Computers & Industrial Engineering 125 (2018) 480–489

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

# Prediction of hard failures with stochastic degradation signals using Wiener process and proportional hazards model

Jianing Man[a,*], Qiang Zhou[b]

[a] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong
[b] Department of Systems and Industrial Engineering, The University of Arizona, USA

ARTICLE INFO

*Keywords:*
Hard failure
Wiener process
Remaining useful life (RUL) prediction
Proportional hazards (PH)

ABSTRACT

In this paper, we propose a method to predict the remaining useful life (RUL) of systems subject to hard failures, which are probabilistically linked to system degradation signals (health indictors). A joint modeling framework is adopted to incorporate both the degradation signals and time-to-event data. In the joint model, a Wiener process with drift is used to model stochastic degradation signals, and the proportional hazards (PH) model with nonparametric baseline hazard is used to model time-to-event data. With proposed joint model and Markovian property of the Wiener process, system RUL could be predicted. Extensive simulations and a case study are conducted to demonstrate the performance of the proposed method.

## 1. Introduction

Factories are developing into smart manufacturing environment with intelligent devices in the 4th industrial revolution (Industry 4.0), and well connected through the Internet of Things (IoT) (Barbosa & Aroca, 2017). One of the important features of smart manufacturing is self-awareness and self-predictiveness ability, in which the remaining useful life (RUL) prediction plays an important role. RUL is also critical for conducting system prognostics and health management (PHM) and conditional-based maintenance (CBM) to ensure system operating or health conditions, which has arisen intensive attentions in academia and industry.

With the development of data acquisition technology, extensive condition monitoring (CM) data can be observed, among them system degradation signals highly related to system working or health conditions are possible obtained. Based on the available data, the evolution path of degradation signals as well as system deterioration process could be well investigated. In the literature, numerous research works have focused on the RUL prediction with observed data. Among them, prediction methods can be categorized as physics-based, data-driven methods, or a hybrid of both. For the data-driven models, data-mining or statistical methods could be utilized on the basis of observed CM data. In this paper, we focus on the data-driven method for RUL prediction which is flexible and requires minimal physical knowledge about the deterioration process.

In the literature, there are two types of failure mechanisms with respect to degradation signals. One is the soft failure, in which a unit is declared to be "failed" when its degradation signal reaches a predefined threshold for the first time, such as Lu & Meeker (1993), Si, Wang, Chen, & Zhou (2013). The other one is the hard failure which occurs as a unit fails to perform its intended function or stops working. In hard failure cases, a threshold does not needed, which is appropriate for applications with unclear threshold, or threshold simply does not exist. Compared with two different types of failure, soft failure time is determined by the degradation signal with its threshold, whereas time-to-event data such as failure times and censoring times must be provided to analyze hard failure problems with the observed degradation signals.

In order to achieve accurate RUL prediction performance, degradation signals which reflect system health status should be accurately modeled. Stochasticity is a major characteristic of degradation signals that differentiates individual units and contributes to the uncertainty in RUL estimation (Si et al., 2013). To consider the stochastic property of degradation signals, the random effects regression (RER) model and stochastic process (SP) models are widely used. The RER models consider the unobservable endogenous factors for a population with unit-to-unit variability, which was studied by Tsiatis, Degruttola, & Wulfsohn (1995), Tsiatis & Davidian (2004), Zhou, Son, Zhou, Mao, & Salman (2014), *etc*. The SP models are used to represent a collection of random variables that change over time; a good review can be found in Ye & Xie (2015).

Different types of stochastic degradation processes are extensively studied for soft failure problems, such as the Wiener process (Ye, Wang,

* Corresponding author.
*E-mail address:* jianinman2-c@my.cityu.edu.hk (J. Man).

https://doi.org/10.1016/j.cie.2018.09.015
Received 6 February 2018; Received in revised form 4 August 2018; Accepted 5 September 2018
Available online 06 September 2018
0360-8352/ 

Tsui, & Pecht, 2013; Wang et al., 2017), Gamma process (Pan & Balakrishnan, 2011; Park & Padgett, 2005), inverse Gaussian process (Barndorff-Nielsen, 1997; Ye & Chen, 2014), geometric Brownian motion (Park & Padgett, 2005) and Ornstein–Uhlenbeck process (Ricciardi & Sato, 1988). The random effects model may also be considered together with stochastic processes, such as the random effects with the Wiener process and Gamma process (Wang, 2010; Lawless & Crowder, 2004).

Unlike the extensive literature on soft failure prediction, there are very limited works for the hard failure analysis and prediction in engineering, such as Liao, Zhao, & Guo (2006), Yu & Fuh (2010), and Zhao & Elsayed (2004). Zhou et al. (2014) proposed a joint modeling framework that uses a mixed-effects model for degradation signals and the Cox proportional hazards (PH) model for time-to-event data. Son, Zhang, Sankavaram, & Zhou (2015) extended it by considering change-points in the degradation path.

Besides the statistical methods, data-mining methods have been studied to predict system RUL dealing with high dimensional CM covariates. Tian (2012) proposed to use an artificial neural network to predict life percentage. Khelif et al. (2017) used support vector machine, and Guo, Li, Jia, Lei, & Lin (2017) used the recurrent neural network with selected variables to conduct the prognosis. Based on the studies of degradation patterns, Li, Ding, & Sun (2018) and Zhao, Liang, Wang, & Lu (2017) utilized different types of neural networks to predict system RUL.

In soft failure prognosis, SP models are extensively studied using statistical methods, and the stochasticity of degradation processes also has been investigated using data-mining algorithms. To the best of our knowledge, SP models have not been considered in hard failure problems, although the same justifications of using them apply equally to degradation signals in both failure types. To fill this gap, we aim to introduce Wiener process into the joint modeling framework which is used to analyze hard failures. In particular, the degradation signal will be described by a Wiener process with drift, and the time-to-event data will be described using a Cox PH model.

The remainder of the paper is organized as follows. In Section 2, the methodology is proposed including the modeling, estimation, and prediction of hard failures. In Section 3, a simulation study is conducted to demonstrate the performance of the method, and a comparison study between our proposed method and degradation pattern learning based method is illustrated. In Section 4, a case study is conducted using real data. Finally, we draw some conclusions and discuss future works in Section 5.

## 2. Methodology

For each unit $i(i = 1, ..., n)$, we observe its event time $V_i = \min(T_i, C_i)$ with the indicator function $\Delta_i = 1_{(T_i \leqslant C_i)}$, where $T_i$ is the lifetime and $C_i$ is the censoring time. Also available are the time-fixed covariates $\mathbf{w}_i$ and time-dependent covariates $\mathbf{y}_i = \{y_{i,0}, ..., y_{i,j}\}$, $(i = 1, ..., n, j = 0, ..., m_i)$ at observation times $\mathbf{t}_i = \{t_{i,0}, ..., t_{i,j}\}$, $(t_{i,j} \leqslant V_i)$. Collectively, the available data for the $i$th unit are $(V_i, \Delta_i, \mathbf{y}_i, \mathbf{w}_i, \mathbf{t}_i)$.

Time-dependent covariate $y_{i,j}$ of $i$th unit observed at $t_{i,j}$ could manifest the system deterioration or operating status, which is also known as degradation signals. For examples, the crack length of bearing, the resistance of battery, and light intensity of light-emitting diode (LED). Time-fixed covariates $\mathbf{w}_i$ could be categorical data or dummy variables related to system conditions, such as different manufacturers or types of components.

### 2.1. Modeling

Assume all the units are sampled from the same underlying population, hence they share some similarities but with unit-to-unit variations, and the variation varies with observation times, which is often observed from stochastic degradation signals. To model this, a Wiener process with drift is used, which is a widely used and well-studied stochastic process in the literature, e.g. Whitmore (1995), Wang, Balakrishnanb, & Guo (2014).

The evolution of the degradation signal is modeled as

$$y_{i,j} = y_{i,0} + W_i(t_{i,j}), \tag{1}$$

$$W_i(t_{i,j}) = \theta^T \boldsymbol{\Lambda}(t_{i,j}) + \sigma_b B_i(t_{i,j}), \tag{2}$$

where $y_{i,j}$ is the observed degradation signal at $t_{i,j}$, $y_{i,0}$ is the initial status of the system at $t_{i,0}$ which follows an *i.i.d.* normal distribution. $N(\mu_0, \sigma_0^2)$ indicating the randomness of initial statuses, $W(t_{i,j})$ is the Wiener process with drift coefficient $\theta^T = [\theta_1 ... \theta_k]$ and diffusion coefficient $\sigma_b$. $\boldsymbol{\Lambda}(t) = [\Lambda_1(t) ... \Lambda_k(t)]^T$ is the drift function and $B_i(t) \sim N(0, t)$ is the standard Brownian motion. For demonstration purpose, we use a quadratic form (without intercept) where $\theta^T = [\theta_1\ \theta_2]$ and $\boldsymbol{\Lambda}(t) = [t\ t^2]^T$. Drift function $\theta^T \boldsymbol{\Lambda}(t_{i,j})$ can take various forms according to different applications, transformations may need to be implemented to use the quadratic form, and higher order regression function could be considered.

We further denote $F_i(t_{i,j}) = \sigma_b B_i(t_{i,j})$ for the $i$th unit, which is a stochastic component of $i$th unit at $t_{i,j}$ with $F_i(t_{i,j} = 0) = 0$. As $F_i(t_{i,j})$ is not observable, it could be estimated using $\widehat{F}_i(t_{i,j}) = y_{i,j} - y_{i,0} - \theta^T \boldsymbol{\Lambda}(t_{i,j})$. Based on the properties of Brownian motion, $F_i(t_{i,j})$ depends only on the latest $F_i(t_{i,j-1})$. The increments $\Delta F_i(t_{i,j}) = F_i(t_{i,j}) - F_i(t_{i,j-1})$ are independent and $\Delta F_i(t_{i,j})$ follows normal distribution $N(0, (t_{i,j} - t_{i,j-1})\sigma_b^2)$. Accordingly, the observations $y_{i,j}$ only depends on the latest observation $y_{i,j-1}$, and the increment follows a normal distribution.

Following the joint modeling framework as in Zhou et al. (2014), both the time-dependent and time-fixed covariates affect degradation process through the hazard function, which can be modeled by the well-studied PH function (Cox, 1972).

$$h(t_{i,j}) = h_0(t_{i,j}) \exp(\beta y_{i,j} + \boldsymbol{\gamma}^T \mathbf{w}_i), \tag{3}$$

$$h_0(t_{i,j}) = \sum_{\varsigma} C_\varsigma 1_{(t_{\varsigma-1} \leqslant t_{i,j} < t_\varsigma)}, \tag{4}$$

where $h_0(\cdot)$ is a nonparametric baseline hazard function using the stepwise function (Zeng & Lin, 2007; Tseng, Su, Mao, & Wang, 2015), $1_{(\cdot)}$ is an indicator function with the time interval $[t_{\varsigma-1}, t_\varsigma)$ for the stepwise baseline hazard rate $C_\varsigma$. $t_\varsigma$ is the ordered observed failure times, and $\varsigma$ is the number of observed failure time in historical data used for estimation, which varies in different applications. $\beta, \boldsymbol{\gamma}$ are the coefficients for the time-dependent covariate (the degradation signal) and time-fixed covariate(s). With the hazard function in Eq. (3), the probability density function (pdf) of lifetime is

$$f_T(t) = h(t)S(t) = h(t)\exp\left(\int_0^t -h(s)ds\right), \tag{5}$$

where $S(t) = \exp\left(\int_0^t -h(s)ds\right)$ is the survival function.

### 2.2. Parameter estimation

Based on the historical data, the unknown parameters $\boldsymbol{\Theta} = (\beta, \boldsymbol{\gamma}, h_0(\cdot), \theta, \sigma_b, \mu_0, \sigma_0)$ can be estimated using the maximum likelihood estimation (MLE) method. For simplicity, we assume there is only one time-fixed covariate, and it is straightforward to be extended to the multiple time-fixed covariates cases if the covariates are linearly independent. In this case, both the time-fixed covariate and its coefficient are scalars, i.e. $w_i$ and $\gamma$.

Since the Brownian motion has independent and normally distributed increments, we define

$$\Delta y_{ij} = \begin{cases} y_{i,j}, & j = 0 \\ y_{i,j} - y_{i,j-1}, & j \geqslant 1 \end{cases}. \tag{6}$$