[gr''(\Theta,\, \hat{\boldsymbol{b}}_{n-1},\, \ddot{X}_n)]^{-1}|).$$

Let $\hat{\Theta}^k = \{\hat{\lambda}^k,\, \hat{\theta}^k,\, \hat{\sigma}^k,\, \hat{\mu}^k,\, \hat{Q}^k\}$ denote the estimates of the model parameters, obtained after the kth iteration ($k = 1,\, 2,\, ...$). Given the current estimates, $\hat{\Theta}^k$, the LA algorithm in the $(k+1)$th iteration evolves as follows.

Step-One: Minimize $gr(\hat{\Theta}^k,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ with respect to $\boldsymbol{b}_{n-1}$ and substitute the minimizer $\hat{\boldsymbol{b}}_{n-1}^{(k+1)}$ into the Laplace approximation: $\ell_{LA}\left(\Theta|\ddot{X}_n,\, \hat{\boldsymbol{b}}_{n-1}^{(k+1)}\right)$.

Step-Two: Maximize the Laplace approximation $\ell_{LA}\left(\Theta|\ddot{X}_n,\, \hat{\boldsymbol{b}}_{n-1}^{(k+1)}\right)$ with respect to Θ and update $\hat{\Theta}^k$ by the maximizer $\hat{\Theta}^{(k+1)}$.

The LA algorithm alternates between the minimizing step and the maximizing step until certain convergence criterion is met. In comparison to directly maximizing $\ell\left(\Theta|\ddot{X}_n\right)$, the approximating procedure is much more computationally efficient and quite accurate.

5 NUMERICAL EXAMPLES

The following data-generating method is to generate degradation data from the non-stationary Wiener process. For later reference, we term it as the DG method. At time $\triangle$, we randomly generate an observation of the degradation, denoted by x_1, from the normal distribution $N(\lambda \triangle^\theta, \sigma^2 \triangle)$.

- If a maintenance action is then performed at time $\triangle$, the value of the scale parameter changes, immediately after the maintenance, from λ to $b_1\lambda$. b_1 is a realization of b, which is randomly generated from the normal distribution $N(\mu, Q)$. At time 2 $\triangle$, we randomly generate a degradation increment, denoted by y_2, from the normal distribution $N\left(b_1\lambda(2^\theta - 1) \triangle^\theta, \sigma^2 \triangle\right)$.
- If no maintenance action is performed at time $\triangle$, at time 2 $\triangle$ we randomly generate a degradation increment, y_2, from the normal distribution $N(\lambda(2^\theta - 1) \triangle^\theta, \sigma^2 \triangle)$.

The degradation measurement obtained at time 2 $\triangle$ is denoted by x_2: $x_2 = x_1 + y_2$. By analogy, at time $i \triangle$ $(i = 2, 3, 4, ...)$, we randomly generate a degradation increment y_i from the normal distribution $N(\prod_{j=1}^{i-1} b_j^{m_j} \lambda[i^\theta - (i-1)^\theta] \triangle^\theta, \sigma^2 \triangle)$. The ith degradation measurement is denoted by x_i: $x_i = x_{i-1} + y_i$. $\{b_j, j = 1, 2, 3, ...\}$ are all randomly generated from the normal distribution $N(\mu, Q)$. $\{m_j, j = 1, 2, 3, ...\}$ are the maintenance indicators defined in Section 2.

5.1 Simulation Study

Maintenance decisions are made based on the current estimate of the scale parameter. Therefore, before implementing the maintenance policy, we make an investigation into the performance of the proposed recursive filter. Simulated degradation data, generated by the DG method, are used to serve the purpose. The degradation parameters are assigned with values $\lambda = 7\text{e-}5$, $\theta = 2.5$ and $\sigma = 2$. The maintenance parameters are assigned with values $\mu = 0.7$ and $Q = 0.1$. The length of the DM interval is 100 units of time, i.e. $\triangle = 100$. The maintenance indicators $\{m_j, j = 1, 2, 3, ...\}$ are randomly generated from the Bernoulli distribution with success probability 0.5. Hence, the probability of taking a maintenance action is 0.5. The starting values for the recursive filter are set to be $\lambda_0 = 1$ and $P_0 = 1$. In comparison to the true value, $\lambda = 7\text{e-}5$, the starting value, λ_0, is assigned with a very big deviation. The prior estimate error variance, P_0, is also assigned with a big value.

Run the recursive filter on simulated degradation data to produce estimates of the scale parameter. At each DM point $i \triangle$ $(i = 1, 2, 3, ...)$, we calculate the value of the deviation $\hat{\lambda}_i - \lambda_i$. To inquire into the performance of the Kalman filter, we plot the true values and estimates of the scale parameter over time in Figure 2. Though the initial estimate of the scale parameter is far away from the true value, the estimate converges to the true value very quickly. After obtaining 10 degradation measurements, the filter is able to produce a fairly accurate estimate. The evolution of the deviation $\{\hat{\lambda}_i - \lambda_i, i = 1, 2, 3...\}$ is depicted in Figure 3 (left). The evolution of the a posterior estimate error variance $\{P_i, i = 1, 2, 3...\}$ is depicted in Figure 3 (right). It is observed that, with the DM number increasing, the deviation and the estimate error variance converge to zero rapidly, showing the competence and efficiency of the proposed algorithm. The filter can return a very accurate estimate even when the data size is small. Moreover, to show the robustness of the algorithm, we run the DG method for 100 times to generate 100

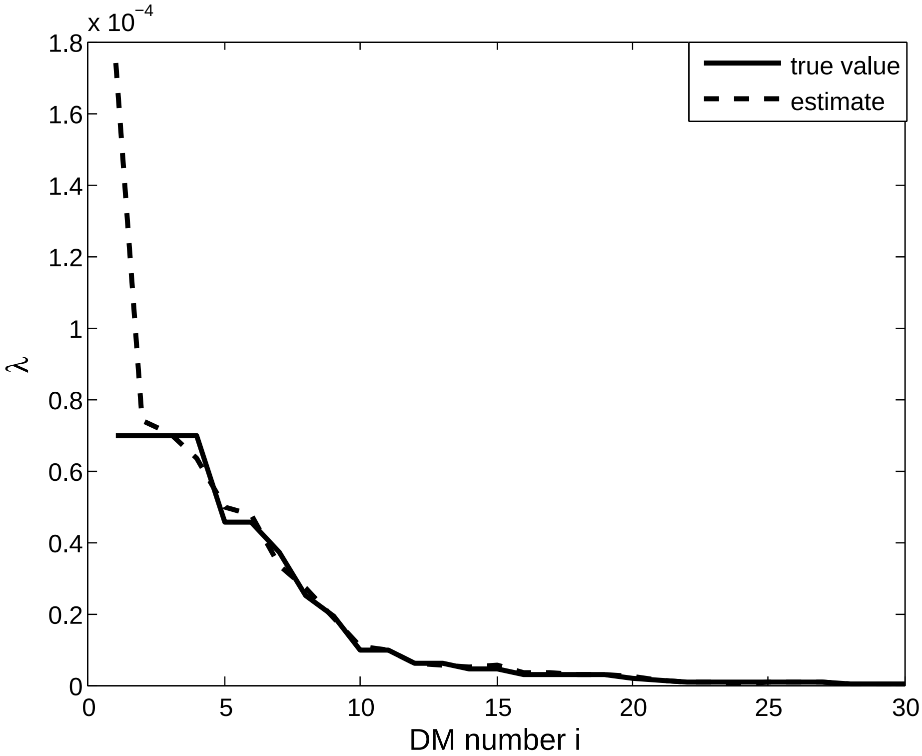

Figure 2: True values and estimates of the scale parameter.

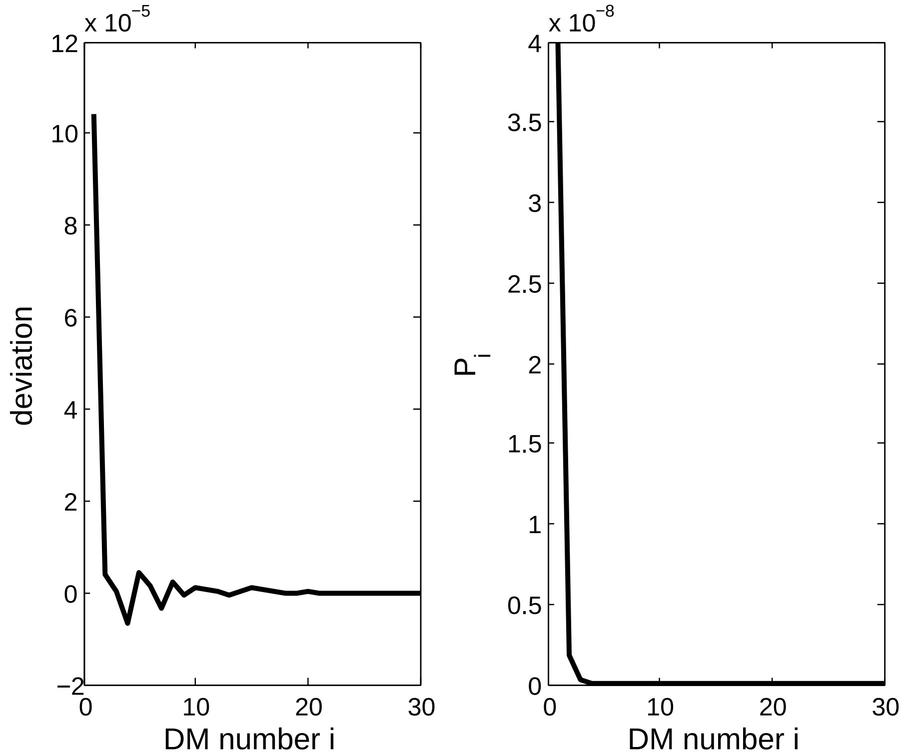

Figure 3: Evolving paths of $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and $\{P_i,\ i = 1,2,3...\}$.

different degradation data sets. Run the recursive filter on the data sets respectively. The 100 evolving paths of the deviation $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and 100 evolving paths of the estimate error variance $\{P_i,\ i = 1,2,3...\}$ are plotted in Figure 4. In Figure 4, all the deviation paths

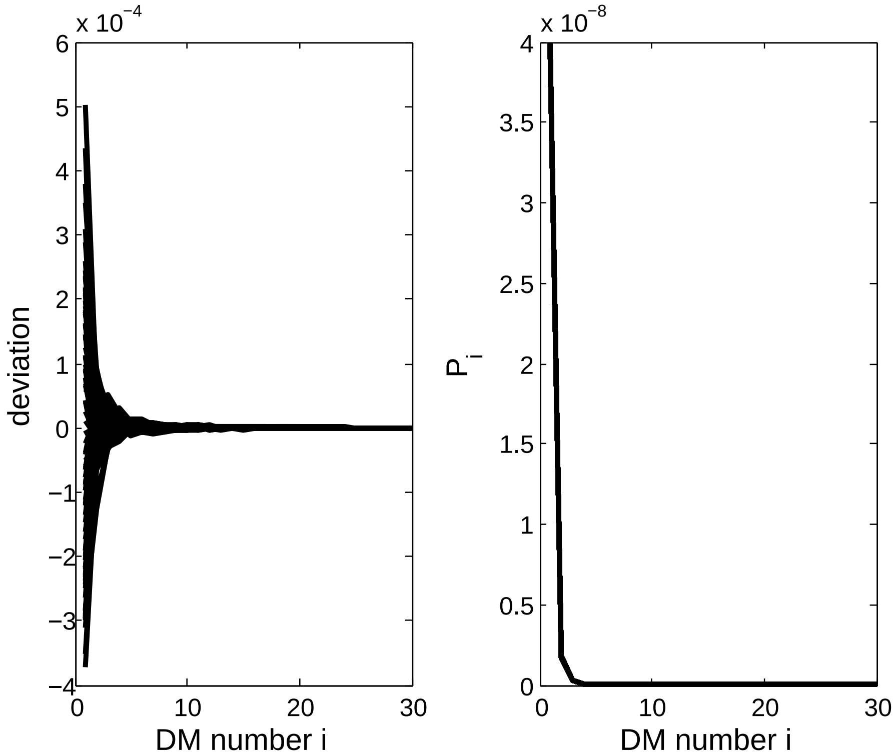

Figure 4: Evolving paths of $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and $\{P_i,\ i = 1,2,3...\}$ with 100 samples.

and variance paths converge rapidly to zero, showing the robustness of the stochastic filtering in assessing the scale parameter.

A sensitivity study is conducted to investigate the impact of the number of maintenance actions on the performance of the Kalman filter. Intuitively, with the maintenance actions being more intensive, the scale parameter should be more difficult to assess. With all the other parameters being the same, we gradually increase the success probability in the Bernoulli distribution from 0.05 to 0.95 with step size 0.05. For each value of success probability, use the DG method to generate a degradation data set and run the recursive filter on the simulated degradation data. The 19 evolving paths of the deviation $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and 19 evolving paths of the estimate error variance $\{P_i,\ i = 1,2,3...\}$ are plotted in Figure 5. In Figure 5, both the deviation and the estimate error variance convergence rapidly to zero. In each of the 19 cases, the Kalman filter can return an accurate estimate of the scale parameter requiring only about 15 degradation measurements. Therefore, the performance of the proposed Kalman filter is insensitive to the frequency of maintenance.

Another potential factor affecting the performance of the Kalman filter is the variance of the degradation-rate-reduction factor, i.e., Q. With the variance being larger, the estimate returned by the filter may be less accurate. Or, with the variance being larger, it may take a longer time for the estimate to be close enough to the true value. Hence, a sensitivity study is conducted to investigate the impact of the degradation parameter Q on the performance of the Kalman filter. Because the degradation-rate-reduction factor is assumed to be normally distributed, the

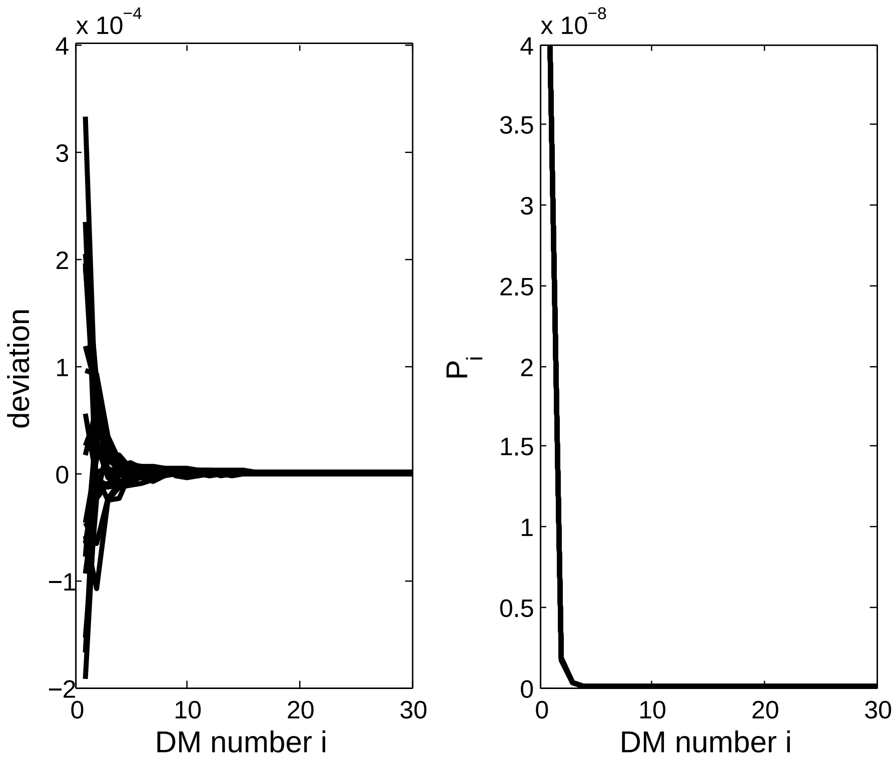

Figure 5: The impact of the number of maintenance actions on the performance of the Kalman filter.

realization of b may be negative or larger than one, which is inconsistent with the fact. If a realization of b is larger than one, then after the maintenance the system deteriorates faster. If a realization of b is negative, then after the maintenance the degradation decreases over time. To make the realization constrained in the interval $(0,1]$, the variance, Q, cannot be too large. Future work can be done to develop a suitable distribution for b with domain $(0,1]$. The topic is beyond the scope of this paper. From the mathematical point of view, for illustrative purpose, we gradually increase the variance in the normal distribution from 0.01 to 1 with step size 0.01. The starting values for the filter are $\lambda_0 = 1$ and $P_0 = 10$. All the other parameters remain the same. The 100 evolving paths of the deviation $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and 100 evolving paths of the estimate error variance $\{P_i,\ i = 1,2,3...\}$ are plotted in Figure 6. Figure 6 shows that the estimate converges rapidly to the true value even if Q and P_0 are relatively large. The convergent sequences, $\{\hat{\lambda}_i - \lambda_i,\ i = 1,2,3...\}$ and $\{P_i,\ i = 1,2,3...\}$, verify the robustness of the Kalman filter. Therefore, we say that the Kalman filter is insensitive to the starting values, the number of maintenance actions and the variance of the degradation-rate-reduction factor. As long as the four parameters, θ, σ, μ and Q, are accurately assessed, the Kalman filter is able to produce an accurate estimate on the scale parameter with small number of observations.

5.2 Real Data Analysis

For inertial navigation systems, the drift of a gyro is served as a performance indicator in evaluating the physical condition of an inertial platform; see Woodman (2007) for a general description of inertial platforms and gyros. Si et al. (2012) provided a set of drift degradation measure-

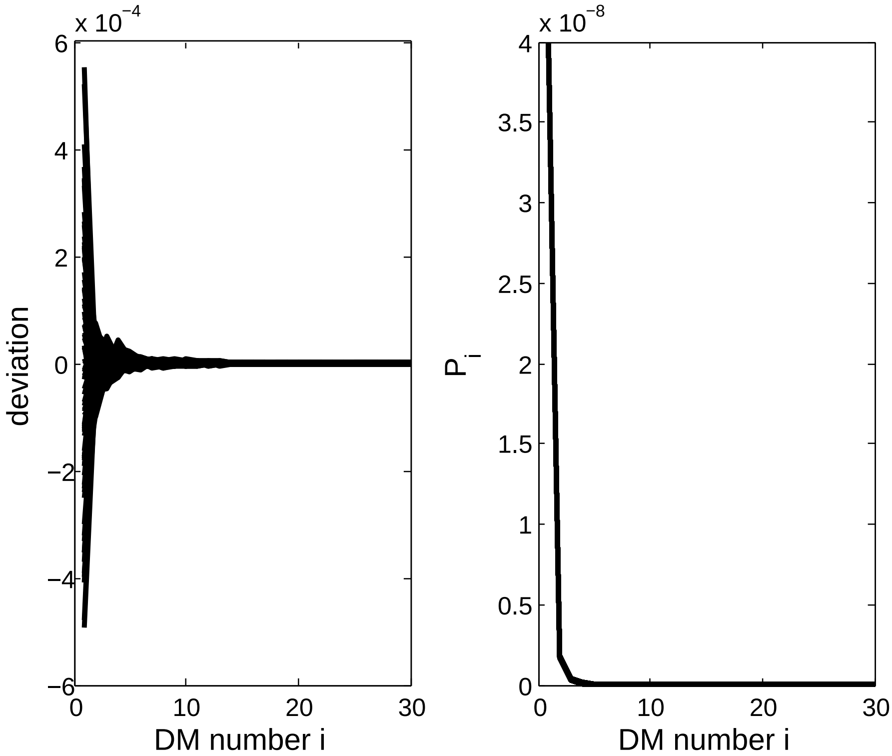

Figure 6: The impact of the variance Q on the performance of the Kalman filter.

ments (along the sense axis). The experimental degradation data from 5 monitored gyros are shown in Table 1. In the experiment, the drift degradation of each gyro was monitored every

Table 1: Drift degradation data from an inertial platform.

	gyroscopic drift ($^\circ/h$)				
time (h)	item 1	item 2	item 3	item 4	item 5
2.5	0.0507	0.1300	0.0135	0.0928	0.1052
5.0	0.1789	0.2554	0.0309	0.2633	0.1996
7.5	0.2059	0.3153	0.0077	0.3010	0.2927
10.0	0.2548	0.3966	0.1510	0.3515	0.3060
12.5	0.3117	0.2424	0.1162	0.2818	0.2860
15.0	0.5236	0.3473	0.0237	0.4375	0.2918
17.5	**0.6114**	0.4994	0.2681	0.3357	0.3540
20.0	**0.6616**	**0.7105**	0.4934	0.4413	0.4326
22.5	**1.8166**	**2.6399**	**0.6513**	**0.8157**	0.5042

2.5 hours. The experiment was terminated after 22.5 hours. It is required that the drift along the sense axis should not exceed 0.6°/hour, beyond which the mission of the inertial platform is no longer fulfilled. The drift degradation paths of the 5 monitored gyros are plotted in Figure 7. It is clearly observed that the degradation paths are non-monotone and non-linear. Therefore, the non-stationary Wiener process is suitable to fit the data set. See Si et al. (2012) for the justification of fitting the drift degradation data by the non-stationary Wiener process with drift function $v(t) = \lambda t^{\theta}$. The maximum likelihood estimates of the degradation parameters are

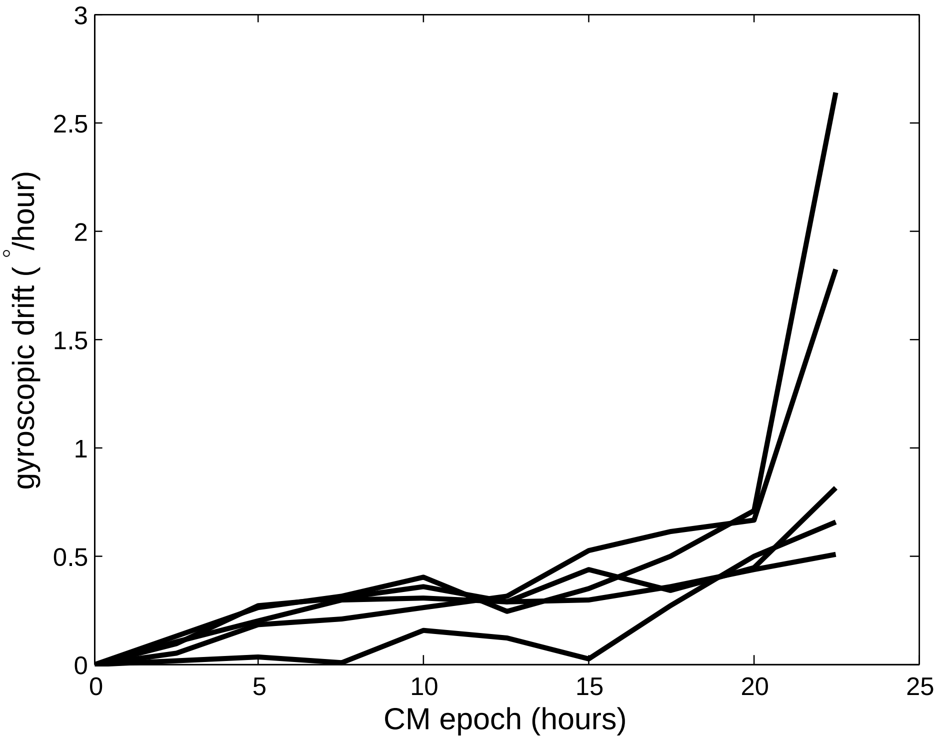

Figure 7: Drift degradation paths of gyros.

obtained as $\hat{\lambda} = 1.3082\text{e-}18$, $\hat{\theta} = 13.2024$ and $\hat{\sigma} = 0.1604$. The maximized log-likelihood is -2.1184. If the drift function takes the exponential form: $v(t) = \lambda \exp(\theta t)$. The corresponding maximum likelihood estimates of the degradation parameters are $\hat{\lambda} = 1.5592\text{e-}6$, $\hat{\theta} = 0.5920$ and $\hat{\sigma} = 0.1602$. The maximized log-likelihood is -2.0684. Hence, in terms of the fitness, the power model, $v(t) = \lambda t^{\theta}$, is practically equal to the exponential model, $v(t) = \lambda \exp(\theta t)$.

We then implement the proposed maintenance policy on the inertial platforms to illustrate the procedure for determining the optimal maintenance actions. At each DM point, having obtained the drift degradation measurement, engineers have to decide whether or not to preventively maintain the target inertial platform. Preventive maintenance on the inertial platform is assumed to be imperfect, e.g., replacing one of the gyros. Unexpected failure of the inertial platform is hazardous and costly. The inertial platform itself is also very expensive. A high-end marine-grade inertial navigation system may cost over 1 million dollars. In the present paper we consider an industrial-grade inertial navigation system, and the cost configurations are specified as follows. The cost of replacing an inertial platform is $c_r = \$2000$. The cost of an unexpected failure of the inertial platform is $c_f = \$10000$. The cost of each inspection is $c_I = \$20$. A PM action on the inertial platform costs $c_p = \$100$. Here, c_p need not be constant during the whole replacement cycle, but can vary according to practical conditions.

Because there was no maintenance in the experiment, we use the DG method to simulate degradation measurements. The degradation parameters assume the values estimated: $\lambda = 1.3082\text{e-}18$, $\theta = 13.2024$ and $\sigma = 0.1604$. To be consistent, the maintenance parameters assume the same values as what have been assigned in Section 5.1: $\mu = 0.7$ and $Q = 0.1$. The starting value, λ_0, is assigned with the estimate obtained based on the drift degradation measurements of the first four gyros: $\lambda_0 = 1.3431\text{e-}18$. Estimate the scale parameter based on the drift degradation measurements of one individual gyro. Calculate the variance of the four estimates which are obtained based on the first four gyros, respectively. The starting value, P_0, is assigned with the variance of the four estimates: $P_0 = 3.5995\text{e-}11$. Once a new degradation measurement is obtained, we immediately update the a posterior estimate of the scale parameter and the a posterior estimate error variance. Under the preceding parameter settings and cost configurations, one realization of the replacement cycle is described in Table 2.

Table 2: An optimal maintenance scheme implemented on the inertial platform.

i	1	2	3	4
$EC^i_{pm}(t^*_p)$	3.7554e+03	2.2660e+03	0.2343e+03	1.0903e+03
$EC^i_{na}(t^*_n)$	3.7879e+03	2.2745e+03	0.2288e+03	1.0973e+03
m_i	1	1	0	1
t^*_p	3.0116	3.0740	2.7820	2.6435
t^*_n	4.0323	2.8474	2.5591	2.2610
λ_i	1.3082e-18	1.0685e-18	7.1367e-19	7.1367e-19
$\hat{\lambda}_i$	2.1445e-08	2.1887e-10	1.5639e-14	6.8281e-15
i	5	6	7	8
$EC^i_{pm}(t^*_p)$	0.8842e+03	0.7564e+03	0.1309e+03	0.1274e+03
$EC^i_{na}(t^*_n)$	0.8898e+03	0.7608e+03	0.1284e+03	0.1237e+03
m_i	1	1	0	0
t^*_p	3.2007	3.1802	3.6786	1.7637
t^*_n	2.7598	2.6083	3.2651	1.5856
λ_i	5.2930e-19	2.9720e-19	1.3367e-19	1.3367e-19
$\hat{\lambda}_i$	4.3790e-16	7.0723e-17	2.1252e-18	7.7187e-19

According to Table 2, the decision-making process involves the following steps. At time $\triangle$, since the minimized cost rate $EC^1_{pm}(t^*_p)$ is smaller than its counterpart $EC^1_{na}(t^*_n)$, the system is preventively maintained at time $t = \triangle$. Since the optimal preventive-replacement time t^*_p is larger than the DM interval, the drift degradation is monitored at the second DM point. Likewise, at the ensuing DM point $2\triangle$, the optimal maintenance decision is to measure the degradation at time $3\triangle$ with a maintenance action taken at time $2\triangle$. At time $3\triangle$, since the minimized cost rate $EC^3_{pm}(t^*_p)$ is larger than its counterpart $EC^3_{na}(t^*_n)$, the system is left intact and the degradation is measured at time $4\triangle$. At time $8\triangle$, the optimal maintenance decision is to leave the system intact. Since the optimal preventive-replacement time t^*_n is smaller than the DM interval, the system is preventively replaced at time $t = 8\triangle + 1.5856$, if it survives to this point. Table 2 also lists the true value and estimate of the scale parameter at each DM point. Because the initial estimate error variance, P_0, is assigned with a large value, the first few estimates, e.g. $\{\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4\}$, have a large error. However, after obtaining 8 measurements, the filter is able to return a good estimate.

5.3 Parameter Estimation

This section gives illustrative examples on estimating the model parameters λ, θ, σ, μ and Q. Simulated degradation data are used to serve the purpose. The true values of the model parameters are set to be $\lambda = 1$, $\theta = 2$, $\sigma = 1$ and $\mu = 1$. The degradation is measured every one unit of time, i.e. $\triangle = 1$. To show the robustness of the Laplace's method, we vary the data size in $\{15,\ 30,\ 50\}$ and the variance of the degradation-rate-reduction factor in $\{0.01,\ 009,\ 0.25\}$. The number 15 corresponds to a small sample; the number 30 corresponds to a moderate sample; the number 50 corresponds to a large sample. The value 0.01 corresponds to a small variance; the value 0.09 corresponds to a moderate variance; the value 0.25 corresponds to a large variance. The approximating maximum likelihood estimates are tabulated in Table 3.

Table 3: Approximating maximum likelihood estimates of the model parameters.

	$Q=0.01$			$Q=0.09$			$Q=0.25$		
	$n=15$	$n=30$	$n=50$	$n=15$	$n=30$	$n=50$	$n=15$	$n=30$	$n=50$
$\hat{\lambda}$	1.0304	1.0170	1.0009	1.0848	1.0268	1.0001	1.0231	1.0059	1.0001
$\hat{\theta}$	1.9963	2.0018	1.9999	2.0296	2.0087	2.0000	2.0504	2.0009	2.0000
$\hat{\sigma}$	0.9412	1.0329	1.0187	1.0989	0.9733	1.0104	1.0449	1.0209	1.0059
$\hat{\mu}$	1.0569	1.0278	0.9940	1.1189	1.0494	1.0228	1.1318	1.0844	1.0267
$\hat{Q}$	0.0167	0.0105	0.0102	0.1075	0.0942	0.0904	0.2725	0.2631	0.2504

Table 3 shows that the general likelihood function can be well approximated by the Laplace approximation. When the sample size is moderate or large, approximating maximum likelihood estimates concentrate near the true values even when Q is large. When the sample size is small, the LA algorithm can still return useful estimates. For a fixed sample size, the variance of the degradation-rate-reduction factor is of no impact on the performance of the LA algorithm. Note that a crucial step in the LA algorithm is the calculation of the minimizer

$$\hat{\boldsymbol{b}}_{n-1} = \arg\min_{\boldsymbol{b}} gr(\Theta,\ \boldsymbol{b},\ \ddot{X}_n).$$

If $\hat{\boldsymbol{b}}_{n-1}$ is close to $\boldsymbol{b}_{n-1}$, then the Laplace approximation is close to the general likelihood function. To take a deeper look into the evolution of the LA algorithm, after obtaining a set of approximating maximum likelihood estimates $\hat{\Theta} = \{\hat{\lambda},\ \hat{\theta},\ \hat{\sigma},\ \hat{\mu},\ \hat{Q}\}$, we calculate the minimizer of $gr(\hat{\Theta},\ \boldsymbol{b},\ \ddot{X}_n)$. For $Q = 0.09$ and $n = 50$, Figure 8 plots the simulated 49 realizations of the degradation-rate-reduction factor, i.e. $\boldsymbol{b}_{n-1}$, and the minimizer of $gr(\hat{\Theta},\ \boldsymbol{b},\ \ddot{X}_n)$. The solid line threads the realizations, and the dashed line threads the estimates. Evidently, the minimizer $\hat{\boldsymbol{b}}_{n-1}$ is close to $\boldsymbol{b}_{n-1}$. Hence, Figure 8 explains to some extent the reason for the good performance of the LA algorithm.

6 CONCLUSIONS

In this paper, we have proposed and studied a degradation-based maintenance policy in which maintenance actions are assumed to be imperfect. There are three main contributions of this paper. The first contribution is the modeling of imperfect maintenance via the degradation rate function instead of the hazard rate function. The method is more suitable for degradation-based maintenance. The second is that the improving factors are considered as random variables and a recursive filter for the dynamical estimation of the improving factors is developed. The third is the development of an approximation to the likelihood to avoid multiple integral. Simulation studies and sensitivity analysis have verified the feasibility, robustness and competence of the proposed methods. The models and numerical study presented here have left a number of practical issues untouched. Many of these can be dealt with by appropriate technical extensions.

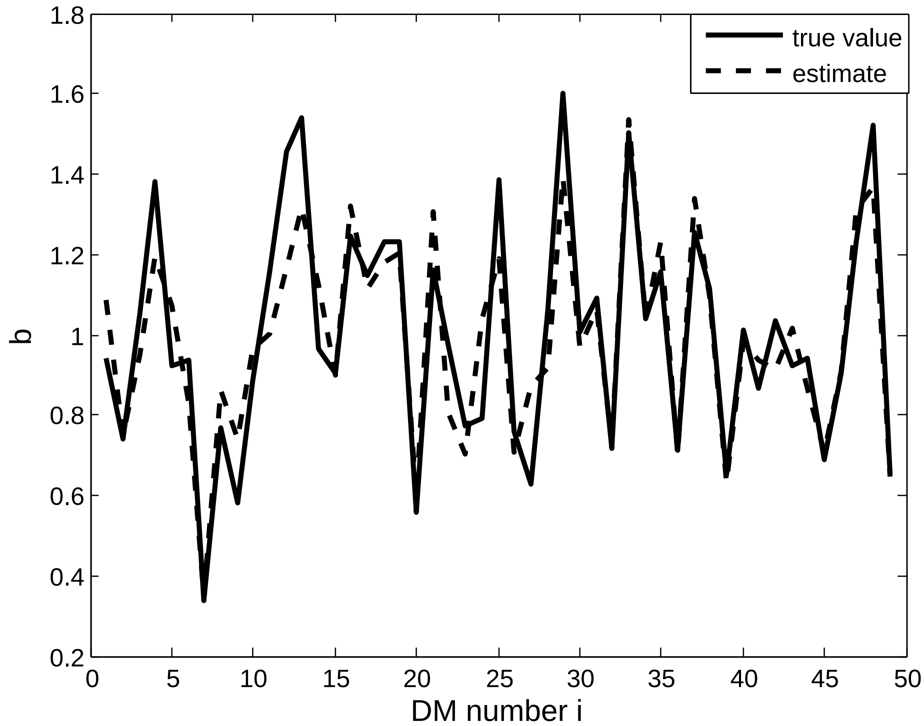

Figure 8: Simulated realizations of the degradation-rate-reduction factor and the corresponding estimates.

- As mentioned before, the extended improvement-factor method involves the degradation-rate-reduction factor and the shifting factor, yet not the age-reduction factor. Future work can be done to develop a model with three improving factors.
- The choice of a distribution for the degradation-rate-reduction factor, b, is of great importance. Therefore, it is of interest to develop a suitable distribution with domain $(0,1]$ for b. An alternative approach, when the variance of b is small, is to assume that b is a constant.
- The non-stationary Wiener process represents a family of continuous yet non-monotone diffusion processes. Other stochastic processes, such as non-stationary Gamma processes and non-stationary inverse Gaussian processes, can be used to characterize monotonically drifted diffusion processes. Developing stochastic filters which are built on monotonically drifted diffusion processes is an issue of great importance.

APPENDIX A Derivation of the Process Equation and the Measurement Equation

Note that the value of the scale parameter at time $(i-1)\,\triangle$ (right after the maintenance, if any) is identical to the value of the scale parameter at time $i\,\triangle$ (before any maintenance action). The degradation rate function over the ith interval is

$$v_i^{'}(t) = \lambda_i \theta t^{\theta-1}, \qquad (i-1)\,\triangle < t < i\,\triangle\,.$$

The drift function over the ith interval is

$$v_i(t) \cong \lambda_i t^{\theta}, \qquad (i-1) \triangle < t < i \triangle .$$

The symbol $\cong$ denotes equality up to an additive constant. With the assumption that the degradation conforms to the non-stationary Wiener process, we have

$$\begin{aligned} X_{i\triangle} - X_{(i-1)\triangle} &= \lambda_i (i \triangle)^{\theta} - \lambda_i [(i-1) \triangle]^{\theta} + \sigma W_{\triangle} \\ &= \rho_i \lambda_i + \sigma W_{\triangle}, \end{aligned}$$

with the constant being eliminated and ρ_i defined by

$$\rho_i = (i \triangle)^{\theta} - [(i-1) \triangle]^{\theta}.$$

Hence, we have the measurement equation

$$y_i = \rho_i \lambda_i + \omega_i,$$

in which ω_i is a realization of $\sigma W_{\triangle}$ and is termed as the measurement noise.

With the operational condition varying wildly, each maintenance action should have a different degree of impact on the degradation process. To characterize the uncertainty, we assume that the degradation-rate-reduction factor, b, is normally distributed with mean μ and variance Q. μ and Q are two unknown model parameters that will be estimated in Section 4. If no maintenance action is taken at time $(i-1) \triangle$, we have $\lambda_i = \lambda_{i-1}$. If a maintenance action is taken at time $(i-1) \triangle$, then $\lambda_i = b_{i-1} \lambda_{i-1}$ is normally distributed with mean $\mu \lambda_{i-1}$ and variance $Q \lambda_{i-1}^2$. b_{i-1} is a realization of b. The relationship between λ_{i-1} and λ_i can be summarized into

$$\lambda_i = \mu_{i-1} \lambda_{i-1} + \delta_{i-1}.$$

If no maintenance action is taken at time $(i-1) \triangle$, then $\mu_{i-1} = 1$ and $\delta_{i-1} = 0$. If a maintenance action is taken at time $(i-1) \triangle$, then $\mu_{i-1} = \mu$ and δ_{i-1} is normally distributed with mean 0 and variance $Q \lambda_{i-1}^2$. δ_{i-1} is termed as the process noise. For ease of exposition, we might say δ_{i-1} is normally distributed with mean 0 and variance Q_{i-1}, in which Q_{i-1} is defined by $Q_{i-1} = Q \lambda_{i-1}^2 m_{i-1}$. Note that, with b_{i-1} being independent of λ_{i-1}, Q_{i-1} is independent of Q_{i-2}. The degradation-rate-reduction factor need not be normally distributed. The topic of non-normal degradation-rate-reduction factor is beyond the topic of this paper.

APPENDIX B Derivation of $gr''(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$

The first order partial derivative of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ with respect to b_k $(1 \le k \le n-1)$ is given by

$$\frac{\partial gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)}{\partial b_k} = \frac{1}{\sigma^2 \triangle} \times \sum_{i=k+1}^{n} \left(x_i - x_{i-1} - \prod_{j=1}^{i-1} b_j \lambda \rho_i \right) \left(-\frac{1}{b_k} \prod_{j=1}^{i-1} b_j \lambda \rho_i \right) + (b_k - \mu)/Q.$$

In general, no closed-form expression for the minimizer $\hat{\boldsymbol{b}}_{n-1}$ can be constructed.

The second order partial derivative of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ with respect to b_k $(1 \le k \le n-1)$ is given by

$$\frac{\partial^2 gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)}{\partial b_k \partial b_k} = \frac{1}{\sigma^2 \triangle} \times \sum_{i=k+1}^{n} \left(-\frac{1}{b_k} \prod_{j=1}^{i-1} b_j \lambda \rho_i \right)^2 + \frac{1}{Q}.$$

For $1 \le k < l \le n-1$, we have

$$\begin{aligned} \frac{\partial^2 gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)}{\partial b_k \partial b_l} &= \frac{1}{\sigma^2 \triangle} \times \sum_{i=l+1}^{n} \left(x_i - x_{i-1} - \prod_{j=1}^{i-1} b_j \lambda \rho_i \right) \left(-\frac{1}{b_k b_l} \prod_{j=1}^{i-1} b_j \lambda \rho_i \right) \\ &\quad + \frac{1}{\sigma^2 \triangle} \times \sum_{i=l+1}^{n} \left(-\frac{1}{b_l} \prod_{j=1}^{i-1} b_j \lambda \rho_i \right) \left(-\frac{1}{b_k} \prod_{j=1}^{i-1} b_j \lambda \rho_i \right). \end{aligned}$$

The Hessian matrix, $gr''(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$, is hence at hand.

Remark 1. *For $1 \le k < l \le n-1$, the mixed partial derivative of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$, i.e. $\partial^2 gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)/\partial b_k \partial b_l$, is composed of two parts. When $\boldsymbol{b}_{n-1} = \hat{\boldsymbol{b}}_{n-1}$, the first part is usually negligible compared to the second part. Hence, the mixed partial derivative of $gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)$ at point $\hat{\boldsymbol{b}}_{n-1}$ can be approximated by its second part:*

$$\left. \frac{\partial^2 gr(\Theta,\, \boldsymbol{b}_{n-1},\, \ddot{X}_n)}{\partial b_k \partial b_l} \right|_{\boldsymbol{b}_{n-1} = \hat{\boldsymbol{b}}_{n-1}} \approx \frac{1}{\sigma^2 \triangle} \times \sum_{i=l+1}^{n} \left(-\frac{1}{\hat{b}_l} \prod_{j=1}^{i-1} \hat{b}_j \lambda \rho_i \right) \left(-\frac{1}{\hat{b}_k} \prod_{j=1}^{i-1} \hat{b}_j \lambda \rho_i \right).$$

REFERENCES

L. Bian and N. Gebraeel. Computing and updating the first-passage time distribution for randomly evolving degradation signals. *IIE Transactions*, 44(11):974–987, 2012.

A. Buonocore, L. Caputo, E. Pirozzi, and L. M. Ricciardi. The first passage time problem for Gauss-diffusion processes: algorithmic approaches and applications to LIF neuronal model. *Methodology and Computing in Applied Probability*, 13(1):29–57, 2011.

M. J. Carr and W. Wang. An approximate algorithm for prognostic modelling using condition monitoring information. *European Journal of Operational Research*, 211(1):90–96, 2011.

T. Chen, J. Li, P. Jin, and G. Cai. Reusable rocket engine preventive maintenance scheduling using genetic algorithm. *Reliability Engineering & System Safety*, 114(0):52–60, 2013.

C. K. Chui and G. Chen. *Kalman filtering with real-time applications*. Springer, New York, 2009.

D. R. Cox and H. D. Miller. *The theory of stochastic processes*. Methuen and Company, London, 1965.

L. Doyen and O. Gaudoin. Modeling and assessment of aging and efficiency of corrective and planned preventive maintenance. *IEEE Transactions on Reliability*, 60(4):759–769, 2011.

D. G. Ece and M. Basaran. Condition monitoring of speed controlled induction motors using wavelet packets and discriminant analysis. *Expert Systems with Applications*, 38(7):8079–8086, 2011.

A. H. Elwany, N. Z. Gebraeel, and L. M. Maillart. Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research*, 59 (3):684–695, 2011.

N. Gebraeel, M. Lawley, R. Li, and J. Ryan. Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions*, 37(6):543–557, 2005.

S. Hudak, A. Saxena, R. Bucci, and R. Malcolm. Development of standard methods of testing and analyzing fatigue crack growth rate data. *Technical Report AFML-TR-78-40*, 1978. Westinghouse R & D Center, Westinghouse Electric Corporation, Pittsburgh, PA 15235.

J. Humpherys, P. Redd, and J. West. A fresh look at the kalman filter. *SIAM Review*, 54(4): 801–823, 2012.

S. Kluge, K. Reif, and M. Brokate. Stochastic stability of the extended kalman filter with intermittent observations. *IEEE Transactions on Automatic Control*, 55(2):514–518, 2010.

M. Kurt and J. P. Kharoufeh. Optimally maintaining a Markovian deteriorating system with limited imperfect repairs. *European Journal of Operational Research*, 205(2):368–380, 2010.

P. Lall, R. Lowe, and K. Goebel. Prognostics health management of electronic systems under mechanical shock and vibration using Kalman filter models and metrics. *IEEE Transactions on Industrial Electronics*, 59(11):4301–4314, 2012.

D. C. Li, C. W. Liu, and T. L. Chen. Recursive operation time maximization model for the maintenance of power generation equipment. *Computers & Operations Research*, 39(5): 1117–1121, 2012.

B. H. Lindqvist. On the statistical modeling and analysis of repairable systems. *Statistical Science*, 21(4):532–551, 2006.

W. Q. Meeker and L. A. Escobar. *Statistical Methods for Reliability Data*. New York: John Wiley & Sons, 1998.

S. Mercier and I. Castro. On the modelling of imperfect repairs for a continuously monitored gamma wear process through age reduction. *Journal of Applied Probability*, 50(4), To appear 2013.

R. Nicolai, J. Frenk, and R. Dekker. Modelling and optimizing imperfect maintenance of coatings on steel structures. *Structural Safety*, 31(3):234–244, 2009.

M. Nikulin, N. Limnios, N. Balakrishnan, W. Kahle, and C. Huber-Carol. *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance*. Statistics for Industry and Technology. Birkh*ä*user, 2010 edition, 2010.

C. Park and W. Padgett. Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. *Lifetime Data Analysis*, 11(4):511–527, 2005.

F. X. Perrin, C. Merlatti, E. Aragon, and A. Margaillan. Degradation study of polymer coating: Improvement in coating weatherability testing and coating failure prediction. *Progress in Organic Coatings*, 64(4):466–473, MAR 2009.

H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

J. C. Pinheiro and D. M. Bates. Approximations to the log-likelihood function in the nonlinear mixed effects models. *Journal of Computational and Graphical Statistics*, 4:12–35, 1995.

K. Rafiee, Q. Feng, and D. W. Coit. Reliability modeling for dependent competing failure processes with changing degradation rate. *IIE Transactions*, 2013. To appear.

A. Ray and S. Phoha. Stochastic modeling of fatigue crack damage for information-based maintenance. *Annals of Operations Research*, 91:191–204, 1999.

X. S. Si, W. Wang, C. H. Hu, D. H. Zhou, and M. Pecht. Remaining useful life estimation based on a nonlinear diffusion degradation process. *IEEE Transactions on Reliability*, 61(1): 50–67, 2012.

X. S. Si, W. Wang, C. H. Hu, and D. H. Zhou. Remaining useful life estimation-A review on the statistical data driven approaches. *European Journal of Operational Research*, 213(1): 1–14, 2011.

K. L. Son, M. Fouladirad, A. Barros, E. Levrat, and B. Iung. Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety*, 112(0):165–175, 2013.

X. Wang. Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis*, 101(2):340–351, 2010.

Y. Wang and W. Wang. Failure prognostic of systems with hidden degradation process. *Journal of Systems Engineering and Electronics*, 23(2):314–324, 2012.

Z. Q. Wang, C. H. Hu, W. Wang, and X. S. Si. An additive wiener process-based prognostic model for hybrid deteriorating systems. *IEEE Transactions on Reliability*, 63(1):208–222, 2014.

G. Welch and G. Bishop. An introduction to the kalman filter. *Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC*, 2006. Technical Report, TR 95-041.

O. Woodman. An introduction to inertial navigation. *University of Cambridge: Cambridge*, 2007. Technical Report UCAM-CL-TR-696, Computer Laboratory.

M. Zhang, M. Xie, and O. Gaudoin. A bivariate maintenance policy for multi-state repairable systems with monotone process. *IEEE Transactions on Reliability*, 62(4):876–886, 2013.

R. Zhou, N. Gebraeel, and N. Serban. Degradation modeling and monitoring of truncated degradation signals. *IIE Transactions*, 44(9, SI):793–803, 2012.

R. R. Zhou, N. Serban, N. Gebraeel, and H. G. Mller. A functional time warping approach to modeling and monitoring truncated degradation signals. *Technometrics*, 56(1):67–77, 2014.

Y. Zhou, Y. Sun, J. Mathew, R. Wolff, and L. Ma. Latent degradation indicators estimation and prediction: A Monte Carlo approach. *Mechanical Systems and Signal Processing*, 25 (1):222–236, 2011.