IIE Transactions

ISSN: 0740-817X (Print) 1545-8830 (Online) Journal homepage: https://www.tandfonline.com/loi/uiie20

Sensor-driven prognostic models for equipment replacement and spare parts inventory

Alaa H. Elwany & Nagi Z. Gebraeel

IIE Transactions (2008) **40**, 629–639

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170701730818

Sensor-driven prognostic models for equipment replacement and spare parts inventory

ALAA H. ELWANY and NAGI Z. GEBRAEEL*

H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, USA
E-mail: nagi.gebraeel@isye.gatech.edu

Received October 2006 and accepted July 2007

Accurate predictions of equipment failure times are necessary to improve replacement and spare parts inventory decisions. Most of the existing decision models focus on using population-specific reliability characteristics, such as failure time distributions, to develop decision-making strategies. Since these distributions are unaffected by the underlying physical degradation processes, they do not distinguish between the different degradation characteristics of individual components of the population. This results in less accurate failure predictability and hence less accurate replacement and inventory decisions. In this paper, we develop a sensor-driven decision model for component replacement and spare parts inventory. We integrate a degradation modeling framework for computing remaining life distributions using condition-based *in situ* sensor data with existing replacement and inventory decision models. This enables the dynamic updating of replacement and inventory decisions based on the physical condition of the equipment.

Keywords: Degradation modeling, replacement models, spare parts inventory, condition monitoring

1. Introduction

Unexpected system failures pose a significant problem in human safety and health care applications, service and manufacturing sectors, national infrastructure (nuclear power plants and civil structures), and national security (military operations). The main challenges associated with unexpected failures are related to characterizing the failure uncertainty and the stochastic nature of the degradation processes. An accurate failure time prediction and a reliability assessment are necessary if the appropriate maintenance resources (personnel, tools, spare parts, etc.) are to be assembled. Reliability distributions have been widely used in replacement modeling and spare parts inventory planning (Venkatesan, 1984; Armstrong and Atkins, 1996; Aronis *et al.*, 2005; Vaughan, 2005). A precise reliability assessment is crucial for making sound maintenance-related logistical decisions. For instance, deciding which component to replace and when, requires a careful balance between the cost associated with premature replacement and the cost of unexpected failure. Furthermore, the ordering time of spare parts and their stocking quantities need to be planned such that holding costs are kept to a minimum while avoiding stockouts. However, reliability functions represent the failure time characteristics of component populations, and are unaffected by the underlying physical failure process.

Thus, the characteristics of failure time distributions remain unchanged over a component's life cycle. Consequently, when reliability functions are used to support replacement and/or inventory models, the decision-making process remains unchanged and is unaffected by different degradation rates among the individual components—even components operating under the same environmental conditions. By excluding the degradation characteristics of the individual components from failure time distributions, we compromise the accuracy of failure predictions.

Recent advances in sensor technology now enable us to continuously monitor the health of operating components. These condition-based sensor signals are, typically, correlated with the severity of the underlying degradation process. For example, continued wear of mechanical components may cause larger clearances between various contacts and linkages, which results in increased vibration. By interpreting these rich sensor data streams, we can acquire a better understanding of the uncertainty and randomness of the physics of the failure. This provides a tremendous opportunity for improving failure predictability and improving the current reliability paradigm.

The goal of this paper is to develop sensor-driven prognostic models for supporting component replacement and spare parts inventory decision making. The paper presents a mathematical framework for integrating

*Corresponding author

degradation-based sensor data streams with high-level logistical decision models. This integration provides an effective "Sense and Respond" architecture for decision making. First, we develop a stochastic degradation modeling framework for predicting the remaining life distributions of partially degraded components. Unlike existing research, we present a sensor-driven updating methodology that combines general population-specific reliability characteristics with the degradation properties of individual components. To do this, we monitor individual components and use real-time sensor information to update the remaining life distributions in a Bayesian manner. Next, the updated remaining life distributions are integrated with replacement and spare parts inventory models in place of the traditional failure time distributions. Each time we monitor a component, the signals are used to update the remaining life distribution and, in turn, update any previous replacement or inventory decisions.

The paper is organized as follows: Section 2 surveys some of the related literature. Section 3 presents conventional replacement and spare parts inventory models developed by Armstrong and Atkins (1996). These models rely on reliability distributions to estimate optimal component replacements and spare parts ordering times. Section 4 presents a sensor-driven updated degradation modeling framework with two base cases, a linear and an exponential case. The integration of the sensor-driven degradation models with the conventional decision models is presented in Section 5. Section 6 presents a case study using real data from degrading rolling element bearings. Final conclusions are drawn in Section 7.

2. Literature review

Most of the existing literature on maintenance-related logistical decision models focuses on equipment replacement and spare parts inventory. The goals are to determine optimal replacement times for single- and multi-component systems, and to determine optimal inventory policies, ordering quantities, and ordering times, separately or simultaneously.

Previously reported studies can be divided into one of two classes. In the first class, replacement and inventory decisions do not rely on failure time distributions to determine the decision policies. Venkatesan (1984) used a Markov decision process model to make production and replacement decisions for a single-product single-equipment deteriorating system. Production and replacement decisions were based on the current amount of on-hand inventory and the state of degradation of the equipment, which was modeled as a finite-state Markov process. Chang *et al.* (2005) presented a stochastic inventory model to determine the ordering quantity and reordering point for spare parts. The demand distribution of the spare parts was assumed to be normal. Replacement decisions were mainly based on the current state of the equipment, whereas inventory decisions were based on given demand distributions (Huiskonen, 2001; Hartman, 2004). Montoro-Cazorla and Perez-Ocon (2006) considered a system with different failure modes. They modeled the states of the system using a Markov process and presented a procedure to determine optimal replacement rules for the system. Other works have based replacement and maintenance decisions on the current state of health of the equipment that is obtained by monitoring the equipment during operation; an approach known as Condition-Based Maintenance (CBM). Akturk and Gurel (2007) derived optimal CBM policies for equipment in unknown conditions by formulating the problem as a partially observed Markov decision process and solving it using Dynamic Programming. Lu *et al.* (2007) used Kalman filters to predict future deterioration levels in systems. These predictions were then used to determine cost-effective CBM policies based on failure probabilities, preventive and corrective maintenance costs, and the profit loss due to deterioration. Pedregal and Camero (2006) determined optimal preventive replacement policies for a turbine in a petrochemical plant based on the expected cost per unit time. System vibration levels were monitored, and the signal was modeled as a bivariate random walk with drift. Other works on CBM include Maillart (2006), Ghasemi *et al.* (2007), and Zhou *et al.* (2007).

The other class of studies relies on failure time distributions to make replacement and inventory decisions. Failure time distributions are primarily used to estimate the expected failure times and number of failures. Aronis *et al.* (2005) presented an approach to determine the parameter of an $(S, S-1)$ inventory policy for the spare parts of electronic equipment. The demand was considered to be generated by random failures of the components, which was assumed to follow an exponential distribution. They determined the demand distribution using a Bayesian approach based on the prior distribution and historical data on failure rates. The demand distribution was then used to determine the optimal inventory policy for the spare parts. In other works, such as Vaughan (2005), the demand was generated both from random failures and regular scheduled preventive maintenance. The proposed model was used to determine the optimal inventory policy for spare parts.

Replacement and spare parts inventory decisions are traditionally performed separately. Replacement models assume that spare parts are always available in stock and base the replacement decisions solely on failure distributions. On the other hand, many inventory models rely on hypothetical demand distributions or demand forecasts instead of replacement times. Other papers in the literature consider sequential or joint replacement inventory decisions. Armstrong and Atkins (1996) considered age replacement and inventory ordering decisions for a simple system with only one component subject to random failure. They proposed cycle-based replacement and inventory cost functions. The objective was to compute optimal replacement and inventory ordering times, which minimized the replacement and inventory cost functions. The authors proposed two alternative models. In the first model, the replacement and inventory decisions were made sequentially, whereas the second model considered jointly optimized replacement and inventory decisions. Cheung and Hausman (1997) formulated an analytical model for the joint implementation of preventive maintenance and spare parts safety stocks in unreliable environments. The objective was to minimize the total expected costs of the system that was subject to random failures. Aka *et al.* (1997) addressed the joint optimization of replacement and spare part inventory decisions for systems with parallel machines.

The majority of the literature on replacement and spare parts inventory decision making relies on lifetime/reliability distributions obtained from reliability testing or manufacturers' specifications. For example, Liao and Elsayed (2006) predicted the reliability of components' populations subject to stochastic stresses while in actual use. Accelerated degradation tests under several constant stress levels were used to make inferences about the in-service reliability. These distributions generally represent measures of the reliability characteristics of a population of components. Consequently, challenges arise due to the fact that different degradation rates among the individual components are typically circumvented. In other words, the component's state of degradation is neglected.

Components usually degrade during their service life. The degradation of identical components can differ drastically. Degradation processes are typically accompanied by specific physical phenomena that evolve over time, such as increased vibration, temperature changes, and increased crack propagation. Generally, such physical phenomena can be observed using sensor-driven condition monitoring technology. Many components exhibit characteristic patterns in their sensor signals known as degradation signals and these signals evolve with respect to the component's state of degradation (Nelson, 1990). If properly modeled, degradation signals can be used to predict a component's remaining life distribution. In degradation modeling, failure is defined in terms of the degradation signal reaching a predetermined failure threshold. Lu and Meeker (1993) developed a random coefficients growth model to estimate the failure time distribution based on degradation information from a population of components. Robinson and Crowder (2000) modeled fatigue crack growth in components subject to cyclic loading as a non-linear regression model with random coefficients. Two-stage least squares, maximum likelihood principles, and Bayesian approaches were used to estimate the model parameters. Doksum and Hoyland (1992) modeled the accumulated decay of cable insulation for units subject to accelerated testing under variable stress levels. Accumulated decay was modeled as a Wiener process with drift and diffusion. The associated failure time followed an inverse Gaussian distribution. Whitmore and Schenkelberg (1997) modeled the degradation of self-regulating heat cables subject to high-stress reliability testing. The degradation process was modeled as a Wiener process that assumes a constant rate of degradation. Chinnam (2002) modeled the degradation signals to estimate the reliability of high-speed drill bits using polynomial regression models. Bae *et al.* (2007) discussed additive and multiplicative degradation models to derive the lifetime distribution of degraded components. The authors pointed out that some distributional assumptions in the degradation models can affect the obtained lifetime distributions and suggested studying these distributions to ensure they reflected the assumptions of the experimenter. Park and Padgett (2006) developed stochastic degradation models that incorporate several accelerating variables. Exact likelihood functions for the degradation paths were determined and used to obtain lifetime distributions. Joseph and Yu (2006) used design of experiments to improve the reliability of products with degradation data being available as the response in the experiment. Kharoufeh and Cox (2005) presented a hybrid approach for estimating the residual life distribution of a single-unit system. The degradation rate of the system was assumed to be random and dependent on external factors in the operating environment. The evolution of these factors was modeled as a continuous-time Markov chain. Their hybrid approach included two models. In the first, real sensor data provided information on the degradation state of the system, and in the second, information on the cumulative degradation up to some point in time was provided.

We note that most degradation modeling research efforts focus either on population-specific degradation characteristics or component-specific degradation signals. Few research efforts target the integration of reliability information with condition-based sensor signals with the goal of predicting the remaining life. The benefit of such integration is that it provides a comprehensive assessment of the current degradation state of the component and its potential evolution. Gebraeel *et al.* (2005) used a random coefficients exponential model to predict the remaining life distributions of partially degraded components. The authors used real-time condition-based degradation signals from individual components to update and continuously revise the remaining life distributions of the individual components. The distributions were updated in a Bayesian manner.

Next, we discuss conventional renewal-theory-based replacement and inventory models developed by Armstrong and Atkins (1996). Our goal is to replace the fixed lifetime distributions used in the models with sensor-driven remaining life distributions that dynamically evolve according to the degradation states of the individual components. This enables the updating of replacement and inventory decisions using real-time condition-based signals. Our hypothesis is that improved predictability of the remaining life distributions using degradation models will result in making better replacement and inventory decision and minimizing their associated costs.

3. Spare parts replacement and inventory models

This section reviews the work by Armstrong and Atkins (1996) that represents the traditional approach for determining replacement and inventory policies. The authors consider a single-unit system with room to store only one spare part. The component is subject to random failure with a failure time density function $f(t)$ and cumulative density function (cdf) $F(t)$. Each time the component fails, the system incurs a failure cost. This cost is usually high and includes the costs of corrective maintenance, labor, lost production, etc. In the event that the component is replaced according to a planned schedule, the system incurs a planned replacement cost, which is the regular cost of preventive maintenance, labor, and cost of the part itself. Typically, the planned replacement cost is less than the cost incurred due to sudden failure. Whether planned or failure replacements occur, it is necessary to have a spare part available in stock in order to perform the replacement action. The system incurs a holding cost per unit time to store. If the spare part is unavailable at the required replacement time, the system incurs a shortage cost per unit time. The objective is to determine the optimal planned replacement time, t_r, and the optimal spare part ordering time, t_o, such that the total cost rate of the system is minimized. The lead time, L, is assumed to be fixed, thus, $t_o + L \leq t_r$. When the component is replaced, the system is restored to its initial, as good as new, condition. Planned and unplanned replacements are considered to be regeneration points of the system, thus the objective is to minimize the system's expected cost per unit time for each cycle, where a cycle is defined by the random time between two successive planned or unplanned replacements.

3.1. *Replacement policy*

Given the failure time distribution of a component, the objective of the replacement model is to find the optimum planned replacement time t_r^*. The optimal replacement time is the time that minimizes the expected costs of preventive replacement and failure replacement. The long-run average cost per cycle is expressed as

$$C_r = \frac{c_p \bar{F}(t_r) + c_f F(t_r)}{\int_0^{t_r} \bar{F}(t)\mathrm{d}t}, \qquad (1)$$

where C_r is the expected long-run replacement cost, c_p is the planned replacement cost, c_f is the failure replacement cost, and $\bar{F}(t) = 1 - F(t)$, where $F(t)$is the cdf of the component's failure time. The numerator represents the expected cost per cycle and the denominator represents the expected cycle length.

3.2. *Inventory ordering policy*

Armstrong and Atkins (1996) consider a sequential decision-making process where the optimal replacement time is first evaluated followed by the optimal ordering time. Once the optimal replacement time, t_r^*, has been computed, it is then used to decide when to order the spare part. Due to the assumption of a single-unit storage capacity, the order quantity is always a single unit. The optimal ordering time is the ordering time that minimizes the holding cost of spare parts and the cost of stockouts. The long-run average inventory cost per cycle is expressed as

$$C_o = \frac{k_s \int_{t_o}^{t_o+L} F(t)\mathrm{d}t + k_h \int_{t_o+L}^{t_r} \bar{F}(t)\mathrm{d}t}{\int_{t_o}^{t_o+L} F(t)\mathrm{d}t + \int_0^{t_r} \bar{F}(t)\mathrm{d}t}, \qquad (2)$$

where C_o is the expected long-run ordering cost, k_h is the holding cost per unit time, k_s is the shortage cost per unit time, and L is the fixed lead time elapsed from the moment of placing the order up till order receipt. We note that the expected cycle length is not the same for the replacement policy case due to the possibility of stockouts occurring and resulting in a longer cycle.

In this paper, we extend this traditional approach for determining replacement and inventory policies. Rather than using the failure time distribution of the components' population, we incorporate sensor-driven Remaining Life Distributions (RLDs) that are obtained using the degradation modeling framework illustrated in the next section. These updated RLDs capture the underlying state of degradation of the components using real-time sensor signals. The impact of doing this is twofold; first the increased accuracy of failure prediction results in more sound decision policies and less costs. Second, it allows for dynamically updating the decision policies based on the health of the component.

4. Degradation modeling framework

This section presents a stochastic degradation modeling framework to compute and update the remaining life distribution of partially degraded components. Our goal is to integrate these updated distributions with the replacement and inventory decision models presented earlier. To do this, we replace the lifetime distributions currently being used in the replacement and inventory models with the sensor-driven RLDs.

Our approach is to model the degradation signal as a continuous-time stochastic process $\mathbf{S} = \{S(t),\ t > 0\}$, where $S(t)$ denotes the value of the degradation signal at time t. This signal evolves according to some functional form that depends on the application. We consider two cases. In the first the signal evolves according to a linear functional form. In the second case, the signal evolves according to an exponential functional form. Both cases are discussed in detail in the following sections.

4.1. *Linear degradation model*

The linear degradation model is used in cases where the rate of degradation is not significantly affected by the cumulative damage. In this model the degradation signal is assumed to follow the linear form:

$$S(t_k) = \phi + \theta\ t_k + \varepsilon\ (t_k), \tag{3}$$

where $S(t_k)$ is the value of the signal at time t_k, ϕ is a constant deterministic parameter, θ is a random variable following a prior normal distribution $\pi(\theta)$ with mean μ_o and variance σ_o^2, and $\varepsilon\ (t_k)$ is the error term to model measurement noise and signal fluctuations and is assumed to follow a Brownian motion process. Modeling the error term as Brownian motion is suitable for applications where correlation exists between successive error fluctuations in sensor readings.

The prior distribution $\pi(\theta)$ is acquired from degradation testing. A Bayesian technique is then used to update this prior distribution using real-time sensor degradation signals. Given that we have observed a partial degradation signal, $S(t_1),\ S(t_2), \ldots,\ S(t_k)$, at times t_1, t_2, $\ldots, t_k$, the posterior distribution of the stochastic parameter, $\gamma(\theta|S(t_1), \ldots, S(t_k))$, is normal with the following parameters:

$$\mu_\theta\ =\ \frac{\mu_o\sigma^2 + (S(t_k) - \phi)\sigma_o^2}{t_k\sigma_o^2 + \sigma^2}, \tag{4}$$

$$\sigma_\theta^2\ =\ \frac{\sigma^2\sigma_o^2}{t_k\sigma_o^2 + \sigma^2}. \tag{5}$$

The posterior distribution of θ is then used to compute and update the RLD of the components as shown in the following section.

4.2. *Exponential degradation model*

The exponential degradation model is used in applications where cumulative damage significantly affects the rate of degradation. In this model, the degradation signal is assumed to follow the exponential form given by:

$$S(t_k) = \phi + \theta e^{\beta t_k + \varepsilon(t_k) - \frac{\sigma^2 t_k}{2}} = \phi + (\theta e^{\beta t_k})\left(e^{\varepsilon(t_k) - \frac{\sigma^2 t_k}{2}}\right) \tag{6}$$

where, $S(t_k)$ is the value of the signal at time t_k, ϕ is a constant deterministic parameter, θ and β are random variables, and $\varepsilon(t_k)$ is the error term that is assumed to follow a Brownian motion process.

For this exponential model, it is easier to work with the logarithm of the degradation signal denoted by $L(t_k)$:

$$L(t_k) = \theta' + \beta'\ t_k + \varepsilon(t_k), \tag{7}$$

where the parameters $\theta' = \ln(\theta)$ and $\beta' = \beta - (\sigma^2/2)$ are assumed to follow a bivariate normal distribution, $\pi(\theta', \beta')$ with means (μ_o, μ_1'), variances (σ_o^2, σ_1^2) and a correlation coefficient, ρ_o. Note that $\mu_1' = \mu_1 - (\sigma^2/2)$ where μ_1 is the mean of the random variable β.

Similar to the linear model, the prior distribution of the stochastic parameters is updated by computing the joint distribution of θ' and β' given the observed signals. The posterior distribution, $\gamma(\theta', \beta'|L(t_1), \ldots, L(t_k))$, is found to follow a bivariate normal distribution with the following parameters:

$$\mu_{\theta'} = \frac{(t_k + \mathrm{B}\sigma^2)((L_1/t_1) + C\sigma^2) - \left(\sum_{i=1}^k L_i + D\sigma^2\right)(1 + E\sigma^2)}{((1/t_1) + A\sigma^2)(t_k + B\sigma^2) - ((1/\sigma^2) + E(E\sigma^2 + 2))}, \tag{8}$$

$$\mu_{\beta'} = \frac{\left(\sum_{i=1}^k L_i + D\sigma^2\right)[(1/t_1 + A\sigma^2) - (1 + E\sigma^2)]}{((1/t_1) + A\sigma^2)\ (t_k + B\sigma^2) - ((1/\sigma^2) + E(E\sigma^2 + 2))}, \tag{9}$$

$$\sigma_{\theta'}^2 = \frac{\left[t_k\sigma_1^2\left(1 - \rho_o^2\right) + \sigma^2\right]\left(\sigma^2\sigma_o^2 t_1\right)}{\sigma^2\left(\sigma_o^2 + \sigma_1^2 t_k t_1 + \sigma^2 t_1 + 2\sigma_o\sigma_1^\rho \mathrm{o}_1^t\right)}, \tag{10}$$
$$-\left[\sigma_o^2\sigma_1^2\left(1 - \rho_o^2\right)(t_1 - t_k)\right]$$

$$\sigma_{\beta'}^2 = \frac{\sigma_1^2\sigma^2\left[\sigma_o^2\left(1 - \rho_o^2\right) + \sigma^2 t_1\right]}{\sigma^2\left(\sigma_o^2 + \sigma_1^2 t_k t_1 + \sigma^2 t_1 + 2\sigma_o\sigma_1\rho_o t_1\right)}, \tag{11}$$
$$-\ [\sigma_o^2\sigma_1^2(1 - \rho_o^2)(t_1 - t_k)]$$

$$\rho = \frac{\sigma^2\rho_1 - \sigma_0\sigma_1\left(1 - \rho_o^2\right)\sqrt{t_1}}{\sqrt{\left[\left(\sigma_0^2\left(1 - \rho_o^2\right) + \sigma^2 t_1\right)\left(t_k\sigma_1^2\left(1 - \rho_o^2\right) + \sigma^2\right)\right]}}, \tag{12}$$

where, L_i is the difference between the value of the logarithm of the signal at times t_i and t_{i-1}, $L_i = L(t_i) - L(t_{i-1})$, for $i\ =\ 1,\ 2,\ 3,\ .\ .$ with $L_1 = L(t_1)$, $A = 1/\sigma_o^2\ (1 - \rho_o^2)$, $B = 1/\sigma_1^2(1 - \rho_o^2)$, $C = (\mu_o/\sigma_o^2(1 - \rho_o^2)) - (\mu_1\rho_o/\sigma_o\sigma_1(1 - \rho_o^2))$, $E = -\rho_o/\sigma_o\sigma_1(1 - \rho_o^2)$, and $\mathrm{D} = (\mu_1/\sigma_1^2(1 - \rho_o^2)) - (\mu_o\rho_o/\ \sigma_o\sigma_1(1 - \rho_o^2))$.

In practice, prior information about the stochastic parameter(s) is obtained from testing a sample of components and analyzing their degradation signals. This can be achieved by fitting the functional form of the model to the degradation data and estimating the model parameters. The sample values of the model parameters are used to estimate the prior distribution.

4.3. *Computing RLDs*

After computing the posterior distributions of the stochastic parameters at each updating epoch, the RLD is obtained by evaluating the distribution of the predicted time at which the degradation signal hits a predetermined failure threshold. Real-time sensor signals are used to update the distribution of the stochastic parameters of the degradation model. The projected path of the degradation signal changes as these parameters are updated. The updated prognostic degradation model is then used to revise the RLD.

4.3.1. *Linear degradation model*

Given an observed partial degradation signal $S(t_1), S(t_2), \ldots, S(t_k)$, up to time t_k, we define the random variable $S(t_k + t)$ as the value of the degradation signal after t time units. $S(t_k + t)$ will be normally distributed with mean and variance:

$$\tilde{\mu}(t+t_k) = S(t_k) + E[\theta]\, t = \mu_\theta t + S(t_k), \tag{13}$$

$$\tilde{\sigma}^2(t+t_k) = t^2\, \mathrm{V}[\theta] + \mathrm{V}[\varepsilon(t+t_k) - \varepsilon(t_k)] = \sigma_\theta^2 t^2 + \sigma^2 t. \tag{14}$$

The RLD is evaluated by computing the distribution of the time until the degradation signals reaches the failure threshold, V. Let T be a random variable that denotes the remaining life of a partially degraded component given that we have observed $S(t_1),\ S(t_2), \ldots, S(t_k)$. Therefore, T satisfies $S(T+t_k) = V$ and its distribution is given by

$$\begin{aligned} P(T \le t | S_1, \ldots, S_k) &= P(S(t+t_k) \ge V | S_1, \ldots, S_k) \\ &= 1 - (P(S(t+t_k)) \le V | S_1, \ldots, S_k) \\ &= 1 - P\left(Z \le \frac{V - \tilde{\mu}(t+t_k)}{\tilde{\sigma}(t+t_k)}\right) \\ &= \Phi\left(\frac{\tilde{\mu}(t+t_k) - V}{\tilde{\sigma}(t+t_k)}\right), \end{aligned} \tag{15}$$

where $\Phi(\cdot)$ is the cdf of a standardized normal random variable Z.

4.3.2. *Exponential degradation model*

The RLD for the exponential model is computed in a similar fashion. Since we work with the logged values of the degradation signal, we define the random variable $L(t_k + t)$ as the logged value of the degradation signal after t time units. The mean and variance of the random variable $L(t_k + t)$ are given as (Gebraeel *et al.*, 2005):

$$\tilde{\mu}(t+t_k) = \sum_{i=1}^{k} L_i + \mu_{\beta'} t = L(t_k) + \mu_{\beta'} t, \tag{16}$$

$$\tilde{\sigma}^2 = \sigma_{\beta'}^2 t^2 + \sigma^2 t, \tag{17}$$

The RLD is expressed as follows:

$$P(T \le t | L_1, \ldots, L_k) = \Phi\left(\frac{\tilde{\mu}(t+t_k) - \ln(V)}{\tilde{\sigma}(t+t_k)}\right). \tag{18}$$

5. Sensor-driven replacement and inventory policies

This section discusses sensor-driven replacement and inventory models capable of revising the decision-making criteria based on the latest degradation states of the components being monitored. To do this, we replace the traditional failure time distributions in the spare parts logistics decision model presented by Armstrong and Atkins (1996) with sensor-driven RLDs obtained from the degradation modeling framework. This is done as a heuristic to approximate the replacement and inventory policies at updating epochs, even though these epochs do not represent regeneration points of the system.

The improved failure predictability obtained by using sensor-based degradation models results in improved decision making. In other words, precise information about a component's failure time results in an accurate optimal replacement time that maximizes the component's utilization without necessarily increasing the risk of unexpected failure. Gebraeel (2006) evaluated the accuracy of failure predictions using the exponential sensor-driven degradation model. It was shown that the average prediction error was around 8% compared to an average of 22% for conventional degradation models without the sensor-driven updating methodology.

Next we consider online sensor signals obtained at a specific updating time t_k. Each time a signal is acquired, the RLD of the degraded component is updated. The updated distribution is then used to compute the optimal replacement and the optimal spare part ordering times. The long-run average replacement and inventory costs can now be expressed in terms of the updated RLD as follows:

$$C_{\mathrm{r}}^k = \frac{c_{\mathrm{p}} \bar{F}^k(t) + c_{\mathrm{f}} F^k(t_{\mathrm{r}})}{\int_0^{t_{\mathrm{r}}^k} \bar{F}^k(t)\mathrm{d}t + t_k}, \tag{19}$$

$$C_{\mathrm{o}}^k = \frac{k_{\mathrm{s}} \int_{t_{\mathrm{o}}^j}^{t_{\mathrm{o}}^k + L} F^j(t)\mathrm{d}t + k_{\mathrm{h}} \int_{t_{\mathrm{o}}^j + L}^{t_{\mathrm{r}}^k} \bar{F}^k(t)\mathrm{d}t}{\int_{t_{\mathrm{o}}^k}^{t_{\mathrm{o}}^k + L} F^k(t)\mathrm{d}t + \int_0^{t_{\mathrm{r}}^k} \bar{F}^k(t)\mathrm{d}t + t_k}, \tag{20}$$

where C_{r}^k and C_{o}^k are the replacement and inventory ordering cost rates per cycle, respectively, at updating time t_k. $F^k(t)$ is the updated cdf of the residual life at the updating time t_k. In other words, given that the component has survived up to time t_k and that we have observed a partial degradation signal up to time t_k, $F^k(t)$ is the cumulative probability that the component fails after an additional t time units. The terms t_{r}^k and t_{o}^k are the optimal replacement and inventory ordering times, respectively, at the given updating epoch. Note that t_k, the updating time, has been added in the denominator to the cycle time. In other words, each cycle is now composed of two components, a fixed term given by the time up to which the component has survived and a random component given by the integral of the RLD.

6. Case study

In this section, we validate the proposed model using real-world data from a practical application. We acquire vibration signals from rolling element bearings using condition monitoring. Rolling element bearings have been chosen for this case study for several reasons. A rolling element bearing is a typical example of a component that undergoes degradation during operation. Degradation of rolling element bearings is typically accompanied by increased levels of vibration due to the formation of spalls on the surface of the raceways (Harris, 2001). The spalls also increase the friction levels between the rolling elements and bearing raceways resulting in an increased temperature. Although there are several physical phenomena that can be used to characterize bearing degradation, we focus our attention on bearing vibration. For the purpose of this paper, bearings are chosen as a source of real-world sensor-based degradation data. An experimental setup is used to perform accelerated degradation testing on several identical thrust ball bearings. The relatively low cost of bearings allows for high volume testing to ensure fidelity of the experimental results. Details of the experimental setup can be found in Gebraeel (2006). Note that this methodology is not specific to bearing applications and can be applied to a wide variety of other applications that show gradual degradation.

Bearing vibration data was continuously acquired during each degradation test. Bearing-specific information was extracted and processed into a vibration-based degradation signal. The degradation signal used in this research represents the evolution of the average amplitude of the defective frequency and its first six harmonics over time (Fig. 1). The evolution of the degradation signal is correlated with the underlying degradation state of the bearing. In other words, the degradation signal indirectly captures the physical transitions associated with the degradation process.

Bearing failure is defined by the bearing-specific degradation signal crossing a prespecified failure threshold. For the degradation signal developed in this paper, we define the failure threshold as the amplitude of the degradation signal corresponding to 2.2 Gs of overall vibration. After observing several degradation signals, the failure threshold of the degradation signals was identified as 0.03 $\mathrm{volt}_{\mathrm{rms}}$. Vibration spectra were acquired every 2 minutes. Each successive vibration spectrum represents a single point on the degradation signal. Next, we discuss how the degradation signals were used to compute RLDs and how these distributions are updated in real-time.

We also demonstrate how the sensor-driven RLDs are used to evaluate replacement and inventory policies for partially degraded bearings.

6.1. *Computing RLDs*

By analyzing several degradation signals, we observed that the degradation signal grows exponentially as the degradation progresses. Indeed, the exponential model has been widely used to model degradation processes where cumulative damage increases the rate of degradation, which is the case with bearing applications (Harris, 2001; Gebraeel *et al.*, 2005).

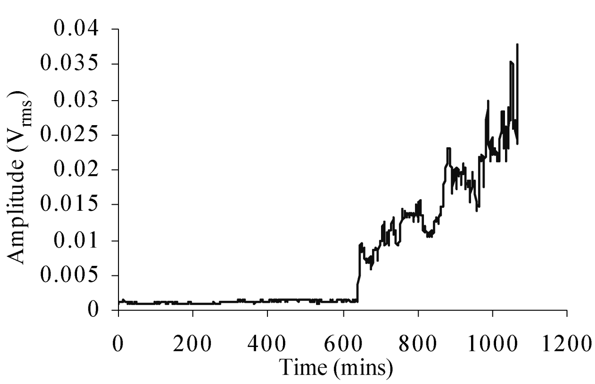

Fig. 1. Vibration-based degradation signal.

Upon identifying the exponential functional form as the most suited for bearings, degradation signals from 25 bearings were first acquired to verify model assumptions and compute the prior distributions of the stochastic model parameters, θ' and β'. The 25 degradation signals were each fitted with an exponential functional form. The deterministic parameter, ϕ, represented by the average value of the intercept of each signal with the Y-axis, was observed to be $\approx 0.002\ \mathrm{V}_{\mathrm{rms}}$. The value of ϕ was set to zero as a simplifying assumption (Gebraeel *et al.*, 2005; Gebraeel, 2006). Next, the natural logarithms of the signal values were evaluated and the resulting signal was fitted using Equation (7). Twenty-five values representing the stochastic parameters, θ' and β', were obtained and used to calculate the values of the parameters of the prior distribution $\pi(\theta', \beta')$. The computed values of these parameters are $\mu_0 = 6.031$, $\mu_1' = 0.008\,06$, $\sigma_{\mathrm{o}}^2 = 0.3464$, $\sigma_1^2 = 1.0347 \times 10^{-5}$, and $\rho_{\mathrm{o}} = -0.3625$. We note that θ' and β' are negatively correlated, justifying the assumption of their joint distribution.

The prior distribution $\pi(\theta', \beta')$ is specific to the population of bearings under consideration. We now consider an individual bearing other than the 25 bearings used for obtaining the prior distribution. Each time, t_k, a degradation signal from that bearing is observed, the values of the prior parameters are updated using Equations (8) to (12). The computed posterior parameters, $\mu_{\theta'}$, $\mu_{\beta'}$, $\sigma_{\theta'}^2$, $\sigma_{\beta'}^2$, and ρ, are used to update the projection of the degradation signal based on the recently acquired signal. The updated degradation model is then used to revise the RLD of the bearing. The updated RLD is computed by evaluating the expression $P(T \le t | L_1, \ldots, L_k)$ in Equation (18), given that the failure threshold $V = 0.03\ \mathrm{V}_{\mathrm{rms}}$ For this case study, updating was performed every 2 minutes because vibration signals can be acquired at no significant cost. This might not always be the case for industrial applications where data acquisition becomes costly.

Figures 2 and 3 show the updated probability density functions (pdfs) and cdfs, respectively, of the bearing at different degradation percentiles. The updated RLDs tend to be skewed as shown in Fig. 2. It is common practice in this case to use the median of the RLD to represent the time to failure (Ebeling, 2005).

6.2. *Replacement and inventory policies*

Each time the RLD is updated, the updated cdf, $F^k(t)$, is used to compute the optimal replacement time, t_{r}^k, at time t_k using Equation (19). The computed optimal replacement time is then incorporated into Equation (20) to compute the optimal inventory ordering time t_{o}^k. The following data was

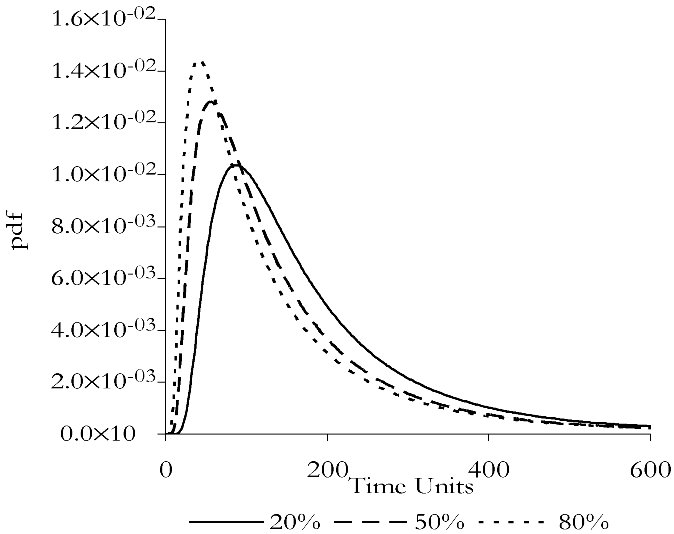

Fig. 2. Updated pdfs at different degradation percentiles.

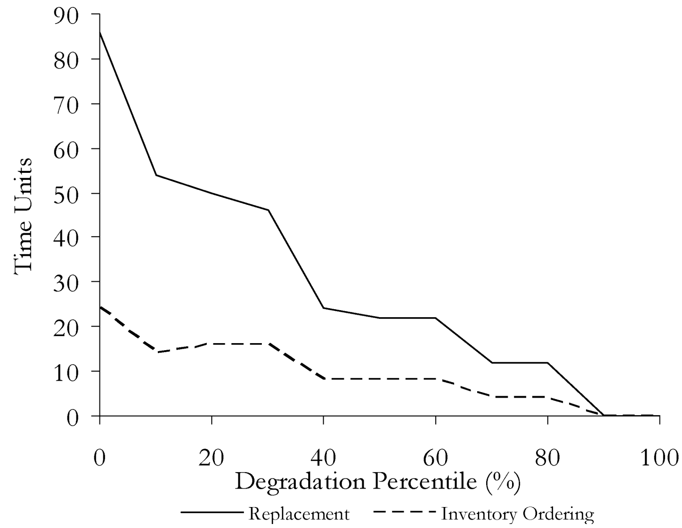

Fig. 4. Evolution of optimal replacement and spare part ordering times.

used: $c_p = \$\ 25$, $c_f = \$100$, $k_h = \$0.10$/unit time, $k_s = \$350$/unit time, and $L = 4$ time units.

Figure 4 shows the evolution of the computed optimal replacement and inventory ordering times at different degradation percentiles, and Table 1 displays numerical results.

We note from the results that the constraint imposed on the inventory ordering time, $t_o + L \leq t_r$, is satisfied for all degradation percentiles and Fig. 5 summarizes the proposed methodology. The first step is to observe the real-time degradation signal from the component during operation. Each time we acquire a signal, the RLD is updated using the degradation modeling framework, as shown in step 2. The type of degradation model to be used (functional form and assumptions concerning the model parameters and error terms) is chosen according to the application. At each updating epoch, the updated RLD is used to compute optimal replacement and inventory ordering times. Steps 3 and 4 show replacement and inventory cost rate curves, respectively, for three different degradation percentiles. When the replacement and inventory ordering policies are jointly optimized, we perform steps 3 and 4 simultaneously.

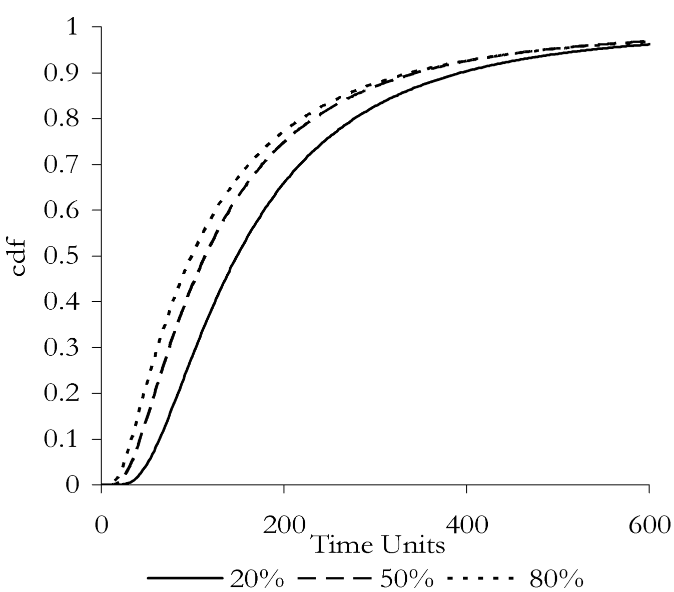

Fig. 3. Updated cdfs at different degradation percentiles.

In practical application, the process continues until some stopping rule is satisfied. In this case study our stopping rule is to stop updating once $t_r \geq t_f - L$, where t_f is the median of the RLDs, t_r is the optimal replacement time, and L is the lead time for ordering a replacement component.

6.3. *Sensor-driven versus traditional replacement and inventory policies*

This section extends the case study to highlight the advantages of implementing sensor-driven replacement and inventory policies. We demonstrate how implementing these policies can result in improved decision policies due to the improved accuracy of failure prediction. To do this, we implement both policies over a fixed time horizon and

Table 1. Numerical results

Degradation percentile (%)	t_r	t_o
0	86	24
10	54	14
20	50	16
30	46	16
40	24	8
50	22	8
60	22	8
70	12	4
80	12	4
90	0	0
100	0	0

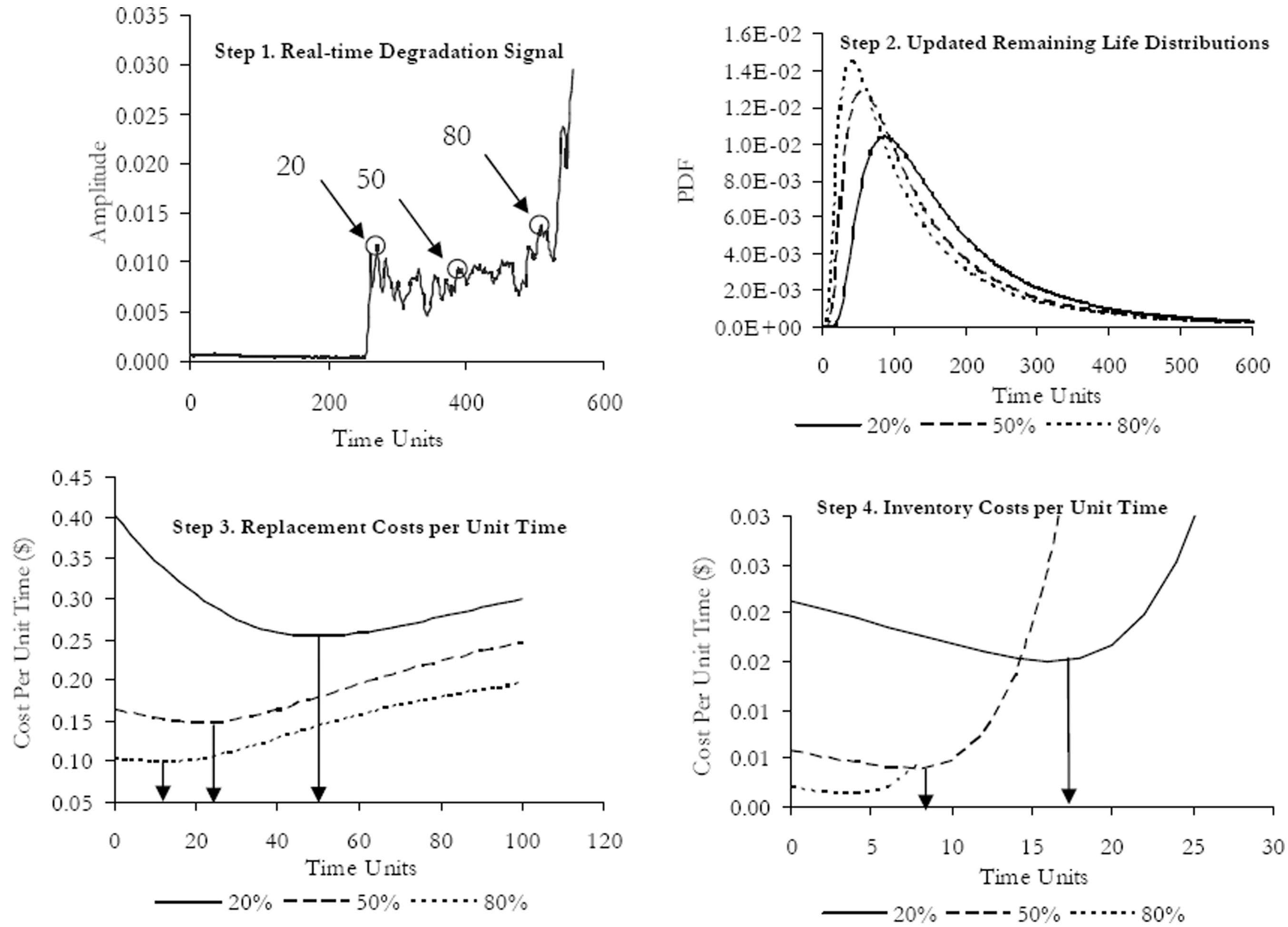

Fig. 5. Summary of the sensor-driven methodology for spare parts decision policies.

compare their performances. We choose two measures to evaluate performance. First, the total number of failure replacements occurring over the planning horizon. The second performance measure is the total maintenance costs, (sum of planned replacement, failure replacement, inventory holding, and stockout costs) associated with each policy. In our bearing monitoring case study, we assume that sensor data can be acquired at no significant cost using vibration monitoring accelerometers, thus data acquisition costs were not considered in the analysis.

6.3.1. *Traditional decision policy*

First, a Weibull distribution was fitted to the actual failure times of the 25 bearings used to obtain values of the prior distribution parameters. The Weibull distribution is commonly used to characterize failure times of component populations. The given failure data yielded a shape parameter $\alpha = 797.48$ and a scale parameter $\lambda = 2.65$. This failure time distribution was used to compute optimal replacement and inventory ordering times using Equations (1) and (2) based on the same cost data used in the previous section. The computed optimal times were $t_r^* = 440$ time units and $t_o^* = 60$.

Next, nine random failure times were generated from the given distribution and the decision policy computed above was implemented. The costs associated with this decision policy over the nine failure cycles were then computed $C_{tot} = \$500.4$. The results are summarized in Table 2.

6.3.2. *Sensor-driven decision policy*

The same generated failure times were considered, but dynamic sensor-driven decision policies were implemented using the methodology demonstrated in the first two sections of the case study. The RLD of the bearing, replacement time, and inventory ordering time were updated every 2 minutes. The process continues according to the simple stopping rule until $t_r^* \geq t_f - L$, after which updating is stopped and the most recently updated decision policy is executed. This results in a planned replacement. Otherwise, if failure occurs prior to this, failure replacement takes place at a higher cost. The same procedure is applied to the newly replaced bearing. Using nine bearings over the same

Table 2. Results summary for the traditional decision policy

Cycle	*Failure time (units)*	*Failure/ planned*	*Replacement time (units)*	*Inventory ordering time (units)*	*Cycle cost ($)*
1	919.62	P	440	380	30.6
2	552.82	P	880	820	30.6
3	430.49	F	1310.49	1250.49	105.6
4	290.58	F	1601.07	1541.07	105.6
5	697.89	P	2041.07	1981.07	30.6
6	432.67	F	2473.74	2413.74	105.6
7	1070.82	P	2913.74	2853.74	30.6
8	448.84	P	3353.74	3293.74	30.6
9	1017.82	P	3793.74	3733.74	30.6
Total					500.4

planning horizon, the total costs associated were computed and found to be $C_{\text{tot}} = \$228.6$. Following the traditional decision policy resulted in three failures over the nine cycles, whereas the sensor-driven decision policy resulted in no failures due to the higher prediction accuracy.

Remark 1. The presence of sensor-based monitoring systems is a critical component of the proposed sensor-driven framework. Although recent advances in sensor technology have resulted in the provision of relatively low cost sensors; the cost of data acquisition and installation of sensor systems can still constitute a significant component of the underlying economics associated with implementing the sensor-driven methodology.

The case study presented in this paper did not consider the fixed costs associated with the acquisition and installation of these sensor-based monitoring systems. Thus, it does not reflect an accurate comparison between the proposed and traditional methodologies from an economics standpoint. Rather, it compares the potential for cost savings in systems that already have existing sensor-based monitoring systems, such as on-board diagnostics in state-of-the-art electronic systems, and other applications where the costs of acquisition and installation are significantly lower than the costs of planned replacement and spare parts holding, needless to say the much higher costs of failure and spare part stockouts.

7. Conclusions

In this paper, we present a sensor-driven decision-making methodology for component replacement and spare parts inventory ordering policies. We focus on a renewal-theoretic single-unit replacement and inventory model, which uses lifetime distributions to compute optimal component replacement and optimal spare part ordering for a system with a single-unit storage capacity (Armstrong and Atkins, 1996). Next, we discuss a degradation modeling framework for computing RLDs of partially degraded components. *In situ* degradation signals from functioning components are then used to continuously/periodically update these distributions in real-time. These sensor-driven RLDs are used to replace the conventional lifetime distributions to approximate optimal replacement and spare parts ordering times. The resulting replacement and inventory policies are therefore driven by the underlying degradation process. The methodology is validated using real-world degradation data from a bearing application. We demonstrate how replacement and inventory ordering decisions evolve with progressive bearing degradation, and then highlight the advantages of using the proposed methodology as opposed to traditional decision policies.

This paper demonstrates the importance of establishing a linkage between low-level degradation-based sensor data with upper-level decision models, which is necessary for improving the decision-making process. Future research directions include the development of sensor-driven replacement models for multi-component systems and systems with variable lead times. Since the updating epochs do not represent regeneration points in the system, further research is also needed to develop exact sensor-driven decision policies.

References

Aka, M., Gilbert, S. and Ritchken, P. (1997) Joint inventory/ replacement policies for parallel machines. *IIE Transactions*, **29**(6), 441–449.

Akturk, M.-S. and Gurel, S. (2007) Machining conditions-based preventive maintenance. *International Journal of Production Research*, **45**(8), 1725–1743.

Armstrong, M. and Atkins, D. (1996) Joint optimization of maintenance and inventory policies for a simple system. *IIE Transactions*, **28**(5), 415–424.

Aronis, K., Magou, I., Dekker, R. and Tagaras, G. (2005) Inventory control of spare parts using a Bayesian approach: a case study. *European Journal of Operational Research*, **154**, 730–739.

Bae, S.J., Kuo, W. and Kvam, P.H. (2007) Degradation models and implied lifetime distributions. *Reliability Engineering and System Safety*, **92**, 601–608.

Chang, P., Chou, Y. and Huang, M. (2005) A (r,r,Q) inventory model for spare parts involving equipment criticality. *International Journal of Production Economics*, **97**, 66–74.

Cheung, K. and Hausman, W. (1997) Joint determination of preventive maintenance and safety stocks in an unreliable production environment. *Naval Research Logistics*, **44**, 257–272.

Chinnam, R.B. (2002) On-line reliability estimation for individual components using statistical degradation signal models. *Quality and Reliability Engineering International*, **18**, 53–73.

Dekker, R. (1996) Applications of maintenance optimization models: a review and analysis. *Reliability Engineering and System Safety*, **51**, 229–240.

Doksum, K.A. and Hoyland, A. (1992) Models for variable-stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution. *Technometrics*, **34**(1), 74–82.

Ebeling, C. (2005) *Introduction to Reliability and Maintainability Engineering*. Waveland Pr. Inc.

Gebraeel, N. (2006) Sensory updated residual life distribution for components with exponential degradation patterns. *IEEE Transactions on Automation Science and Engineering* (in press).

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

Ghasemi, A., Yacout, S. and Ouali, M.S. (2007) Optimal condition based maintenance with imperfect information and the proportional hazards model. *International Journal of Production Research*, **45**(4), 989–1012.

Harris, T. (2001) *Rolling Bearing Analysis*. Wiley, New York, NY.

Hartman, J. (2004) Multiple asset replacement analysis under variable utilization and stochastic demand. *European Journal of Operational Research*, **159**, 145–165.

Huiskonen, J. (2001) Maintenance spare parts logistics: special characteristics and strategic choices. *International Journal of Production Economics*, **71**, 125–133.

Joseph, V.-R. and Yu, I. (2006) Reliability improvement experiments with degradation data. *IEEE Transactions on Reliability*, **55**(1), 149–157.

Kharoufeh, J.-P. and Cox, S.M. (2005) Stochastic models for degradation-based reliability. *IIE Transactions*, **37**, 533–542.

Liao, H. and Elsayed, E. (2006) Reliability inference for field conditions from accelerated degradation testing. *Naval Research Logistics*, **53**, 576–587.

Lu, J.C. and Meeker, W. (1993) Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**, 161–174.

Lu, S., Tu, Y. and Lu, H. (2007) Predictive condition-based maintenance for continuously deteriorating systems. *Quality and Reliability Engineering International*, **23**, 71–81.

Maillart, L.M. (2006) Maintenance policies for systems with condition monitoring and obvious failures. *IIE Transactions*, **38**(6), 463–475.

Montoro-Cazorla, D. and Perez-Ocon, R. (2006) Replacement times and costs in a degrading system with several types of failure: the case of phase-type holding times. *European Journal of Operational Research*, **175**, 1193–1201.

Nelson, W. (1990) *Accelerated Testing Statistical Models, Test Plans, and Data Analysis*. Wiley, New York, NY.

Park, C. and Padgett, W.J. (2006) Stochastic degradation models with several accelerating variables. *IEEE Transactions on Reliability*, **55**(2), 379–390.

Pedregal, D.J. and Carnero, M.-C. (2006) State space models for condition monitoring: a case study. *Reliability Engineering and System Safety*, **91**, 171–180.

Robinson, M. and Crowder, M. (2008) Bayesian methods for a growth-Curve degradation model with repeated measures. *Lifetime Data Analysis*, **6**, 357–374.

Vaughan, T. (2005) Failure replacement and preventive maintenance spare parts ordering policy. *European Journal of Operational Research*, **161**, 183–190.

Venkatesan, M. (1984) Production-inventory with equipment replacement – PIER. *Operations Research*, **32**(6), 1286–1295.

Whitmore, G.A. and Schenkelberg, F. (1997) Modeling accelerated degradation using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, **3**(1), 27–45.

Zhou, X., Xi, L. and Lee, J. (2007) Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation. *Reliability Engineering and System Safety*, **92**, 530–534.

Biographies

Alaa Elwany is a pursuing a Ph.D. in the School of Industrial and Systems Engineering at Georgia Institute of Technology. He received his B.Sc. in Production Engineering and MS in Industrial Engineering from the University of Alexandria, Egypt. His primary research interests are in integrating degradation models with sensor-driven decision models for maintenance operations and logistics. His secondary research interests are focused on applications of operations research in manufacturing systems and supply chain management.

Nagi Gebraeel is an Assistant Professor in Industrial and Systems Engineering. He received his Ph.D. from Purdue University in 2003. His research interests are in prognostics and sensor-driven logistics, degradation modeling and reliability. His research is mostly targeted towards improving the accuracy of predicting unexpected failures of engineering systems by leveraging sensor-based data streams. In addition, he is interested in studying the impact of these developments on maintenance operations and spare parts logistics. He is a member of IIE and INFORMS.