Prognostics of slurry pumps based on a moving-average wear degradation index and a general sequential Monte Carlo method

Dong Wang, Peter W. Tse * 

Smart Engineering Asset Management Laboratory (SEAM) and Croucher Optical Non-destructive Testing and Quality Inspection Laboratory (CNDT), Department of Systems Engineering and Engineering Management, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, China

Abstract

Slurry pumps are commonly used in oil-sand mining for pumping mixtures of abrasive liquids and solids. These operations cause constant wear of slurry pump impellers, which results in the breakdown of the slurry pumps. This paper develops a prognostic method for estimating remaining useful life of slurry pump impellers. First, a moving-average wear degradation index is proposed to assess the performance degradation of the slurry pump impeller. Secondly, the state space model of the proposed health index is constructed. A general sequential Monte Carlo method is employed to derive the parameters of the state space model. The remaining useful life of the slurry pump impeller is estimated by extrapolating the established state space model to a specified alert threshold. Data collected from an industrial oil sand pump were used to validate the developed method. The results show that the accuracy of the developed method improves as more data become available.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Slurry pumps play a vital role in pumping mixtures of abrasive and erosive liquids and solids in oil sand mining operations. Such activity causes constant wear of slurry pump impellers. The failure of the slurry pump impellers is the main reason for the breakdown of slurry pumps, which results in significant economic losses. The current and future health conditions of the slurry pump impellers must be assessed immediately to prevent unexpected downtime.

Walker and Bodkin [1] investigated the empirical wear relationship of a slurry pump impeller and revealed that solid particle sizes, slurry concentration and pump speeds greatly influence wear rate. Li et al. [2] conducted a failure analysis of a slurry pump impeller and demonstrated that duplex stainless steel with an equal austenite/ferrite ratio can resist corrosive wear. Bross and Addiei [3] proposed a model to predict the influence of different impeller design parameters on wear behavior. Xing et al. [4] used a finite element analysis tool, namely, ANSYS, to simulate the wear process of flow components and found that pits on the surface of flow components are caused by particle impact. A systematic study on the failure
analysis of slurry pump impellers revealed that weight loss of corrosive wear is influenced by impact velocity, and that impeller failure is mainly caused by wear \[5\]. These results aid in improving the design of slurry pumps and in predicting wear degrees under steady working conditions. However, the wear relationships established by using the aforementioned methods may not be useful in evaluating the current and future health conditions of slurry pump impellers in practice because of uncontrollable working conditions. Therefore, developing online methods for evaluating impeller health condition is necessary.

The University of Alberta recently collaborated with Canadian oil sand mining industry on a series of research that assesses impeller health condition. An experimental system designed by Wang et al. \[6\] was used at the early stage to provide controllable working variables in studying the wear process of slurry pump impellers. Different damage modes with different wear degrees were produced artificially on the slurry pump impellers. Intelligent impeller fault diagnosis and prognosis methods were developed by using data collected from the experimental system. A support vector machine, a novel data cleaning algorithm and a classical sequential backward feature selection were combined to classify four different impeller damages: hole-through damage, vane trailing edge damage, vane leading edge damage, and expeller vane damage \[7\]. Their results illustrated that the data cleaning algorithm is effective in improving identification accuracy. Qu and Zuo \[8\] then developed a least squares support vector regression-based fault diagnosis method to evaluate impeller wear degrees and to provide a quantitative description for wear degrees. Zhao et al. \[9\] developed a modified neighborhood rough set model to select useful features for impeller fault identification. The results found that the selected features can be used to achieve a higher classification rate than the features generated by the original neighborhood rough set model. The combination of half and full spectra, fuzzy preference-based rough sets and principle component analysis was then developed to generate a monotonic health indicator to describe impeller health condition. However, these developed online impeller health condition evaluation methods were validated by data collected from the experimental system with some artificial damages. The data do not fully reflect the natural wear propagation of slurry pump impellers. Our literature review reveals that a health evaluation of slurry pump impellers that uses natural wear data remains lacking \[10\]. In this paper, we developed a prognostic method to analyze industrial slurry pump data, which facilitates an assessment of the natural wear of the slurry pump impellers.

The developed method consists of two steps. The first step aims to assess the performance degradation of a slurry pump impeller. Such an assessment monitors the current condition of a component or system. This step aims to assess the deviation of the current condition of the component or system from its normal condition. Our literature review shows that numerous methods were recently developed to evaluate the health condition of bearings and gears. Qiu et al. \[11\] used an optimal wavelet filter to enhance weak bearing fault signatures and employed a self-organizing map to track bearing defect development. Wang et al. \[12\] developed a rapid performance degradation assessment method based on discrete wavelet transform to evaluate gears. Wang et al. \[13\] employed a complex Morlet wavelet transform to analyze gear motion residual signal to assess gear health condition under various load conditions. Lin et al. \[14\] developed a weighted fault growth parameter based on gear residual error signal to track gear condition. Ocak et al. \[15\] used wavelet packet node energies to train a normal hidden Markov model. The probabilities of the trained hidden Markov model were used to track bearing health condition. A similar idea was used by Miao et al. \[16\] to describe gear health evolution, but, with the application of an adaptive signal processing method, namely, empirical mode decomposition \[17\], to extract gear fault features. Liao and Lee \[18\] proposed a novel degradation assessment method based on data collected from transient periods of different working loads. Pan et al. \[19\] combined wavelet packet transform and a fuzzy c-means to assess bearing health condition and then developed a hybrid method \[20\], that consists of a support vector data description and a fuzzy c-means, to evaluate bearing health condition. Wang et al. \[21\] used a series of wavelet filters to extract gear fault features and employed a support vector data description to track the current health condition of a gear. In the work, two health indicators were developed to identify an early gear fault and to assess gear degradation. The use of these two health indicators is reasonable because gear performance degradation assessment is insignificant until an early gear fault is detected. Zhu et al. \[22\] developed an incremental rough support vector data description method for assessing the performance degradation of a bearing. Yu \[23\] developed locality preserving projections-based Gaussian mixture models to track bearing health condition. Miao et al. \[24\] constructed a wavelet filter bank to extract bearing fault features in describing fault propagation in fan bearings. To explore the performance degradation assessment of the slurry pump impeller, a health indicator called moving-average wear degradation index (MAWDI) is proposed in this paper to describe the current health condition of the slurry pump impeller.

Based on the proposed MAWDI, the second step aims to estimate the remaining useful life (RUL) of the slurry pump impeller. RUL estimation is the prediction of the period from the current time until the component or system no longer satisfies its functionality \[25–27\]. RUL estimation aids in conducting maintenance activities, providing spare parts on time, and preventing accidents. A general sequential Monte Carlo method, particularly a general particle filter, has recently been applied to derive the posterior probability functions of state parameters of a state space model given known measurements. The established state space model is then used for component or system prognosis \[28\]. For example, Sun et al. \[29\] employed a particle filter to estimate the RUL of a gas turbine. He et al. \[30\] used the Dempster–Shafer theory to initialize the sum of two exponential functions and employed a particle filter to estimate the RUL of lithium-ion batteries. Following the work of He et al., Miao et al. \[31\] used an unscented particle filter for RUL estimation of lithium-ion batteries. Xing et al. \[32\] developed an ensemble lithium-ion battery state space model and used a particle filter to estimate the parameters of the state space model and the RUL of the lithium-ion batteries. Zio et al. \[33,34\] applied a
particle filter to estimate the parameters of a fatigue crack growth model and to infer RUL of the fatigue crack. Chen et al. [35] developed a high-order particle filter based on a high-order Markov assumption to predict the RUL of carrier plates and bearings. In the current paper, the state space model of the MAWDI is constructed. The parameters of the state space model are derived by the general particle filter given some pump vibration measurements. An extrapolation of the state space model to a specified alert threshold is used to estimate the RUL of the slurry pump impeller.

The rest of this paper is organized as follows. In Section 2, the principle of the general particle filter is introduced. The prognostic method of the slurry pump impeller is developed in Section 3. An industrial oil sand pump prognostic case is studied in Section 4. Conclusions are drawn in Section 5.

2. Introduction of a general particle filter

2.1. Nonlinear Bayesian tracking

The evolution of a system state sequence \( x_k \) is represented by

\[
x_k = f_k(x_{k-1}, v_k), \tag{1}
\]

where \( f_k(\cdot) \) is a state evolution function, which may be linear or nonlinear, and \( v_k \) is an independent and identically distributed (i.i.d.) process noise sequence.

The aim of tracking is to recursively estimate \( x_k \) from a measurement sequence \( z_k \). The measurement function \( h_k(\cdot) \), which may be linear or nonlinear, is given as follows:

\[
z_k = h_k(x_k, n_k), \tag{2}
\]

where \( n_k \) is an independent and identically distributed (i.i.d.) measurement noise sequence.

The tracking based on a Bayesian view requires the construction of a probability density function (PDF) \( p(x_k|z_{1:k}) \), given measurements \( z_{1:k} \) up to the time \( k \). For deriving an estimate of the PDF, \( p(x_k|z_{1:k}) \), three assumptions are made as follows. First, an initial PDF \( p(x_0|z_0) = p(x_0) \) is known. Second, the evolution of the system state sequence is a Markov process of order one. Third, the measurements are conditionally independent of the measurements and the other states.

A recursive estimate of the PDF \( p(x_k|z_{1:k}) \) can be obtained by the prediction and update steps [36].

(1) Prediction step. Suppose the PDF \( p(x_{k-1}|z_{1:k-1}) \) is available. The Chapman–Kolmogorov equation shows a prior PDF of the state \( x_k \) as follows:

\[
p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}. \tag{3}
\]

(2) Update step. When a new measurement \( z_k \) becomes available, an update of the prior PDF of the state \( x_k \) can be calculated via a Bayes rule:

\[
p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} = \frac{p(z_k|x_k) \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}}{\int \int p(z_k|x_k)p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1}dx_k}. \tag{4}
\]

Three points should be noted. First, the PDF \( p(x_k|x_{k-1}) \) is determined by the characteristics of the \( f_k(\cdot) \) and the known statistics of the \( v_k \). Second, the PDF \( p(z_k|x_k) \) is determined by the characteristics of the \( h_k(\cdot) \) and the known statistics of the \( n_k \). Third, the recursive estimate of the PDF \( p(x_k|z_{1:k}) \) is not analytically determined in general. If Eqs. (1) and (2) are restricted to linear functions and Gaussian distribution, Kalman filter and grid-based filters can be used to analytically derive the optimal Bayesian state estimate. However, if the linear functions and the Gaussian distribution are not satisfied, an analytic solution is difficult to be tractable. The optimal Bayesian state estimate must be approximated by other methods, such as extended Kalman filters, approximate grid-based filters and particle filters [36].

Once the PDF \( p(x_k|z_{1:k}) \) is determined, all statistical inferences of a future state \( x_{k+1} \) can be drawn. For example, the PDF \( p(x_{k+1}|z_{1:k}) \) of a state \( x_{k+1} \) can be calculated by

\[
p(x_{k+1}|z_{1:k}) = \int p(x_{k+1}|x_{k+1-1})p(x_{k+1-1}|z_{1:k})dx_{k+1-1} = \int \cdots \int p(x_i|x_{i-1})p(x_i|z_{1:k})dx_i. \tag{5}
\]

A point estimate of the state \( x_{k+1} \) can be calculated by taking the conditional mean of Eq. (5):

\[
E\{x_{k+1}|z_{1:k}\} = \int x_{k+1}p(x_{k+1}|z_{1:k})dx_{k+1}
\]
2.2. A general particle filter for nonlinear Bayesian state tracking

A particle filter is a method to implement a recursive Bayesian filter by using Monte Carlo simulations. Suppose $N_s$ random particles $\{x^i_k\}_{i=1}^{N_s}$ with their associated weights $\{\omega^i_k\}_{i=1}^{N_s}$ characterize the posterior PDF $p(x_k|z^{1:k})$ introduced in Section 2.2. Additionally, the weights are normalized so that the summation of the weights is equal to 1. The key idea of the particle filter is that the random particles with their associated weights are used to represent the true posterior density function $p(x_k|z^{1:k})$ via the following equation:

$$p(x_k|z^{1:k}) \approx \sum_{i=1}^{N_s} \omega^i_k \delta(x_k - x^i_k),$$

(7)

where $\delta(\cdot)$ is the Dirac delta function. As the number of the random particles increases, the posterior density function approximated by the random particles is gradually equal to the true posterior density function [37,38]. If the random particles are directly drawn from the true density function $p(x_k|z^{1:k})$, Eq. (7) can be expressed by

$$p(x_k|z^{1:k}) \approx \frac{1}{N} \sum_{i=1}^{N_s} \delta(x_k - x^i_k).$$

(8)

However, in many practical cases, it is intractable to directly draw the random particles from the true posterior density function $p(x_k|z^{1:k})$. Suppose the true posterior density function $p(x_k|z^{1:k})$ is proportional to a function $\pi(x_k|z^{1:k})$ which can be analytically evaluated. The random particles $x^i_k$, $i = 1, \ldots, N_s$ can be generated from another proposal distribution $q(x_k|z^{1:k})$, which is also known as an importance function. Then, the density function $p(x_k|z^{1:k})$ can be approximated by Eq. (7) and the weight $\omega^i_k$ is expressed by

$$\omega^i_k \propto \frac{\pi(x^i_k|z^{1:k})}{q(x^i_k|z^{1:k})} \propto \frac{p(x^i_k|z^{1:k})}{q(x^i_k|z^{1:k})}$$

(9)

Assume the proposal distribution can be factorized as

$$q(x^i_k|z^{1:k-1}) = q(x^i_k|z^{1:k-1}, z_k)q(z_k|z^{1:k-1}).$$

(10)

Consider the fact that $p(x^i_k|z^{1:k})$ can be represented by [37,38]

$$p(x^i_k|z^{1:k}) = p(z_k|x^i_k)p(x^i_k|z^{1:k-1})p(x^{i-1}_k|z^{1:k-1}).$$

(11)

Substituting Eqs. (10) and (11) into Eq. (9), the weight $\omega^i_k$ can be iteratively updated by

$$\omega^i_k \propto \frac{p(x^i_k|z^{1:k})}{q(x^i_k|z^{1:k})} \propto \frac{p(z_k|x^i_k)p(x^i_k|z^{1:k-1})p(x^{i-1}_k|z^{1:k-1})}{q(x^i_k|z^{1:k-1}, z_k)q(z_k|z^{1:k-1})} \propto \omega_k^{i-1} \frac{p(z_k|x^i_k)p(x^i_k|z^{1:k-1})}{q(x^i_k|z^{1:k-1}, z_k)}.$$

(12)

If $q(x_k^i|z_k^{i-1}, z_k) = q(x_k^i|z_k^{i-1}, z_k)$ is satisfied, Eq. (12) can be simplified as

$$\omega^i_k = \omega_k^{i-1} \frac{p(z_k|x^i_k)p(x^i_k|z_k^{i-1})}{q(x^i_k|z_k^{i-1}, z_k)} \left( \frac{\sum_{i=1}^{N_s} \omega_k^{i-1} p(z_k|x^i_k)p(x^i_k|z_k^{i-1})}{\sum_{i=1}^{N_s} q(x^i_k|z_k^{i-1}, z_k)} \right).$$

(13)

From Eq. (13), the proposal distribution only depends on the previous state and the current measurement. Therefore, computational memory storage can be reduced. However, the major problem of Eq. (13) is that most of the weights become negligible and the variances of the weights increase over time after a few iterations. In other words, many computational efforts are used to update the random particles which have little contribution to Eq. (7). This phenomenon is called a degeneracy problem. In order to relieve the effects of the degeneracy problem, a resampling is often used to redraw the random particles from the approximate distribution of $p(x_k|z^{1:k})$ according to the size of $\omega^i_k$. After the resampling is conducted, Eq. (7) is transformed to Eq. (8). Here, the random particles used in Eq. (8) are directly drawn from the approximate posterior distribution of Eq. (7) rather than the true posterior distribution. In order to enhance computing efficiency, the resampling is necessary to be conducted only when the weights have a large variance. An effective sample size (ESS) criterion is frequently used to judge when the resampling is required. The formula of the ESS is given by [37,38]

$$\text{ESS} = \left( \frac{1}{\sum_{i=1}^{N_s} \omega^i_k} \right)^{-1}.$$  

(14)
The ESS varies from 1 to $N_r$. The resampling is necessary only when the ESS is below a threshold, such as the half of the $N_r$.

3. A prognostic method for estimating the remaining useful life of a slurry pump impeller

When a slurry pump operates over time, its impeller undergoes constant wear caused by mixture of abrasive and erosive liquids and solids in the slurry pump. Severe wear of the slurry pump impeller is the main cause of slurry pump breakdowns. Estimation of remaining useful life of the slurry pump impeller can enhance the reliability of the slurry pump and prevent unexpected downtime. A slurry pump impeller prognostic method is developed in this section and is applied to estimate the remaining useful life of an oil sand pump impeller. The procedure of the developed slurry pump impeller prognostic method is described in Fig. 1. The details of the procedure are illustrated in the following sections.

3.1. Performance degradation assessment of a slurry pump impeller

Spectrum analysis of pump vibration measurements is a simple and effective method to assess pump health condition [39] because the frequency spectra of a normal pump and an abnormal pump have fundamental differences, which can be potentially used to indicate the health evolution of a slurry pump. In this paper, vibration data collected from an industrial oil sand pump were used to test our developed slurry pump impeller prognostic method. The industrial oil sand pump was driven by a motor with a rotation frequency $f_m$ equal to 26 Hz and was stepped down through a gearbox. The pump rotation frequency $f_p$ was calculated as 6.62 Hz. The vane-passing frequency $f_{vpf}$ was calculated as 26.48 Hz by multiplying the pump rotation frequency by four (four impeller blades). The gear meshing frequency $f_{gmf}$ was calculated as 362 Hz. These estimated frequencies changed over time because the data collected from the industrial oil sand pump were influenced by some uncontrollable conditions. The pump vibration measurements were collected by using the smart asset management system (SAMS). The data acquisition equipment, which consisted of a National Instrument (NI) DAQ 9172 and a DAQ module NI 9234, was used. Four accelerometers mounted at four different locations are shown in Fig. 2. The PCB 352A60 accelerometers

![Procedure for estimating the remaining useful life of a slurry pump impeller](image-url)
The data collected from location C3 were used for the analyses in this paper. Sand pump measurements, each measurement is normalized by Eq. (17) as the value of the parameter smoothness and the computing time, the parameter smoothness. In Fig. 4(a), the fluctuation of the EE is small at the beginning, and then becomes large as the document number increases.

The sampling frequency was set to 51,200 Hz. The vibration data length \( L \) for each measurement was equal to 51,200 samples. The data collected from location C3 were used for the analyses in this paper.

Denote \( N \) successive slurry pump vibration measurements as \( y_k(t), \ k = 1, 2, \ldots, N \). To remove statistical error from the oil sand pump measurements, each measurement is normalized by

\[
y_k(t) = \left( y_k(t) - \frac{\sum_{t=1}^{L} y_k(t)}{L} \right) \sqrt{\frac{\sum_{t=1}^{L} \left( y_k(t) - \frac{\sum_{t=1}^{L} y_k(t)}{L} \right)^2}{L-1}}, \quad k = 1, 2, \ldots, N. \tag{15}
\]

The Fourier transform of Eq. (15) is given by

\[
y_k(f) = \sum_{t=1}^{L} y_k(t) e^{-2\pi i t (t - 1) \frac{1}{L}}, \quad k = 1, 2, \ldots, N. \tag{16}
\]

According to references [6–9,39], it is found that the vane-passing frequency and its harmonics are highly related to impeller wear evaluation. Therefore, it is possible to use the summation of the amplitudes of these frequencies as a fault feature for reflecting the impeller health evolution. To validate this point, oil sand slurry pump measurements at different document numbers were analyzed. The frequency spectra of the oil sand pump measurements at three different document numbers were plotted in Fig. 3(a), (b) and (c) respectively, where it is obviously found that the amplitude of the vane-passing frequency increases over time.

Because the vane-passing frequency varies over time, it is more convenient to use the summation of the amplitudes of the frequency band that covers the vane-passing frequency and its harmonics as a fault feature to track the health condition of the oil sand pump impeller. The fault feature is named as energy evolution (EE) and is defined as follows:

\[
EE(k) = \sum_{f=f_1}^{f_2} \sum_{k=K}^{K+1} \frac{y_k(f)}{K}, \quad k = K, K+1, \ldots, N, \tag{17}
\]

where \( f_1 \) and \( f_2 \) are the lower and higher cut-off frequencies of the frequency band. By inspecting the frequency spectra shown in Fig. 3, the lower and higher cut-off frequencies were set to 20 Hz and 80 Hz, respectively, to sufficiently cover the vane-passing frequency and its harmonics. This frequency band is highlighted by the rectangle with the dotted line in Fig. 3. \( K \) is a moving-average number and controls the smoothness of Eq. (17). The larger the parameter \( K \) is, the smoother Eq. (17) is. As the value of the parameter \( K \) increases, the computing time increases. Considering the trade-off between the smoothness and the computing time, the parameter \( K \) was empirically set to 5. Fig. 4(a) shows the evolution of the EE. In Fig. 4(a), the fluctuation of the EE is small at the beginning, and then becomes large as the document number increases.

In order to extract the central tendency of the EE, a moving-average wear degradation index (MAWDI) is proposed as follows:

\[
MAWDI(j) = \log \left( \frac{\sum_{k=K}^{K+1} EE(k)}{J-K+1} \right) = \log \left( \frac{\sum_{k=K}^{K+1} \sum_{f=f_1}^{f_2} \sum_{k=K}^{K+1} \frac{y_k(f)}{K}}{J-K+1} \right), \quad k = K, K+1, \ldots, N. \tag{18}
\]
Fig. 3. Frequency spectra of the oil sand pump vibration measurements (a) at document number 33; (b) at document number 338; and (c) at document number 561.

Fig. 4. Health assessment of a slurry pump impeller by using (a) the energy evolution and (b) the moving-average wear degradation index.
Fig. 4(b) shows the evolution of the MAWDI. The MAWDI used for the performance degradation assessment of the oil sand impeller only becomes meaningful as the fluctuation of the EE gradually increases. It means that the performance degradation assessment only becomes significant after the inspected component or system enters in an abnormal condition [21]. Therefore, it is reasonable to assume that the performance degradation assessment of the oil sand pump impeller begins at document number 63, which is observed in Fig. 4(b). In other words, the remaining useful life estimation of the oil sand pump impeller starts at document number 63.

3.2. A state space model and its parameters updating by using a particle filter

In Section 3.1, it is discovered that the MAWDI can reflect the central tendency of wear of the oil sand pump impeller. In order to reflect the evolution of the MAWDI, a state space model is constructed as follows:

\[
\begin{align*}
    x_k &= x_{k-1} + v_{k-1}, \quad v_{k-1} \sim N(0, \sigma_1^2), \\
    y_k &= y_{k-1} + u_{k-1}, \quad u_{k-1} \sim N(0, \sigma_2^2), \\
    z_k &= x_k + a_0^k + n_k, \quad n_k \sim N(0, \sigma_3^2).
\end{align*}
\]

The exponential function used in Eq. (21) assumes the monotone evolution of the MAWDI. This construction is inspired by the accumulative property of impeller wear. Eqs. (19) and (20) are the parameter evolution of Eq. (21). In order to illustrate the feasibility of the exponential function used for the MAWDI fitting, the MAWDI was processed by a nonlinear least squares regression [40]. Since the MAWDI tends to be stable at the value of 10.3 at the final stage of the MAWDI, an alert threshold of the MAWDI was set to 10.3 in this paper. Because the performance degradation assessment begins at document number 63 and the alert threshold is 10.3, only the MAWDI from document numbers 63 to 824 is necessary to be regressed by the exponential function. A fitted curve by using the specific data and the exponential function is plotted with the thick line in Fig. 4(b). The x and y were estimated as 9.8360 and 6.252 × 10⁻⁶, respectively. Goodness of fit statistics was used to quantify the fitting performance of the exponential function. A root mean squared error (RMSE), A R², and an adjusted R² were calculated as 0.01363, 0.9904 and 0.9904. The three statistical values demonstrate that the exponential function is capable of fitting the MAWDI. The closer the value of the RMSE is to 0, the better the performance of the model is. In addition, the closer the values of the R² and the adjusted R² are to 1, the better the performance of the exponential function is. When MAWDI(k), k = K, ..., M, here M < N, are available, the nonlinear least squares regression can be also used to provide an initial estimate of the x₀ and y₀. According to the principle of the particle filter introduced in Section 2.2, the two state parameters x₀ and y₀ can be iteratively updated by the following steps.

**Step 1.** Draw N₀ initial random particles \( \{x_0^i\}_{i=1}^{N_0} \) and \( \{y_0^i\}_{i=1}^{N_0} \) of parameters x₀ and y₀ from the distributions of q(x₀|z₁) = N(x₀, σ₁²) and q(y₀|z₁) = N(y₀, σ₂²), respectively. Because these random particles are iteratively used to approximate a posterior probability density function, the accuracy of such Monte Carlo approximation is determined by the number of random particles. The larger the number of random particles is, the closer the Monte Carlo approximation is to a true probability density function. However, lots of random particles increase the computing burden. Here, the N₀ was set to 3000 and its sufficiency can be seen in Section 4 by matching the values predicted by the general particle filter with MAWDI. σ₁ and σ₂ were empirically set to 0.002 and 0.0002, respectively, according to the scales of the initial estimates of x₀ and y₀ introduced in Section 4. Initialize their associated weights \( \{a_0^i\}_{i=1}^{N_0} \) to an equal value of 1/N₀. Because the importance density is often conveniently chosen to be the prior q(x₀|x₀⁻¹, zₖ) = p(x₀|x₀⁻¹), the weight updating becomes the following equation:

\[
a_0^k = a_{k-1}^i p(z_k|x_k^i, y_k^i) / \left( \sum_{i=1}^{N_0} a_{k-1}^i p(z_k|x_k^i, y_k^i) \right).
\]

Then, according to Eqs. (21) and (22), the associated weights \( \{a_0^i\}_{i=1}^{N_0} \) can be updated by

\[
a_0^i = a_0^i \times \frac{1}{\sigma_3^2} \times 2\pi e^{(-z_k-x_0 \times e_0^i y^i)/2\sigma_3^2} / \left( \sum_{i=1}^{N_0} a_0^i \times \frac{1}{\sigma_3^2} \times 2\pi e^{(-z_k-x_0 \times e_0^i y^i)/2\sigma_3^2} \right),
\]

where \( \sigma_3 \) was empirically set to 0.1 because the scale of MAWDI is small and MAWDI is ranged from 9.7 to 10.3. If the ESS calculated by Eq. (14) is below the half of N₀, then resample the random particles. In this paper, a systematic resampling algorithm [41] was employed because it is the most efficiency and popularity among other resampling algorithms, such as residual resampling and multinomial resampling. Its principle is simply reviewed as follows.

First, construct the cumulative distribution function of the weights. Let \( c_1 = 0 \) and \( c_i = c_{i-1} + a_0^i \). Draw a starting point \( a_1 \) from a uniform distribution U[0, \( N_0^{-1} \)]. For each increased point \( a_j = a_1 + N_0^{-1}(j-1) \), \( j = 1, 2, ..., N_0 \), moving along the cumulative distribution function of the weights, if \( a_j \geq c_i \) is satisfied, \( i = i + 1 \). Then, \( x^i = x^i \) and \( y^i = y^i \). After the resampling is completed, all of the weights are set to 1/N₀.
Step 2. Draw $N_s$ new random particles $\{x_k^i\}_{i=1}^{N_s}$ and $\{y_k^i\}_{i=1}^{N_s}$ of parameters $x_k$ and $y_k$ from the distributions of $q(x_{k-1}|z_{1:k}) = N(x_{k-1}, \sigma^2_x)$ and $q(y_{k-1}|z_{1:k}) = N(y_{k-1}, \sigma^2_y)$, respectively. The associate weights are updated by

$$
\omega_k^i = \omega_{k-1}^i \times \frac{1}{\sigma_3} \exp\left(-\frac{|x_k^i - x_M^i|}{\sigma_3}\right) / \sum_{i=1}^{N_s} \omega_{k-1}^i \times \frac{1}{\sigma_3} \exp\left(-\frac{|x_k^i - x_M^i|}{\sigma_3}\right) \sigma_3^2.
$$

(24)

If the ESS is below the half of $N_s$ then resample the random particles by the systematic resampling algorithm introduced in Step 1.

Step 3. Increase $k = k + 1$ and repeat Step 2 until $k > M$. The posterior probability density functions of $x_M$ and $y_M$ can be expressed as follows:

$$
p(x_M|z_{1:M}) \approx \sum_{i=1}^{N_s} \omega_M^i \delta(x_M - x_M^i),
$$

(25)

$$
p(y_M|z_{1:M}) \approx \sum_{i=1}^{N_s} \omega_M^i \delta(y_M - y_M^i).
$$

(26)

3.3. Remaining useful life estimation of a slurry pump impeller

Given the values of the MAWDI$(k)$, $k = K, \ldots, M$, the posterior probabilities of the two unknown parameters $x_M$ and $y_M$ can be established by Eqs. (25) and (26). Then, the future document numbers are input into the measurement equation defined in Eq. (21) to predict future MAWDI values, extrapolating the measurement equation. The probability density functions of the predicted MAWDI values at the future document numbers are derived as follows:

$$
p(z_k|z_{1:M}) = \sum_{i=1}^{N_s} \omega_M^i \delta(z_k - x_M^i \times e^{y_M^i x_M^i}), \quad k = M + 1, M + 2, \ldots, N.
$$

(27)

The means of the predicted MAWDI values at the future document numbers can be expressed as

$$
z_k = \sum_{i=1}^{N_s} \omega_M^i x_M^i \times e^{y_M^i x_M^i}, \quad k = M + 1, M + 2, \ldots, N.
$$

(28)

Before estimation of remaining useful life, a specified alert threshold should be established. Because only a few works has been done for impeller degradation evaluation and we lack of enough historical data, it is difficult to establish a standard alert threshold. As illustrated in Section 3.2, an alert threshold of the MAWDI was assumed to be equal to 10.3. Once the predicted future MAWDI values reach the specified alert threshold $z_{\text{threshold}}$, the probability density function of the remaining useful life of the slurry pump impeller at document $M$ can be derived as

$$
p(\text{RUL}|z_{1:M}, z_{\text{threshold}}) = \sum_{i=1}^{N_s} \omega_M^i \delta(\text{RUL} - \text{RUL}^i(k)),
$$

(29)

where $\text{RUL}^i(k)$ is obtained by solving the following equation:

$$
\text{RUL}^i(k) = \inf(k \in N: x_M^i \times e^{y_M^i x_M^i} \geq z_{\text{threshold}}) - M, \quad i = 1, 2, \ldots, N_s.
$$

(30)

A predicted RUL value at document $M$ can be taken as the median (the 50th percentile of the RUL) of Eq. (29)

$$
\frac{\sum_{\text{RUL}^i(k) \leq \text{RUL}^j}}{\sum_{\text{RUL}^i(k) \geq \text{RUL}^j}} p(\text{RUL} = \text{RUL}^i(k)|z_{1:M}, z_{\text{threshold}}) \geq \frac{1}{2},
$$

$$
\sum_{\text{RUL}^i(k) \geq \text{RUL}^j} p(\text{RUL} = \text{RUL}^i(k)|z_{1:M}) \geq \frac{1}{2}, \quad k = M + 1, M + 2, \ldots, N, \quad i = 1, 2, \ldots, N_s.
$$

(31)

where $\text{RUL}^j$ is the median of the RUL.

The alert probability density function (APDF) of the slurry pump impeller at document number $M$ can be derived as

$$
p(\text{APDF}|z_{1:M}, z_{\text{threshold}}) = \sum_{i=1}^{N_s} \omega_M^i \delta(\text{APDF} - \text{APDF}^i(k)),
$$

(32)

where $\text{APDF}^i(k)$ is obtained by solving the following equation:

$$
\text{APDF}^i(k) = \inf(k \in N: x_M^i \times e^{y_M^i x_M^i} \geq z_{\text{threshold}}), \quad i = 1, 2, \ldots, N_s.
$$

(33)

A predicted APDF value at document $M$ can be taken as the median of Eq. (32):

$$
\frac{\sum_{\text{APDF}^i(k) \leq \text{APDF}^j}}{\sum_{\text{APDF}^i(k) \geq \text{APDF}^j}} p(\text{APDF} = \text{APDF}^i(k)|z_{1:M}, z_{\text{threshold}}) \geq \frac{1}{2},
$$

$$
\sum_{\text{APDF}^i(k) \geq \text{APDF}^j} p(\text{APDF} = \text{APDF}^i(k)|z_{1:M}, z_{\text{threshold}}) \geq \frac{1}{2}, \quad k = M + 1, M + 2, \ldots, N, \quad i = 1, 2, \ldots, N_s.
$$

(34)

where $\text{APDF}^j$ is the median of the APDF.
The alert cumulative distribution (ACD) of the slurry pump impeller at document number $M$ can be derived as

$$ ACD(x) = \sum_{x_j \leq x} \ln \left( \frac{\text{PDF}}{\text{APDF}} \right) = \sum_{x_j \leq x} \sum_{i=1}^{N_i} a_{ij} \delta(x_j - \text{APDF}(k)). $$

(35)

4. Prognostic results

In this section, the developed prognostic method is used to analyze the data collected from the industrial oil sand pump that is introduced in Section 3.1. Based on the data from document numbers 63 to 200, the initial estimates of $x_0$ and $y_0$

![Predicted results obtained by using the developed method at document number 300 for slurry pump impeller prognosis.](image-url)
were calculated as 9.8250 and $5.412 \times 10^{-5}$, respectively, by using the nonlinear least squares regression. The error of predicted RUL values, the 5th and 95th percentiles of predicted RUL values (5 statistical significance level for 90% intervals of predicted RUL values) were used to quantify the performance of the developed prognostic method.

The prognostic results at document number 300, in which document numbers 100–300 were used to update the initial estimates of $x_0$ and $y_0$, are shown in Fig. 5. The updated parameters $x_{300}$ and $y_{300}$ were obtained by taking the medians of Eqs. (25) and (26) and are equal to 9.8280 and $7.1136 \times 10^{-5}$, respectively. The APDF and its corresponding ACD were obtained by using Eqs. (33) and (35), and are plotted in Fig. 5(b). The PDF of the RUL is plotted in Fig. 5(c), where the PDF of the RUL has the same shape as the PDF of the APDF. The 5th, 50th (the predicted RUL value) and 95th percentiles of the RUL

![Fig. 5. Prognostic results obtained by using the developed method at document number 300 for slurry pump impeller prognosis.](image-url)

![Fig. 6. Predicted results obtained by using the developed method at document number 400 for slurry pump impeller prognosis.](image-url)
are 279, 414 and 622, respectively. The error between the predicted RUL 414 and the actual RUL 524 is 110 documents. The error is large because only limited data were available to update the parameters of the state space model.

Fig. 6 presents the prognostic results obtained at document number 400 by using the developed method. The medians of the updated parameters $x_{400}$ and $y_{400}$ were derived as 9.8254 and $7.1057 \times 10^{-5}$, respectively. The 5th, 50th and 95th percentiles of the RUL are 205, 324 and 498, respectively. The error between the actual RUL 424 and the predicted RUL 324 is 100 documents. The prognostic results obtained by using the developed method at document number 500 are shown in Fig. 7. The medians of the updated parameters $x_{500}$ and $y_{500}$ were derived as 9.8329 and $6.4807 \times 10^{-5}$, respectively. The prediction error of the predicted RUL 274 is 50 documents away from the actual RUL 324, which largely improves the prognostic accuracy. The 5th and 95th percentiles of the RUL are 148 and 436, respectively.

Fig. 7. Predicted results obtained by the developed method at document number 500 for slurry pump impeller prognosis.
The prognostic results at document number 600 are plotted in Fig. 8, where the predicted RUL is 178 documents. The medians of the updated parameters $x_{600}$ and $y_{600}$ were derived as 9.8281 and $6.5397 \times 10^{-3}$, respectively. The predicted error was calculated as 46 documents. The 5th and 95th percentiles of the RUL are 80 and 318, respectively. The prognostic results at document number 700 are shown in Fig. 9. The medians of the updated parameters $x_{700}$ and $y_{700}$, and the predicted RUL are 9.8240, $6.5047 \times 10^{-3}$ and 91 documents, respectively. The error between the predicted RUL and the actual RUL is 33 documents. The 5th and 95th percentiles of the RUL are 1 and 202, respectively. The last prediction is conducted at document number 800. The prognostic results are shown in Fig. 10, where the APDF and the PDF of the RUL become sharpened. It reveals that the distribution of the PDF of the RUL concentrates around the actual RUL 24 documents. The medians of the updated parameters $x_{800}$ and $y_{800}$ were obtained by the developed method as 9.8314 and $6.0382 \times 10^{-3}$.

![Fig. 8. Predicted results obtained by using the developed method at document number 600 for slurry pump impeller prognosis.](image-url)
respectively. The 5th, 50th and 95th percentiles of the RUL are 1, 32 and 149, respectively. The prediction error is 8 documents.

The results obtained in Figs. 5–10 indicate the predicted accuracy improves as more data become available to update the parameters of the state space model that was used in the developed prognostic method. The predicted RUL values, their 90% intervals and the actual RUL values at more document numbers from 300 to 800 with an increment of 50 are summarized in Fig. 11, where the developed prognostic method is concluded to estimate the RUL of the oil sand pump impeller well.

In this paper, all computations were conducted using a MATLAB installed on a desktop with 3.1 GHz CPU and 4 GB (3.24 GB usable) RAM. For each inspection document, the average time used for prediction is 0.29 s. The two reasons why the calculation time is short are given as follows. First, because the parameters of the state space model used in the developed prognostic method are recursively updated when a new degradation observation is available. Second, a desktop with a high performance was used to further speed up the calculation time.

![Graph A](image1.png)

**Fig. 9.** Predicted results obtained by using the developed method at document number 700 for slurry pump impeller prognosis.
5. Conclusions

A prognostic method for a slurry pump impeller is developed in this paper. The developed method consists of two steps. The first step aims to assess the performance degradation of the slurry pump impeller because performance degradation assessment is the basis of RUL estimation. First, the vibration components of slurry pump vibration data are analyzed. The frequency components that are unrelated to the slurry pump impeller have no contribution in evaluating the health condition of the slurry pump impeller. Therefore, the energy evolution is constructed by summarizing the amplitudes of the frequency band that covers the vane passing frequency. The moving-average wear degradation index is then proposed based on the energy evolution to track the underlying trend of the energy evolution to reflect pump impeller degradation. The second step aims to predict the remaining useful life of the slurry pump impeller based on the proposed moving-average wear degradation index. The state space model of the moving-average wear degradation index is developed for this step. The parameters are updated by the general particle filter, which is a
special case of the general sequential Monte Carlo method. The predicted alert and the remaining useful life probability density functions are approximately derived by using the weights and the random particles generated from the posterior probability density functions of the state parameters. The data collected from industrial oil sand pumps were used to illustrate how the developed method worked. The prognostic results were obtained from 11 different document numbers. These results showed that the prognostic accuracy improves as more data become available to update the parameters of the state space model.

The contributions of this paper can be summarized in the following four points. First, the performance degradation assessment of the slurry pump impeller is realized by the proposed moving-average wear degradation index. Second, the state space model of the moving-average wear degradation index is constructed to describe possible future degradation evolution. Third, the predicted alert and the remaining useful life probabilities are approximated by the weights and the random particles generated by the posterior probability density functions of the state parameters. The predicted alert and the remaining useful life probabilities can be mathematically described. Finally, numerous other potential statistics used for prognosis can be constructed based on the posterior probability density functions of the state parameters.

Acknowledgments

The work described in this paper is fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 122513) and a grant from City University of Hong Kong (Project No.7004251).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymssp.2014.10.010.

References
