

ORIGINAL ARTICLE

Open Access



# Sequential Monte Carlo Method Toward Online RUL Assessment with Applications

Ya-Wei Hu<sup>1\*</sup>, Hong-Chao Zhang<sup>1,2</sup>, Shu-Jie Liu<sup>1</sup> and Hui-Tian Lu<sup>3</sup>

## Abstract

Online assessment of remaining useful life (RUL) of a system or device has been widely studied for performance reliability, production safety, system conditional maintenance, and decision in remanufacturing engineering. However, there is no consistency framework to solve the RUL recursive estimation for the complex degenerate systems/device. In this paper, state space model (SSM) with Bayesian online estimation expounded from Markov chain Monte Carlo (MCMC) to Sequential Monte Carlo (SMC) algorithm is presented in order to derive the optimal Bayesian estimation. In the context of nonlinear & non-Gaussian dynamic systems, SMC (also named particle filter, PF) is quite capable of performing filtering and RUL assessment recursively. The underlying deterioration of a system/device is seen as a stochastic process with continuous, nonreversible degrading. The state of the deterioration tendency is filtered and predicted with updating observations through the SMC procedure. The corresponding remaining useful life of the system/device is estimated based on the state degradation and a predefined threshold of the failure with two-sided criterion. The paper presents an application on a milling machine for cutter tool RUL assessment by applying the above proposed methodology. The example shows the promising results and the effectiveness of SSM and SMC online assessment of RUL.

**Keywords:** Sequential Monte Carlo method, Remaining useful life, Stochastic processes, State-space model, Bayesian estimation, Particle filter, Milling cutter lifetime

## 1 Introduction

Serious losses happen often in practice due to accidental system failure and the lack of online message of remaining useful life (RUL) and the performance reliability. Nowadays, high reliability and complexity of equipment make the operational reliability more than ever an outstanding concern. RUL online estimation, after many years of practice in the field of maintenance, is one facet in condition-based maintenance (CBM) [1, 2] and Prognostic and Health Management (PHM) engineering [3, 4]. Furthermore, RUL plays a crucial role in decision making on reuse or recycling of an old product in remanufacturing engineering.

The International Standard Organization (ISO) defines failure prognostics as “the Estimation of the Time to

Failure (ETTF) and the risk of existence or later appearance of one or more failure modes.” Note that most of the definitions cited in the literature use the terminology of RUL instead of ETTF [5]. The methods of residual (remaining) life prediction has get in-depth research for a long time, and there has been a growing interest in monitoring the ongoing “health” of products and systems in order to predict failures [6]. The degradation process of an operating device is a process of gradual deterioration and the underlying degradation can be detected to a certain extent through the measurement of covariate variables. The RUL can be expressed as  $X_t = T - t | T > t, \mathbf{Z}(t)$ , where  $T$  represents the random variable of lifetime,  $t$  the current time and vector  $\mathbf{Z}(t)$  the observations of the covariate variables available up to the current time  $t$ . There are a few comments about the RUL online assessment: (1) The degradation of the system/device is assumed to be monotonic and not-reversible, which causes the lifetime of the device to be limited to the RUL; (2) In practice, most of the time, the underlying

\*Correspondence: yaweihu@126.com

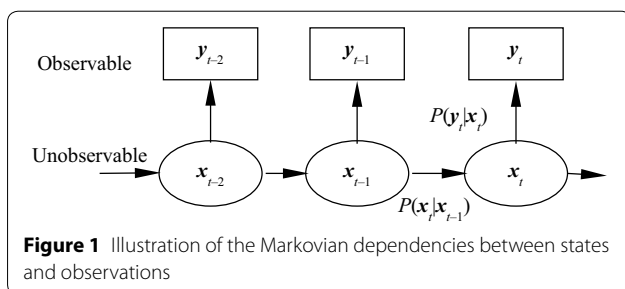
<sup>1</sup> Department of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China

Full list of author information is available at the end of the article

degradation of a device is not directly measurable, especially not measurable directly online; (3) Indirect measurement of system performance by covariate variable is used to model the degradation of the system performance, which are observable and measurable online [7]. There are two stochastic processes: the underlying degradation and the performance observation. To assess the non-observable degradation of the device from measurable observations of the performance, hidden Markov modeling is adopted.

The Hidden Markov model (HMM) with the assumption that the system states can be seen as a first order Markov process as the state  $\mathbf{x}_t \sim P(\mathbf{x}_t|\mathbf{x}_{0:t-1}) = P(\mathbf{x}_t|\mathbf{x}_{t-1})$ , where  $\mathbf{x}_{0:t-1} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{t-1}\}$ , and the observations under the given system states are independent from each other, i.e.,  $\mathbf{y}_t \sim P(\mathbf{y}_t|\mathbf{x}_t)$ . State  $\mathbf{x}_t$  presents the inherent characteristics of the dynamics at time  $t$ . The observation vector with model noise is  $\mathbf{y}_t$  and  $\mathbf{y}_{1:t} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t\}$ . For a dynamic Bayesian network, HMM is a proved valid methodology to model the remaining life assessment [8, 9]. The explicit expression of HMM is shown in Figure 1 [10].

The state space model consist of state and observation equations, presenting a first order HMM, is convenient for modeling multivariate data and nonlinear/non-Gaussian processes, with significant advantages over traditional time-series techniques [5, 11, 12]. For online predictions based on the state-space model (SSM), the recursive assessment of the posteriori distributions of  $\mathbf{x}_t$ , which modeling the degradation, is of a great concern. With updates of the state estimation and prediction, the RUL can be estimated by the system failure through a predefined degradation threshold [13]. This paper is organized as follows. A brief review based on the Bayesian estimation method from Markov chain Monte Carlo (MCMC) to Sequential Monte Carlo (SMC) is summarized in Section 2. The state estimation methodology for a given SSM with time invariant parameters and the RUL online assessment from the degrading process is detailed in Section 3. A case study is introduced and the results are discussed in Section 4. The final conclusion is presented in Section 5.



**Figure 1** Illustration of the Markovian dependencies between states and observations

## 2 From MCMC to SMC

### 2.1 A Brief Review

MCMC (Markov chain Monte Carlo) implements sampling distribution with dynamic simulation in accordance with system dynamics [14, 15]. The basic idea is a Markov chain is constructed that makes the stationary distribution to be a priori distribution of the dynamics; produce the samples of the priori distribution through the Markov chain, and implement MC (Monte Carlo) integration for the samples of the stationary distribution [15–17]. The MCMC procedure is shown in the following steps:

- (1) Establish a Markov chain that is converged to a stationary distribution  $\pi(\mathbf{x})$ ;
- (2) Generate samples from the initial point  $\mathbf{x}(0)$ . The simulation of the sampling was carried out using the Markov chain in Step (1), and produce the sequence  $\mathbf{x}(1), \dots, \mathbf{x}(n)$ ;
- (3) The estimation of expectation of a given function  $f(\mathbf{x})$  is  $\frac{1}{n-m} \sum_{t=m+1}^n f(\mathbf{x}(t))$ , from sequence points  $m$  to  $n$ .

The construction of the MCMC transition kernel, 1-step transition probability from state  $i$  to state  $j$ , notated as  $P(i, j) = P(i \rightarrow j) = P(\mathbf{x}_{t+1} = s_j | \mathbf{x}_t = s_i)$ , is important in order to iterative algorithm converges speedily. Different MCMC methods in applications have been proposed by selecting different transition kernels, and two main types of commonly used in MCMC methods are Gibbs sampler [18] and Metropolis–Hastings algorithm [19]. Since MCMC estimation for the state vector is required at every sampling time point when the new observations are available, historical data set could be huge. To avoid the problem of storing massive data and repeating the process on existing data in every sampling interval, a recursive mechanism in filtering, which means processing only on received new data sequentially if a new observation becomes available, is required to update the estimate results. SMC (Sequential Monte Carlo) method has been successfully developed and applied in many different fields [20]. SMC method approximates the actual posteriori probability density by a group of random samples with their associated weights. When the number of samples is large enough, the estimated probability can be close enough to the actual posteriori probability density function, reaching the result of the optimal Bayesian estimation [21].

In general, the actual posteriori probability density function cannot be sampled, instead using another distribution  $q(\mathbf{x}_{0:t}^{(i)} | \mathbf{y}_{1:t})$ , the importance distribution, or instrumental distribution. Importance sampling tries to use the limited number of sampling points covering the great contribution to the integral by a similar density, thus obtaining a higher computational efficiency.

Hammersley, Morton, et al. developed the sequential importance sampling (SIS) method in the 1950s [22]. However, SIS algorithm has a serious flaw. The importance weight variance increases gradually over time, which soon causes only a small number of samples (particles) that are associated with larger weights. This phenomena is called particle degrading in SIS [23]. The degradation with iteration process results in that there are no enough particles meaningful to present the posteriori density distribution eventually. Until 1993, Gordon proposed the concept of re-sampling [24], which laid the base for the particle filter and its practical applications. The resampling method ensures a stable total particle number by enlarging the sample numbers with higher weight values and reducing the number of particles with small weight so as to avoid much computation cost on those faded particles. Since then, SIR (Sampling Importance Resampling filter) technology also called as particle filtering has been gradually matured and widely used in applications of online state estimation.

### 2.2 Particle Filter Algorithm

The schematic of the standard particle filter algorithm process with re-sampling function [25] is shown in Figure 2, modified from Ref. [24].

Figure 3 shows the flow chart of particle filter algorithm. Re-sampling process taken in particle filtering is to overcome the degradation of the particle weights with iteration processing. In other words, re-sampling can improve the degradation to keep a stable total number of the particles in filtering.

The concept of effective re-sampling scales is defined as

$$N_{eff} = \frac{n}{\sum_{i=1}^n (w_t^{(i)})^2},$$

where  $w_t^{(i)}$ ,  $i = 1, \dots, n$ , are unnormalized weights of particles. This formula is difficult to determine the actual calculation instead the approximation is generally used as

$$N_{eff} = \frac{1}{\sum_{i=1}^n (\tilde{w}_t^{(i)})^2},$$

where  $\tilde{w}_t^{(i)}$ ,  $i = 1, \dots, n$ , are normalized weights, and

$$\tilde{w}_t^{(i)} = w_t^{(i)} / \sum_{i=1}^N w_t^{(i)}.$$

If  $N_{eff} < N_{th}$ , re-sampling should be adopted, and  $N_{th} = 2n/3$  in general. Note that it is not always necessary to have resampling in every iteration step. The judgment on the need of resampling or not is dependent on particle degradation in SIS process. The resampling starts only in the iteration when the particle degradation is beyond a certain predefined limit.

### 2.3 State Assessment in SSM

An SSM is built by two equations: the state (translation) equation and the observation (transformation) equation. The state equation presents the relationship between the next state and the current state, while the observation equation reflects the intrinsic relationship between the

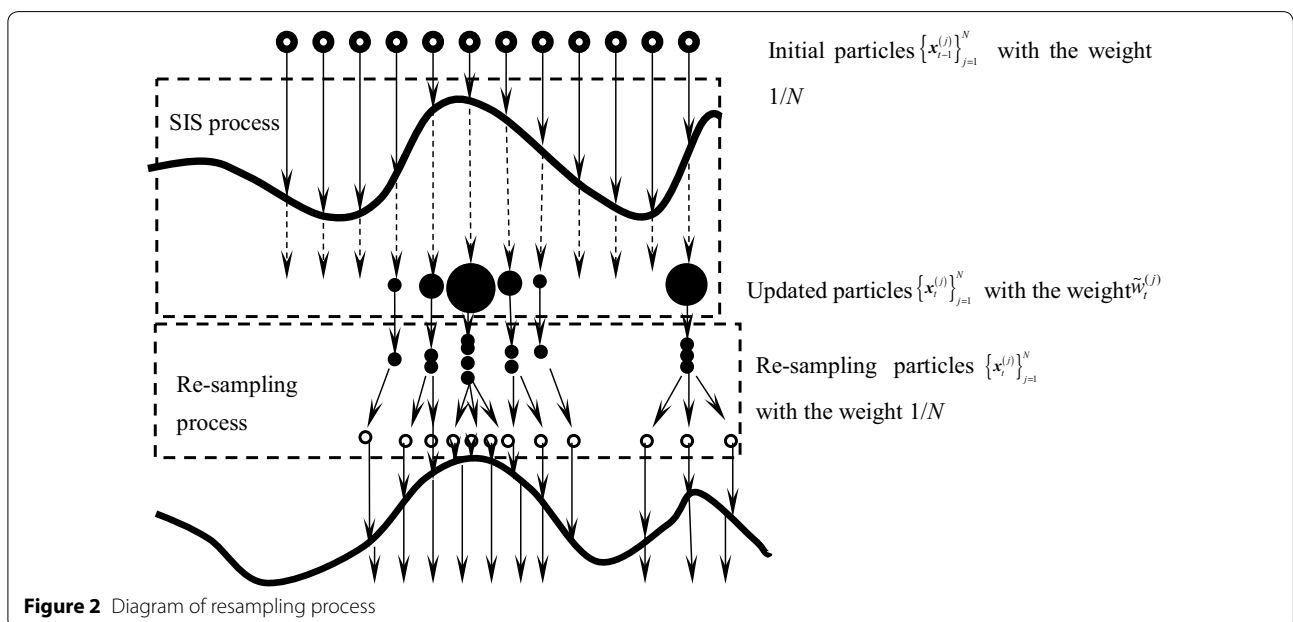


Figure 2 Diagram of resampling process

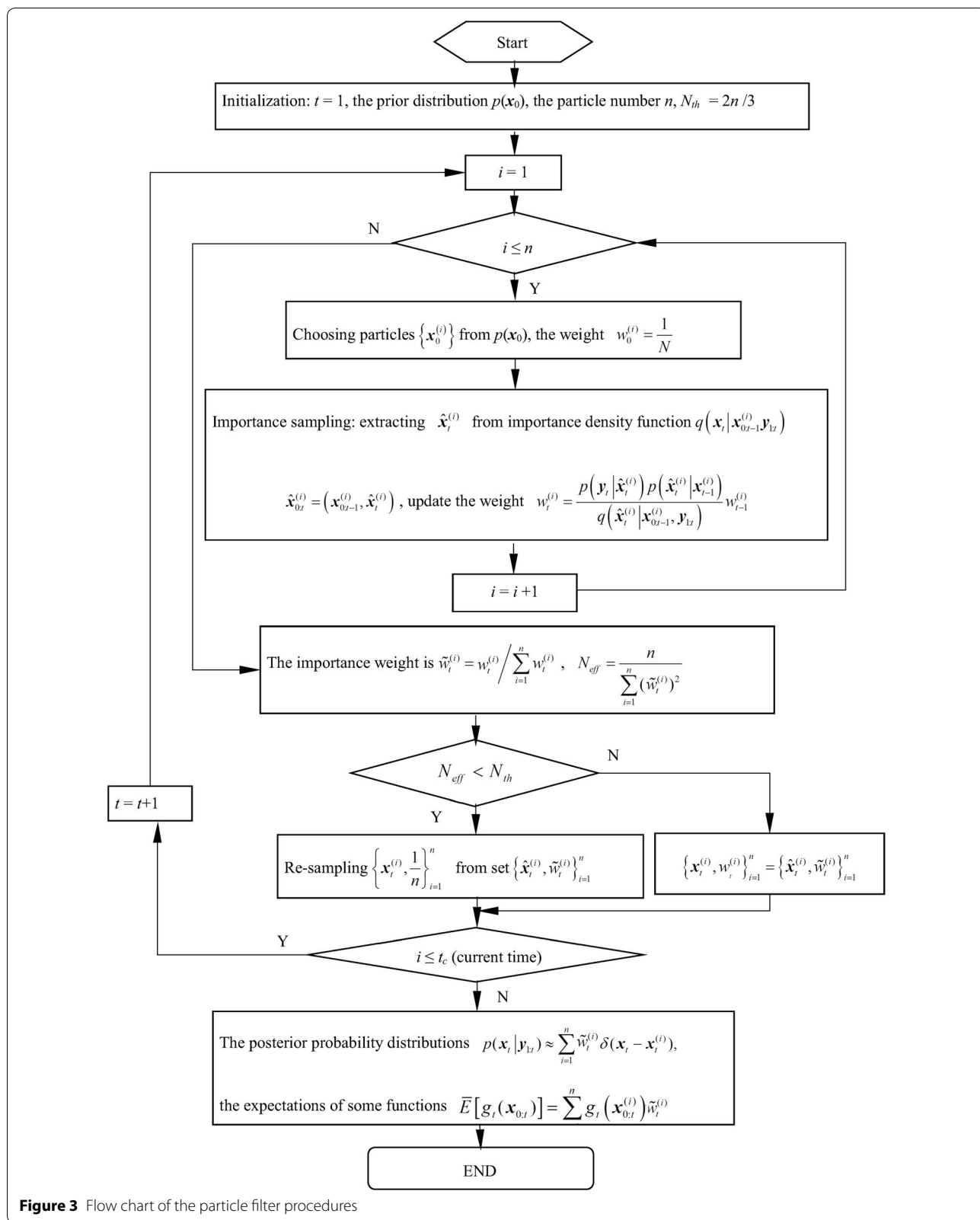


Figure 3 Flow chart of the particle filter procedures

observations and the state of the system. SSM is a discrete-time presentation of an HMM. The state in SSM is the hidden stochastic process in HMM.

The SSM of a system (in general a nonlinear and non-Gaussian system) can be written as

$$\mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{u}_t, \boldsymbol{\varepsilon}_t, \boldsymbol{\theta}), t \in T, \quad (1)$$

$$\mathbf{y}_t = h(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\eta}_t, \boldsymbol{\theta}), t \in T, \quad (2)$$

where Eqs. (1), (2) are the state and observation equations respectively.  $\mathbf{x}_t \in \mathbf{R}^{N_x}$  is the system state vector,  $\mathbf{y}_t \in \mathbf{R}^{N_y}$  is observation vector,  $\mathbf{u}_t \in \mathbf{R}^{N_x}$  is the input vector of the system,  $\boldsymbol{\theta}$  is the static parameters of the model.  $\boldsymbol{\eta}_t$  ( $\boldsymbol{\eta}_t \in \mathbf{R}^{N_y}$ ) and  $\boldsymbol{\varepsilon}_t$  ( $\boldsymbol{\varepsilon}_t \in \mathbf{R}^{N_x}$ ) denote the observation noise and states noise respective, which are independent with each other.  $\mathbf{f} : \mathbf{R}^{N_x} \times \mathbf{R}^{N_x} \mapsto \mathbf{R}^{N_x}$  is the state function and  $h : \mathbf{R}^{N_x} \times \mathbf{R}^{N_x} \mapsto \mathbf{R}^{N_y}$  is the observation function. Assuming these two functions known and depend on  $\mathbf{u}_t$  (sometime it is omitted for simplicity). Priori distribution of the initial state  $\mathbf{x}_0$  is assumed to be  $p(\mathbf{x}_0)$ .

The purpose of constructing an SSM is that SSM provides a convenient way to estimate the state  $\mathbf{x}_t$  recursively (Bayesian estimation). To this end, two sequential steps are performed. The prediction step estimates the priori probability density of next time state vector using both the process model and the information of the previous state estimate, expressed by the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1}, \quad (3)$$

The filtering step, set up the recursive relation for the forecast, considers the current observation  $\mathbf{y}_t$  into the estimator of the state vector to correct the obtained priori probability density, and get the posteriori probability density of the state vector by the Bayesian formula:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t, \mathbf{y}_{1:t-1} | \mathbf{x}_t) p(\mathbf{x}_t)}{p(\mathbf{y}_t, \mathbf{y}_{1:t-1})} = \frac{p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t-1})}{p(\mathbf{y}_t | \mathbf{y}_{1:t-1})}, \quad (4)$$

where  $p(\mathbf{y}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) d\mathbf{x}_t$  is called as a normalized factor.

The posteriori probability density can be expressed as

$$p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_{1:t} | \mathbf{x}_{0:t}) p(\mathbf{x}_{0:t})}{p(\mathbf{x}_{1:t})}, \quad (5)$$

and with all observations are independent with each other,  $p(\mathbf{y}_{1:t} | \mathbf{x}_{0:t}) = \prod_{i=1}^t p(\mathbf{y}_i | \mathbf{x}_{0:t})$ . Under the assumption that the observations are independent with all status of other moments  $p(\mathbf{y}_{1:t} | \mathbf{x}_{0:t}) = \prod_{i=1}^t p(\mathbf{y}_i | \mathbf{x}_i)$ , the posteriori probability density is gotten as

$$p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) = p(\mathbf{x}_0) \frac{\prod_{i=1}^t p(\mathbf{y}_i | \mathbf{x}_i) p(\mathbf{x}_i | \mathbf{x}_{i-1})}{p(\mathbf{y}_{1:t})}, \quad (6)$$

where  $p(\mathbf{x}_{0:t}) = p(\mathbf{x}_0) \prod_{i=1}^t p(\mathbf{x}_i | \mathbf{x}_{i-1})$  because of the system following Markov process. So the recursive formula of joint probability density is

$$p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) = p(\mathbf{x}_{0:t-1} | \mathbf{y}_{1:t-1}) \frac{p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{x}_{t-1})}{p(\mathbf{y}_t, \mathbf{y}_{1:t-1})}. \quad (7)$$

For the recursive computation of the posteriori state probability density function above, it is more conceptual than practical, and in most cases, the integrals Eq. (4) do not have an analytical solution, so it is hard to describe the probability density functions using a closed analytical form. Numerous researchers want to obtain approximate methodologies which could minimize the variance of these integral estimates in the mid-1960s, such as the extended Kalman filter, the Gaussian sum filter, and grid-based method [20, 26]. Although in certain settings these methods have been applied successfully, they are invalid if the posteriori distribution cannot be approximated by a Gaussian distribution. In particular, that it remains unimodal, which is typically not true in many nonlinear state-space scenarios [24]. Monte Carlo (MC) methods however, in which the posteriori distribution is represented by a collection of random points, play a central role in the 40 s of the 20th century along with advanced development of technology and digital computer. MC simulations, not only relevant for simulating models of interest, but constitute a valuable tool for approaching statistics [27].

The idea of Monte Carlo simulation is to draw an i.i.d. set of samples  $\{\mathbf{x}_{0:t}^{(i)}\}_{i=1}^N$  from the posteriori distribution of the state  $p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t})$  [28]. Using these  $N$  samples to approximate the target density with the following empirical function:

$$\hat{p}(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{(\mathbf{x}_{0:t}^{(i)})}(\mathbf{x}_{0:t}), \quad (8)$$

where the  $\delta_{(\mathbf{x}_{0:t}^{(i)})}(\mathbf{x}_{0:t})$  denotes the Dirac delta mass located at  $\mathbf{x}_{0:t}$ . So the expectation for any  $f(\mathbf{x}_{0:t})$  is  $E(f(\mathbf{x}_{0:t})) = \int f(\mathbf{x}_{0:t}) p(\mathbf{x}_{0:t} | \mathbf{y}_{1:t}) d\mathbf{x}_{0:t}$ . The integral operation can be estimated by  $\bar{E}(f(\mathbf{x}_{0:t})) = \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_{0:t}^{(i)})$ , where  $\bar{E}(f(\mathbf{x}_{0:t}))$  is the estimate of the expectation for  $f(\mathbf{x}_{0:t})$ .

The convergence of the above MC calculations can be obtained by the law of large numbers, and independent with the dimension of the state, therefore Monte Carlo simulation provides an effective method to find solutions

for applications in high-dimensions. Due to the error of MC approximation has nothing to do with dimension when increasing, so it is particularly suitable to solve high dimensional integral problem [29]. In recent decade, this SMC technology has been widely used to estimate the posteriori probability of the state of nonlinear, non-Gaussian dynamic systems.

### 3 RUL Online Assessment from Performance Degradation

#### 3.1 Degradation Prediction

As discussed in previous sections, RUL online assessment is based on the underlying performance degradation of the system and its predefined failure threshold. The degradation process is modeled in a hidden stochastic process as the state in SSM with constant parameters. For the  $l$ -step ahead recursive estimation of the degrading state  $(x_{t+1:t+l}|y_{1:t})$ , it can be obtained by following steps:

For  $j = 1, \dots, l$ ,

For  $i = 1, \dots, N$ , sample  $x_{t+j}^{(i)} \sim p(x_{t+j}|x_{t+j-1}^{(i)}, \theta_t^{(i)})$  and  $x_{t+1:t+j}^{(i)} = (x_{t+1:t+j-1}^{(i)}, x_{t+j}^{(i)})$ ,

At the end, a sample set  $\{x_{t+1:t+j}^{(i)}\}_{i=1}^N$  available, and an estimation of  $p(x_{t+1:t+l}|y_{1:t})$  can be estimated as

$$p(x_{t+1:t+l}|y_{1:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{(x_{t+1:t+l}^{(i)})} (x_{t+1:t+l}), \quad (9)$$

The estimation of the degradation state (the mean value of the state) based on the predicted distribution at time  $t + l$  can be calculated as:

$$\bar{x}_{t+l} = \frac{1}{N} \sum_{i=1}^N x_{t+l}^{(i)}, \quad (10)$$

#### 3.2 RUL Definition and the Prediction

The remaining useful life (RUL), defined as  $X_t = T_i - t | T_i > t, Z(t), i = 1, 2, \dots, n$ , is a random variable, the result of RUL estimation is presented in the form of a discrete probability distribution. For assessment of the RUL, a specified threshold value  $\lambda$  for the degradation state is pre-defined for failure. In deriving the distribution of the remaining lifetime,  $\lambda$  is as the failure criterion compared with each sample path to obtain  $T_1, \dots, T_n$  total  $n$  different time points.  $T_i$  represents the least time at which the sample path  $x_{t+1:t+k}^{(i)}$  will equal or exceed the threshold value  $\lambda$ . In this paper, the two-sided criterion for system failure is cited as  $C_f = \{H_{low} \leq x \leq H_{up}\}$ , where,  $H_{low}$  and  $H_{up}$  are the upper and lower boundaries of the failure interval respectively. Figure 4 shows the concept of derivation.

As shown in Figure 4 (modified from Ref. [24]),  $H_{low}$  and  $H_{up}$  are set symmetrically on both sides of  $\lambda$ . The

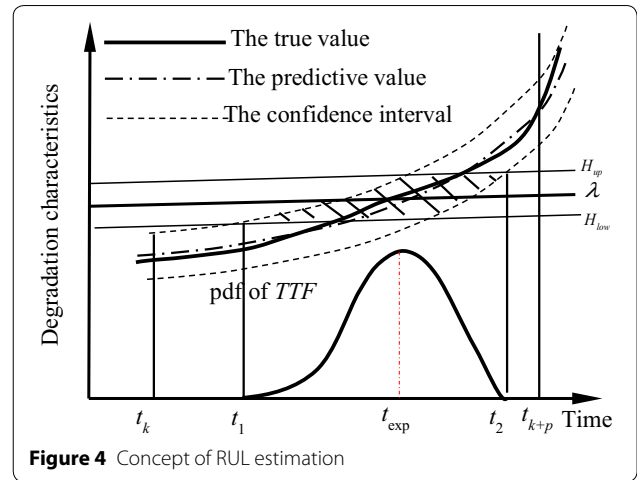


Figure 4 Concept of RUL estimation

sample particle swarms of state estimation from  $t_k$  to  $t_{k+p}$  overlap with the hazard zone (the light shaded area). The sum of the normalized weights of all sample particles which locate in the light shaded area at any time step between  $t_k$  and  $t_{k+p}$  represents the probability of system failure occurring at the corresponding time step. The normalizing constant is the sum of weights of total sample particles which locate in the light shaded area from time  $t_k$  to  $t_{k+p}$ . Therefore, an approximation of the probability distribution (probability density of time to failure,  $Pr_{TTF}$ ) can be obtained through a set of an equal interval discrete samples with their corresponding probabilities, that is

$$Pr_{TTF}(t_{k+j}) = \sum_{i=1}^N P(H_{low} \leq x_{k+j}^i \leq H_{up}) \tilde{w}_{k+j}^i, \quad (11)$$

$i = 1, 2, \dots, N; j = 1, 2, \dots, p,$

where  $\tilde{w}_{k+j}^i$  is the normalized weight of each particle at each prediction time,  $P(\cdot)$  is the probability to failure when the particle value is within the range of the defined failure band. The mean remaining life is therefore estimated to be

$$E_{RUL} = t_{exp} - t_k, \quad (12)$$

where  $t_{exp}$  is the system expected failure time estimated at time  $t_k$ .

### 4 An Numerical Example of Cutter Lifetime Assessment

A cutter is the key component of machine cutting process. Its state influences the total manufacturing effectiveness and stability of machining. Cutter life in milling machine is studied using this proposed methodology. With an accurate estimate of cutter lifetime, worn tools

can be changed in time to reduce nonconforming products and tooling costs. It therefore guarantees surface quality in manufacturing. In assessment of the RUL using data driven method, observation signal of the cutters were obtained from the sensors (acceleration sensors, force sensors, acoustic emission sensors, etc.) mounted on the high speed milling test rig, and the experimental parameters shown in Table 1. Six milling cutters (C1, C2, C3, C4, C5 and C6) with three blades were experimented in 315 cycles under the same working conditions. The cutter wearing of C1, C4 and C6 were measured online and the data was download from the web [30] with more details. Figure 5 is the experimental data of milling cutters' wear.

**4.1 Modeling and the Priori Distribution of the Model Parameters**

**4.1.1 A Discrete-time State Space Model**

A Wiener process has found application as a degradation model in many studies due to its good properties [31]. The degradation of cutters wear in the study is assumed at a constant rate, which meets Wiener process. Thus the state equation, which expresses the underlying degradation of the milling tools (wearing), can be presented as follows:

$$X(t) = \beta t + \sigma_B \cdot B(t), \tag{13}$$

where  $\beta$  is a constant for drifting,  $B(t)$  is a Brownian motion;  $\sigma_B$  is the corresponding diffusion coefficients; both  $\beta$  and  $\sigma_B$  are unknown.

**Table 1 Milling machine working conditions**

Rotate speed (r/min)	Feed speed (mm/min)	Cutting depth of Y direction (mm)	Cutting depth of Z direction (mm)	Sample frequency (kHz)
10400	1555	0.125	0.2	50

In general, the degree of tool wear can determine the tool failure. Therefore, tool wearing process can be considered as the underlying tool degradation. The observation  $Y(t)$  contains noise errors, the observation equation is expressed by

$$Y(t) = X(t) + \sigma_R \cdot \varepsilon(t), \tag{14}$$

where  $\sigma_R$  is the measurement error;  $\varepsilon(t)$  is assumed as white noise with the mean 0 and the variance 1. Because the operational precision of the measurement system is usually known, the  $\sigma_R$  in the model is assumed given.

Since digital computer DAQ (data acquisition) system is used for signal acquisition, a discrete-time state space model, both state and observation equations, is developed for the process.

$$X_{n+1} = X_n + \beta_n + \sigma_{B,n} \cdot W_n, \tag{15}$$

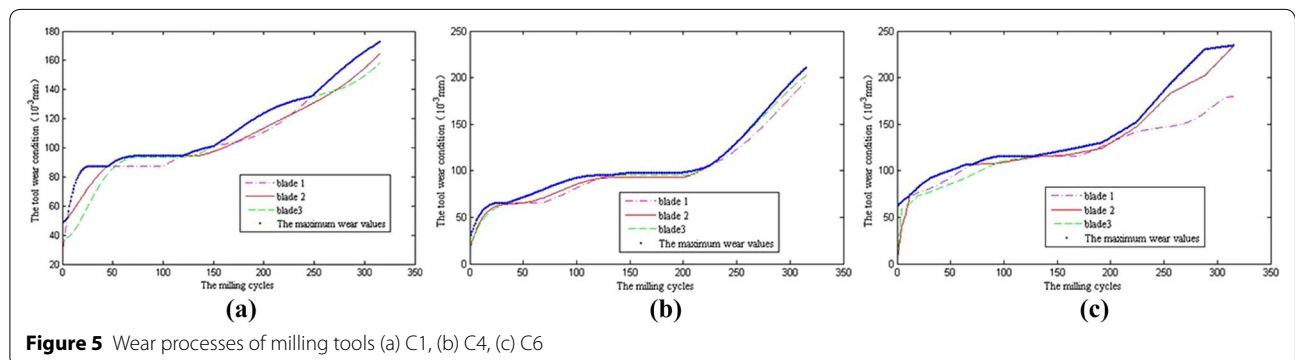
$$Y_n = X_n + \sigma_{R,n} \cdot V_n, \tag{16}$$

where  $\beta_n$  and  $\sigma_{B,n}$  are the unknown model parameters at time  $n$ , and  $W_n, V_n$  are noise terms in  $N(0, 1)$ .  $\sigma_{B,n}$  is given from the DAQ system.

**4.1.2 Determination of the Priori Distribution of the Model Parameters**

There are two estimates needed toward online RUL assessment. They are online state (degradation) estimate and the model parameter estimate ( $\beta$  and  $\sigma_B$ ). In the Bayesian inference framework, a reasonable priori distribution is first assigned for the unknown parameters of the model. The joint distribution  $\pi_0(\beta, \sigma_B^2)$  of  $\beta$  and  $\sigma_B^2$  is considered as the priori distribution of the unknown parameters in the model, and there is likely some dependency existing between  $\beta$  and  $\sigma_B$ .

According to the theory of Bayesian parameter estimation, the conjugate priori distribution for these two parameters is applied, i.e., a normal inverse-Gamma distribution as  $\pi_0(\beta|\sigma_B^2) \sim N(m_0, \sigma_B^2/n_0)$ ,



**Figure 5** Wear processes of milling tools (a) C1, (b) C4, (c) C6

$\pi_0(\sigma_B^2) \sim IG(a_B, \lambda_B)$ . And the density function of inverse Gamma distribution is

$$f(x|a, \lambda) = \frac{\lambda^a}{\Gamma(a)} x^{-(a+1)} e^{-\lambda/x}, x > 0, a > 0, \lambda > 0, \quad (17)$$

where  $m_0 = \hat{\mu}_\beta$ ,  $v_0 = \hat{\sigma}_\beta^2, v_B = \hat{\sigma}_B^2$ ,  $n_0 = v_B/v_0$ ,  $IG(a, \lambda)$  inverse-Gamma distribution,  $a$  is the shape parameter,  $\lambda$  is the scale parameter.

To obtain the hyper-parameters of the priori distribution, the bootstrap method [32], an important estimate method through statistical variance for interval estimation, is applied in this paper.

### 4.2 Posteriori Analysis Based on SMC

The SMC algorithm is adopted in the assessment of both the state and the model parameters ( $\beta$  and  $\sigma_B^2$ ) online. After the priori distributions of the parameters  $\beta$  and  $\sigma_B^2$  were determined, the parameters and the states should be estimated at each iteration step using 'sufficient statistic' to carry out the joint estimation to the parameters and the states. Based on the conjugate nature, the updating process of the corresponding posteriori distribution of the parameters  $\beta$  and  $\sigma_B^2$ , with  $n$  observations of tool wear available, are:

$$m_n = \frac{n}{n + n_0} m_0 + \frac{n}{n + n_0} \bar{\delta}_n, \quad (18)$$

$$\sigma_n^2 = (n + n_0) v_B^2, \quad (19)$$

$$a_{B,n} = a_{B,0} + \frac{n}{2}, \quad (20)$$

$$\lambda_{B,n} = \lambda_{B,0} + \frac{1}{2} S_{\delta,n}^2 + \frac{n_0 n}{2(n + n_0)} (\delta_i - \bar{\delta}_n), \quad (21)$$

where,  $\bar{\delta}_n = \frac{1}{n} \sum_{i=1}^n \delta_i$ ,  $S_{\delta,n}^2 = \sum_{i=1}^n (\delta_i - \bar{\delta}_n)^2$ .

Therefore, the parameters and the states in the SSM are estimated recursively based on the particle filter and using sufficient statistic [33]. The flow chart of the assessment showed in Figure 6.

### 4.3 Results

An experiment for model validation was done on one of the individual milling tools based on the wearing data. The online measurements of the tool's wear are shown in Figure 7 with total of 350 data points. It is a non-decreasing process. In practice, residual life prediction is considered in the middle and later periods of the cutter life. The data points from 125 to 315 milling cycle times were used from the total 350 data points separated as the validation data.

Using the bootstrap method resampling, 900 groups of samples are generated with different forecast origins

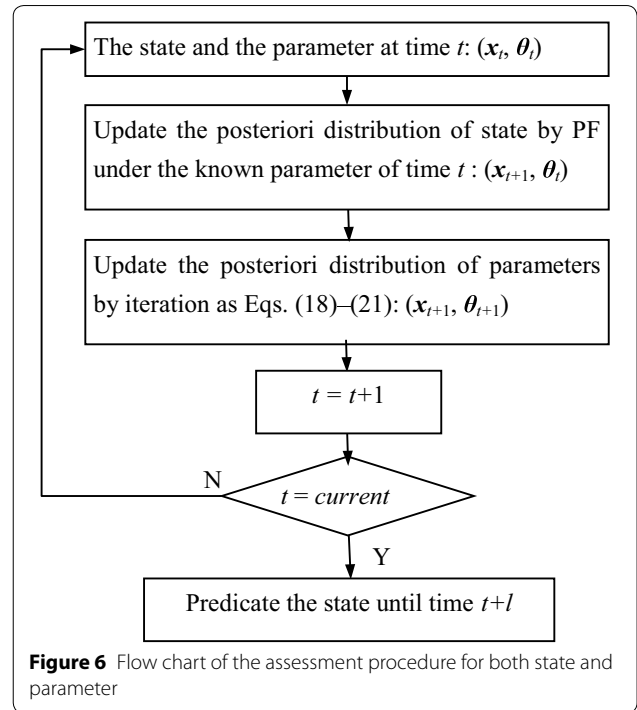


Figure 6 Flow chart of the assessment procedure for both state and parameter

at  $K1 = 35$ ,  $K2 = 55$ ,  $K3 = 75$ ,  $K4 = 95$ ,  $K5 = 115$  and  $K6 = 135$  (in milling cycles) respectively. The results obtained from applying the discussed method on the model parameter estimation are listed in Table 2, where the particle number is 30 and the priori probability density function is selected as the importance function.

Using 900 samples generated in each milling cycle, the estimated model parameter are adopted for RUL assessment. The pseudo codes for PF and RUL online assessment in Matlab is shown as follows:

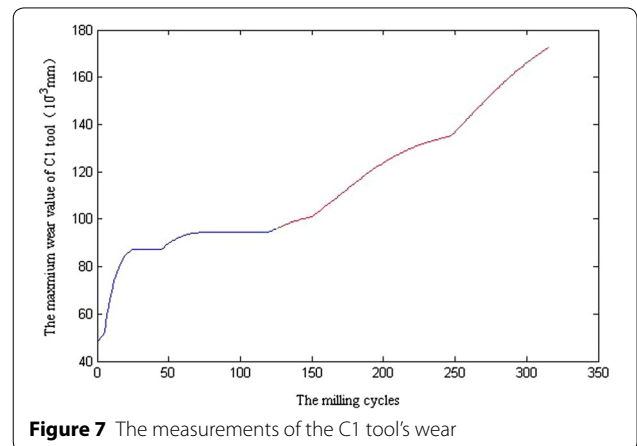


Figure 7 The measurements of the C1 tool's wear



**Pseudocode for the RUL online assessment algorithm**

1. **function** RUL ONLINE ASSESSMENT
2. Input:  $y_{1:k}$ ;  $\theta_1 = (\beta, \sigma_B, \sigma_R)$ ; The state threshold  $w$
3. Initialization:  $x_1^{(i)} \sim p(x_1)$
4. **for**  $t = 2: k$  **do**

Draw samples:  $\hat{x}_t^{(i)} \sim q(x_t | x_{t-1}^{(i)}, \theta_t)$  for  $i=1, 2, \dots, N$

$\hat{y}_t^{(i)} \sim p(y_t | x_t^{(i)}, \theta_t)$  for  $i = 1, 2, \dots, N$

Calculate weights  $w_t^{(i)} = p(y_t | \hat{x}_t^{(i)}, \theta_t) w_{t-1}^{(i)}$ , and

$$\tilde{w}_t^{(i)} = w_t^{(i)} / \sum_{i=1}^n w_t^{(i)}$$

if  $N_{eff} < N_{th}$ ,

Re-sampling particles  $\{x_t^{(i)}, \frac{1}{n}\}_{i=1}^n$  from  $\{\hat{x}_t^{(i)}, \tilde{w}_t^{(i)}\}_{i=1}^n$

else  $\{x_t^{(i)}, w_t^{(i)}\}_{i=1}^n = \{\hat{x}_t^{(i)}, \tilde{w}_t^{(i)}\}_{i=1}^n$

update  $\theta_t = (\beta_t, \sigma_{Bt}, \sigma_{Rt})$

**end**

# prediction

**for**  $j = k+1: t$

$\hat{x}_j^{(i)} \sim p(x_j | x_{j-1}^{(i)}, \theta_j)$  for  $i = 1, 2, \dots, N$

$$\bar{x}_j = \sum_{i=1}^N \hat{x}_j^{(i)} \bar{x}_j$$

if  $\bar{x}_j \geq w$

$$T_{RUL}(j) = t - j$$

**end**

output:  $T_{RUL}(j)$

**end function**

The multistep prediction results can be obtained based on the formula according to the degradation model. The critical threshold of the tool wear is defined as  $\lambda = 0.15$  mm, and the  $H_{low} = \lambda - 1.96\sigma_\lambda$ ,  $H_{up} = \lambda + 1.96\sigma_\lambda$ , respectively, where  $\sigma_\lambda$  is standard deviation of the threshold which assumed as 1.5 in the paper. The predicted results under the various time origins are shown in Figure 8. The thin solid lines represent the degradation processes of each particle; the two horizontal lines are the threshold band; the heavy lines are the predicted values; the real values are also provided by the thick dash-and-dot point line. The particles, when total number is large enough, are assumed following a normal

**Table 2 Model parameter estimation results**

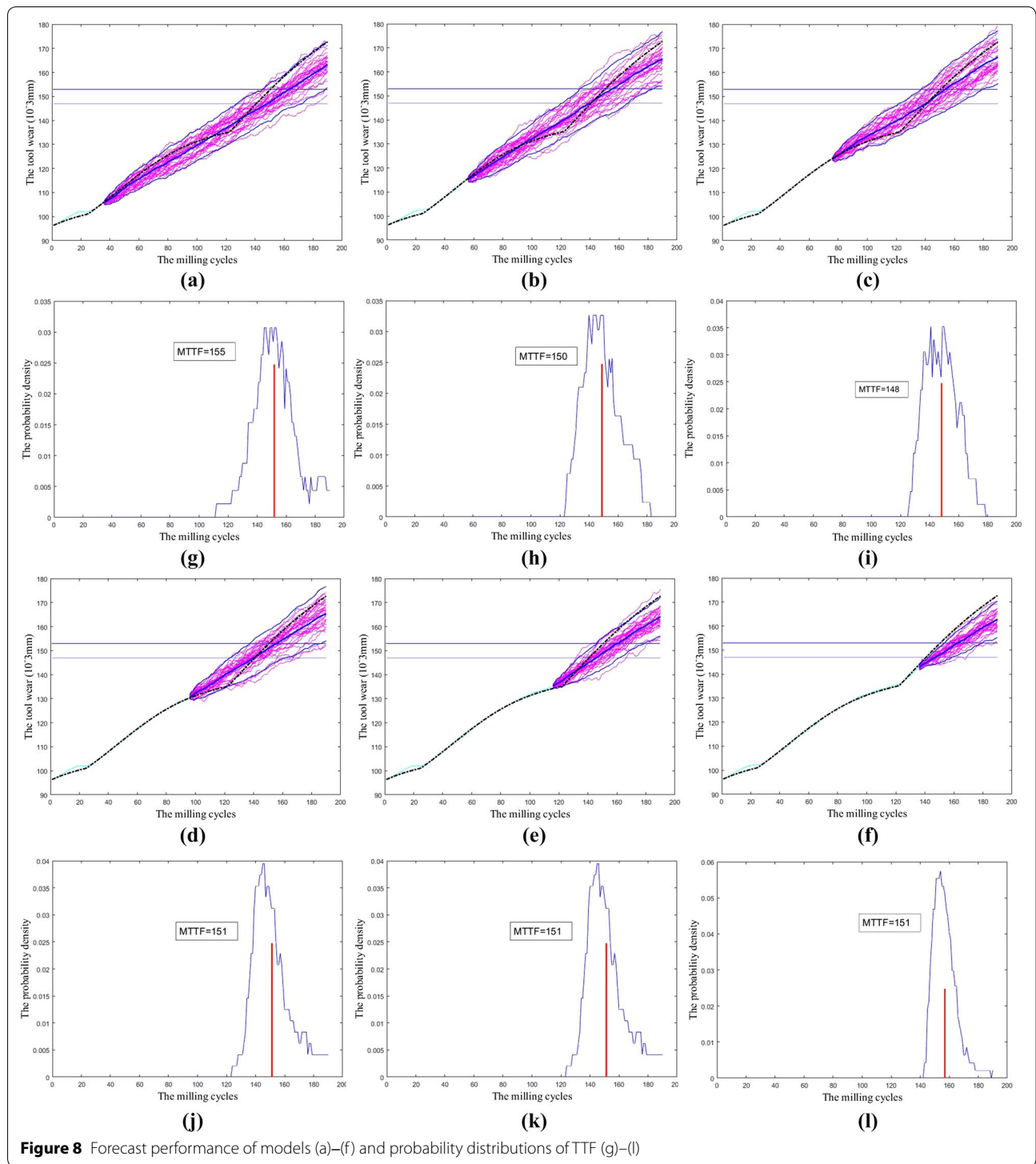
Milling cycles	$\mu_\beta$	$a$	$\lambda$	$\sigma_B$	$\sigma_R$
K1 = 35	0.3768	45.55	0.421	0.0094	0.5
K2 = 55	0.3629	276.64	0.342	0.0076	0.5
K3 = 75	0.3779	328.23	0.252	0.00072	0.5
K4 = 95	0.3811	423.33	0.194	0.00036	0.5
K5 = 115	0.3456	533.97	0.113	0.00015	0.5
K6 = 135	0.3505	743.83	0.128	0.00024	0.5

distribution, and the 95% confidence interval is given as  $(\bar{x}_{t+l} - 1.96\sigma/\sqrt{n}, \bar{x}_{t+l} + 1.96\sigma/\sqrt{n})$ . The RUL probability density function curves at each time origin of assessment are shown in Figure 8. The vertical lines at the right column are expectation failure times based on the estimated failure probability density functions.

Table 3 lists the predicted failure times, the true failure times, the predicted residual life and actual residual life at different forecast origins. Figure 9 shows the probability distribution functions of RUL at different forecast origins. It can be seen in Figure 10 from the comparison of the predicted values and the real values of RUL in the experiment that proposed model and SMC method for online assessment are valid and promising.

**5 Conclusions**

RUL assessment and modeling have become increasingly important in system reliability and PHM. System health management involves determining the system performance status and the RUL of critical systems used for the maintenance plan, decision making, and system global optimization. RUL assessment becomes a science in assessing the degradation based on observations of covariate performance variables due to the complexity of the system, failure mechanism, and un-observability of the physical degradation. Naturally, there are two stochastic processes in RUL assessment. One is a hidden degradation process; the other is observation of the measurable process. The state space model, as a first order hidden Markov, provides a desired format for SMC based on Bayesian estimation of the posteriori distribution. A nonlinear state space model using online recursive particle filter is the research focus for online RUL assessment. In this paper, the online assessment of milling tool life from the degradation of wearing was used to establish a discrete-time SSM. The SMC algorithm was applied to estimate the state and model parameters simultaneously through sufficient statistic. Predicted results of the RUL distribution and its expectation show the plausibility and effectiveness of this approach. There are a few points worth noting, which might lead to future research.



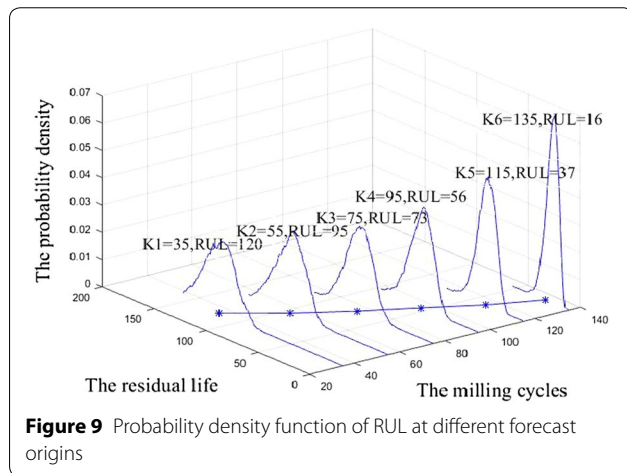
**Figure 8** Forecast performance of models (a)–(f) and probability distributions of TTF (g)–(l)

The estimated SSM in the case study was built up based on few assumptions without off-line performance analysis for physics mechanism. Model building methods are application oriented, varying from case to case.

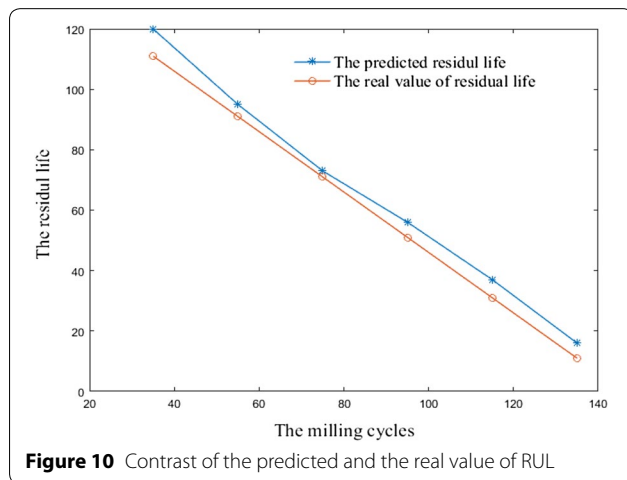
The sufficient statistic method exhibited good effect on sequential parameter learning in the text. However, the measurement error was assumed to be of a fixed value to simplify the calculation. The topic of how to assess the

**Table 3 RUL and TTF results at different assessment origins**

Milling cycles	The predicted failure time	True failure time	The predicted residual life	The actual residual life
K1 = 35	155	146	120	111
K2 = 55	150	146	95	91
K3 = 75	148	146	73	71
K4 = 95	151	146	56	51
K3 = 115	152	146	37	31
K4 = 135	151	146	16	11



**Figure 9** Probability density function of RUL at different forecast origins



**Figure 10** Contrast of the predicted and the real value of RUL

unknown parameters with variable measurement error online is still an open to research.

For the RUL prediction in the paper, the two-sided criterion for the failure threshold was cited directly without the support from reliability theory and the degradation threshold for a defined soft failure has all along been a controversial topic.

**Authors' contributions**

Y-WH designed the study, performed the assays and prepared the manuscript, contributed in its application part. H-CZ conducted the optimization and assay validation studies. H-CZ and S-JL participated in discussing the results, and revised the manuscript. H-TL did the final sequence alignment in the manuscript and drafted the manuscript. All authors read and approved the final manuscript.

**Author details**

<sup>1</sup> Department of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China. <sup>2</sup> Department of Industrial Engineering, Texas Tech University, Lubbock, TX 79409-3061, USA. <sup>3</sup> Department of Construction and Operations Management, South Dakota State University, Brookings, SD 57007, USA.

**Authors' Information**

Ya-Wei Hu, born in 1989, is currently a PhD candidate at *Department of Mechanical Engineering, Dalian University of Technology, China*. Her research interests include reliability theory, stochastic filtering theory, remaining useful life prediction.

Hong-Chao Zhang, an Interim Chair & E.L Derr Endowed Professor at Texas Tech University, USA. He is now working in Texas Tech University and Dalian University of Technology, China. Dr. Hong-Chao Zhang's research groups have been redesigning end-of-life (EOL) strategies for many products such as heavy-duty equipment remanufacturing, and developing new materials technologies to make sustainable products. Zhang's groups developed new processes for delaminating, recycling of printed circuit boards using a supercritical carbon dioxide process. The groups carried out microstructural transformation of materials such as shape memory polymer nanocomposites for active disassembly (AD) applications. Zhang's groups have also extended literature on product innovation and sustainable manufacturing by developing sustainability index and metric for 3D assessment of product's sustainability.

Shu-Jie Liu, born in 1977, is currently a lecturer at *Dalian University of Technology, China*. She received her PhD degree from *the University of Tokyo, Japan*, in 2007. Her research interests include remaining useful life assessment, precision engineering and sustainable engineering.

Hui-Tian Lu is a professor at *Department of Engineering Technology and Management, College of Engineering, South Dakota State University, USA*.

**Acknowledgements**

Supported by Basic Research and Development Plan of China (973 Program, Grant Nos. 2011CB013401, 2011CB013402), and Special Fundamental Research Funds for Central Universities of China (Grant No. DUT14QY21).

**Competing interests**

The authors declare that they have no competing interests.

**Ethics approval and consent to participate**

Not applicable.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 28 December 2016 Accepted: 14 January 2018

Published online: 28 February 2018

## References

- P Do, A Voisin, E Levrat, B Lung. A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. *Reliability Engineering & System Safety*, 2015, 133: 22–32.
- H Hong, W Zhou, S Zhang, et al. Optimal condition-based maintenance decisions for systems with dependent stochastic degradation of components. *Reliability Engineering & System Safety*, 2014, 121: 276–288.
- J Lee, F Wu, W Zhao, M Ghaffari, et al. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 2014, 42(1): 314–334.
- M Esteves, E P Nunes. Prognostics health management: perspectives in engineering systems reliability prognostics. *Safety and Reliability of Complex Engineered Systems*, 2015: 2423–2431.
- K Medjaher, D A Tobon-Mejia, N Zerhouni. Remaining useful life estimation of critical components with application to bearings. *Reliability, IEEE Transactions on*, 2012, 61(2): 292–302.
- G Niu. *Data-driven technology for engineering system health management: design approach, feature construction, fault diagnosis, prognostics, fusion and decisions*. Springer, 2016.
- J Scheerens, H Luyten, S M van den Berg, et al. Exploration of direct and indirect associations of system-level policy-amenable variables with reading literacy performance. *Educational Research and Evaluation*, 2015, 21(1): 15–39.
- K Medjaher, D A Tobon-Mejia, N Zerhouni. Remaining useful life estimation of critical components with application to bearings. *IEEE Transactions on Reliability*, 2012, 61(2): 292–302.
- H C Zhang, S Liu, H Lu, Y Zhang, et al. Remanufacturing and remaining useful life assessment. In: *Handbook of Manufacturing Engineering and Technology*, Springer, 2015: 3137–3193.
- O Cappé, S J Godsill, E Moulines. An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE*, 2007, 95(5): 899–924.
- X Zhang, X Chen, B Li, et al. Review of life prediction for mechanical major equipments. *Journal of Mechanical Engineering*, 2011, 47(11): 100–116. (in Chinese)
- J Sun, H Zuo, W Wang, et al. Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance. *Mechanical Systems and Signal Processing*, 2012, 28: 585–596.
- X-S Si, W Wang, C-H Hu, et al. A wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. *Mechanical Systems and Signal Processing*, 2013, 35(1): 219–237.
- D Meimaroglou, C Kiparissides. Review of Monte Carlo methods for the prediction of distributed molecular and morphological polymer properties. *Industrial & Engineering Chemistry Research*, 2014, 53(22): 8963–8979.
- C Andrieu, N De Freitas, A Doucet, et al. An introduction to MCMC for machine learning. *Machine Learning*, 2003, 50(1-2): 5–43.
- C Pooley, S Bishop, G Marion. Using model-based proposals for fast parameter inference on discrete state space, continuous-time Markov processes. *Journal of The Royal Society Interface*, 2015, 12(107): 20150225.
- J L Beck, A A Taflanidis. Prior and posterior robust stochastic predictions for dynamical systems using probability logic. *International Journal for Uncertainty Quantification*, 2013, 3(4): 271–288.
- G Casella, E I George. Explaining the Gibbs sampler. *The American Statistician*, 1992, 46(3): 167–174.
- C P Robert, G Casella. *Monte Carlo statistical methods*. Springer, 2004.
- A Doucet, N De Freitas, N Gordon. An introduction to sequential Monte Carlo methods. In: *Sequential Monte Carlo methods in practice*, Springer, 2001: 3–14.
- P Bunch, S Godsill. *The progressive proposal particle filter: Better approximations to the optimal importance density*. arXiv preprint [arXiv:1401.2791](https://arxiv.org/abs/1401.2791). 2014.
- J L Zhang, J S Liu. A new sequential importance sampling method and its application to the two-dimensional hydrophobic–hydrophilic model. *The Journal of Chemical Physics*, 2002, 117(7): 3492–3498.
- M E Orchard, G J Vachtsevanos. A particle-filtering approach for on-line fault diagnosis and failure prognosis. *Transactions of the Institute of Measurement and Control*, 2009, 31(3/4): 221–246.
- M E Orchard. A particle filtering-based framework for on-line fault diagnosis and failure prognosis. *Georgia Institute of Technology*, 2007.
- J S Liu, R Chen. Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 1998, 93(443): 1032–1044.
- J R Stroud, M Katzfuss, C K Wikle. *A Bayesian adaptive ensemble Kalman filter for sequential state and parameter estimation*. arXiv preprint [arXiv:1611.03835](https://arxiv.org/abs/1611.03835). 2016.
- B A Berg. Introduction to Markov chain Monte Carlo simulations and their statistical analysis. *Markov Chain Monte Carlo Lect Notes Ser Inst Math Sci Natl Univ Singap*, 2005, 7:1–52.
- W K Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 1970, 57(1): 97–109.
- J Hammersley. *Monte carlo methods*. Springer Science & Business Media, 2013.
- P Society. PHM data challenge <https://www.phmsociety.org/competition/phm/10>, 2010.
- Z Pan, N Balakrishnan. Reliability modeling of degradation of products with multiple performance characteristics based on gamma processes. *Reliability Engineering & System Safety*, 2011, 96(8): 949–957.
- G Jin, D E Matthews, Z Zhou. A Bayesian framework for on-line degradation assessment and residual life prediction of secondary batteries inspacecraft. *Reliability Engineering & System Safety*, 2013, 113: 7–20.
- N G Polson, J R Stroud, P Müller. Practical filtering with sequential parameter learning. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2008, 70(2): 413–428.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)