IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 July 2022, accepted 29 July 2022, date of publication 4 August 2022, date of current version 10 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3196016

RESEARCH ARTICLE

Evolving Connectionist System and Hidden Semi-Markov Model for Learning-Based Tool Wear Monitoring and Remaining Useful Life Prediction

MUQUAN LIN[1,3], SONG WANQING[1,3], DONGDONG CHEN[1,2,3], AND ENRICO ZIO[4,5], (Senior Member, IEEE)

[1]School of Electronic and Electrical Engineering, Minnan University of Science and Technology, Quanzhou 362700, China
[2]Jiangxi New Energy Technology Institute, Xinyu 338000, China
[3]Key Laboratory of Industrial Automation Control Technology and Application, Fujian University, Quanzhou 362700, China
[4]MINES ParisTech, CRC, PSL Research University, 06904 Sophia Antipolis, France
[5]Department of Energy, Politecnico di Milano, 20156 Milano, Italy

Corresponding author: Dongdong Chen (cdd1911@qq.com)

This work was supported in part by the Science and Technology Project of Quanzhou City under Grant 2020C011R and Grant 2019CT003, and in part by the Science Nature Fund of Fujian Province under Grant 2021J01530.

ABSTRACT Tool wear can cause dimensional accuracy and poor surface quality in milling process. During the operation of tool wear, it can also cause breakage and damage of the workpieces. To prevent these conditions, it's important that the tool wear is monitored and the remaining useful life (RUL) is predicted in real time. In this paper, time domain and frequency domain statistical features are firstly extracted using multi-sensory fusion method, including the cutting force, vibration and acoustic emission sensor. Seven eigenvectors are selected as the input of the prediction model based on the distance correlation coefficient between 140 feature vectors and the wear value, which provide the most sensitive features to wear faults. The paper establishes a nonlinear relationship between high-dimensional feature vectors and tools wear based on the evolving connectionist system (ECoS), which uses the incremental learning algorithm to realize real-time prediction of the tools wear. Finally, using the wear value predicted by ECoS as hidden state sequence of Hidden Semi-Markov Model (HSMM), the RUL prediction of the tool based on HMM is established. The 2010 PHM challenge data were used to train the model. The experimental result shows that in comparison with artificial neural network, the ECoS model has higher prediction accuracy, and its mean RMSE error for three tools is 14.8. In comparison with the RUL prediction of HMM model, Probability-based RUL prediction of HSMM is more stable.

INDEX TERMS Multi-sensor fusion, evolving connectionist system, incremental learning algorithm, HSMM, remaining useful life prediction.

ABBREVIATIONS

RUL	Remaining useful life.
HMM	Hidden markov model.
ECoS	Evolving connectionist system.
PHM	Prognostic and health management.
HSMM	Hidden semi-markov model.
RMSE	Root mean square error.

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li.

I. INTRODUCTION

A. BACKGROUND AND SIGNIFICANCE

At present, the number of workpieces, which requires high-precision processing is getting larger. This especially applied toaerospace industry. As the most important component of the milling process, the quality of the tool directly affects the quality of the processed workpiece, and the tools wear is the direct cause of poor quality of the workpieces and high rejection rate. If tool wear cannot be accurately measured, it is likely to bring out two issues: Firstly, Premature

replacement of tools that are still in their life span will lead to waste of tools and increase unnecessary costs. Secondly, the tool is not replaced in time, and the failed tool is used to process the workpiece, resulting in poor surface quality of the workpiece. [1] Hence real-time monitoring and accurate RUL prediction of tool wear are much important for manufacturing companies to improve the stability of the processing environment and the precision of the workpiece, to protect the machine tool and processing safety, and to enhance the production efficiency of the enterprise and cut down production costs. Some researches show that accurate prediction of tool wear is helpful for timely replacement of tool and doing necessary predictive maintenance, which will effectively reduce unplanned downtime by 75%, increase production efficiency by 10%-60% [2], and reduce production costs by 10%-40 % [3], [4].

B. LITERATURE REVIEW

At present, there exist two main monitoring methods of tool wear, including direct monitoring methods and indirect monitoring methods. The direct method mainly monitors the reflective intensity of the tool wear surface, e.g., changes of the blade displacement, radioactivity of cutting surface, contact resistance and workpiece size to measure the wear of the tool [5]. The indirect method mainly monitors signals to indirectly reflect the wear of the tool, such as workpieces temperature, ultrasonic signal, acceleration vibration signal, cutting force, the change of torque, motor power or current. Then statistical features are extracted through signal processing technology, which are not sensitive to different tool cutting conditions and noise, but to a certain degree of sensitivity to tool wear status, such as initial wear, stable wear and severe wear. Therefore, by monitoring the changes of these feature parameters, the status of the tool wear can be monitored. In terms of the existing monitoring technology of tool wear, the direct method has strong operability, obtaining intuitive data, avoiding complicated data analysis, but it is easily affected by cutting temperature, cutting fluid, and machine tool movement accuracy, with poor real-time performance. However, the indirect method can relatively collect accurate data, but it can't directly reflect the wear status of the tool. Therefore, it is a new development trend to establish a nonlinear relationship between direct and indirect methods by a complex mathematical model.

Current AI technology is extensively introduced to monitoring the tool wear and RUL. Fig.1 shows different monitoring method of tool wear [6]. Zhao *et al.* [12] introduced the Convolutional Neural Network (CNN) based on the one-dimensional signal converted into image to research fault diagnosis. Liu *et al.* [13], [14] used stochastic degradation model to predict machine RUL. Sohyung *et al.* [15] used the maximum value extracted from the cutting force as the eigenvector, and applied SVM to predict RUL. Letot *et al.* [16] selected four kinds of sensor signals to predict RUL of the tool wear with cutting force, vibration, acoustic emission (AE) as labelled data. The result indicated that model prediction accuracy using force signal was the best. Jáuregui *et al.* [17] used the force and vibration signal to define the state of tool wear. They discovered that compared with based on the vibration signal, the prediction results based on the force signal were better. Dong *et al.* [18] selected 16 eigenvalues extracted from the force signal, the predicted the tool wear based on the ANN model and obtained good experimental results. Shi and Gindy [19] used principal component analysis (PCA) to process a variety of sensor signals, and utilizing a LS-SVM model to achieve a better result. Yu *et al.* [20] introduced a weighted hidden Markov model (WHMM) to predict tool wear. Niaki *et al.* [21] introduced extended Kalman filter (EKF) to monitor tool wear under machining conditions of different feed rate. In comparison with estimating tool wear area, it can improve the estimation accuracy. He *et al.* [22] built a probability model based on Particle Learning (PL). This model has perfect robustness and computationally efficient. Penedo *et al.* [23] applied mixed approach of fuzzy KNN and least squares regression to monitor tool wear. This approach has better performance than the traditional fuzzy neural network. Traditional ANN has the following disadvantages: (1) unchanging topology structure during training, which indicates that their network structure is rigid and predefined; (2) requesting all data samples to offline train; and (3) requires multiple passes during training process. However, ANN is unstable to remember information about old samples if new samples are learned. The trained neural network topology is a supervised learner, which extracts features from a network whose ground truth values are known. If new data become available, a new offline training process must be carried out to add new samples to the training data set. Therefore, adding the new samples to the training dataset is of great importance, when processing tasks with complex non-stationary data is a typical condition [24].

C. PROBLEM STATEMENT

At present, tool wear monitoring technology is only focused on theory. A single sensor monitoring cannot fully and accurately reflect the characteristic information of tool wear. Therefore, it cannot have the advantages of high sensitivity, high fault tolerance and high reliability. At the same time, it cannot be adapted to the actual processing environment. Due to the changes of the working conditions in the actual machining process, there may be multiple failure modes for tool wear. The traditional non-evolutionary neural network performs well in predicting performance where the tool wear pattern does not significantly change during the test. However, when the fault status changes suddenly, the predictive performance will decrease. This article introduces an online detection method of tool wear. The main goal is to find new fault information in training sample based on online incremental learning algorithm [25]. Evolving neural networks, e.g., evolving connectionist system (ECoS), do not have fixed network topology structure different from traditional neural network [26]. The ECoS network can adjust the structure online, and the evolving layer allows to change

its connectionist structure so as to dynamically adapt to new sample information. The advantages of the ECoS model are listed as follows: (1) Fast learning through a single incremental training; (2) Strong resistance to catastrophic forgetting; (3) Strong generalization ability. Since the state of the actual signal can change over time, the predictive performance of "non-evolved" traditional neural networks may be greatly reduced [24], [27].

This paper sets up an evolving neural network prediction model of tools wear based on online incremental learning algorithm. The result shows that the proposed ECoS is a quite appropriate approach to monitor tool wear online. The process steps of tool wear monitoring and RUL prediction are as follows: collecting three kinds of sensor signals, extracting 140 eigenvectors from time domain, frequency domain and time-frequency features. Then seven eigenvectors sensitive to the tool wear are selected as the input of the prediction model based on the distance correlation coefficient between 140 eigenvectors and the wear. Finally, apply an incremental learning algorithm of ECoS to establish predictive model of tools wear. according to the wear value predicted, the RUL prediction of the tool based on the Hidden Semi-Markov Model (HSMM) is established.

The structure of the paper is arranged as follows. Section 2 addresses the tool wear prediction based on evolving connectionist system (ECoS). In Section 3, Remaining Useful Life prediction of HMM is given. In Section 4, Experiment set-up and procedure are provided to demonstrate the applicability and superiority of the proposed algorithm. In Section 5, comparative results and discussion are given. Then, the paper is concluded in Section 7.

II. TOOL WEAR PREDICTION USING EVOLVING NEURAL NETWORK

A. EVOLVING NEURAL NETWORK

Evolving connectionist system (ECoS) represent a family of constructive ANN algorithms, which were proposed by Kasabov [28], [29]. ECoS consists of three-layer neural network. Fig.2 shows its structure.

Layer 1 and layer 3 represent input and output, respectively.

The k_{th} neuron activation in the evolving layer is calculated by

$$A_{\mathrm{k}} = 1 - D_{jk} \tag{1}$$

where $D_{jk} \in [0, 1] \forall j, k$, indicates a normalized distance, considering the input vector and, a weight vector, both m dimensional column vectors. The Manhattan distance between and can be calculated as follows:

$$D_{jk} = \frac{\sum_{i=1}^{m} \left| W_{ik} - I_{(j)} \right|}{\sum_{i=1}^{m} \left| W_{ik} + I_{(j)} \right|} \tag{2}$$

where m indicates the number of input features. Activation functions based on equations (1) and (2) suggest that samples that match a pattern can lead to complete activation of a neuron in contrast, samples that are far away from a pattern can cause an activation value closing to zero.

B. ONLINE INCREMENTAL LEARNING ALGORITHM

Online incremental learning of ECoS includes in fitting new data samples as weights of the EcoS evolving layer [27]. The connection weights are modified or a new neuron is added to the evolving layer. The decision about establishing a new neuron and adjusting weights depends on the novelty of the input sample data. The degree of novelty of the data is determined by comparing the activation level of the most active neuron with a predefined threshold. Or by comparing output error with predefined error threshold, new neurons are added to the evolving layer of ECoS. No neurons are added to network structure if the activation values or error are less than predefined threshold values. In this case, online increment learning can adjust the connection weights of neuron to fit new sample information, which is the most activated.

The input weights are calculated by

$$W_{ik(t+1)} = W_{ik(t)} + \eta_1 (I_{i(t+1)} - W_{ik(t)}) \tag{3}$$

The output weight is calculated as follows:

$$W_{k1(t+1)} = W_{k1(t)} + \eta_2 (A_k E_1) \tag{4}$$

Equations (3) and (4) represent recursive calculation. Parameters, denote the learning rates.

$$E_1 = y_1 - \hat{y}_1 \tag{5}$$

where is the estimation errors.

Firstly, a new neuron added can be placed using minimum distance (MD) approaches to facilitate potential further aggregations, which means that new neurons are placed in the position. Output weight vector of the neuron is the closest to the real output vector.

Since new sample data can lead to a large number of neurons to be added to the evolving, ECoS can increase rapidly in size, which can lead to memory issues if aggregation algorithm is not used. Neuron aggregation means that two or more neurons are aggregated into one [27] Neurons are aggregated if the calculated distance is less than the given threshold. Input and output distances among two neurons in evolving layer during aggregation, namely, neurons o and p, are calculated as

$$D_{op}^{in} = \frac{\sum_{i=1}^{m} \left| W_{io} - W_{ip} \right|}{\sum_{i=1}^{m} \left| W_{io} + W_{ip} \right|} \tag{6}$$

and

$$D_{op}^{out} = \frac{\sum_{i=1}^{m} \left| W_{o1} - W_{p1} \right|}{\sum_{i=1}^{m} \left| W_{o1} + W_{p1} \right|} \tag{7}$$

Before training starts, because of no Neurons existence in the evolving layer, the network structure does not need initialization and no weights need to be set. Online incremental learning algorithm of ECoS can be summarized as follows

Set A_{thr}, E_{thr}, D_{thr} and learning rates η_1, η_2
  for each $(I(j), y(j)), j = 1, 2, \ldots,$ **do**
  Propagate $I(j)$ through the ECoS
  Find the most activated neuron k in the evolving layer
    Calculate the error E1 between $y(j)$ and $\hat{y}(j)$
    if $A_k < A_{the} \forall k$ or $E_j < E_{the}$ **then**
      Add a neuron based on the MD approach
    else
      Update weights $W_{i,k} \forall i$, $W_{k,1}$
    end if
    if $\max\left(D_{op}^{in}, D_{op}^{out}\right) < D_{\text{thr}}$ **then**
    Aggregate neurons
    end if
  end for

C. PERFORMANCE EVALUATION OF PREDICTIVE MODEL

The following four error indicators are selected to calculate the accuracy of ECoS prediction result:

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{8}$$

Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{9}$$

Normalized Mean Squared Error (NMSE)

$$\text{NMSE} = \frac{n-1}{n\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{10}$$

SCORE

$$S_i = \begin{cases} e^{y_i - \hat{y}_i/13} - 1, & \hat{y}_i \le y_i \\ e^{\hat{y}_i - y_i/10} - 1, & \hat{y}_i > y_i, \end{cases}$$

$$\text{SCORE} = \sum_{i=1}^{n} S_i \tag{11}$$

where $\hat{y}_i$ is the predictive value, y_i is the real value and $\bar{y}$ is the mean value. The higher the value, the lower the prediction accuracy of the model, n is number of samples.

III. RUL PREDICTIVE ALGORITHM OF HSMM

If it is assumed that the tool wear degradation process needs to go through N states, $s = \{s_1, s_2, \ldots, s_N\}$ is the hidden state set as constitutes the HSMM model; The multi-sensor fusion feature is composed of a vector ([Fx-Ku, Fy-RMS, Fz-RMS, Vx-PI, Vy-SD, Vz-KF, VE-RMS]) as the observation sequence of the HSMM model. The structure of the HSMM-based tool RUL prediction model is shown in Figure 3. For each wear state, multiple cutting experiments may be required. The resulting observations form a sub-sequence $\{o_{i1}, o_{i2}, \ldots, o_{ik}\}$, all observation sub-sequences of the wear state constitute an observation sequence set.

According to this model, the tool wear migration process is random, which is represented by state transition probability matrix A. The process of generating observation values for each tool wear state is also random. Because the observation values change continuously, the observation probability matrix B cannot be used to description, using a multivariate Gaussian distribution $\left(\mu_i, \sum_i\right)$ description. Among them, since tool wear is a one-way process, its state transition probability matrix A can be expressed as follows:

$$\begin{bmatrix} a_{11} & a_{12} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & a_{23} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & a_{(N-1)(N-1)} & a_{(N-1)N} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Tool wear always occurs from the first wear state, and its initial state probability vector can be expressed as: $\pi = (1, 0, \ldots, 0)$.

The tool RUL algorithm steps of HSMM are as follows:

Step 1. Tool wear discretization. According to the size of the tool wear value, the tool wear process is divided into N discrete states, denoted as $s_1, s_2, \ldots, s_N$, And record the tool wear value at the beginning of each state $w_1, w_2, \ldots, w_N$.

Step 2. Model training with labelled data. Analyzing the monitoring signal of each sensor, the characteristic vector sensitive to the change of tool wear is extracted as the observation value of HMM. A tool has a wear states sequence $S^i = (s_1, s_2, \ldots, s_T,)$, and a observation sequence of signal feature $O^i = (o_1, o_2, \ldots, o_T)$. Carrying out M times of tool life experiment, M observation sequences and corresponding state sequences are got $\left(O^1, S^1\right), \left(O^2, S^2\right), \ldots, \left(O^M, S^M\right)$. The maximum likelihood estimation method is applied for training the model parameters $\lambda = (\pi, A, \mu_i, \Sigma_i)$.

Step 3. Model testing. For the new tool, the observation sequence at the current time t and model λ is known, and we need to calculate the probability distribution $\gamma_t(i)$ of the tool state at t time. Then calculate the expectation of the current wear value.

Step 4. According to the state transition matrix A, the probability distribution $\gamma_{t+1}(i)$ of the tool state for the next moment can be obtained.

Step 5. Computation of the tool wear value $w\left(t'\right)$ for next moment in time t'

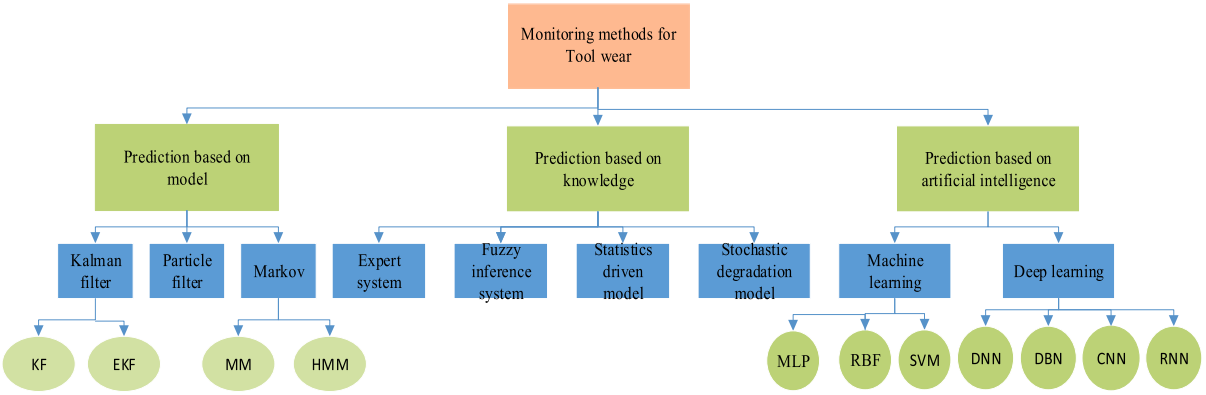


FIGURE 1. Classification of tool wear prediction.

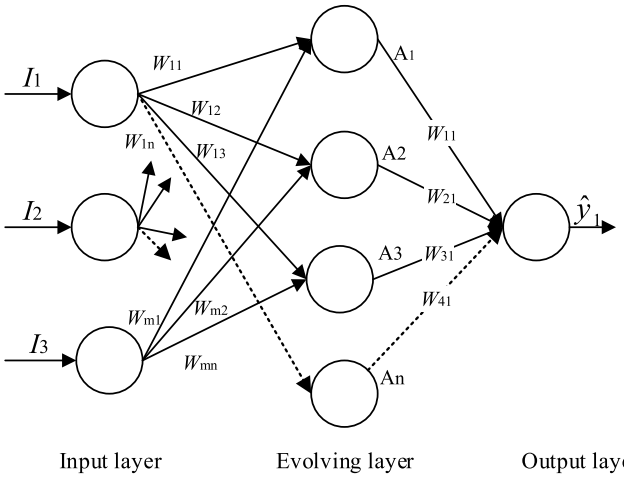


FIGURE 2. ECoS network structure.

Step 6. Estimating of the time t_f to the failure threshold. The RUL can be calculated as follows: RUL $= t_f - t$.

IV. EXPERIMENTAL VALIDATION

A. EXPERIMENTAL SETUP

Experimental sample data were obtained from Roders Tech RFM760 milling cutting center [30]. Schematic representation of the experimental platform if depicted in Fig. 4. The cutting parameters of each tool during the processing are unchanged. Experimental parameters in the milling process are shown in Table 1. Flowchart of tool wear prediction is shown in Fig.2 [31]. The dynamometers for measuring cutting force of three directions were installed on the machining table. Three Kistler Piezo accelerometers were installed on the workpiece so as to measure the tool vibration in three directions. DAQ NI PCI1200 was used to save sensor data with a 50HZ sampling frequency. The sensor data includes: cutting force (Fx, Fy, Fz), acceleration vibration in three directions (Vx,Vy,Vz) and acoustic emission (AE). The LEICA MZ12 microscope is used to measure VB value of three flutes offline after finishing each surface, which will be the target output. The goal is to predict actual VB value The test tools were recorded named c1, c4, and c6 as our dataset. Each tool includes 315 data samples, while each data sample has a corresponding wear value.

B. TOOL WEAR MEASUREMENT

Each tool needs complete 315 processing and the length of each pass is 108 mm. The flank wear value of each cutting is measured and recorded. Analyzing the experimental full-life wear data of Tools, the VB wear curves of Tool 1, 4 and 6 are shown in Fig.5, showing the relationship between the VB value and the number of cuts. Fig. 5 shows three stages change of tool VB value. In the initial stage, due to the sharpness of the new cutting edge, tool wear rises rapidly. As the contact area between the tool and the workpiece gets larger, the wear enters a steady stage and starts to increase slowly. When the threshold is achieved, the tool wear enters the Sharp stage. In this study we use the maximum value of the wear among the three flutes.

Fig.6 shows a flowchart of the tools condition detection and RUL prediction in the milling process. The flowchart mainly includes the following steps: (1) obtaining multi-sensor original signals, including cutting force, vibration and acoustic emission signals; (2) feature extraction methods and dimension reduction algorithm; and (3) detection of tools wear and RUL prediction.

C. SIGNAL ANALYSIS

1) MULTISENSOR FUSION SIGNAL ANALYSIS

Analyzing the experimental full-life sensors data of Tool 1, Tool 4, Tool 6 with 315 cuts, Visual analysis of the cutting force signals in z-directions based on the whole life data of tool 1 are shown in Fig.7(a), which indicates that the cutting force signal is more sensitive to the steady wear stage and severe wear of the tool. Original signals of acceleration vibration in x-directions are shown in Fig.7(b), which reveals that vibration signal is the most sensitive to sharp wear of the tool. Acoustic emission (AE) signal is shown in Fig.7(c), which shows that AE signal is more sensitive to the initial wear and severe wear of the tool. Due to the various forms of tool status failures in the milling process and the limited monitoring conditions, this paper uses multi-sensor data

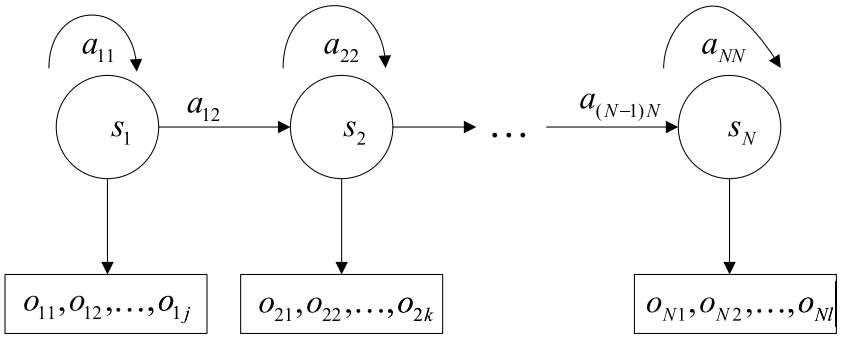
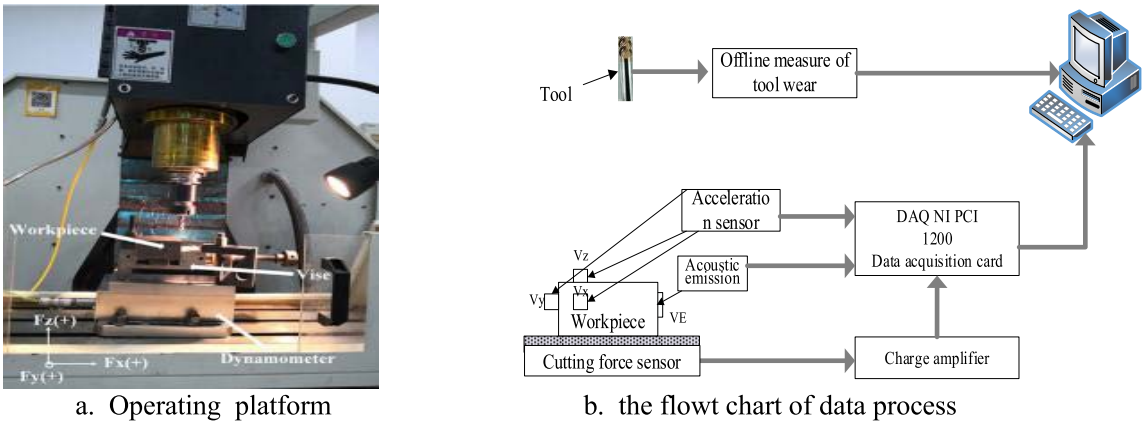


FIGURE 3. Tool wear monitoring structure based on HMM.



a. Operating platform b. the flowt chart of data process

FIGURE 4. PHM 2010 competition platform.

TABLE 1. Experimental parameters of PHM 2010 competition.

Parameters	Values	Parameters	Values
Machine type	Roders Tech RFM 760	Radial depth in y-direction	0.125mm
Workpiece material	Radial depth	Axial depth in z-direction	0.2mm
Tool	tungsten carbide ball head	Class of sensors	3
Cutting speed	10,400 rpm/min	Number of sensor channels	7
Feed rate in x-direction	1555 mm/min	Sampling frequency	50kHz
Milling mode	Down milling	Number of cuts	315

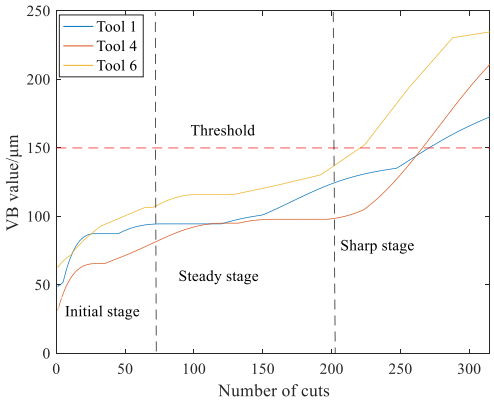


FIGURE 5. The relationship between tool wearVB and number of cuts.

fusion method to improve the accuracy and reliability of tool condition monitoring.

2) MULTIDOMAIN CANDIDATE FEATURE PARAMETERS

In order to avoid the shortcomings of single-domain feature parameters containing incomplete tool status information, multi-domain combination method was used to construct a candidate feature parameter sets. Let A total of 20 features are extracted from each channel signal, including 8 dimensional features of time domain, 8 dimensionless time domain features, and 4 frequency domain features. Therefore, a total of 140 features from 7 sensor signals were extracted through frequency domain and time domain analysis. The dimensional features of time domain include max, mean, mean of absolute values(AM), Peak-to-peak value(P2P), variance (Va), standard deviation(SD), root-mean square(RMS)and peak index(PI) [30]; the dimensionless features of time domain have skewness (SK), kurtosis(KU), margin factor(MF), crest factor(CF), impulse factor(IF), waveform factor(WF), skewness factor(SF)and

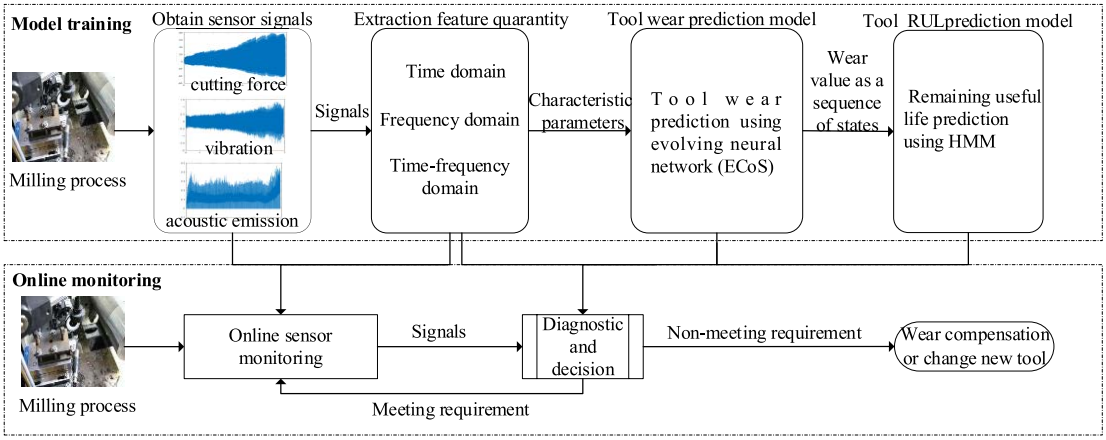


FIGURE 6. Flowchart of the tools condition detection in milling process.

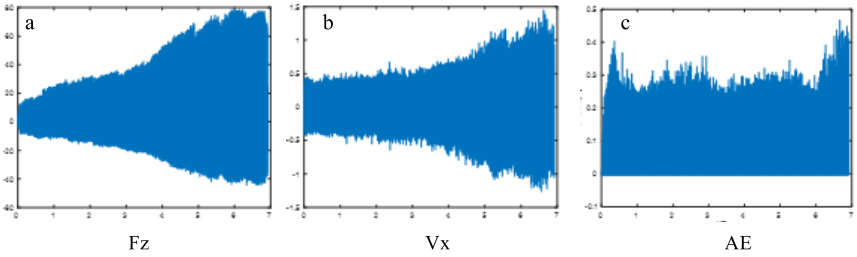


FIGURE 7. Full life data of Fz, Vx, AE sensors of tool 1.

jumping factor(JF) [32]; the frequency domain features are vibration velocity energy(VE), vibration intensity(VI), the amplitude at the pass 520Hz frequency of the power spectrum(PSA_520Hz) and the mean of the power spectrum amplitude(PSAM). Analyzing the features of the cutting force signals in y-directions of tool 1, tool 4 and tool 6, Fig.8 shows the relationship between dimensional features and number of cuts; Fig.9 shows the relationship between dimensionless features and number of cuts; Fig.10 shows the relationship between frequency domain features and number of cuts.

It can be seen from Figs.8-10 that different statistical features in the time domain reflect tool health conditions from different stage. In different wear levels, it is different for each indicator to contain wear information. Even many indicators cannot reflect the changes of wear to some extent. Therefore, more features should be extracted to effectively monitor the wear of tool and moreover sensitive features should also be further screened out from them.

3) SELECTION OF CHARACTERISTIC QUANTITY

The above candidate feature parameters are not all useful for predicting wear value and further feature selection is required. Distance correlation coefficient (DC) was used to measure the degree of correlation between 140 features quantity of tool 1 and VB value. Features with large correlation coefficients are selected as predictor variables. [33] defined DC concept. DC overcomes disadvantages of the traditional Pearson correlation coefficient, which can measure the relationship between nonlinear correlation variables. The DC calculation formula is as follow between random vectors x and y:

$$R^2(x,y) = \frac{v^2(x,y)}{\sqrt{v^2(x,x)\,v^2(y,y)}} \tag{12}$$

where $v^2(x,y)$ is defined as follows:

$$v^2(x,y) = \frac{1}{n}\sum_{j,j-1}^{n} A_{i,j}B_{i,j} \tag{13}$$

$$\begin{aligned} A_{i,j} = {} & \|x_i - x_j\|_2 - \frac{1}{n}\sum_{k=1}^{n}\|x_k - x_j\|_2 \\ & - \frac{1}{n}\sum_{l=1}^{n}\|x_i - x_l\|_2 \\ & + \frac{1}{n^2}\sum_{k,l=1}^{n}\|x_k - x_l\|_2 \end{aligned} \tag{14}$$

$$\begin{aligned} B_{i,j} = {} & \|y_i - y_j\|_2 - \frac{1}{n}\sum_{k=1}^{n}\|y_k - y_j\|_2 \\ & - \frac{1}{n}\sum_{l=1}^{n}\|y_i - y_l\|_2 \\ & + \frac{1}{n^2}\sum_{k,l=1}^{n}\|y_k - y_l\|_2 \end{aligned} \tag{15}$$

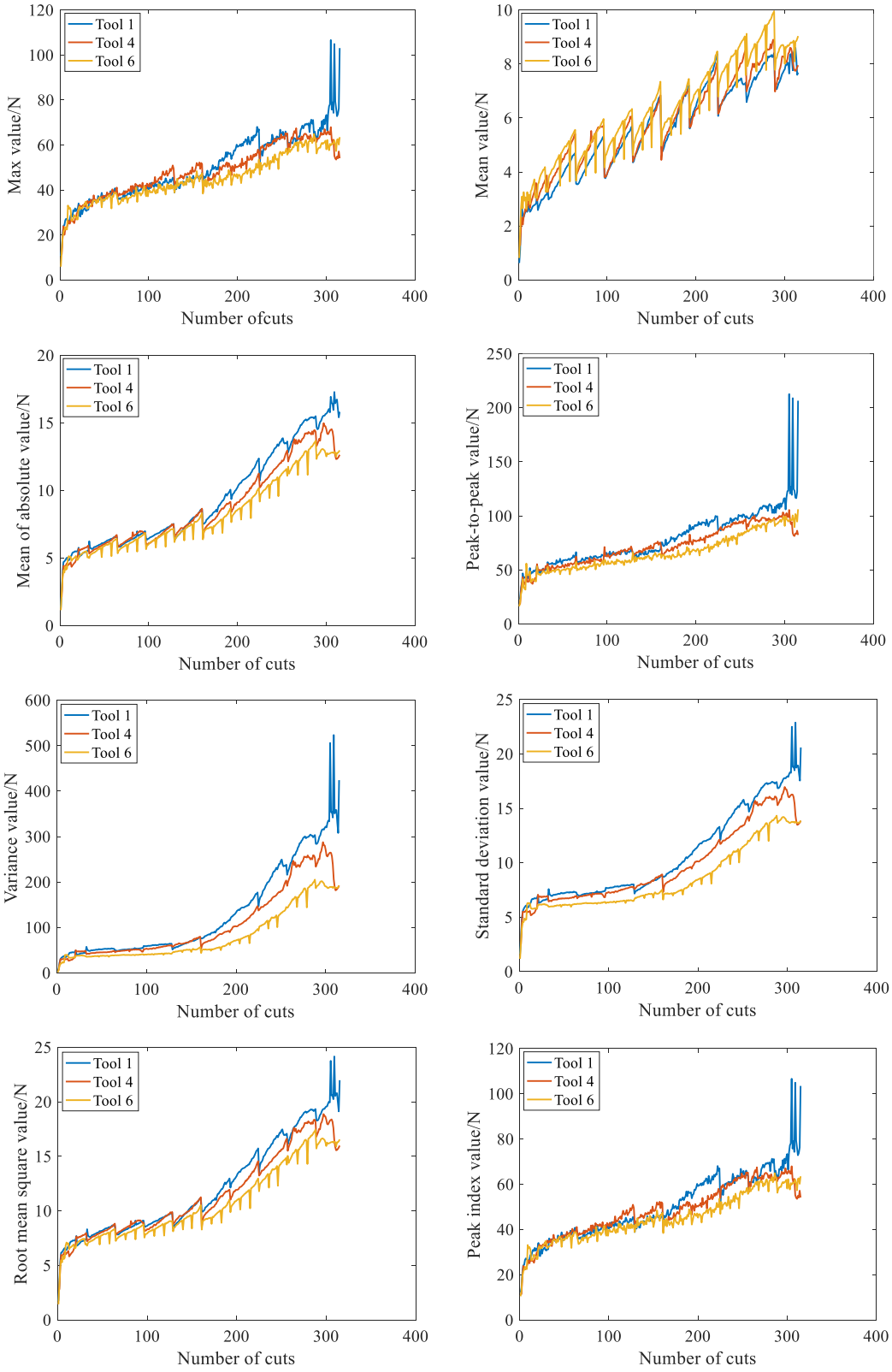


FIGURE 8. The relationship between dimensional characteristics and number of cuts.

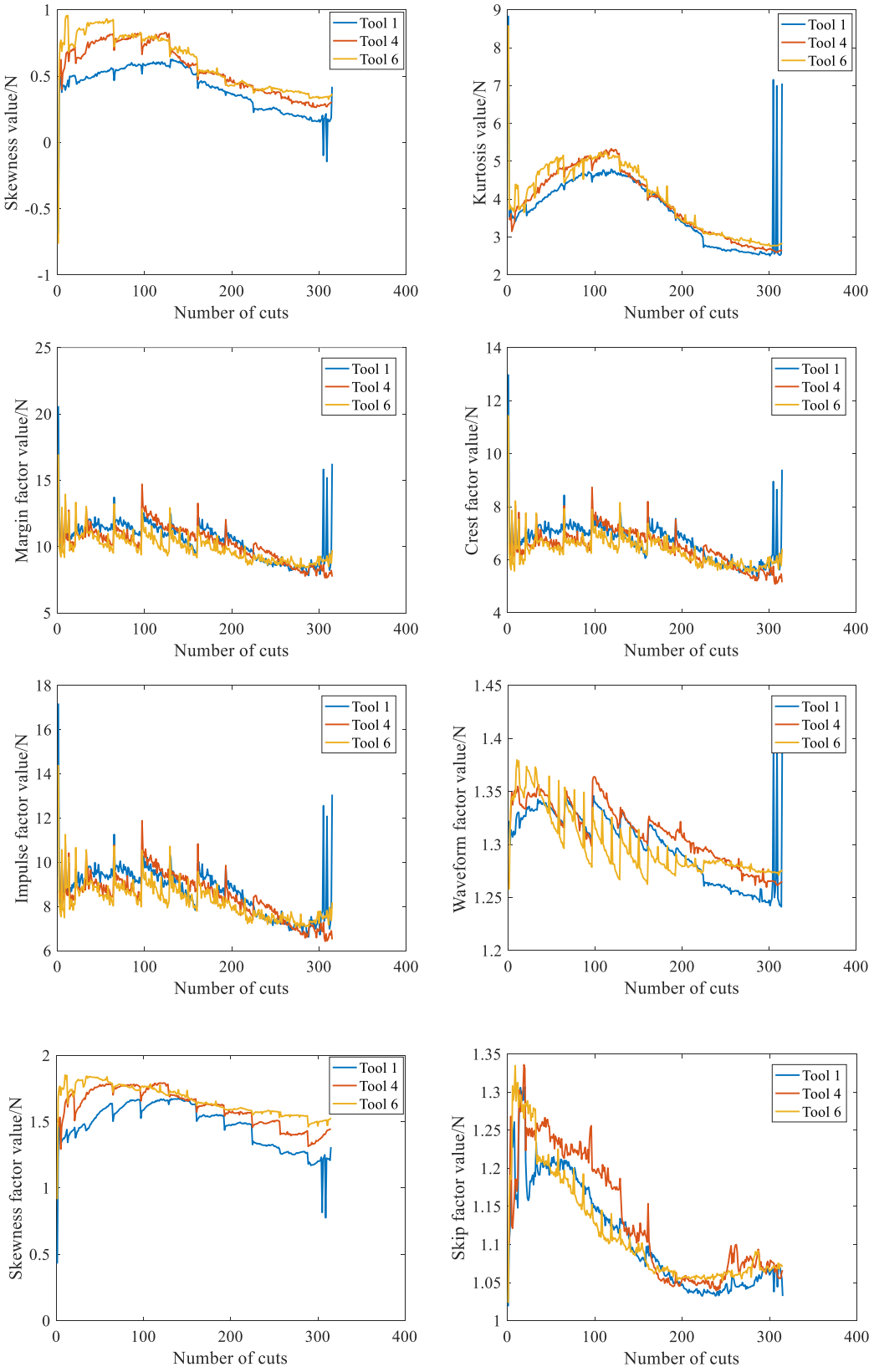


FIGURE 9. The relationship between frequency characteristics and number of cuts.

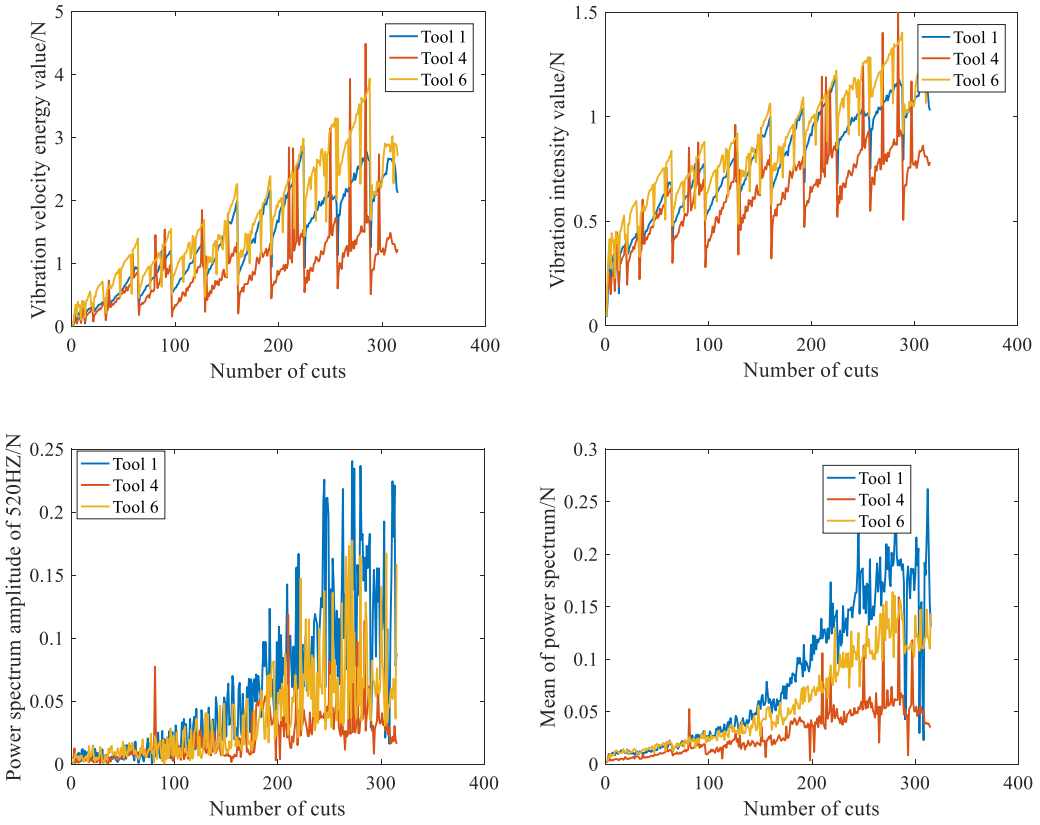


FIGURE 10. The relationship between dimensionless parameters and number of cuts.

Similarly, ones have:

$$v^2(x,x) = \frac{1}{n^2}\sum_{i,j=1}^{n} A_{i,j}^2 \tag{16}$$

$$v^2(y,y) = \frac{1}{n^2}\sum_{i,j=1}^{n} B_{i,j}^2 \tag{17}$$

where x indicates a certain feature vector of sensor signal and y indicates wear value (VB) of the tool. R^2 indicates DC value which is between 0 and 1. The larger the R^2 is, the stronger the correlation is. There is no correlation when $R^2 = 0$; It represents a weak correlation when $R^2 \in (0, 0.6]$; It represents a strong correlation when $R^2 \in [0.6, 1]$.

It is critical for improving prediction accuracy to ensure a strong correlation between eigenvectors and the wear value. Table 2 shows eigenvectors with the strong correlation coefficient for tool 1, x, y, and z indicate three different directions. RMS-Fy is root mean square feature extracted from the cutting force sensor of y direction.

V. RESULTS AND DISCUSSION

A. PERFORMANCE EVALUATION OF THE ECoS MODEL

Fig. 11 shows RMSE result, tool 1 as training data. The signal characteristic vectors of multi-sensor fusion [Fx-KU, Fy-RMS, Fz-RMS, Vx-PI, Vy-SD, Vz-JF, VE-VE] are used to predict wear value in this section. The ECoS can keep on learning new samples signals so as to increase the

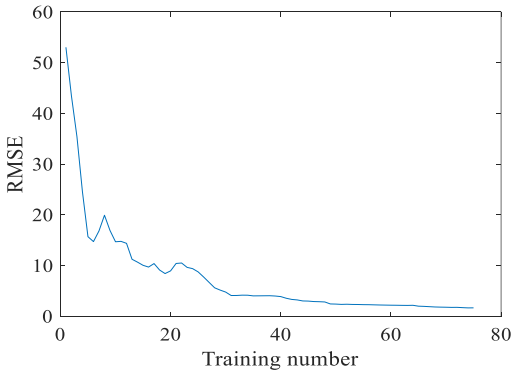


FIGURE 11. The results of training RMSE for tool 1.

prediction accuracy of trained model. The fast learning of online increment can adaptively update the weight of input and output based on the novelty of input sample.

Cross-validation method was used to verify the prediction accuracy of the ECoS model. The max wear value is 150 when the tool wear achieves the threshold. Firstly, using tools 4 and 6 as the training tools, tool 1 as the test tool to predict the corresponding tool wear values using seven input signal eigenvectors. The prediction results are shown in Fig. 12 (a). Then the tool 4 is used as the test set, and the other two tools as the training set for verification. The results are shown in Fig. 12(b) and Fig. 12(c). Optimal parameters of in the training ECoS model are respectively set to 0.9, 0.001,

TABLE 2. R^2 between characteristic quantity and wear value for tool 1.

Characteristic	R^2	Characteristic	R^2	Characteristic	R^2	Characteristic	R^2
Fx-KU	**0.812**	Fy-P2P	0.728	Vx-SD	0.761	Vz-SD	0.821
Fx-PSAM-520Hz	0.757	Fz-RMS	**0.884**	Vx-AM	0.722	Vz-RMS	0.741
Fx-RMS	0.732	Fz-AM	0.793	Vy-SD	**0.977**	Vz-VA	0.695
Fx-SK	0.712	Fz-SD	0.732	Vy-AM	0.826	VE-VE	**0.785**
Fy-RMS	**0.854**	Fz-VA	0.651	Vy-RMS	0.791	VE-VI	0.607
Fy-AM	0.811	Vx-PI	**0.851**	Vy-VA	0.754	VE-AM	0.571
Fy-SD	0.741	Vx-Max	0.848	Vz-JF	**0.857**	VE-RMS	0.568

The bold portion is a larger distance correlation coefficient as input eigenvector of tool wear monitoring model.

TABLE 3. The RMSE error for the testing tools using ECoS and traditional ANN.

Testing tools	Tool 1	Tool 4	Tool 6	Mean RMSE
ECoS	18.13	14.15	12.14	14.80
ANN	51.56	30.23	42.31	41.37

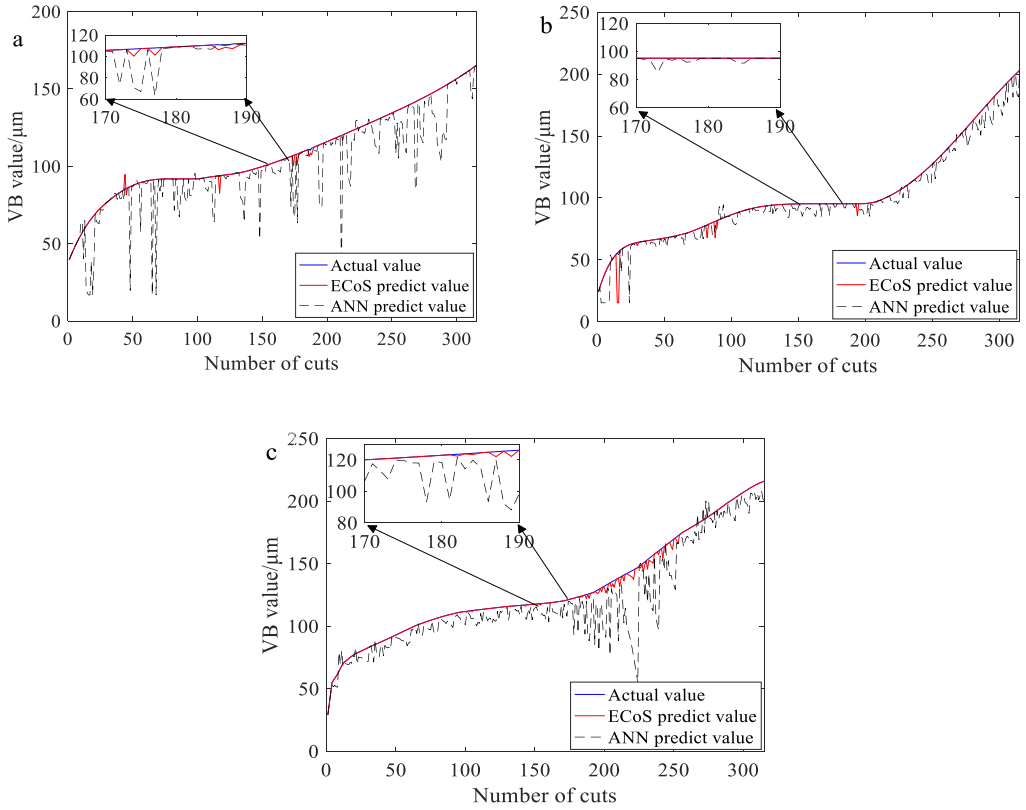


FIGURE 12. Predicted results of tool wear VB: a) wear estimation for tool 1 using tools 4 and 6 as training tools; b) wear estimation for cutter 4 using tools 1 and 6 as training tools; c) wear estimation for tool 6 using tools 1 and 4 as training tools.

and 0.5. When the result of ECoS training model keep stable and not change, the number of neurons in the evolving layer is 340. Compared with ANN predict model, the prediction accuracy of ECoS model is higher. The error results on different test tools based on the ECoS prediction model are shown in Figure 13. The good generalization performance of the prediction model is verified. Since the new tool is very unstable in the initial stage of wear, the prediction accuracy is not high. In the stable wear stage, the cutting force increases rapidly, resulting in a rapid increase of the tool wear rate. Table 3 shows the RMSE error results of different prediction models.

To verify the predictive influence of different sensor characteristics on the ECoS model, the tool 4 is used as the test set, and the other two tools as the training set for verification. Compared with using the force, vibration and VE single sensor features, using multi-sensor fusion feature as the input vectors of ECoS can improve prediction accuracy. Seven feature parameters from 60 candidate parameters of single force sensor signal were selected based on larger distance correlation coefficient as input vectors, including [Fz-RMS, Fy-RMS, Fx-Ku, Fy-AM, Fz-AM, Fx-PSAM, Fy-SD]. Correspondingly, the feature parameters of single vibration sensor signal selected are [Vy-SD, Vy-JF, Vx-PI,

TABLE 4. Evaluation index of VB value prediction results in different sensor signals for tool 1 using tools 4 and 6 as training tools.

Feature classification	MAE	RMSE	AMSE	SCORE
Force	16.56	**19.45**	2.73	653.9
Vibration	18.43	40.23	3.56	856.1
VE	30.65	53.45	9.21	970.3
Multi-sensor fusion	**15.23**	20.13	**1.11**	**580.2**

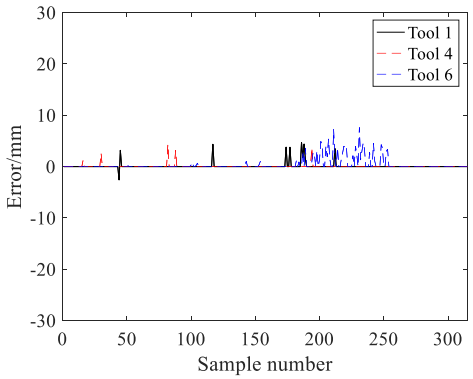


FIGURE 13. The error results of wear prediction.

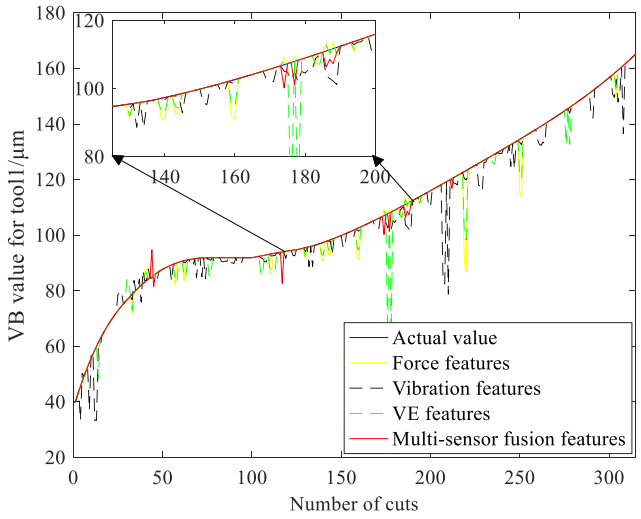


FIGURE 14. Predicted results of ECoS using single sensor and multi-sensor fusion.

Vx-Max, Vy-AM, Vy-RMS, Vx-SD]. The results are shown in Fig. 14. The prediction results based multi-sensor fusion features prove the high-precision prediction of the ECoS. The prediction result based vibration sensor features is the worst because vibration signals are susceptible to noise, therefore the prediction accuracy reduces. The prediction result based force features is better than based VE sensor because the force signal is more sensitive to changes of tool wear in the milling process. Table 4 shows comparison of ECoS prediction results using different features.

B. RUL PREDICTION WITH HSMM

The predicted value of RUL is the estimated number of time steps that the degradation path first passes the critical threshold. The wear threshold in this article is set to using

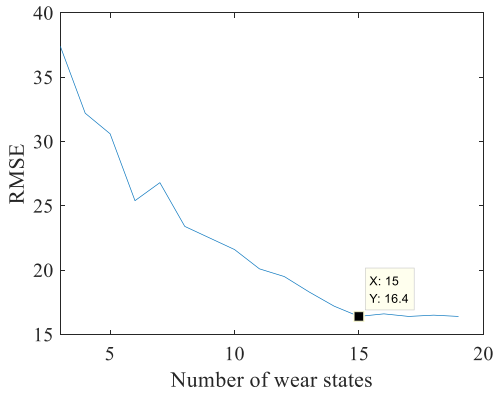


FIGURE 15. Relationship between training of RMSE error and number of wear states.

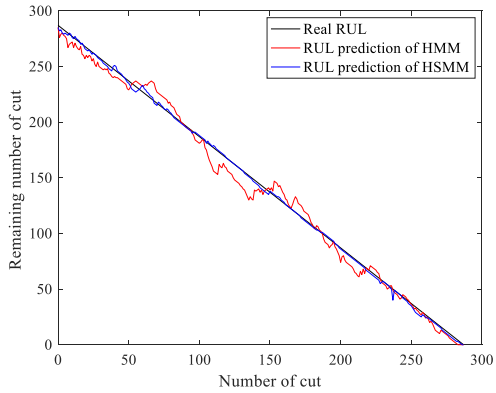


FIGURE 16. RUL prediction results.

tools 4 and 6 as the training tools, tool 1 as the test tool to predict the tool RUL. The number of hidden states of the HSMM model has a greater impact on the model prediction accuracy. It is vital to set the number of tool wear states before training. The training errors (RMSE) of the HMM model are calculated under different numbers of wear states. The results are shown in Fig. 15. When the number of wear states reaches 15, the training error decreases slowly. When the number of wear cycles is large, the sample points in each wear state also decrease. Therefore, the least tool wear state is selected within the allowable error range. From this, it is determined that the number of tool wear states is 15.

The failure threshold of VB wear is set to in this paper. Firstly, using the tools 4 and 6 as the training set, the tool 1 as the test is set to predict the RUL sequence of the tool at the current moment. The prediction results of the HSMM-based model and the ECoS-based model are shown in Fig. 16. using cross-validated method verifies the accuracy of the above models. Then using the tools 4 and 6 as the test set and the

TABLE 5. Comparison of model prediction accuracy using cross-validated method.

	RUL Prediction model	MSE	MRE
Verification I	HSMM	150.45	23.42%
	HMM	160.58	24.21%
Verification II	HSMM	73.41	6.31%
	HMM	120.23	14.36%
Verification III	HSMM	130.54	20.12%
	HMM	170.62	17.53%

other two tools as the training is set for verification. The prediction results show that when the observation sequence is less, the accuracy of the model is lower. As the time sequence increases, the probability-based RUL prediction of the HMM is more accurate. The predicted RUL curve is close to the actual RUL curve. The prediction accuracy in the later period is higher than that in the early period, which has greater significance for predictive maintenance of tool. Compared with RUL prediction of ECoS model, Probability-based RUL prediction of HSMM is more stable and quantify the uncertain effects of noise and state change to some extent. In order to compare the accuracy of the two tool remaining prediction methods, the mean square error MSE and the average absolute percentage error MAPE are used as the criteria to compare the results of the three cross-validation. The results are shown in Table 5.

VI. CONCLUSION

This paper uses multi-sensor information fusion technology to establish online tool wear monitoring based on ECoS neural network, and establishes the dynamic complex non-linearity between the signal characteristics of each sensor that indirectly reflects tool wear and the flank wear value VB of direct tool wear. Relationship, continuous tool wear values can be obtained, which more intuitively reflects the law of performance degradation of actual tools. Then a residual life prediction model based on HSMM is established. Through accurate prediction of the number of remaining passes of the tool, it provides decision-making theoretical support for tool maintenance and the optimal replacement time of the tool. The main conclusions of this paper are highlighted as follows:

1) ECoS builds a complex non-linearity relationship between seven characteristic quantities (Fx-KU, Fy-RMS, Fz-RMS, Vx-PI, Vy-SD, Vz-JF, VE-VE) of multi-sensor fusion signals and VB values so as to timely monitor tool wear online. The above seven features can distinctly reflect the severe degree of the wear as tool wear becomes more and more sever. and the wear rate in the sharp stage is significantly larger than the wear rate in the initial stage. When the wear rate changes significantly, there are sudden changes for ECoS prediction result, which can reflect the current state of the tool to some extent. Compared with conventional ANN, ECoS prediction model has higher prediction accuracy and its mean RMSE is 14.8. The evolving neural network (ECoS) provides a new modeling method for non-linear complex systems with multiple outputs and single output.

2) Four error indicators (MAE/RMSE/NMSE/SCORE) were used to measure the prediction accuracy of the ECoS with single sensor features and multi-sensor fusion features. The prediction performance of the tool wear based multi-sensor fusion method is better than single sensor.

3) Considering the uncertainty impact of environmental changes on tool wear, our study proposes RUL prediction model of tool based on HSMM. The wear output information of the ECoS evolutionary neural network is used as the hidden state sequence of the HSMM model for training model parameters. Different from the HMM model, the HSMM model considers the residence time of the historical state into the current life prediction. The life prediction is more realistic and the prediction result is more accurate than the HMM model. Accurate prediction of RUL provides a decision-making analysis basis for tool predictive maintenance or changing new tool.

REFERENCES

[1] A. H. Musfirah, J. A. Ghani, and C. H. C. Haron, "Tool wear and surface integrity of inconel 718 in dry and cryogenic coolant at high cutting speed," *Wear*, vols. 376–377, pp. 125–133, Apr. 2017, doi: 10.1016/j.wear.2017.01.031.

[2] H. Zeng, R. Yan, P. Du, M. Zhang, and F. Peng, "Notch wear prediction model in high speed milling of AerMet100 steel with bull-nose tool considering the influence of stress concentration," *Wear*, vols. 408–409, pp. 228–237, Aug. 2018, doi: 10.1016/j.wear.2018.05.024.

[3] A. Kumar, S. Kaminski, S. N. Melkote, and C. Arcona, "Effect of wear of diamond wire on surface morphology, roughness and subsurface damage of silicon wafers," *Wear*, vols. 364–365, pp. 163–168, Oct. 2016, doi: 10.1016/j.wear.2016.07.009.

[4] M. Aramesh, Y. Shaban, S. Yacout, M. H. Attia, H. A. Kishawy, and M. Balazinski, "Survival life analysis applied to tool life estimation with variable cutting conditions when machining titanium metal matrix composites (Ti-MMCs)," *Machining Sci. Technol.*, vol. 20, no. 1, pp. 132–147, Jan. 2016, doi: 10.1080/10910344.2015.1133916.

[5] N. Ambhore, D. Kamble, S. Chinchanikar, and V. Wayal, "Tool condition monitoring system: A review," *Mater. Today, Proc.*, vol. 2, nos. 4–5, pp. 3419–3428, 2015, doi: 10.1016/j.matpr.2015.07.317.

[6] N. Ghosh, Y. B. Ravi, A. Patra, S. Mukhopadhyay, S. Paul, A. R. Mohanty, and A. B. Chattopadhyay, "Estimation of tool wear during CNC milling using neural network-based sensor fusion," *Mech. Syst. Signal Process.*, vol. 21, no. 1, pp. 466–479, Jan. 2007, doi: 10.1016/j.ymssp.2005.10.010.

[7] X.-C. Cao, B.-Q. Chen, B. Yao, and W.-P. He, "Combining translation-invariant wavelet frames and convolutional neural network for intelligent tool wear state identification," *Comput. Ind.*, vol. 106, pp. 71–84, Apr. 2019, doi: 10.1016/j.compind.2018.12.018.

[8] D. M. D'Addona, A. M. M. S. Ullah, and D. Matarazzo, "Tool-wear prediction and pattern-recognition using artificial neural network and DNA-based computing," *J. Intell. Manuf.*, vol. 28, no. 6, pp. 1285–1301, Aug. 2017, doi: 10.1007/s10845-015-1155-0.

[9] R. Corne, C. Nath, M. El Mansori, and T. Kurfess, "Study of spindle power data with neural network for predicting real-time tool wear/breakage during inconel drilling," *J. Manuf. Syst.*, vol. 43, pp. 287–295, Apr. 2017, doi: 10.1016/j.jmsy.2017.01.004.
[10] C. Drouillet, J. Karandikar, C. Nath, A.-C. Journeaux, M. El Mansori, and T. Kurfess, "Tool life predictions in milling using spindle power with the neural network technique," *J. Manuf. Processes*, vol. 22, pp. 161–168, Apr. 2016, doi: 10.1016/j.jmapro.2016.03.010.
[11] P. B. Huang, C. C. Ma, and C. H. Kuo, "A PNN self-learning tool breakage detection system in end milling operations," *Appl. Soft Comput.*, vol. 37, pp. 114–124, Dec. 2015, doi: 10.1016/j.asoc.2015.08.019.
[12] W. Zhao, Z. Wang, W. Cai, Q. Zhang, J. Wang, W. Du, N. Yang, and X. He, "Multiscale inverted residual convolutional neural network for intelligent diagnosis of bearings under variable load condition," *Measurement*, vol. 188, Jan. 2022, Art. no. 110511.
[13] H. Liu, W. Song, Y. Zhang, and A. Kudreyko, "Generalized Cauchy degradation model with long-range dependence and maximum Lyapunov exponent for remaining useful life," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2021.3063749.
[14] H. Liu, W. Song, and E. Zio, "Generalized Cauchy difference iterative forecasting model for wind speed based on fractal time series," *Nonlinear Dyn.*, vol. 103, no. 1, pp. 759–773, Jan. 2021, doi: 10.1007/s11071-020-06150-z.
[15] S. Cho, S. Asfour, A. Onar, and N. Kaundinya, "Tool breakage detection using support vector machine learning in a milling process," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 3, pp. 241–249, Mar. 2005, doi: 10.1016/j.ijmachtools.2004.08.016.
[16] C. Letot, R. Serra, M. Dossevi, and P. Dehombreux, "Cutting tools reliability and residual life prediction from degradation indicators in turning process," *Int. J. Adv. Manuf. Technol.*, vol. 86, nos. 1–4, pp. 495–506, Sep. 2016, doi: 10.1007/s00170-015-8158-z.
[17] J. C. Jauregui, J. R. Resendiz, S. Thenozhi, T. Szalay, A. Jacso, and M. Takacs, "Frequency and time-frequency analysis of cutting force and vibration signals for tool condition monitoring," *IEEE Access*, vol. 6, pp. 6400–6410, 2018, doi: 10.1109/ACCESS.2018.2797003.
[18] J. Dong, K. V. R. Subrahmanyam, Y. S. Wong, G. S. Hong, and A. R. Mohanty, "Bayesian-inference-based neural networks for tool wear estimation," *Int. J. Adv. Manuf. Technol.*, vol. 30, nos. 9–10, pp. 797–807, Oct. 2006, doi: 10.1007/s00170-005-0124-8.
[19] D. Shi and N. N. Gindy, "Tool wear predictive model based on least squares support vector machines," *Mech. Syst. Signal Process.*, vol. 21, no. 4, pp. 1799–1814, May 2007, doi: 10.1016/j.ymssp.2006.07.016.
[20] J. Yu, S. Liang, D. Tang, and H. Liu, "A weighted hidden Markov model approach for continuous-state tool wear monitoring and tool life prediction," *Int. J. Adv. Manuf. Technol.*, vol. 91, nos. 1–4, pp. 201–211, Jul. 2017, doi: 10.1007/s00170-016-9711-0.
[21] F. A. Niaki, M. Michel, and L. Mears, "State of health monitoring in machining: Extended Kalman filter for tool wear assessment in turning of IN718 hard-tomachine alloy," *J. Manuf. Processes*, vol. 24, pp. 361–369, Oct. 2016, doi: 10.1016/j.jmapro.2016.06.015.
[22] X. He, Z. Wang, Y. Li, S. Khazhina, W. Du, J. Wang, and W. Wang, "Joint decision-making of parallel machine scheduling restricted in job-machine release time and preventive maintenance with remaining useful life constraints," *Rel. Eng. Syst. Saf.*, vol. 222, Jun. 2022, Art. no. 108429, doi: 10.1016/j.ress.2022.108429.
[23] F. Penedo, R. E. Haber, A. Gajate, and R. M. Del Toro, "Hybrid incremental modeling based on least squares and fuzzy and K-NN for monitoring tool wear in turning processes," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 811–818, Nov. 2012, doi: 10.1109/TII.2012.2205699.
[24] D. Leite, P. Costa, and F. Gomide, "Evolving granular neural networks from fuzzy data streams," in *Proc. Int. Joint Conf. Neural Netw. IEEE*, 2010.
[25] D. Leite, R. Ballini, P. Costa, and F. Gomide, "Evolving fuzzy granular modeling from nonstationary fuzzy data streams," *Evolving Syst.*, vol. 3, no. 2, pp. 65–79, Jun. 2012, doi: 10.1007/s12530-012-9050-9.
[26] P. Angelov, D. P. Filev, and N. Kasabov, *Evolving Intelligent Systems: Methodology and Applications*. Hoboken, NJ, USA: Wiley, 2010, doi: 10.1002/9780470569962.ch14.
[27] N. Kasabov, *Evolving Connectionist Systems: The Knowledge Engineering Approach*. New York, NJ, USA: Springer, 2007.
[28] M. Watts, "A decade of Kasabov's evolving connectionist systems: A review," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 39, no. 3, pp. 253–269, May 2009, doi: 10.1109/TSMCC.2008.2012254.
[29] A. Ghobakhlou, M. Watts, and N. Kasabov, "Adaptive speech recognition with evolving connectionist systems," *Inf. Sci.*, vol. 156, pp. 71–83, Nov. 2003, doi: 10.1016/S0020-0255(03)00165-8.
[30] (2010). *2010 PHM Data Challenge*. [Online]. Available: https://www.phmsociety.org/competition/phm/10
[31] J. Wang, J. Xie, R. Zhao, L. Zhang, and L. Duan, "Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing," *Robot. Comput.-Integr. Manuf.*, vol. 45, pp. 47–58, Jun. 2017, doi: 10.1016/j.rcim.2016.05.010.
[32] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism Mach. Theory*, vol. 70, pp. 441–453, Dec. 2013, doi: 10.1016/j.mechmachtheory.2013.08.014.

MUQUAN LIN is currently a Lecturer at the Minnan University of Science and Technology. He is also a Graduate Student in instrument science and technology at East China Jiaotong University. He is also with Fujian Normal University, majoring in electronic information engineering. He is mainly engaged in the research of Internet of Things applications, reliability analysis, and life prediction. His research interests include power electronics system design, power converter digital control, and optimal design.

SONG WANQING received the B.Sc. degree from the Inner Mongolia University of Science and Technology, Inner Mongolia, China, in 1983, the M.Sc. degree from the University of Science and Technology Beijing, Beijing, China, in 1990, and the Ph.D. degree from Donghua University, Shanghai, China, in 2010. He is currently a Professor with the Shanghai University of Engineering Science, Shanghai. His main research interests include condition monitor and fault diagnosis, big data analysis, health, and reliability.

DONGDONG CHEN received the B.S. degree in electronic information engineering from Zhejiang University, Hangzhou, China, in 2013, and the Ph.D. degree in electrical engineering from the College of Electrical Engineering, Zhejiang University, in 2018.

He is currently a Vice Professor at the Minnan University of Science and Technology. His research interests include power electronics system design, power converter digital control, and optimal design.

ENRICO ZIO (Senior Member, IEEE) received the M.Sc. degree in nuclear engineering from the Politecnico di Milano, in 1991, the M.Sc. degree in mechanical engineering from UCLA, in 1995, the Ph.D. degree in nuclear engineering from the Politecnico di Milano, in 1996, and the Ph.D. degree in probabilistic risk assessment from MIT, in 1998. He has authored and coauthored seven books and more than 300 articles on international journals. His research interests include modeling of the failure-repair-maintenance behavior of components and complex systems, for the analysis of their reliability, maintainability, prognostics, safety, vulnerability, resilience and security characteristics, and on the development and use of Monte Carlo simulation methods, artificial techniques, and optimization heuristics. He is the chairman and co-chairman of several international conferences, an associate editor of several international journals, and a referee of more than 20 journals.