Wear 376-377 (2017) 1227–1233

Contents lists available at ScienceDirect

Wear

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

Parameter estimation and remaining useful life prediction of lubricating oil with HMM

Ying Du [a,b], Tonghai Wu [a,*], Viliam Makis [b]

[a] Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[b] Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Ontario, Canada, M5S 3G8

ARTICLE INFO

Article history:
Received 2 September 2016
Received in revised form
23 November 2016
Accepted 25 November 2016

Keywords:
Lubricating oil deterioration
Hidden Markov model
Parameter estimation
Remaining useful life

ABSTRACT

Lubricating oil plays a vital role in the full life-span performance of the machine. Lubricating oil deterioration, which leads to the attenuation of oil performance and severe wear afterwards, is a slow degrading process, which can be observed by condition monitoring, but the actual degree of the oil degradation is often very difficult to examine. The main purpose of lubricating oil degradation prediction is to estimate the failure time when the oil no longer fulfills its functions. We suppose that the state process evolution of lubricating oil degradation can be modeled using a hidden Markov model (HMM) with three states: healthy state, unhealthy state, and failure state. Only the failure state is observable. While the lubricating oil is in service, vector data that are stochastically related to the deterioration state are obtained through on-line condition monitoring by an OLVF (On-line Visual Ferrograph) sensor at regular sampling epochs. A method of Time Series Analysis (TSA) is applied to the healthy portions of the oil data histories to get the residuals as the observable process containing partial information to fit the hidden Markov model. The unknown parameters of the fitted hidden Markov model are estimated by the Expectation-Maximization (EM) algorithm. The remaining useful life (RUL) of lubricating oil can be evaluated through explicit formulas of the characteristics such as the conditional reliability function (CRF) and mean residual life (MRL) function in terms of the posterior probability.

1. Introduction

Lubricating oil is used to reduce wear and friction from the mobile components, eliminate contamination, remove heat from friction surfaces, and avoid machine failure and reduce the cost for unscheduled maintenance afterwards. Therefore, lubricating oil condition should be monitored and the oil should be replaced regularly to extend the period when the machine is in good state [1]. Recently, condition monitoring (CM) of lubricating oil has attracted a considerably attention in research and it plays a vital role in industries [2]. The oil data obtained from CM have been used to assess the actual condition of the operating machine in [3], but to our knowledge, HM models for lubricating oil deterioration and replacement when the machine is in the healthy state have not been developed in the literature. Taking into account that the time period when a machine is in the healthy state is usually considerably longer than the length of time between oil replacements, it is assumed in this paper that the machine condition is stable and will not affect considerably the speed of oil deterioration.

Wear debris level in lubricating oil has been proved to be one of the most common degradation features to evaluate lubricating oil degradation [2]. Relative wear debris concentration has been obtained from an image captured by an on-line sensor (OLVF) at a sampling epoch [4]. By this, the wear debris presenting in a lubricating oil sample can be categorized as a large and small group according to their sizes by controlling the oil flow rate and magnet field intensity [4]. When the machine is in operation, wear debris accumulate in the lubricating oil and the concentration increases, which leads to the lubricating oil degradation [5].

Although analysis on wear debris in lubricating oil has been utilized in practice for many years to estimate machine condition, little work has been done using statistical approaches to model and analyze oil data for the purpose of assessing the lubricating oil degradation and predicting its remaining useful life. The prediction with the CM data can be obtained by the conditional reliability function (CRF) and mean residual life (MRL) function [3], which indicates the failure time when the oil cannot fulfill its functions anymore, and should be changed. In condition monitoring area, RUL prediction was applied for particle contaminated lubricating oil by applying physical models using a particle filtering technique [5], as well as the application on rotational bearings with two-phase threshold model using Bayesian methods [6].

* Corresponding author.
E-mail address: wt-h@163.com (T. Wu).

http://dx.doi.org/10.1016/j.wear.2016.11.047
0043-1648/

To our knowledge, no statistical models have been developed in the literature which could be applied to the RUL prediction of lubricating oil.

In this paper, we present a statistical approach to predict the RUL of lubricating oil subject to deterioration using oil data obtained by an OLVF sensor. Statistical methods of deteriorating systems, which are utilized in industry to model the degradation process for early fault detection using CM information, are categorized into one of three main approaches: the proportional hazard modeling (PHM) [7], stochastic recursive filtering [8], and hidden Markov modeling (HMM) [9]. HMM, which has been proved to be efficient in gradual degradation system modeling [10,11] and the RUL prediction [11], and applied in many areas such as speech recognition, econometrics and condition-based maintenance [12], is employed in this paper with two types of stochastic processes, namely hidden state process and observation vector process in order to model the degradation process and predict the RUL of lubricating oil. The state and observation parameter estimations of the fitted HMM under partial observations of lubricating oil deterioration can be obtained by the Expectation-Maximization (EM) algorithm [11].

The residuals are obtained using a vector autoregressive (VAR) model as the observation process in the hidden Markov framework. Once the parameters of the HMM are estimated, we use the explicit formulas for the conditional reliability function (CRF) and the mean residual life (MRL) function in terms of the posterior probability [13], which can be used for RUL prediction of lubricating oil. The CRF indicates the probability that the oil can survive during a period of time and has not failed yet, and the MRL can be calculated by using the posterior probability [14]. It is very new in the Tribology area to apply the EM algorithm and HMM to model the degradation process and estimate the RUL of lubricating oil focusing on the lubricating oil condition.

The HMM-based procedure is shown in Fig. 1. The rest of the paper is organized as follows. In Section 2, a VAR model is fitted to the real 2-dimensional oil data from an OLVF sensor collected at regular time epochs, and the residuals for both the healthy and unhealthy portions of the oil data histories are obtained. In Section 3, the state and residual process are modeled as an HMM, and the estimation procedure of the unknown parameters is developed using the EM algorithm. In Section 4, the formulas for the conditional RF and MRL have been applied, which can be used to predict the RUL of lubricating oil. Finally, the conclusions and future research are summarized in Section 5.

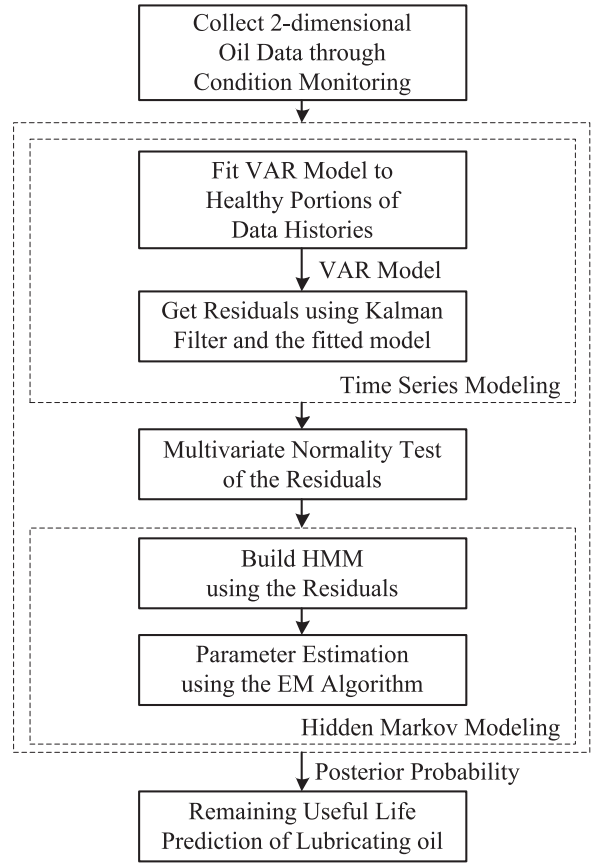

Fig. 1. HMM-based procedure for RUL prediction.

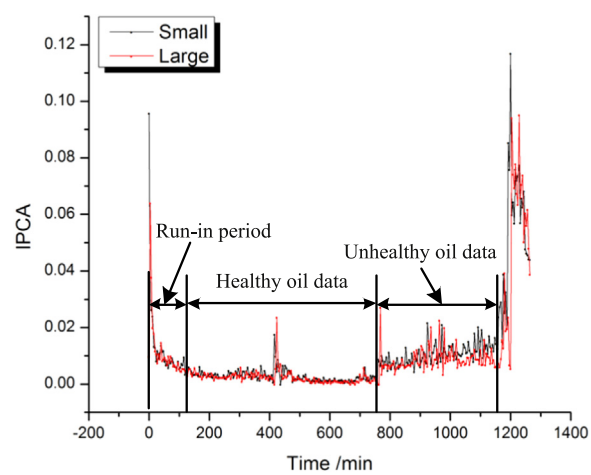

Fig. 2. Wear debris concentration of small and large particles.

2. Vector autoregressive modeling and computation of residuals

The real condition monitoring data were obtained for the detection of lubricating oil deterioration from a four-ball test rig [15] in order to predict the RUL of the lubricating oil afterwards. During the operational life of the tribo-pairs, oil data were collected every $\triangle = 4$ minutes by an OLVF sensor, which is an on-line ferrographic sensor based on Image Technology, and it provided wear debris concentrations that came from the direct wear during these 4 minutes. The total number of data histories recorded is 27, which consist of $N=11$ failure histories and $M=16$ suspension histories. The failure history is defined as the history that ends with observable failure, which indicates that the lubricating oil is out of use at that moment, and the suspension history is defined as the history that ends when the lubricating oil is still in operation and has not lost its functions.

To avoid over-parameterization, we use the 2-dimensional monitoring data consisting of small wear particles and large wear particles obtained from the OLVF sensor for analysis. A typical data history is given in Fig. 2, and the working conditions are listed in Table 1. IPCA shows three stages including run-in, normal, and severe stages, which agrees with the typical "Bathtub Curve".

Table 1
Working conditions for the test.

Test no.	Load/N	Rotated rate/rpm	Time/min	Downtime duration/min
1	1500	1000	360	0
2	1500	1000	360	240
3	2000	1000	210	720
4	2000	1000	240	60
5	2000	2000	60	480

At the beginning of the sampling, the tribo-pairs are in the run-in stage, which should be removed from the data set in the modeling process. After the run-in period, the tribo-pairs are operating under normal conditions, which are regarded as the healthy portion of the data history. Severe wear dramatically appears at 190*th* sampling epoch after 760 operational minutes, which is regarded as the beginning of the unhealthy portion of the data history. For our experiment, the four-ball test rig was stopped and restarted when working conditions changed, which explains the peak points at the beginning of each section in Fig. 2. Besides, this particular oil sample failed at 290*th* sampling epoch after 1156 operational minutes.

As described by Kim and Makis [11], one should first fit a model using only the healthy portions of the oil data histories that account for cross and autocorrelation in the data. Since there is no agreed upon criterion for selecting the 'optimal' segmentation, in order to partition a non-stationary time series data, we simply divide the data histories into two portions (healthy and unhealthy) via graphical examination. The purpose of segmentation is to identify the healthy portions of the data histories so that a stationary time series model can be fitted and the residuals can be computed using the fitted model. For each of the $N + M = 27$ data histories, the healthy portions of the data histories are denoted as $\left\{ z_1^i, z_2^i, \ldots, z_{t_i}^i, i = 1, 2, \ldots, N + M \right\}$.

The healthy portions of the data histories are assumed to follow a common stationary vector auto-regressive (VAR) process [16] given by

$$(Z_n - \mu_0) - \sum_{r=1}^{p} \Phi_r (Z_{n-r} - \mu_0) = \varepsilon_n,\ n \in \mathbf{Z} \tag{1}$$

where ε_n are i.i.d. $N_2(\mathbf{0}, \mathbf{C})$, the model order $p \in \mathbf{N}$, the autocorrelation matrices $\Phi_r \in \mathbf{R}^{2\times 2}$, and the mean and covariance model parameters $\mu_0 \in \mathbf{R}^2$ and $\mathbf{C} \in \mathbf{R}^{2\times 2}$. All the model parameters are unknown and need to be estimated.

In order to get the standard form of VAR Model, set $\mu = \mu_0 - \sum_{r=1}^{p} \Phi_r \mu_0$. Then, Eq. (1) can be written as

$$Z_n = \mu + \sum_{r=1}^{p} \Phi_r Z_{n-r} + \varepsilon_n,\ n \in \mathbf{Z} \tag{2}$$

Therefore, the regression representation $\mathbf{W} = \mathbf{VA} + \mathbf{E}$ is obtained for the observed healthy portions of the data histories $\left\{ z_1^i, z_2^i, \ldots, z_{t_i}^i, i = 1, 2, \ldots, N + M \right\}$, where

$$\mathbf{W} = \left[z_{t_{N+M}}^{N+M}, \ldots, z_{p+1}^{N+M}, \ldots, z_{t_1}^1, \ldots, z_{p+1}^1 \right]' \tag{3}$$

$$\mathbf{V} = \begin{bmatrix} 1 & z_{t_{N+M}-1}^{N+M} & \ldots & z_{t_{N+M}-p}^{N+M} \\ 1 & \vdots & \vdots & \vdots \\ 1 & z_p^{N+M} & \ldots & z_1^{N+M} \\ 1 & \vdots & \vdots & \vdots \\ 1 & z_{t_1-1}^1 & \ldots & z_{t_1-p}^1 \\ 1 & \vdots & \vdots & \vdots \\ 1 & z_p^1 & \ldots & z_1^1 \end{bmatrix} \tag{4}$$

$$\mathbf{A} = \left[\mu, \Phi_1, \ldots, \Phi_p \right]' \tag{5}$$

$$\mathbf{E} = \left[\varepsilon_{t_{N+M}}^{N+M}, \ldots, \varepsilon_{p+1}^{N+M}, \ldots, \varepsilon_{t_1}^1, \ldots, \varepsilon_{p+1}^1 \right]' \tag{6}$$

Using the method of least squares [16], the estimates for $\mathbf{A}$ and $\mathbf{C}$ are given by

$$\hat{\mathrm{A}} = (\mathbf{V}'\mathbf{V})^{-1}\mathbf{V}'\mathbf{W} \tag{7}$$

$$\hat{\mathbf{C}} = \frac{S_p}{T - (2p + 1)} \tag{8}$$

where $T = \sum_{i=1}^{N+M} (t_i - p)$ is the total number of the available healthy portions of the data histories, and $S_p = (\mathbf{W} - \mathbf{V}\hat{\mathbf{A}})'(\mathbf{W} - \mathbf{V}\hat{\mathbf{A}})$ is the residual sum of squares matrix. The estimation of model order $p \in \mathbf{N}$ is obtained by testing H_0: $\Phi_p = \mathbf{0}$ against H_a: $\Phi_p \neq \mathbf{0}$ using the likelihood ratio statistic given by

$$M_p = -\left(T - 2p - 1 - \frac{1}{2} \right) \ln\left(\frac{\det(S_p)}{\det(S_{p-1})} \right) \tag{9}$$

For our real 2-dimensional oil data, we obtained $M_4 = 54.33$ and $M_5 = -25.64$. From the chi-square distribution with 2 degrees of freedom and $\alpha = 0.05$, $\chi^2_{2,0.05} = 5.99$. Since $M_4 > \chi^2_{2,0.05}$ and $M_5 < \chi^2_{2,0.05}$, we reject H_0: $\Phi_4 = \mathbf{0}$ and fail to reject H_0: $\Phi_5 = \mathbf{0}$. Therefore, $\hat{p} = 4$ is an adequate model order, and the VAR model parameter estimates are given by

$$\hat{p} = 4,\ \hat{\mu} = \begin{pmatrix} 0.0974 \\ 0.0236 \end{pmatrix},\ \hat{\mathbf{C}} = \begin{pmatrix} 3.4290e-3 & 1.2452e-4 \\ 1.2452e-4 & 1.8464e-3 \end{pmatrix},$$

$$\hat{\Phi}_1 = \begin{pmatrix} 0.2492 & -0.0675 \\ 0.2318 & -0.1982 \end{pmatrix},\ \hat{\Phi}_2 = \begin{pmatrix} 0.1887 & -0.1083 \\ 0.4168 & -0.0157 \end{pmatrix},$$

$$\hat{\Phi}_3 = \begin{pmatrix} 0.0011 & 0.0155 \\ -0.0045 & 0.0161 \end{pmatrix},\ \hat{\Phi}_4 = \begin{pmatrix} 0.0143 & -0.0042 \\ 0.0129 & 0.0026 \end{pmatrix}$$

Using estimates $\hat{\gamma} = (\hat{\mu}, \hat{p}, \hat{\Phi}_1, \hat{\Phi}_2, \hat{\Phi}_3, \hat{\Phi}_4, \hat{\mathbf{C}})$, we define the residual process $(Y_{n\Delta})$ by

$$Y_{n\Delta} := Z_n - E_{\hat{\gamma}}(Z_n | \vec{\mathbf{Z}}_{n-1}) \tag{10}$$

where $\vec{\mathbf{Z}}_{n-1} = (Z_1, Z_2, \ldots, Z_{n-1})$.

For $n > \hat{p}$, $Y_{n\Delta} = Z_n - (\hat{\mu} + \sum_{r=1}^{\hat{p}} \Phi_r Z_{n-r})$, and for $n < \hat{p}$, we recursively compute $Y_{n\Delta}$ using the Kalman filter, applying the procedure used in reference [17]. Therefore, the residuals can then be computed for both the healthy and unhealthy portions of all monitoring oil data histories, provided graphically in a 2-dimensional scatter plot shown in Fig. 3. The crosses in blue are residuals computed from the healthy portions of the oil data histories, and the circles in red are residuals computed from the unhealthy portions of the oil data histories.

The statistical test of the normality assumption was performed using the Henze-Zirkler Multivariate Normality Test [18] with a given significance level $\alpha = 0.05$, and the obtained results are shown in

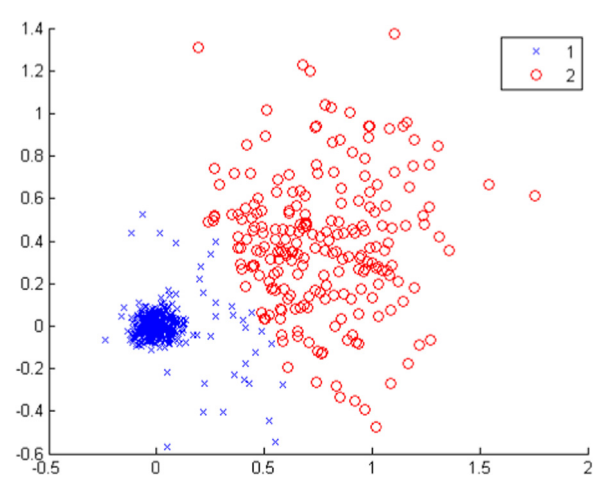

Fig. 3. Scatter plot for all the residuals. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 2
P-value of the residual normality test.

Test	Healthy data set	Unhealthy data set
Normality(Henze-Zirkler)	0.8232	0.2983

Table 2, which show that the calculated residuals follow the multivariate normal distribution and satisfy the assumption of multivariate normality for both healthy and unhealthy data sets [19].

In the next section, we will apply the HMM framework using the obtained residuals as the observation process, and estimate the state and observation parameters by applying the EM algorithm.

3. Hidden Markov modeling and parameter estimation using EM algorithm

Lubricating oil deterioration is a slow degrading process, and it is difficult to be examined. In order to investigate the RUL of the lubricating oil, a degradation model of the process should be considered. We assume that the deterioration state of lubricating oil characterized by the two monitoring indexes follows a continuous time homogeneous Markov process $\{X_t: (t \in \mathbf{R}_+)\}$, with the state space $S = \{0, 1\} \cup \{2\}$. In general, states 0 and 1 are always unobservable, representing the healthy and unhealthy operational states of lubricating oil, respectively. Only the failure state (state 2) is observable. It is assumed that the lubricating oil starts in a "good as new" state, i.e. $X_0 = 0$, and the transition rate matrix is given by

$$\Lambda = \begin{pmatrix} -(\lambda_{01}+\lambda_{02}) & \lambda_{01} & \lambda_{02} \\ 0 & -\lambda_{12} & \lambda_{12} \\ 0 & 0 & 0 \end{pmatrix} \tag{11}$$

where $\lambda_{01}, \lambda_{02}, \lambda_{12}, \in (0, +\infty)$ are unknown model parameters.

Let $\xi = inf\{t \in \mathbf{R}_+: X_t = 2\}$ be the observable failure time when the lubricating oil no longer fulfills its functions. The residual process $(Y_{n\triangle}: n \in \mathbf{N})$ defined in Eq. (10) is assumed to be conditionally independent given the state of the lubricating oil, and for each $n \in \mathbf{N}$, we assume that $Y_{n\triangle}$, conditional on $X_{n\triangle} = x$, has bivariate normal distribution $N_2(\mu_x, \Sigma_x)$, where x=0,1, with the density given by

$$f_{Y_{n\Delta}|X_{n\Delta}}(y|x) = \frac{1}{\sqrt{(2\pi)^2 \det(\Sigma_x)}} \exp\left(-\frac{1}{2}(y-\mu_x)'\Sigma_x^{-1}(y-\mu_x)\right) \tag{12}$$

where $\mu_x \in \mathbf{R}^2$ and $\Sigma_x \in \mathbf{R}^{2\times 2}$, $(x = 0, 1)$ are unknown parameters of the observation process.

When the oil fails, $P(Y_{n\triangle} = \eta | X_{n\triangle} = 2) = 1$, $\eta \notin \mathbf{R}^2$ is the failure signal obtained by the OLVF sensor. We denote the N=11 failure oil data histories as $\{F_1, F_2, \ldots F_N\}$, with the form of $\mathbf{Y}_i = \left(y_1^i, y_2^i, \ldots y_{T_i}^i\right)$, $(i = 1, 2, \ldots N)$ for each failure history F_i. The observable failure time ξ_i for each history is $\xi_i = t_i$, $(T_i\triangle < t_i \le (T_i+1)\triangle)$. The sampling history $\mathbf{Y}_i$ consists of the residuals $y_t^i \in \mathbf{R}^2$, $t \le T_i$, calculated from the collection of oil data until the lubricating oil fails at time t_i.

Similarly, we denote the M=16 suspension oil data histories as $\{S_1, S_2, \ldots S_M\}$, with the form of $\mathbf{Y}_j = \left(y_1^j, y_2^j, \ldots y_{T_j}^j\right)$, $(j = 1, 2, \ldots M)$ for each suspension history S_j. The failure time ξ_j is unobservable, where $\xi_j = t_j > T_j\triangle$. The sampling history $\mathbf{Y}_j$ consists of the residuals $y_t^j \in \mathbf{R}^2$, $t \le T_j$, calculated from the collection of oil data but the lubricating oil is still in operation at time t_j, and has not failed yet.

Let $O = \{F_1, F_2, \ldots F_N, S_1, S_2, \ldots S_M\}$ represents a vector of all the observable lubricating oil data, and $L(\lambda, \theta|O)$ be the associated likelihood function, where $\lambda = (\lambda_{01}, \lambda_{02}, \lambda_{12})$ are the unknown state parameters and $\theta = (\mu_0, \mu_1, \Sigma_0, \Sigma_1)$ are the unknown observation parameters. Expectation-Maximization (EM) algorithm is adopted by iteratively maximizing the so-called pseudo likelihood function to estimate the unknown parameters.

EM algorithm consists of two steps: E-step and M-step, where E-step computes the expectation of the associated likelihood function and M-step obtains the maximization of the unknown states and observation parameters. The E-steps and M-steps are repeated until the Euclidean distance $\left|(\lambda^*, \theta^*) - \left(\hat{\lambda}, \hat{\theta}\right)\right| < \varepsilon$, where $\hat{\lambda}, \hat{\theta}$ are the estimates of the unknown parameters from the former M-step, λ^*, θ^* are the new parameters estimated from the current M-step, and ε is the selected stopping criterion value. More specifically, the EM algorithm works as follows.

E-step: the pseudo likelihood function is defined by

$$Q\left(\lambda, \theta|\hat{\lambda}, \hat{\theta}\right) := E_{\hat{\lambda},\hat{\theta}}(\ln L(\lambda, \theta|C)|O) \tag{13}$$

where $C = \{\bar{F}_1, \ldots \bar{F}_N, \bar{S}_1, \ldots \bar{S}_M\}$ represents the complete lubricating oil data histories.

M-step: compute λ^*, θ^* by

$$\lambda^*, \theta^* \in \underset{\lambda,\theta}{\operatorname{argmax}} Q(\lambda, \theta|\hat{\lambda}, \hat{\theta}) \tag{14}$$

Explicit formulas for the computation of the pseudo likelihood function defined in Eq. (13) and the unique maximums of the unknown state and observation parameters can be found in [11,17].

With the stopping criterion $\left|(\lambda^*, \theta^*) - \left(\hat{\lambda}, \hat{\theta}\right)\right| < 10^{-4}$, we obtained the results shown in Table 3.

Thus, the deterioration state of lubricating oil is modeled as a continuous time homogeneous Markov chain $(X_t: t \in \mathbf{R}_+)$ with

Table 3
Iterations of the EM algorithm.

Parameters	Initial values	First iteration	Second iteration	Final iteration
λ_{01}	0.01	0.0118	0.0132	0.0189
λ_{02}	0.0005	0.00056	0.00059	0.000047
λ_{12}	0.5	0.4884	0.4572	0.1812
μ_0	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -0.0011 \\ 4.31e-4 \end{pmatrix}$	$\begin{pmatrix} -9.64e-4 \\ 3.99e-4 \end{pmatrix}$	$\begin{pmatrix} 1.25e-4 \\ -8.16e-4 \end{pmatrix}$
μ_1	$\begin{pmatrix} 0.05 \\ 0.02 \end{pmatrix}$	$\begin{pmatrix} 0.0036 \\ -0.0020 \end{pmatrix}$	$\begin{pmatrix} 0.0030 \\ -0.0067 \end{pmatrix}$	$\begin{pmatrix} 0.0197 \\ -0.0131 \end{pmatrix}$
Σ_0	$\begin{pmatrix} 0.02 & 0.02 \\ 0.02 & 0.05 \end{pmatrix}$	$\begin{pmatrix} 0.003088 & 0.000266 \\ 0.000266 & 0.001828 \end{pmatrix}$	$\begin{pmatrix} 0.003068 & 0.000264 \\ 0.000264 & 0.001829 \end{pmatrix}$	$\begin{pmatrix} 0.002767 & 0.000288 \\ 0.000288 & 0.001680 \end{pmatrix}$
Σ_1	$\begin{pmatrix} 0.03 & 0.02 \\ 0.02 & 0.05 \end{pmatrix}$	$\begin{pmatrix} 0.003893 & 0.000823 \\ 0.000823 & 0.001858 \end{pmatrix}$	$\begin{pmatrix} 0.003604 & 0.001222 \\ 0.001222 & 0.001984 \end{pmatrix}$	$\begin{pmatrix} 0.002966 & 0.002096 \\ 0.02096 & 0.002077 \end{pmatrix}$
Q	−81.3594	−168.3331	−169.7593	−329.8362
Time/s	0.25	0.4063	0.5313	2.3281

three states $\{0, 1, 2\}$. The transition rate matrix is given by

$$\Lambda = \begin{pmatrix} -0.01894 & 0.018893 & 0.000047 \\ 0 & -0.18116 & 0.181160 \\ 0 & 0 & 0 \end{pmatrix} \tag{15}$$

and the observations $Y_{n\Delta}$ conditional on $X_{n\Delta} = x$, $x = 0, 1$, have bivariate normal distribution $N_2(\mu_0, \Sigma_0)$ for healthy portions of the oil data histories and $N_2(\mu_1, \Sigma_1)$ for unhealthy portions, where

$$\mu_0 = \begin{pmatrix} 0.000125 \\ -0.000816 \end{pmatrix}, \Sigma_0 = \begin{pmatrix} 0.002767 & 0.000288 \\ 0.000288 & 0.001680 \end{pmatrix}$$
$$\mu_1 = \begin{pmatrix} 0.0197 \\ -0.0131 \end{pmatrix}, \Sigma_1 = \begin{pmatrix} 0.002966 & 0.002096 \\ 0.002096 & 0.002077 \end{pmatrix} \tag{16}$$

In the next section, we use the estimated parameters of the model to compute the CRF and MRL functions for RUL prediction of lubricating oil based on the posterior probability that the oil is in state 1 (warning state), which is not observable.

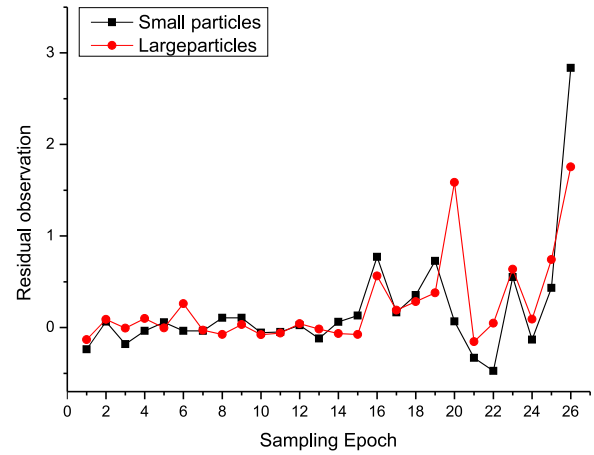

Fig. 4. The residual observation process of real failure history. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

4. CRF and MRL functions for RUL prediction

Explicit formulas of the CRF and the MRL functions of the model proposed in Section 3 for the RUL prediction of lubricating oil will be considered in this section, and as it is shown in [3], both formulas are functions of the posterior probability of the lubricating oil being in the warning state (state 1).

We assume that the deteriorating state process of lubricating oil is described by a continuous-time homogeneous Markov chain $(X_t: t \in \mathbf{R}_+)$, with state space $Z = \{0, 1, 2\}$, where state 0 denotes the state that the lubricating oil is working in a healthy condition, state 1 denotes the state that the lubricating oil is operating in a warning condition, and state 2 represents the failure or absorbing state of lubricating oil, which indicates that it requires an immediate replacement. The transition rate matrix of the state process is given by Eq. (15). Moreover, lubricating oil starts in state 0 and runs on a continuous basis. While the lubricating oil is in service, the residual observation process $(Y_{n\Delta}: n \in \mathbf{N})$, which is obtained from Eq. (10) using the collected data through condition monitoring at equidistant sampling times $\Delta, 2\Delta, \ldots n\Delta$, for $\Delta = 40$ min, has a state-dependent multivariate normal distribution defined in Eqs. (15,16). By solving the Kolmogorov backward differential equations [20], the transition probability matrix for the state process X_t has the following form:

$$\mathbf{P}(t) = P_{ij}(t) = \begin{bmatrix} e^{-(\lambda_{01}+\lambda_{02})t} & \dfrac{\lambda_{01}(e^{-\lambda_{12}t} - e^{-(\lambda_{01}+\lambda_{02})t})}{\lambda_{01}+\lambda_{02}-\lambda_{12}} & 1 - e^{-(\lambda_{01}+\lambda_{02})t} - \dfrac{\lambda_{01}(e^{-\lambda_{12}t} - e^{-(\lambda_{01}+\lambda_{02})t})}{\lambda_{01}+\lambda_{02}-\lambda_{12}} \\ 0 & e^{-\lambda_{12}t} & 1 - e^{-\lambda_{12}t} \\ 0 & 0 & 1 \end{bmatrix} \tag{17}$$

where the transition probabilities $P_{ij}(t) = P(X_t = j | X_0 = i)$, $i, j \in \{0, 1\}$.

We will show that the CRF and MRL functions can be expressed in terms of the posterior probability statistic Π_n, denoting the posterior probability that the lubricating oil is in warning state (state 1) given all available information until time $n\Delta$, which is defined as

$$\Pi_n = P(X_{n\Delta} = 1 | \xi > n\Delta, Y_\Delta, Y_{2\Delta}, \ldots, Y_{n\Delta}) \tag{18}$$

where $\Pi_0 = P(X_0 = 1) = 0$. Using Bayes' rule for $n \geq 1$, we have

$$\Pi_n = \frac{g(\mathbf{y}_n | X_{n\Delta} = 1, \xi > n\Delta, \mathbf{y}_{n-1}) \times P(X_{n\Delta} = i | \xi > n\Delta, \mathbf{y}_{n-1})}{\sum_j g(\mathbf{y}_n | X_{n\Delta} = j, \xi > n\Delta, \mathbf{y}_{n-1}) \times P(X_{n\Delta} = j | \xi > n\Delta, \mathbf{y}_{n-1})} \tag{19}$$

where

$$g(\mathbf{y}_n | X_{n\Delta} = i, \xi > n\Delta, \mathbf{y}_{n-1}) = \begin{cases} f(\mathbf{y}_n | \mu_0, \Sigma_0), & i = 0 \\ f(\mathbf{y}_n | \mu_1, \Sigma_1), & i = 1 \end{cases} \tag{20}$$

Therefore, we have

$$\Pi_n = \frac{(P_{01}(\Delta)(1 - \Pi_{n-1}) + P_{11}(\Delta)\Pi_{n-1})}{\dfrac{f(\mathbf{y}_n | \mu_0, \Sigma_0)}{f(\mathbf{y}_n | \mu_1, \Sigma_1)} P_{00}(\Delta)(1 - \Pi_{n-1}) + (P_{01}(\Delta)(1 - \Pi_{n-1}) + P_{11}(\Delta)\Pi_{n-1})} \tag{21}$$

where P_{ij} are given in Eq. (17), and the ratio of normal densities has the following representation

$$\frac{f(\mathbf{y}_n | \mu_0, \Sigma_0)}{f(\mathbf{y}_n | \mu_1, \Sigma_1)} = \sqrt{\frac{|\Sigma_1|}{|\Sigma_0|}} \cdot \frac{\exp\left(-\frac{1}{2}(y_n - \mu_0)'\Sigma_0^{-1}(y_n - \mu_0)\right)}{\exp\left(-\frac{1}{2}(y_n - \mu_1)'\Sigma_1^{-1}(y_n - \mu_1)\right)} \tag{22}$$

For more details, see reference [17].

Suppose that at decision epoch n, the lubricating oil has not failed, i.e. $\xi > n\Delta$. For any $t \in [0, \Delta]$, the conditional reliability function (CRF), which denotes the probability that the lubricating oil will not fail by $n\Delta + t$, is defined as

$$\begin{aligned} R(t | \Pi_n) &= P(\xi > n\Delta + t | \xi > n\Delta, Y_\Delta, \ldots, Y_{n\Delta}, \Pi_n) \\ &= P_{00}(t) + P_{01}(t) + (P_{11}(t) - P_{00}(t) - P_{01}(t))\Pi_n \end{aligned} \tag{23}$$

Using the CRF in Eq. (23), the MRL function at the nth epoch can be obtained by the following formula:

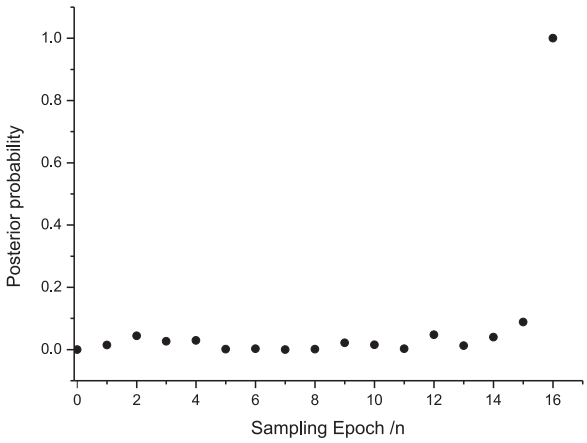

Fig. 5. Posterior probability of the typical failure oil history.

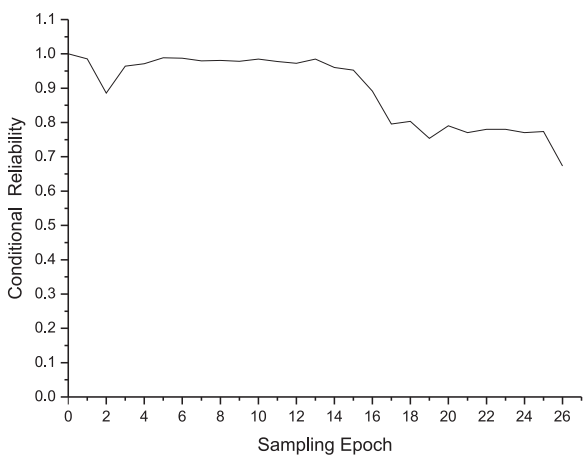

Fig. 6. The conditional reliability function of real failure history.

$$\mu_{n\Delta} = E(\xi - n\Delta | \xi > n\Delta, \vec{\Pi}_n) = \int_0^\infty R(t|\Pi_n)dt = \frac{\lambda_{12} + \lambda_{01} + \Pi_n(\lambda_{02} - \lambda_{12})}{\lambda_{12}(\lambda_{01} + \lambda_{02})} \quad (24)$$

The CRF and MRL functions have been widely used for RUL prediction [3]. (See e.g., [3] and the references in that paper for more details).

We apply the above formulas to one of our failure histories with bivariate residual observations, which is shown in Fig. 4. This failure history ended between 25*th* and 26*th* sampling epoch. The black solid line denotes the residual observations of small particles, and the red solid line denotes the residual observations of large particles.

Fig. 5 plots the posterior probability of lubricating oil data history presented in Fig. 4. The posterior probability was calculated using Eqs. (21,22). The plot shows that the lubricating oil started in the healthy state (state 0) at time 0, ran under the healthy condition till sometime between 1*st* and 15*th* sampling epoch, and after 16*th* sampling epoch, the lubricating oil was working in the warning state (state 1).

Fig. 6 provides the estimated conditional reliability at each sampling epoch for this failure oil data history shown in Fig. 4. It can be seen that the value of the reliability dropped at the 15*th* sampling epoch when the state changed to the warning state (state 1), and after 25*th* sampling epoch, the lubricating oil shifted to the failure state (state 2). It is shown that there is a high probability at that time that the state of the lubricating oil has changed, and the full inspection and a subsequent oil replacement action should be taken.

5. Conclusions and future researches

In this paper, we have considered the situation where the lubricating oil degradation is driven by a continuous time homogeneous Markov chain and the observation process is represented by a 2-dimensional on-line monitoring oil data obtained from a four-ball test rig, where the actual oil states are unobservable except the failure state. A vector autoregressive model has been fitted to the healthy portions of all oil data histories classified as failure histories and suspension histories. The residuals calculated from the VAR model are then used as the observation process in the HMM framework, and the state and the observation process parameter estimates of the lubricating oil have been obtained using the EM algorithm. With the unknown parameters estimated, RUL prediction of lubricating oil has been developed by deriving the explicit formulas of the CRF and MRL functions expressed in terms of the posterior probability for the purpose of future decision-making.

The HMM considered in this paper assumes that the sojourn times in both the healthy and warning states are exponentially distributed, which may not be realistic in some situations. Also, this paper studies only the degradation process of lubricating oil when the machine is in healthy condition, and the RUL from the 'good as new' state to the oil failure state. In future research, more general distributions of the sojourn times will be considered, such as Erlang of phase-type distributions, which will cover more real situations and provide better estimates of RUL when exponentially distributed sojourn times are not appropriate.

Acknowledgment

The financial support for the present research was provided by the National Science Foundation of China (Grant Nos. 51675403, 51275381), which is very much appreciated. Special thanks are also due to the China Scholarship Council (Grant No. 201506280150) for their financial support during the term of this project.

References

[1] Y. Du, T. Wu, J. Cheng, Age detection of lubricating oil with on-line sensors, in: SENSORS, 2015 IEEE, IEEE, 2015a, pp. 1–4.
[2] J. Zhu, J. Yoon, D. He, B. Qiu, E. Bechhoefer, Online condition monitoring and remaining useful life prediction of particle contaminated lubrication oil, in: Proceedings of the Prognostics and Health Management (PHM), 2013 IEEE Conference on, IEEE, 2013, pp. 1–14.
[3] A. Khaleghei, V. Makis, Model parameter estimation and residual life prediction for a partially observable failing system, Nav. Res. Logist. (NRL) 62 (3) (2015) 190–205.
[4] T. Wu, J. Mao, J. Wang, J. Wu, Y. Xie, A new on-line visual ferrograph, Tribol. Trans. 52 (5) (2009) 623–631.
[5] J. Zhu, J.M. Yoon, D. He, E. Bechhoefer, Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines, Wind Energy 18 (6) (2015) 1131–1149.
[6] N. Chen, K.L. Tsui, Condition monitoring and remaining useful life prediction using degradation signals: revisited, IIE Trans. 45 (9) (2013) 939–952.
[7] M.-Y. You, L. Li, G. Meng, J. Ni, Two-zone proportional hazard model for equipment remaining useful life prediction, J. Manuf. Sci. Eng. 132 (4) (2010) 041008.
[8] X.-S. Si, W. Wang, C.-H. Hu, M.-Y. Chen, D.-H. Zhou, A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation, Mech. Syst. Signal Process. 35 (1) (2013) 219–237.
[9] D. Lin, V. Makis, Filters and parameter estimation for a partially observable system subject to random failure with continuous-range observations, Adv. Appl. Probab. (2004) 1212–1230.
[10] R. Jiang, System availability maximization and residual life prediction under partial observations, (Ph.D. thesis), University of Toronto, 2011.

[11] M.J. Kim, V. Makis, R. Jiang, Parameter estimation for partially observable systems subject to random failure, Appl. Stoch. Model. Bus. Ind. 29 (3) (2013) 279–294.
[12] Z.-J. Zhou, C.-H. Hu, D.-L. Xu, M.-Y. Chen, D.-H. Zhou, A model for real-time failure prognosis based on hidden Markov model and belief rule base, Eur. J. Oper. Res. 207 (1) (2010) 269–283.
[13] V. Makis, Multivariate Bayesian control chart, Oper. Res. 56 (2) (2008) 487–496.
[14] A. Ghasemi, S. Yacout, M.-S. Ouali, Evaluating the reliability function and the mean residual life for equipment with unobservable states, IEEE Trans. Reliab. 59 (1) (2010) 45–54.
[15] Y. Du, T. Wu, J. Cheng, R. Gong, Lubricating oil deterioration on a four-ball test rig via on-line monitoring, in: Proceedings of Malaysian International Tribology Conference 2015, Malaysian Tribology Society, vol. 2015, 2015b, pp. 185–186.
[16] G.C. Reinsel, Elements of Multivariate Time Series Analysis, Springer Science & Business Media, Springer, New York, 2003.
[17] M.J. Kim, Optimal control and estimation of stochastic systems with costly partial information, (Ph.D. thesis), University of Toronto, 2012.
[18] N. Henze, B. Zirkler, A class of invariant consistent tests for multivariate normality, Commun. Stat.-Theory Methods 19 (10) (1990) 3595–3617.
[19] J. Yang, V. Make, Dynamic response of residuals to external deviations in a controlled production process, Technometrics 42 (3) (2000) 290–299.
[20] G. Grimmett, D. Stirzaker, Probability and Random Processes, Oxford University Press, Oxford, 2001.