European Journal of Operational Research 226 (2013) 53–66

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

# A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution

Xiao-Sheng Si [a,b], Wenbin Wang [c,*], Mao-Yin Chen [b], Chang-Hua Hu [a], Dong-Hua Zhou [b,*]

[a] Department of Automation, Xi'an Institute of High-Tech, Xi'an, Shaanxi 710025, PR China
[b] Department of Automation, TNLIST, Tsinghua University, Beijing 100084, PR China
[c] Dongling School of Economics and Management, University of Science and Technology, Beijing, PR China

ARTICLE INFO

*Article history:*
Received 26 January 2012
Accepted 22 October 2012
Available online 3 November 2012

*Keywords:*
Replacement
Remaining useful life
First passage time
Expectation maximization
Prognostics

ABSTRACT

Remaining useful life (RUL) estimation is regarded as one of the most central components in prognostics and health management (PHM). Accurate RUL estimation can enable failure prevention in a more controllable manner in that effective maintenance can be executed in appropriate time to correct impending faults. In this paper we consider the problem of estimating the RUL from observed degradation data for a general system. A degradation path-dependent approach for RUL estimation is presented through the combination of Bayesian updating and expectation maximization (EM) algorithm. The use of both Bayesian updating and EM algorithm to update the model parameters and RUL distribution at the time obtaining a newly observed data is a novel contribution of this paper, which makes the estimated RUL depend on the observed degradation data history. As two specific cases, a linear degradation model and an exponential-based degradation model are considered to illustrate the implementation of our presented approach. A major contribution under these two special cases is that our approach can obtain an exact and closed-form RUL distribution respectively, and the moment of the obtained RUL distribution from our presented approach exists. This contrasts sharply with the approximated results obtained in the literature for the same cases. To our knowledge, the RUL estimation approach presented in this paper for the two special cases is the only one that can provide an exact and closed-form RUL distribution utilizing the monitoring history. Finally, numerical examples for RUL estimation and a practical case study for condition-based replacement decision making with comparison to a previously reported approach are provided to substantiate the superiority of the proposed model.



## 1. Introduction

Prognostics and health management (PHM) is an efficient and systematic approach for evaluating the reliability of a system in its actual operating conditions, predicting failure progression, and mitigating operating risks via management actions (Pecht, 2008). In PHM, prognostics can yield an advance warning of impending failure in a system, thereby helping in making maintenance decisions and executing preventive actions. The past decade has witnessed a constant research interest on various aspects of PHM due primarily to the fact that PHM has been extensively applied in a variety of fields including electronics, smart grid, nuclear plant, power industry, aerospace and military application, fleet-industrial maintenance, and public health management (Smith et al., 1997; Wang and Zhang, 2005; Nikhil and Pecht, 2006; Lall et al., 2006; Wang, 2007; Mazhar et al., 2007; Tsui et al., 2011).

In each of these applications and documents, one critical quantity during prognostics for a system is the prognostic distance within which management decisions and repair actions can be planned effectively prior to failure occurrence to extend system life (Derman et al., 1984; Wang and Christer, 2000; Si et al., 2011). This prognostic distance is closely associated with the definition of the remaining useful life (RUL) which is the length of the time from the present to the end of useful life. In fact, RUL estimation is always a key part in any PHM program and management can make use of RUL information in condition-based maintenance (CBM) to produce economic benefits in engineering, maintenance, logistics, and operations. Therefore, over the past few decades, significant advances have been made in developing RUL estimation approaches (Si et al., 2011).

Stochasticity is one of the main characteristics in system operations, that contributes to the uncertainty in estimating the RUL of the system. Therefore, one fundamental issue in RUL estimation is to find the probability density function (PDF) of the RUL. However this also leads to the main difficulty of RUL estimation since how to make full use of condition monitoring (CM) information to infer a

* Corresponding authors. Tel.: +86 010 62794461; fax: +86 010 62786911.
*E-mail addresses:* wangwb@ustb.edu.cn (W. Wang), zdh@mail.tsinghua.edu.cn (D.-H. Zhou).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.10.030

RUL distribution is a not-well-solved problem. So far, RUL estimation has been regarded as one of the most central components in PHM (Pecht, 2008; Camci and Chinnam, 2010). Thus, our primary interest of this paper is to utilize CM information to adaptively estimate the RUL of the system and then apply it to support decision-related applications.

The current RUL estimation approaches can be generally classified as physics of failure, data driven and fusion. Physics of failure approaches rely on the physics of underlying failure mechanisms. Data driven approaches achieve RUL estimation via data fitting mainly including machine learning and statistics based approaches. The fusion approaches are the combination of the physics of failure and data driven approaches. However, for complex or large-scale engineering systems, it is typically difficult to obtain the physical failure mechanisms in advance or cost-expensive and time-consuming to capture the physics of failure. In contrast, data-driven approaches attempt to derive models directly from collected CM and life data, and thus are more appealing and have gained much attention in recent years (Si et al., 2011).

In conventional data-based approaches, estimating the RUL is achieved by evaluating the conditional lifetime distribution given that a system has survived up to a specific time, e.g. $T - t|T > t$, where $T$ denotes the lifetime (Maguluri and Zhang, 1994; Alam and Suzuki, 2009). The obtained RUL distributions from these approaches are generally based on the life characteristics of a population of identical systems and lifetime data are required. However, such data are scarce in reality or even non-existent at all for systems which are costly or time-consuming to collect the life data (Ma and Krings, 2011). With the advances in CM technologies, degradation data can be obtained from routine CM as feasible and low-cost alternatives to estimate the RUL. These data are usually correlated with the underlying physical degradation process. If they are properly modelled, degradation data can be used to predict unexpected failures and accurately estimate the lifetime of gradually degraded systems (Escobar and Meeker, 2006; Joseph and Yu, 2006). In general, degradation data based methods for RUL estimation can be classified into the models based on indirectly observed degradation processes and the models based on directly observed degradation processes (Si et al., 2011). The former models considered that the degradation state was hidden and assumed that the available CM data were stochastically related to the underlying degradation state. In this case, lifetime data must be available to establish the relationship between the CM data and failure. The latter models utilized the observed degradation data directly to describe the underlying degradation state of the system. In this paper, we mainly focus on the directly observed degradation processes.

One common definition of RUL in the directly observed case is related to the concept of the first passage time (FPT) of the degradation process crossing a pre-defined threshold level. The use of the FPT concept as the definition of failure or a terminating event has a long history of application in diverse fields, including medicine, environmental science, engineering, business, economics and sociology (Whitmore, 1986; Lee et al., 2004; Lee and Whitmore, 2006; Balka et al., 2009). It is also acknowledged as a mainstream definition of failure in reliability literature based on degradation data (Singpurwalla, 1995; Park and Padgett, 2006; Aalen et al., 2008; Peng and Tseng, 2009; Pandey et al., 2009; Park and Bae, 2010; Li and Ryan, 2011). Thus, in this paper, we pay particular attention to a type of degradation-data-based models and derive the RUL distribution based on the concept of the FPT. Since degradation data are part of CM data, throughout this paper, we use terms 'CM information' and 'degradation data' inter-changeably.

In most of degradation-data-based models for RUL estimation, an exact and closed-form of the RUL distribution in the FPT sense is only available for some special cases. Frequently, a stepwise approximation or numerical simulation has to be used for finding an approximated RUL (Yuan and Pandey, 2009; Park and Bae, 2010; Wang and Zhang, 2005, 2008; Xu et al., 2008; Carr and Wang, 2011). In addition, most of these models either do not use the in situ degradation data during lifetime inference or only use information contained at the current observation point. However, the degradation data over the path collected up to date could contain more useful information to make the RUL estimation more accurate.

The type of models we specifically consider in this work follows the idea in Gebraeel et al. (2005) where two exponential-like degradation models were proposed. In their models, stochastic parameters were updated via a Bayesian approach to incorporate real-time CM information. Following Gebraeel et al. (2005), many variants and applications have been reported in prognostics, maintenance and inventory management (Li and Ryan, 2011; You et al., 2010; Gebraeel, 2006; Elwany and Gebraeel, 2008, 2009). However, in these papers, they estimated the RUL distribution as the distribution of the time that it takes the trajectory of the degradation signal to cross the failure threshold based on an approximated method. In reality, this is not the FPT since the signal may have already crossed the failure threshold, signifying failure, prior to predicting the RUL. In extreme cases where the degradation fluctuations are large, this approximation could be significantly crude from the FPT concept. Even when the Brownian motion (BM) used as an error term, the availability of the explicit distribution of the FPT from the BM with a drift, i.e. the inverse Gaussian distribution, was not utilized in their models. Elwany and Gebraeel (2009) used the FPT to approximate the mean RUL but the distribution of the RUL was still evaluated by their approximate approach. As such, the results of RUL estimation in Gebraeel et al. (2005) and the followed works in applications are approximations as opposed to the FPT concept. Furthermore, in above works, the obtained RUL distributions belong to a family of Bernstein distributions. Consequently, the moments of the RUL do not exist. But in maintenance practice, the expectation of the RUL is required to be existent sometimes (Derman et al., 1984; Shechter et al., 2008). Also, the stochastic coefficients in Gebraeel et al. (2005) and other following works had some prior distributions but no elaborated method is presented to select the hyperparameters of the prior distributions. Typically, several systems' historical degradation data of the same type are required to determine the deterministic coefficient and the unknown parameters in the prior distributions of the stochastic coefficients. But the scarcity of such historical degradation data of multiple systems is a commonly encountered case in practice, particularly for newly armed systems. As shown in Section 5 of this paper, an inappropriate selection of these parameters can result in an incorrect estimate of the RUL.

Driven by the above survey over the related works, the purpose of this paper is to develop a degradation path-dependent approach for RUL estimation that allows the estimated RUL distribution to be dependent on a system's degradation data history and to be adaptively updated, at the moment that a newly observed data is available. In particular, our goal is to shed light on three fundamental issues: (i) RUL estimation for an individual fielded device without the need of offline data of other similar systems, (ii) parameter estimation/updating of the degradation model from the observed degradation data, and (iii) an exact yet closed-form expression of the RUL distribution given (i) and (ii).

In response to the above issues, the dependency of RUL estimation with a system's past degradation path is presented through the combination of Bayesian updating and expectation maximization (EM) algorithm. This is a novel contribution of the paper and is not fully explored in the conventional RUL modeling paradigms. As such, the deterministic coefficient and the unknown hyperparameters in the prior distributions of the stochastic coefficients can be