The International Journal of Advanced Manufacturing Technology (2018) 96:2509–2523
https://doi.org/10.1007/s00170-018-1768-5

ORIGINAL ARTICLE

Review of tool condition monitoring methods in milling processes

Yuqing Zhou [1,2] · Wei Xue [2]

Received: 3 October 2017 / Accepted: 11 February 2018 / Published online: 23 February 2018

Abstract
Accurate tool condition monitoring (TCM) is essential for the development of fully automated milling processes. However, while considerable research has been conducted in industrial and academic settings, the complexity of milling processes continues to complicate the implementation of TCM. This paper presents a review of the state-of-the-art methods employed for conducting TCM in milling processes. The review includes three key components: (1) sensors, (2) feature extraction, and (3) monitoring models for the categorization of cutting tool states in the decision-making process. In addition, the primary strengths and weaknesses of current practices are presented for these three components. Finally, this paper concludes with a list of recommendations for future research.

Keywords Tool condition monitoring · Milling process · Sensors · Feature extraction · Monitoring model

1 Introduction

Milling is a very common and efficient cutting operation that uses a rotary milling cutter with one or more teeth to intermittently cut workpieces into flat surfaces, grooves, threads, and many geometrically complex components. Highly efficient milling processes are suitable for mass production, and have been employed widely in industrial manufacturing. Cutting tools are considered to be the primary component of the milling process [1], and tool breakage is a major cause of unscheduled stoppage in industrial settings employing milling. Tool breakage is generally the result of an accumulation of tool damage over time, and has negative indirect (time loss) and direct (capital) effects. In terms of these effects, tool failure accounts for 7–20% of total milling-machine downtime [2, 3], and the costs of tools and tool changes account for 3–12% of the total processing cost [4]. In an effort to limit the indirect effects of tool breakage, milling tools employed in industrial milling processes are replaced prior to breakage. However, conventional tool replacement strategies employ uniform time periods that are determined by the subjective experience of operators. Such strategies inevitably result in either the early replacement of workable tools, which increases tools costs and downtime, or the late replacement of worn tools, which results in lower workpiece quality and increased production costs [5]. In fact, research [6, 7] has determined that only 50–80% of the effective life of milling tools is typically used. As such, monitoring the varying conditions of milling tools over time to facilitate a timely detection of tool damage is critical for limiting the indirect effects of tool breakage while maximizing the usable life of milling tools. Thus, tool condition monitoring (TCM) systems have been developed to generate better workpiece surface quality and extend tool life by diagnosing cutting tool deficiencies using appropriate signal processing and pattern recognition techniques. An accurate and reliable TCM system can reduce costs by 10–40% by reducing downtime and maximizing the usable life of milling tools [7, 8].

The application of TCM in milling processes has been studied for over 30 years, and has been based on two types of methods: direct monitoring and indirect monitoring. Direct monitoring methods employ optical equipment and machine vision technology to directly monitor tool condition. For example, optical microscopes are used to capture tool images and the tool condition is evaluated with image analysis technology [9]. Direct methods are advantageous because they do not affect the machining process and offer high recognition accuracy under ideal conditions. However, direct methods are generally unsuitable for manufacturing settings because (1)

✉ Wei Xue
xw_wzu@163.com

1 College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou, China

2 College of Mechanical Engineering, Wenzhou University, Wenzhou, China

the high expense of the required equipment and software can unacceptably increase manufacturing costs, and (2) the recognition accuracy is easily disturbed by the presence of cutting fluid and cutting chips on tool surfaces [10, 11]. Therefore, indirect monitoring methods, which estimate tool condition based on an analysis of signals derived from one or more sensors, have been widely adopted in manufacturing settings. These signals can be representative of numerous characteristics, such as cutting force, vibration, motor current, and acoustic emission (AE). The signal analysis conducted seeks to extract significant features from a signal that are indicative of tool condition, and includes numerous methods, such as those based on the time domain, frequency domain, wavelet transform (WT), empirical mode decomposition (EMD), and multi-domain analysis. Finally, the cutting tool condition could be evaluated with extracted feature parameters based on one certain pattern recognition method, such as artificial neural network (ANN), hidden Markov model (HMM), and support vector machine (SVM). Compared with direct methods, indirect methods are less expensive and more adaptable to practical applications.

Indirect TCM is data-driven, and can be divided into the two phases illustrated in Fig. 1: model training and online monitoring. The model training phase consists of three modules: sensor configuration, feature extraction, and monitoring model. The sensor configuration module provides a sensor signal, the feature extraction module extracts features in the sensor signal that are related to tool condition (e.g., wear, damage, and breakage), and the monitoring model module builds a decision support model for online monitoring. The online monitoring phase consists of three modules: online sensor monitoring and decision-making. The sensor configuration employed during online monitoring is based on that employed during the model training phase, and is generally equivalent. However, if the sensor configuration is altered (it is possible when using multiple sensors initially), the online monitoring phase will not consider irrelevant sensors when evaluating tool condition. In the decision-making module, the sensor signals are firstly subjected to an equivalent feature extraction as that employed during the model training phase, and then the cutting tool condition can be evaluated in real time based on the trained monitoring model obtained during the model training phase.

The key to developing a successful indirect TCM approach for the milling process is the model training phase. Therefore, this paper discusses (1) sensors, (2) feature extraction, and (3) monitoring models in detail. The remainder of this paper is organized as follows. Section 2 presents an overview of the sensors employed in TCM. Section 3 reviews the feature extraction methods employed in TCM. Section 4 describes the advantages and disadvantages of several monitoring models commonly employed in TCM. Section 5 discusses some of the deficiencies in current TCM research applied to milling processes, and suggests a number of future research directions.

2 Sensors

A number of sensors have been employed in TCM to obtain signals for tool condition monitoring. The sensor configuration can be divided according to the types of sensors employed into two categories: single sensor and multi-sensor configurations. In terms of single-sensor configurations, the present study focuses on cutting force, vibration, motor current, and AE sensors. While other types of single sensors have been employed for TCM in milling processes, such as sound [12] and temperature [13] sensors, these sensors are rarely adopted alone because they are significantly affected by environmental conditions. Therefore, they are generally employed in conjunction with other sensors in multi-sensor configurations.

2.1 Single sensor

(1) Cutting force

During milling processes, tools wear increases surface roughness, and leads to a corresponding increase in cutting force. Many studies [1, 14–16] have demonstrated that the cutting force is very sensitive to changes in tool condition, and can therefore accurately estimate the tool state. For example, Wang et al. [17] determined that the cutting force signal is the most stable and reliable signal among commonly employed sensor signals that are closely related to tool wear. Huang et al. [18] employed a piezoelectric dynamometer to monitor the tool state of an end milling operation according to cutting force. Bulent et al. [19] adopted a rotary dynamometer to capture the cutting forces in three dimensions and the torque of the drive moment on a rotating tool. However, cutting force sensors are difficult to apply in industrial environments because their physical properties limit the physical size of a workpiece, which is not appropriate when milling medium and large workpieces [20]. In addition, Koike et al. [21] established that cutting force monitoring interferes with the motion control of the spindle and stage in a milling machine, and reduces its rigidity. Moreover, the expense of commercial dynamometers can unacceptably increase manufacturing costs [22, 23].

(2) Vibration

Vibration sensors are widely employed in TCM because they are inexpensive, easy to install, and provide a similar periodic signal shape to that of the cutting force [24–26]. Besmir et al. [27] established that low levels of vibration are generated with sharp cutting tools, while the levels of vibration increase with increasing deterioration in the tool condition. Numerous studies have demonstrated the feasibility of adopting vibration signals for TCM in milling processes [28–31]. For example, Hsieh et al. [32] demonstrated that the spindle vibration acceleration signal can distinguish

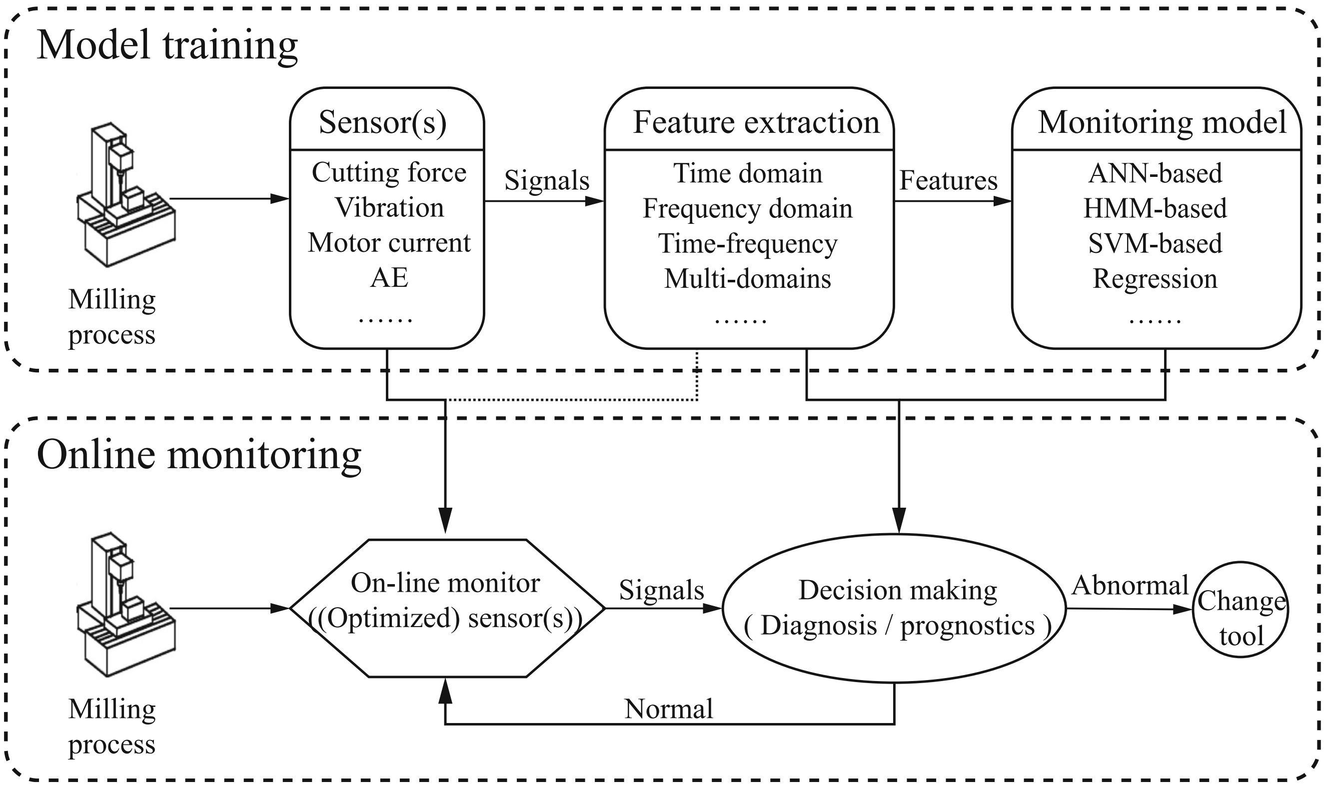

Fig. 1 Basic process flow of tool condition monitoring (TCM) in milling processes

different tool conditions during micro-milling when used in conjunction with appropriate feature extraction and classifiers. Madhusudana et al. [33] installed a tri-axial integrated electronic piezoelectric (IEPE) accelerometer on the spindle housing to capture the spindle vibration acceleration signal during face milling. Gao et al. [34] achieved positive tool condition diagnostic accuracy by adopting a laser vibrometer to acquire the vibration displacement of a tool holder. However, the characteristics of milling processes limit the accuracy of TCM employing vibration signals. First, vibrations are generated during machine operation even when the tool is not engaged in cutting, as during an air-cut operation. In fact, effectively distinguishing between entity-cut and air-cut operations remains an open challenge. Second, vibration signals are difficult to filter, and are therefore prone to providing erroneous data [27]. Finally, the position of sensor installation and the use of cutting fluid can affect the vibration signal [11].

(3) Motor current

Because the cutting force increases with increasing tool wear, the current drawn by the machine motor undergoes a corresponding increase [35]. Motor current sensors are considered to be more suitable for manufacturing settings than cutting force sensors due to their relatively simple application and lack of installation effects on machining operations [36–38]. Ghosh et al. [10] demonstrated that current sensors provide TCM results that are fairly comparable to that of cutting force sensors in actual industrial TCM applications. Stavropoulos et al. [23] demonstrated that motor current signals correlate more strongly with tool wear than do vibration signals, and the motor current signal suffers less sensitivity to environmental noise, resulting in more accurate tool condition diagnoses. Ammouri et al. [35] established a TCM index based on the measured current values of the spindle and drive motors of a machine tool. However, use of the motor current signal for TCM also has a few disadvantages. First, motor current signals contain a considerable amount of noise, obstructing the detection of small fluctuations in the cutting force, and high-frequency components are lost by filtering [21]. Second, the motor current signal is significantly influenced by the viscous damping of the feed system and friction in the mechanical system [39]. Finally, experiments conducted by Lee et al. [40] demonstrated that the motor current is not sensitive to changes in the cutting force at high motor frequencies, which means that the motor current signal is not suitable for TCM conducted at high spindle speeds.

(4) Acoustic emission

Sensors based on AE are particularly suitable for conducting TCM in milling processes because the resulting signals are not mechanically disturbed, have a superior sensitivity to the those of cutting force and vibration signals, and propagate at a frequency much greater than the characteristic frequency caused by cutting, which reduces interference [41–43]. A study conducted by Vetrichelvan et al. [2] demonstrated that the AE signals obtained from sensors located on the top surface of the tool holder can effectively monitor crater wear. Mathew et al. [44] conducted experiments with 1-tooth, 2-tooth, and 3-tooth milling cutters, and demonstrated that AE signals exhibit marked responses to changes in tool condition such as tool breakage and tool chipping. Ren et al. [45] established that AE signals are easily recorded, and provide very rapid responses to changing conditions in the contact between the tool and workpiece; thus, AE sensors are well suited for TCM in micro-milling processes. However, intermittent cutting during milling processes results in AE signal spikes when individual teeth enter or exit the workpiece, which complicates the analysis of AE signals [20]. In addition, AE sensors are highly sensitive to environmental noise [46], which increases the difficulty of extracting valid signal feature information.

2.2 Multiple sensors

Due to the limitations associated with the single-sensor TCM applications discussed above, multi-sensor TCM has been increasingly pursued [20]. Here, different sensor signals are more strongly correlated to the tool condition at different levels of tool wear, and the loss of sensitivity in one particular sensor signal may be offset by the increased sensitivity of another sensor signal [10]. Thus, multiple sensors can enhance the richness of information indicative of potential tool wear levels [47]. Hence, although multi-sensor TCM applications provide redundant information, they can reduce the overall uncertainty of tool wear measurements and improve the resolution and accuracy of TCM systems, achieving better performance and greater robustness [4, 39, 48].

The sensor configurations employed in multi-sensor-based TCM research for milling processes published from 2000 to 2016 are listed in Table 1. We note from the table that, beginning with only one studies published in 2000 and 2007, and ending with 8 studies published in 2016, the number of studies using multi-sensor-based TCM methods in milling processes has gradually increased over the time period considered. For example, Ghosh et al. [10] investigated a sensor configuration composed of all five sensor types, and determined that cutting force and motor current signals contain more useful feature information than four sensors (cutting forces, spindle vibration, spindle motor current, and sound) considered. In this regard, we note that the number of sensors employed in TCM cannot be excessive. First, the costs of production and maintenance and the difficulty of maintenance both increase with an increasing number of sensors. Second, the interference in the milling process caused by sensors increases with an increasing number of sensors. Finally, the redundant information provided by multiple sensors can deteriorate the accuracy of the TCM model. Here, the use of a few key sensor signal features is more effective than the use of all possible sensor features [31, 49]. Therefore, the types and numbers of sensors employed in the TCM of milling processes cannot be ignored.

3 Feature extraction

After the analog sensor signals are captured, they are then preprocessed, typically by analog-to-digital conversion and low-pass or band-pass digital filtering, in preparation for the extraction of features. The primary purpose of the feature extraction module is to extract the feature parameters of the signals that are closely related to the tool state, and to significantly reduce the dimensions of the original information in the time and/or frequency domains. The resulting signal feature parameters output by the feature extraction module will be the input, or part of the input, of the monitoring model module. However, the feature parameters input into the monitoring model module must be carefully defined. An excessive number of feature parameters will greatly increase the computational burden of model construction, and affect the timeliness of online monitoring. Also, irrelevant and redundant feature parameters have a negative impact on the performance of the monitoring model and subsequent TCM. In this regard, a relatively small number of highly appropriate feature parameters can generate a more accurate and robust model, and ensure accurate TCM in real time [31, 62].The present study focuses on feature extraction methods based on the time domain, frequency domain, WT, EMD, and multi-domain analysis.

3.1 Time domain and frequency domain methods

Methods based on the time domain extract feature information related to the tool state from the time dimension of the signal using time series analyses in conjunction with several statistical parameters. Time series analysis includes methods such as the autoregressive (AR) process, the AR moving average (ARMA) process, and time domain averaging (TDA) [3]. Statistical parameters include the root-mean-square, maximum/minimum, average, standard deviation, and kurtosis of time series data [18]. Methods based on the frequency domain extract feature information related to the tool state from the frequency dimension of the signal based on the frequency structure and harmonic components of the signal. These methods first convert sensor signals from the time domain into the frequency domain using the fast Fourier transform (FFT), and then extract feature parameters such as the power spectrum, peak-to-peak amplitude, and tooth frequency [27]. Both methods based on the time and frequency domains can only

Table 1 Sensor configurations employed in multi-sensor-based TCM research for milling processes (2000–2016)

	Sensors				
Author	Cutting force	Vibration	Motor current	AE	Sound
Chen et al. (2000) [48]	•	•			
Ghosh et al. (2007) [10]	•	•	•	•	•
Jemielniak et al. (2008) [41]	•			•	
Binsaeid et al. (2009) [49]	•	•			
Kuljanic et al. (2009) [50]	•	•			
Cho et al. (2010) [51]	•	•	•	•	
Geramifard et al. (2012) [52]	•	•		•	
Lamraoui et al. (2014) [53]	•	•			
Liu et al. (2015) [4]	•	•			
Salehi et al. (2015) [54]		•	•		
Stavropoulos et al. (2016) [23]		•	•		
Wang et al. (2016) [55]	•	•			
Hong et al. (2016) [56]	•	•			
Jahromi et al. (2016) [57]	•	•		•	
Wang et al. (2016) [58]		•	•	•	
Torabi et al. (2016) [58]	•	•		•	
Yu et al. (2016) [59]	•	•		•	
Downey et al. (2016) [60]	•	•		•	
Besmir et al. (2016) [27]	•	•	•		•
Harris et al. (2016) [61]	•	•		•	•

provide feature information from a single perspective, and both assume that the signal is stationary, which is not suitable for non-stationary signals obtained in milling processes [63].

3.2 Wavelet transform

To overcome the shortcomings of time and frequency domain-based methods, time-frequency analysis methods based on the WT have been adopted for feature extraction in TCM for milling processes. Various WT-based methods, such as the continuous WT (CWT), discrete WT (DWT), and wavelet packet transform (WPT), have been applied to extract series of wavelet coefficients as features reflecting the cutting tool state.

A wavelet $\{\psi_{b,a}(t)\}$ is defined according to time t as a function family derived through the scaling and translation of a wavelet basis function $\psi_{b,a}(t)$ as

$$\psi_{b,a}(t) = a^{-1/2}\psi\left(\frac{t-b}{a}\right), \quad (1)$$

where $a(>0)$ is a scaling parameter that represents reduction if $a<1$ and expansion if $a>1$, and b is a position parameter. The term $a^{-1/2}$ guarantees that equal wavelet energy is maintained during the scaling process of $\psi(t)$ for different values of a. Correspondingly, the CWT of a signal $x(t)$ is defined as

$$CWT_x(b,a) = a^{-1/2}\int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt = \langle x(t), \psi_{b,a}(t)\rangle. \quad (2)$$

It can be shown from Eq. (2) that the ideal CWT is an inner product operation that substitutes the wavelet basis function $\psi_{b,a}(t)$ for the basis function e^{iwt} in the Fourier transform, and decomposes $x(t)$ into a number of sub-signals of different frequency bands, which can be employed for seeking the most relevant component or most similar components of $\psi_{b,a}(t)$.

In TCM applications, the coefficients of the WT represent variations in the signal energy with respect to time and frequency. Sevilla et al. [24] applied CWT to extract the wavelet features of a vibration signal, and established that the feature patterns of normal and damaged tools are different. The CWT is recognized as an effective tool for both stationary and non-stationary signals. However, the CWT involves considerable redundant information, and is computationally very slow. Therefore the DWT was introduced for TCM applications. For example, Madhusudana et al. [12] compared several feature extraction methods and determined that DWT is more accurate than either an analysis of statistical features or EMD methods. To define the DWT, let $a = 2$ and $b = k2^j$, such that the DWT of $x(t)$ is defined as

$$DWT_x(j,k) = \left(\sqrt{2^j}\right)^{-1/2}\int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-k2^j}{2^j}\right)dt, \quad (3)$$

where j is the scale and k is the translation variation. The scaling and wavelet functions at different scales are generated from the following two-scale equations.

$$\begin{cases} \phi(t) = \sqrt{2} \sum\limits_{n=-\infty}^{+\infty} h_n \phi(2t-n) \\ \psi(t) = \sqrt{2} \sum\limits_{n=-\infty}^{+\infty} g_n \phi(2t-n) \end{cases} \tag{4}$$

Here, n is an integer, $h_n = <2^{-1/2}\varphi(t), \varphi(2t-n)>$ is a scale coefficient, and $g_n = <2^{-1/2}\psi(t), \varphi(2t-n)>$ is a wavelet coefficient. The DWT decomposes $x(t)$ through every pair of extraction operations, and the two-scale relationship in Eq. (4) can be given as follows.

$$\begin{cases} \phi_{j-1,n} = \sum\limits_{k \in Z} \langle \phi_{j-1,n}, \phi_{j,k} \rangle \phi_{j,k} = \sum\limits_{k \in Z} h_{k-2n} \phi_{j,k} \\ \psi_{j-1,n} = \sum\limits_{k \in Z} \langle \psi_{j-1,n}, \phi_{j,k} \rangle \phi_{j,k} = \sum\limits_{k \in Z} g_{k-2n} \phi_{j,k} \end{cases} \tag{5}$$

Here, the summation occurs over the set of all integer numbers Z, and $h_{k-2n}$and g_{k-2n} are the filter coefficients of low-pass and high-pass filters, respectively. According to Eq. (5), the scale function $\varphi_{j-1,n}$ and wavelet function $\psi_{j-1,n}$ with a resolution of $j-1$ can be denoted by the scale function $\varphi_{j,k}$ with a resolution of j. The DWT decomposition expression can be obtained by the inner product operation of Eq. (4) with $x(t)$ as follows.

$$\begin{cases} a_{j-1,n} = \sum\limits_{k \in Z} h_{k-2n} a_{j,k} \\ d_{j-1,n} = \sum\limits_{k \in Z} g_{k-2n} a_{j,k} \end{cases} \tag{6}$$

Here, $a_{j,k}$ is a scale coefficient and $d_{j,k}$ is a wavelet coefficient that are derived from the projection of $x(t)$ onto the space of scale function $\varphi_{j,k}$ and wavelet function $\psi_{j,k}$, respectively.

It can be shown from Eq. (6) that the DWT decomposition of $x(t)$ orients toward the low-frequency approximate signal $a_{j,k}$ rather than to the high-frequency signal $d_{j,k}$, which leads to high temporal resolution and poor frequency resolution in high-frequency band signals, while leading to low temporal resolution and good frequency resolution in low-frequency band signals. Therefore, the DWT leads to a loss of useful information at high frequency. To overcome this drawback, the WPT was introduced to enhance the frequency resolution of high-frequency band signals. The WPT conducts multi-level band division over the entire signal band, which not only inherits the advantages of good time-frequency localization from the WT, but also further decomposes the high-frequency band for increasing the frequency resolution. Therefore, the WPT is more suitable for TCM applications. Wang et al. [17] employed the WPT to realize noise reduction in the extraction of signal energy features. Hong et al. [56] decomposed signals with the WPT, and applied the Fisher discriminant ratio to select wavelet coefficients that were significantly related to cutting tool wear.

The key aspect of the WPT still lies in selecting a wavelet basis function that matches the fault characteristic waveform. However, it is difficult to select the appropriate wavelet basis function for the recognition of an unknown milling cutter tool fault. This has serious drawbacks because the use of an inappropriate wavelet basis function will dilute the fault feature information, which will complicate the processes of feature extraction and fault diagnosis. To overcome this drawback, the multi-wavelet transform can be introduced into TCM applications. Multi-wavelets are wavelets generated by two or more scaling functions. The idea is to expand the two-scale equation (4) into a multi-dimensional equation by vectorization. This in turn expands the multi-resolution analysis space generated by a single scale function with a single wavelet into a multi-resolution space generated by multiple scale functions, which greatly increases the number of degrees of freedom of the wavelet functions [64]. However, the application of multi-wavelet-based methods to TCM has not yet been reported.

3.3 Empirical mode decomposition

Huang et al. [65] developed EMD as a self-adaptive time-frequency signal analysis technique in 1998. Here, EMD decomposes complex signals into several intrinsic mode functions (IMFs), where each IMF component contains the local characteristic signal of the original signal at different time scales. The decomposition is conducted according to the time-scale features of the signal itself without the pre-selection of any basis functions. As such, the method is essentially different from Fourier transform and WT decomposition methods that are based on a priori harmonic basis functions and wavelet basis functions, respectively. Accordingly, EMD can theoretically be applied to decompose any signal, which is obviously advantageous when working with non-stationary and non-linear signals.

The EMD process seeks to find all the maximum points of the original signal sequence $x(t)$, and fits the upper envelope of the original data using an interpolation algorithm (e.g., cubic spline interpolation function). Similarly, all the minimum points of $x(t)$ forming the lower envelope are fit by the interpolation algorithm above. Then, the mean envelope, representing an average of the upper envelope and the lower envelope, is subtracted from $x(t)$ to obtain a new data sequence $x'(t)$. The presence of negative local maxima and/or positive local minima in $x'(t)$ indicates that it is not yet an IMF, and further sifting is conducted [66].

In recent years, EMD has attracted considerable attention, and has been widely employed in TCM research. For example, Shi et al. [67] proposed an EMD-based tool breakage diagnosis algorithm, which decomposed the sound signal obtained during machining to separate cutting-related sound information from the noise using EMD. Babouri et al. [68] decomposed the vibration signals of cutting tools by combining CWT and EMD, where the CWT was first applied to the signal as a pretreatment, and EMD was subsequently applied to the pretreated signal. The effectiveness of their proposed method for TCM was experimentally verified.

However, classical EMD has some problems related to pattern aliasing and computational inefficiency in the extraction of composite and weak characteristic signals. To overcome these problems, Huang et al. [69] proposed a noise-driven faint-feature signal extraction method, denoted as ensemble EMD (EEMD). In 2014, Wang et al. [70] theoretically demonstrated that EEMD has an equivalent computational efficiency as the FFT, which creates an entirely new route for the online signal processing of non-stationary and non-linear signals. While EEMD has been applied for gearbox wear analysis [71], and achieved good results, it has not yet been reported in TCM research for milling processes.

3.4 Multi-domain-based methods

Multi-domain-based methods select feature parameters from more than a single domain, including the time domain, frequency domain, and/or time-frequency domain, to compose a candidate feature parameters set, and also apply feature selection or dimensional reduction methods to obtain feature parameters that are strongly related to the tool state. Multi-domain-based methods have received considerable attention in TCM research for milling processes. For example, Cho et al. [51] constructed 135 candidate feature parameters by extracting several feature parameters in the time and frequency domains of multiple sensory signals. They then applied the entropy correlation algorithm to select 25 feature parameters most indicative of tool state. Geramifard et al. [52] selected 38 feature parameters using the Fisher discriminant ratio algorithm from 434 candidate feature parameters obtained from the time domain and the WT based on the Daubechies8 (db8) wavelet of multiple sensory signals. Wang et al. [28] decomposed vibration signals by WPT, and calculated 5 time domain and 4 frequency domain feature parameters in each sub-band signal. Then, they used the local preserving projection (LPP) algorithm to reduce the dimension of the feature parameters. Liu et al. [4] selected 138 feature parameters from the time domain, frequency domain, and wavelet energy as a candidate feature parameters set, and then applied the fast correlation filter (FCBF) algorithm to establish the smallest non-redundant feature set, which included 19 feature parameters. Wang et al. [72] established a basic feature parameters set with 54 parameters from the time domain, frequency domain, and wavelet coefficients, which was then reduced to 11 feature parameters using the kernel principle component analysis (KPCA) algorithm.

During the model training phase for TCM, limited knowledge and experience is available for guiding the selection of feature parameters. As such, the model training phase may select feature parameters for the candidate feature set that are not closely related to tool states, and may fail to select feature parameters that are closely related, which results in decreased TCM performance. The advantages of multi-domain methods here are that they provide more candidate feature parameters related to tool state than single-domain methods, and also reduce the risk of losing significant information, which improves the performance of TCM. Moreover, although multi-domain-based methods clearly increase the number of feature parameters in the candidate feature set, the feature set is reduced to a low dimension via feature selection or by the application of a dimension reduction algorithm. This is particularly beneficial during the online monitoring phase, because feature parameters are then determined after dimension reduction, which has little impact on the operation speed of the model.

4 Monitoring models

With the rapid development of artificial intelligence (AI) technology, a large number of AI methods have come to be employed for constructing monitoring models in TCM applications. Figure 2 shows the distribution of AI methods employed in 56 published articles related to TCM in milling processes over the past 10 years. These methods include artificial neural network (ANN), hidden Markov model (HMM), support vector machine (SVM), fuzzy logic, regression, principle component analysis (PCA), control chat, and other methods applied in only a single study, which include C-means clustering, K-star, relevance vector machine (RVM), extreme learning machine (ELM), and others. The present work focuses on the three most commonly employed AI methods, which includes ANN, HMM, and SVM, as supervised learning algorithms, and fuzzy C-means (FCM) clustering (as unsupervised learning algorithm) to establish categories.

(1) Artificial neural network

An ANN represents a series of hierarchical networks that are composed of artificial neurons (basic units or nodes) connected by what are analogous to synapses. A typical ANN includes an input layer, one or more hidden layers, and an output layer, where each layer contains several neurons, and the neurons between adjacent layers are fully interconnected. The strength of the connection between two neurons is denoted as the weight, and the values of the weights of all neuron connections are adjusted by minimizing the output error obtained with training samples. As such, the decision-making

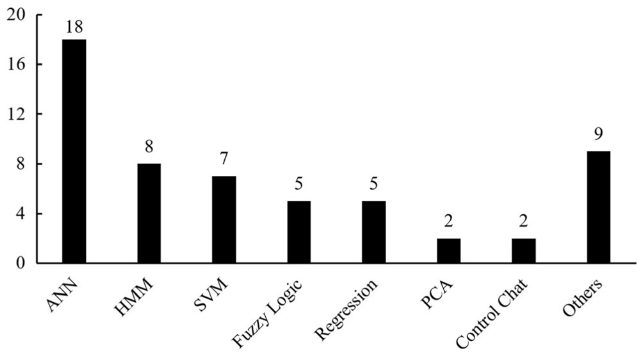

Fig. 2 Frequency of artificial intelligence (AI) methods employed in 56 articles published in the past 10 years related to TCM in milling processes (ANN: artificial neural network; HMM: hidden Markov model; SVM: support vector machine; PCA: principle component analysis). Note: The category "Others" represents the sum of AI methods applied in only a single study, which include C-means clustering, K-star, relevance vector machine (RVM), extreme learning machine (ELM), and others

capability of an ANN is not based on an understanding of the system. This black-box approach is very useful for the TCM applications of milling processes because the relationships between sensor signals and tool wear in milling processes are non-linear and difficult to express using analytical formulas [73–75]. Therefore, ANNs provide monitoring models with strong fault tolerance, adaptability, and noise suppression. Numerous studies have applied ANNs to the TCM of milling process, and have achieved outstanding results [4, 10, 18, 19, 32, 43, 76] [77]. For conducting TCM in milling processes, the input layer of an ANN includes the features extracted from the feature extraction module and several cutting parameters (such as cutting speed, cutting depth, and feed rate), and the output layer represents the tool wear state or tool wear value. However, ANN-based monitoring models suffer from several drawbacks [17, 78, 79]. (a) ANNs require a large number of training samples to achieve high performance, which is time-consuming and costly for milling processes. (b) The selection of the number of hidden layers and neurons in the network structure is a critical factor contributing to ANN performance, but this selection depends on experience, and is not directly related to the tool wear process. Although the large number of existing network structures available can provide some degree of guidance in this regard, an optimal network structure for the TCM of milling processes remains largely unsolved. (c) Because the operation mechanism of an ANN is like a black-box, it yields no information regarding system dynamics, and is therefore not beneficial for conducting further exploration of the correlation between sensor signals and tool wear states.

(2) Hidden Markov model

An HMM is a Markov process with implied unknown parameters. An HMM includes two stochastic processes: one is a Markov process, describing the transition sequence of hidden states, and the other is a random process building the observation sequence of hidden states. Although an HMM was employed early on in 1997 for TCM research [80], it has not been reevaluated until recent years [17, 52, 59, 81–83]. A typical HMM assumes that the state of a system is affected only by the previous state, and is independent of all other states, which is consistent with the additive rule of gradual tool wear as a progressive accumulation process. One of the advantages of HMMs is that they can represent the behavior of milling processes as a dynamic model rather than a static model; thus, HMMs allow for the consideration of feature changes occurring over time, or for the consideration of other features. Therefore, in contrast to ANNs, HMMs are beneficial for exploring tool wear evolution mechanisms. However, HMM-based methods also suffer from some disadvantages. (a) Similar to ANN-based methods, HMM-based methods require considerable training data to achieve high performance. (b) The state duration distribution is a critical factor contributing to HMM performance, but the state duration distributions of tool states in milling processes remain unclear. (c) Although the application of HMM-based methods has been demonstrated to be very useful for monitoring tool flank wear, which is the dominant type of tool wear in milling processes that occurs gradually over the life of a cutting tool, these methods are problematic for other types of tool wear, such as breakage, crater wear, and chipping, and require additional study. It is noteworthy that HMM is looking forward to outperforms ANN if the studies of HMM are more thorough in TCM research for milling processes [84].

(3) Support vector machine

The SVM approach is based on statistical learning theory. Accordingly, SVM non-linearly maps the input samples in the original space to a high-dimensional feature space using a kernel function, and then constructs a linear algorithm in the high-dimensional feature space that corresponds to the solution of the non-linear problem in the original space [85]. Because of its superior performance with small sample sizes, SVM is suitable for the TCM of milling processes, which involve small sample sizes and high-dimensional non-linear signals. As such, SVM has gradually generated considerable interest in TCM research [28, 49, 86]. Several studies have demonstrated that SVM-based methods are superior to ANN-based methods in the classification of tool wear states [12, 51]. However, SVM faces several problems [87–89]. (a) The selection of the kernel function and its parameters strongly impacts the performance of SVM, but the kernel function must be selected in the absence of theoretical criteria. (b) The performance of SVM is very sensitive to the selected value of the penalty parameter, but the penalty parameter value can only be determined by trial-and-error, which greatly complicates its selection. (c) The SVM approach classifies tool state according to the signal features of the current moment, and is therefore unable to make full use of tool state information obtained before and after the current state, which severely limits its capability for exploring the correlation between sensor signals and tool wear states.

(4) Fuzzy C-means clustering

As discussed, the ANN, HMM, and SVM methods are supervised learning algorithms, which require the use of the existing cutter tools category information of training samples. However, when training samples include only fault data, and the corresponding cutter tool category information is missing, clustering methods become the most effective means for model training. In this regard, the FCM clustering algorithm is the most common cluster analysis method applied in fault diagnosis [90–92]. However, few studies have applied FCM clustering in the field of TCM for milling processes [93].

The FCM clustering algorithm is based on the optimization of an objective function. Here, FCM clustering employs the extreme value of the squared error function to obtain a set of center vectors that minimize the sum of squares of sample-center distances, and iteratively determines the optimum fuzzy classification matrix and clustering centers. The essence of FCM clustering is to fuse the redundant or complementary information of multiple feature parameters according to a given criterion, and thus to obtain potential feature information in data. The FCM clustering algorithm provides many advantages. (1) An initially established fault monitoring system does not include a large number of training samples, which results in poor training results for ANN, HMM, or other supervised learning methods. Here, FCM clustering can directly conduct fault analysis with a limited number of samples. As such, initially established fault monitoring system can be reasonably functional, and a considerable number of additional useful samples can be accumulated for subsequent system improvement. (2) Clustering analysis based on FCM clustering is easy to learn and to simulate the experience of experts, and can improve the intelligence of fault monitoring system. (3) The FCM clustering algorithmic highly adaptable, and can complete tasks, even if the knowledge or rules governing the actual conditions change, by making simple modifications in the fault monitoring system.

Due to the excellence of FCM clustering for accommodating uncertainties, and its independence from prior knowledge, FCM clustering is suitable for conducting TCM in milling processes owing to the limited number of samples and the inherently fuzzy tool state categories involved. As such, efforts to combine the FCM clustering algorithm with advanced feature extraction technology for conducting the TCM of milling processes are eagerly anticipated. Of course, FCM clustering requires a preset number of clusters, and the selected number of clusters is a crucial factor affecting clustering performance. Therefore, the means of automatically determining the number of clusters based on sample data must be further studied.

5 Discussion

Monitoring changing tool conditions in the milling process is exceedingly complicated owing to the influence of many factors, such as the cutting conditions, workpiece material, and environmental parameters involved in specific milling processes (e.g., face milling or end milling). This challenge has so far interfered with the development of a general reference model for TCM appropriate for all milling processes. Although much progress has been made in TCM research for milling processes, the following important questions remain open, and require further study.

(1) Inexpensive development of TCM models for specific milling conditions

In the absence of a general reference monitoring model for TCM in milling processes, individual TCM models must be developed for specific milling conditions. However, it is costly and time-consuming to collect sufficient sample data appropriate to very specific milling conditions with which to train TCM models for those conditions [76, 94, 95]. While this can potentially be addressed by developing learning strategies for TCM models that provide high monitoring performance with limited training sample sizes, this approach is very challenging. As a practical means of solving this problem, the numerical simulation + design of experiment (DoE) process is worth investigating. Here, numerical simulation is employed for analyzing changing tool conditions in a specific milling process, and the experimental data in DoE are used to correct the simulation model. Once a highly precise simulation model is constructed, it can be used to develop a new model under altered cutting parameters (e.g., cutting speed, cutting depth, and feed rate), and, thereby, the experimental cost and time required to develop simulation models can be reduced. Although it is difficult to establish a highly precise simulation model for a specific milling process, several studies [36, 96, 97] have established positive foundations for further exploration.

(2) Optimizing multi-sensor/ feature configurations

As discussed in Subsection 2.2, the use of multiple sensors can achieve better results than those obtained with a single sensor; however, the additional cost of employing multiple sensors, the increased interference in the milling process associated with multiple sensors, and the impact of accommodating redundant signals on the subsequent feature extraction cannot be ignored. Consequently, it is necessary to optimize multi-sensor/feature configurations to obtain a tradeoff between monitoring cost and TCM model performance with consideration for the types, numbers, installation locations, and costs of sensors. Here, optimum multi-sensor/feature configurations can be obtained by combining advanced feature extraction technologies (e.g., multi-wavelet and EEMD) through multi-objective optimization.

(3) Use of monitoring models for prognosis rather than diagnosis

Diagnosis determines the current system state based on observed data, such as in tool fault diagnosis. In contrast, prognosis predicts future system states based on observed data, such as for predicting the remaining useful life (RUL) of a tool [52, 98]. Predicting the RUL rather than diagnosing tool condition is of particular interest in industrial settings [59] because the RUL and failure probability of a milling tool are more meaningful than the diagnosis of tool wear [17, 99]. However, predicting the RUL of a tool is a more difficult task than the diagnosis and classification of tool condition. The monitoring models discussed in Section 4 have been generally applied for diagnosis rather than prognosis (shown in Table 2). In these models, the degree of tool wear is adopted as the classification criterion with which to construct a tool state classifier based on an AI method. However, a simple classification of tool condition cannot provide an accurate progression of tool state from one state to the next, and can therefore scarcely provide an accurate prediction of the RUL of a tool. Therefore, additional research focused on methods for tool state prognosis is needed, particularly for providing quantitative assessments of tool condition and modifications of monitoring models to the task of prognosis.

(4) Application of more advanced AI-based monitoring methods

The AI field is rapidly developing, resulting in the emergence of new advanced algorithms. Although ANN, HMM, SVM, and other algorithms have been widely applied with reasonably good success in TCM, the latest research in the AI field (e.g., fuzzy clustering, ensemble learning, incremental learning, transfer learning, and depth learning) should be applied to achieve monitoring models with outstanding performance. These latest developments are discussed as follows.

- Fuzzy clustering: Due to the excellence of Fuzzy clustering for accommodating uncertainties, and its independence from prior knowledge, FCM clustering is suitable for conducting TCM in milling processes owing to the limited number of samples and the inherently fuzzy tool state categories involved. As such, efforts to combine the FCM clustering algorithm with advanced feature extraction technology are eagerly anticipated. Of course, FCM clustering requires a preset number of clusters, and the selected number of clusters is a crucial factor affecting clustering performance. Therefore, the means of automatically determining the number of clusters based on sample data must be further studied.
- Ensemble learning: Due to the complexity of the milling process and the uncertainty in tool condition evolution, it is very difficult to adapt a single classifier to all samples without sacrificing generalization capabilities [95, 102]. Here, ensemble learning integrates multiple classifiers in an effort to make use of their respective advantages to improve the overall performance of a monitoring model. In ensemble learning methods (e.g., Boosting, Bagging, Random Forest), some individual learners are constructed through training several existing classification algorithms with training samples, and then are combined with a certain strategy (e.g., voting).
- Incremental learning: Due to the incompleteness of TCM training samples, the diagnostic and predictive capabilities of a monitoring model would greatly benefit from additional training samples after the model has been initially trained and put into use. However, most classifier models rely solely on pre-defined training samples, and previously trained classifier models must be entirely reconstructed if new training samples are to be introduced [4, 77]. The incremental learning strategy allows for trained models to be updated in an incremental manner without deteriorating the performance obtained with older samples [101].
- Transfer learning: As shown in Table 2, considerable articles studied TCM in milling with different cutting conditions (e.g., tool type, the number of tooth on tool, workpiece material), which is costly and time-consuming in total. The experimental cost and model training time required for the TCM of specific milling processes could be reduced if knowledge specific to an existing TCM model could be applied in the development of another similar TCM model. The goal of transfer learning is to apply the knowledge learned from one environment to tasks conducted in a new environment. This is expected to be greatly beneficial to TCM research for milling processes.
- Deep learning: This is currently one of the most prevalent algorithms studied in the fields of AI and machine learning. Moreover, deep learning algorithms have been widely applied to many areas, such as in image recognition, speech recognition, and natural language processing. However, deep learning has not yet been extensively applied in TCM research. Here, the fact that the size of the training samples available for TCM applications is much less than those available in other areas presents the greatest challenge for applying deep learning algorithms in TCM.

Table 2 Publications for TCM in milling process from 2000

Authors	Cutting tool	Workpiece material	Monitoring method	Monitoring target	Tool condition categories (VB: the average flank wear width)
Chen et al. (2000) [48]	TPMN322-CH550	S45C carbon steel	Neural network	Diagnosis	Light wear, middle wear, severe wear, tool breakage
Shao H et al. (2004) [36]	1-flute and 5-flutes face mill	Cast iron	Threshold	Diagnosis	Tool normal and tool breakage
Ghoshet al. (2007) [10]	1-fluteend mill	C-60 steel	Back propagation neural network	Prognosis	VB value (200 times)
Hsueh et al. (2009) [86]	Face mill cutter	7075 aluminum	Support vector machine	Diagnosis	Tool normal and tooldamahe
Binsaeidet al.(2009) [49], (2010) [51]	Two-flutes end mill	AISI 4340 steel	Multi-class SVM	Diagnosis	Slight wear, medium wear, severe wear, chipping, breakage
Zhu et al. (2009) [82]	Two-flutes micro-mill	Copper, steel	continuous HMM	Diagnosis	Initial wear, progressive wear accelerated wear
Girardin et al.(2010) [100]	4-flutesmill	Titanium alloy TA6V	Rotational frequency analysis	Diagnosis	Tool normal and tool breakage
Sevilla et al. (2011) [38]	3-flutes face mill	Aluminum alloy 6061-T6	Arithmetic mean values of asymmetry	Diagnosis	Tool normal and tool breakage
Wang M et al. (2012) [17]	EGD 4440R mill	ASSAB718HH	CHMM	Diagnosis	Good, medium wear, bad wear, worst (ISO8688-1)
Geramifard et al. (2012) [52]	3-flutes ball nose end mill	Inconel 718	Physically segmented HMM	Prognosis	VB value (320 times)
Hsieh et al. (2012) [32]	2-flutes WC micro-end mill	SK2 steel	Backpropagation neural network	Diagnosis	Tool normal and tool breakage
Grasso et al. (2013) [94]	4-flutes porcupine mill	Titanium	Adaptive SPC	Diagnosis	Tool normal and tool abnormal
Lu et al. (2013) [81]	Micro-end mill	SK2 steel	Hidden Markov model	Diagnosis	Sharp tool and worn tool
Wang et al. (2013a) [76]	4-flutes helical mill	45 steel	Auto associative neural network	Diagnosis	New tool, small wear, middle wear, severe wear
Wang Get al. (2013b) [77]	APMT1604PDER-H2	Titanium alloy	Gaussian ARTMAP network	Diagnosis	New tool, small wear, middle wear, severe wear
Yen et al. (2013) [42]	WC 2-flutes steel mill	SK2 steel	Learning vector quantification	Diagnosis	Sharp tool and worn tool
Lamraoui et al. (2014) [53]	2 and 3flutes end mill	Aluminum 7075-T6	Cyclostationary stationary analysis	Diagnosis	New, flank wear, flank wear and a broken tooth, flank wear and two broken teeth
Wang et al. (2014a) [87]	4-flutes end mill	Titanium alloy	Relevance vector machine	Diagnosis	New tool, middle wear, severe wear
Wang et al. (2014 b) [95]	3-flutes end mill	Titanium alloy	Ensemble learning model	Diagnosis	New tool, small wear, middle wear, severe wear
Wang et al. (2014c) [101]	End mill	Titanium alloy Ti-6Al-4V	Fuzzy ARTMAP network	Diagnosis	New tool, small wear, middle wear, severe wear
Ammouri et al. (2014) [35]	HSS 4-flutes end mill	AISI 1020low carbon steel	Current rise index	Diagnosis	Sharp, slightly worn, workable, dull
Zhu et al. (2014) [46]	2- flutes mcrio mill	Copper, steel	Sparse decomposition	Diagnosis	Initial wear, progressive wear accelerated wear
Ren et al. (2014) [45]	2-flutes WC ball end mill	Steel	Type-2 TSK fuzzy logic system	Diagnosis	Sharp tool and worn tool
PoTsang et al. (2015) [18]	4-flutes HSS mill	Steel 1018	Probabilistic neural network	Diagnosis	Tool normal and tool breakage
Karandikar et al. (2015) [7]	1-flute end mill	1018 steel	Naïve Bayes classifiers	Diagnosis	New tool, middle wear, severe wear
Sevilla et al. (2015) [24, 25]	3-flutes face mill	Aluminum alloy 6061-T6	Threshold	Diagnosis	Healthy, partial worn, severe worn, missing
Liu et al. (2015) [4]	4-flutes mill	Ti-6Al-4V titanium alloy	Ellipsoid ARTMAP network	Diagnosis	New cutter, small wear, middle wear, severe wear
Zhang et al. (2015) [16]	1-flute seco milling cutter	TC21 titanium alloy	Statistical parameters of cutting force	Diagnosis	Small wear, middle wear, severe wear (ISO 8688-2:1989)

Table 2 (continued)

Authors	Cutting tool	Workpiece material	Monitoring method	Monitoring target	Tool condition categories (VB: the average flank wear width)
Drouillet et al. (2016) [11]	4-flutes end mill	Stainless steel SS403	ANN	Prognosis	VB value
Stavropoulos et al. (2016) [23]	5-flutes face milling	CGI 450	Third degree regression model	Diagnosis	Low wear, medium wear, high wear
Madhusudana et al. (2016) [33]	6-flutes face mill	Steel alloy 42CrMo4	k-star classification	Diagnosis	New, flank wear, breakage and chipping
Yu et al. (2016) [59]	3-flutes face mill	HRC52 stainless steel	Weighted HMM	Prognosis	VB value (315 times)
Jahromi et al. (2016) [57]	Ball nose end mill	Inconel 718	Sequential fuzzy clustering dynamic	Prognosis	VB value (500 times)
Wang et al. (2016) [55]	4-flutes mill	Ti–6Al–4V titanium alloy	Multi-scale PCA	Diagnosis	Tool normal and tool breakage (ISO 3685-1977)
Torabi et al. (2016) [58]	2-flutes and 3-flutes ball nose end mill	Inconel 718	Fuzzy C-means clustering	Diagnosis	New tool, half worn, worn
Harris et al. (2016) [61]	End mill	Ti-6Al-4V titanium alloy	Multivariate control chart	Diagnosis	Negligible wear and worn out
Zhang et al. (2016) [31]	2-flutes micro-end mill	Steel HRC52	Neuro-fuzzy network	Prognosis	VB value
Javed et al. (2016) [1]	3-flutes ball nose mill	Inconel 718	Summation wavelet-extreme learning machine	Prognosis	VB value
Hong et al. (2016) [56]	2-flutes micro-end mill	Al-5052 aluminum alloy	Hidden Markov model	Diagnosis	Normal, medium wear, excessive wear
Cuka et al. (2017) [27]	Two-flutes end mill	AISI 1045 steel	Fuzzy logic inference	Diagnosis	Normal, small wear, medium wear, accelerated wear, tool breakage
Wang et al. (2017) [72]	3-flutes ball nose mill	Stainless steelHRC52	Support vector regression	Prognosis	VB value (300 times)
Madhusudana et al. (2017) [12]	6-flutes face mill	Steel alloy 42CrMo4	Support vector machine	Diagnosis	Healthy, flank wear, cutting tip breakage (breakage), chipping on rake face near cutting tip (chipping)

Fortunately, some work has been conducted for developing feasible deep learning algorithms with small training samples [103], which offers the potential for adopting TCM monitoring models based on deep learning for milling processes in the future.

Acknowledgments The authors are grateful for support from the National Science Foundation of China (Grant No. 51405346), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY17E050005), and the Wenzhou City Public Industrial Science and Technology Project of China (Grant Nos. G20160015 and G20170009). We also thank LetPub (www.LetPub.com) for its linguistic assistance during the preparation of this manuscript.

References

1. Javed K, Gouriveau R, Li X, Zerhouni N (2016) Tool wear monitoring and prognostics challenges: a comparison of connectionist methods toward an adaptive ensemble model. J Intell Manuf 30(4):1–18
2. Vetrichelvan G, Sundaram S, Kumaran S, Velmurugan P (2014) An investigation of tool wear using acoustic emission and genetic algorithm. J Vib Control 21(15):3061–3066
3. Bhattacharyya P, Sengupta D, Mukhopadhyay S (2007) Cutting force-based real-time estimation of tool wear in face milling using a combination of signal processing techniques. Mech Syst Signal Process 21(6):2665–2683
4. Liu C, Wang G, Li Z (2015) Incremental learning for online tool condition monitoring using ellipsoid ARTMAP network model. App Soft Comput 35:186–198
5. Aliustaoglu C, Ertunc H, Ocak H (2009) Tool wear condition monitoring using a sensor fusion model based on fuzzy inference system. Mech Syst Signal Process 23(2):539–546
6. Konstantinos S, Athanasios K (2014) Reliability assessment of cutting tool life based on surrogate approximation methods. Int J Adv Manuf Technol 71(5–8):1197–1208
7. Karandikar J, Mcleay T, Turner S, Schmitz T (2015) Tool wear monitoring using naïve bayes classifiers. Int J Adv Manuf Technol 77(9–12):1613–1626
8. Rehorn A, Jiang J, Orban P (2005) State-of-the-art methods and results in tool condition monitoring: a review. Int J Adv Manuf Technol 26(7–8):693–710
9. Dutta S, Kanwat A, Pal S, Sen R (2013) Correlation study of tool flank wear with machined surface texture in end milling. Measurement 46(10):4249–4260
10. Ghosh N, Ravi Y, Patra A, Mukhopadhyay S, Paul S, Mohanty A (2007) Estimation of tool wear during CNC milling using neural network-based sensor fusion. Mech Syst Signal Process 21(1):466–479
11. Drouillet C, Karandikar J, Nath C, Journeaux A, Mansori M, Kurfess T (2016) Tool life predictions in milling using spindle power with the neural network technique. J Manuf Process 22:161–168
12. Madhusudana C, Kumar H, Narendranath S (2017) Face milling tool condition monitoring using sound signal. Int J Syst Assu Eng Manag 4:1–11
13. Kim S, Lee C, Lee D, Kim J, Jung Y (2001) Evaluation of the thermal characteristics in high-speed ball-end milling. J Mate Processing Technol 113(1):406–409
14. Nouri M, Fussell B, Ziniti B, Linder E (2015) Real-time tool wear monitoring in milling using a cutting condition independent method. Int J Mach Tools Manuf 89:1–13
15. Azmi A (2015) Monitoring of tool wear using measured machining forces and neuro-fuzzy modeling approaches during machining of GFRP composites. Adv Eng Softw 82:53–64
16. Zhang H, Zhao J, Wang F, Li A (2015) Cutting forces and tool failure in high-speed milling of titanium alloy tc21 with coated carbide tools. J Eng Manuf 229(1):20–27
17. Wang M, Wang J (2012) CHMM for tool condition monitoring and remaining useful life prediction. Int J Adv Manuf Technol 59(5–8):463–471
18. Huang P, Ma C, Kuo C (2015) A PNN self-learning tool breakage detection system in end milling operations. Appl Soft Comput 37:114–124
19. Kaya B, Oysu C, Ertunc H (2011) Force-torque based on-line tool wear estimation system for CNC milling of inconel 718 using neural networks. Adv Eng Softw 42(3):76–84
20. Prickett P, Johns C (1999) An overview of approaches to end milling tool monitoring. Int J Mach Tools Manuf 39(1):105–122
21. Koike R, Ohnishi K, Aoyama T (2016) A sensorless approach for tool fracture detection in milling by integrating multi-axial servo information. CIRP Ann Manuf Technol 65(1):385–388
22. Ghani J, Rizal M, Nuawi M, Ghazali M, Haron C (2011) Monitoring online cutting tool wear using low-cost technique and user-friendly GUI. Wear 271(9–10):2619–2624
23. Stavropoulos P, Papacharalampopoulos A, Vasiliadis E, Chryssolouris G (2016) Tool wear predictability estimation in milling based on multi-sensorial data. Int J Adv Manuf Technol 82(1–4):509–521
24. Sevilla P, Robles J, Jauregui J, Jimenez D (2015a) FPGA-based reconfigurable system for tool condition monitoring in high-speed machining process. Measurement 64:81–88
25. Sevilla P, Robles J, Muñiz J, Lee F (2015b) Tool failure detection method for high-speed milling using vibration signal and reconfigurable bandpass digital filtering. Int J Adv Manuf Technol 81(5–8):1–8
26. Sevilla P, Jauregui J, Herrera G, Robles J (2013) Efficient method for detecting tool failures in high-speed machining process. J Eng Manuf 227(4):473–482
27. Cuka B, Kim D (2017) Fuzzy logic based tool condition monitoring for end-milling. Robot Comput Integr Manuf 47(10):22–36
28. Wang G, Yang Y, Zhang Y, Xie Q (2014) Vibration sensor based tool condition monitoring using ν, support vector machine and locality preserving projection. Sensors Actuators A Phys 209:24–32
29. Zhou Y, Liu X, Li F, Sun B, Xue W (2015) An online damage identification approach for numerical control machine tools based on data fusion using vibration signals. J Vib Control 21(15):2925–2936
30. Chen B, Chen X, Li B, He Z, Cao H, Cai G (2011) Reliability estimation for cutting tools based on logistic regression model using vibration signals. Mech. Syst. Signal Process 25(7):2526–2537
31. Zhang C, Yao X, Zhang J, Jin H (2016) Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations. Sensors 16(795):1–20
32. Hsieh W, Lu M, Chiou S (2012) Application of backpropagation neural network for spindle vibration-based tool wear monitoring in micro-milling. Int J Adv Manuf Technol 61(1–4):53–61
33. Madhusudana C, Kumar H, Narendranath S (2016) Condition monitoring of face milling tool using k-star algorithm and histogram features of vibration signal. Int J Eng Sci Technol 19(3):1543–1551
34. Gao C, Xue W, Ren Y, Zhou Y (2017) Numerical control machine tool fault diagnosis using hybrid stationary subspace analysis and least squares support vector machine with a single sensor. Applied Sci 7(4):346 1-12

35. Ammouri A, Hamade R (2014) Current rise criterion: a process-independent method for tool-condition monitoring and prognostics. Int J Adv Manuf Technol 72(1–4):509–519
36. Shao H, Wang H, Zhao X (2004) A cutting power model for tool wear monitoring in milling. Int J Mach Tools Manuf 44(14):1503–1509
37. Ritou M, Garnier S, Furet B, Hascoet J (2014) Angular approach combined to mechanical model for tool breakage detection by eddy current sensors. Mech Syst Signal Process 44(1–2):211–220
38. Sevilla P, HerreraG RJ, Jáuregui J (2011) Tool breakage detection in cnc high-speed milling based in feed-motor current signals. Int J Adv Manuf Technol 53(9–12):1141–1148
39. Rizal M, Ghani J, Nuawi M, Che H (2014) A review of sensor system and application in milling process for tool condition monitoring. Rese J Applied Sci Eng Technol 7(10):2083–2097
40. Lee B (1999) Application of the discrete wavelet transform to the monitoring of tool failure in end milling using the spindle motor current. Int J Adv Manuf Technol 15(4):238–243
41. Jemielniak K (2008) Arrazola P (2008) application of AE and cutting force signals in tool condition monitoring in micro-milling. CIRP J Manuf Sci Technol 1:97–102
42. Yen C, Lu M, Chen J (2013) Applying the self-organization feature map (som) algorithm to ae-based tool wear monitoring in micro-cutting. Mech Syst Signal Proces 34(1–2):353–366
43. Pechenin V, Khaimovich A, Kondratiev A, Bolotov M (2017) Method of controlling cutting tool wear based on signal analysis of acoustic emission for milling. Procedia Eng 176:246–252
44. Mathew M, Pai P, Rocha L (2008) An effective sensor for tool wear monitoring in face milling: acoustic emission. Sadhana 33(3):227–233
45. Ren Q, Balazinski M, Baron L, Jemielniak K, Botez R, Achiche S (2014) Type-2 fuzzy tool condition monitoring system based on acoustic emission in micromilling. Info Sci 255(1):121–134
46. Zhu K, Vogel B (2014) Sparse representation and its applications in micro-milling condition monitoring: noise separation and tool condition monitoring. Int J Adv Manuf Technol 70(1–4):185–199
47. Snr D (2000) Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. Int J Mach Tools Manuf 40(8):1073–1098
48. Chen S, Jen Y (2000) Data fusion neural network for tool condition monitoring in CNC milling machining. Int J Mach Tool Manuf 40(3):381–400
49. Binsaeid S, Asfour S, Cho S, Onar A (2009) Machine ensemble approach for simultaneous detection of transient and gradual abnormalities in end milling using multisensor fusion. J Mater Process Technol 209(10):4728–4738
50. Kuljanic E, Totis G, Sortino M (2009) Development of an intelligent multisensor chatter detection system in milling. Mech Syst Signal Process 23(5):1704–1718
51. Cho S, Binsaeid S, Asfour S (2010) Design of multisensor fusion-based tool condition monitoring system in end milling. Int J Adv Manuf Technol 46(5–8):681–694
52. Geramifard O, Xu J, Zhou J, Li X (2012) A physically segmented hidden markov model approach for continuous tool condition monitoring: diagnostics and prognostics. IEEE Trans Ind Inf 8(4):964–973
53. Lamraoui M, Thomas M, Badaoui M (2014) Cyclostationarity approach for monitoring chatter and tool wear in high speed milling. Mech Syst Signal Process 44(1–2):177–198
54. Salehi M, Albertelli P, Goletti M, Ripamonti F, Tomasini G, Monno M (2015) Indirect model based estimation of cutting force and tool tip vibrational behavior in milling machines by sensor fusion. Procedia CIRP 33(1):239–244
55. Wang G, Zhang Y, Liu C, Xie Q, Xu Y (2016) A new tool wear monitoring method based on multi-scale PCA. J Intell Manuf:1–10
56. Hong Y, Yoon H, Moon J, Cho Y, Ahn S (2016) Tool-wear monitoring during micro-end milling using wavelet packet transform and fisher's linear discriminant. Int J Prec Eng Manuf 17(7):845–855
57. Jahromi A, Meng J, Li X, Lim B (2016) Sequential fuzzy clustering based dynamic fuzzy neural network for fault diagnosis and prognosis. Neurocomputing 196:31–41
58. WangP GR (2016) Stochastic tool wear prediction for sustainable manufacturing. Procedia CIRP 48:236–241
59. Yu J, Liang S, Tang D, Liu H (2016) A weighted hidden markov model approach for continuous-state tool wear monitoring and tool life prediction. Int J Adv Manuf Technol 91(1–4):201–111
60. Downey J, O'Sullivan D, Nejmen M, Bombinski S, O'Leary P, Raghavendra R (2016) Real time monitoring of the CNC process in a production environment- the data collection and analysis phase. Procedia CIRP 41:920–926
61. Harris K, Triantafyllopoulos K, Stillman E, Mcleay T (2016) A multivariate control chart for autocorrelated tool wear processes. Qual Reliab Eng Int 32(6):2093–2106
62. Christopher A, John T (2007) Methods for on-line directionally independent failure prediction of end milling cutting tools. Mach Sci Technol 11(1):1–43
63. Zhu K, Wong Y, Hong G (2009) Wavelet analysis of sensor signals for tool condition monitoring: a review and some new results. Int J Mach Tools Manuf 49(7):537–553
64. Sun H, He Z, Zi Y, Yuan J, Wang X, Chen J, He S (2014) Multiwavelet transform and its applications in mechanical fault diagnosis—a review. Mech Syst Signal Process 43(1–2):1–24
65. Huang N, Shen Z, Long S, Wu M, Shih M, Zheng Q, Yen N, Tung C, Liu H (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond 454:903–995
66. LeiY, LinJ, He Z, Zuo M (2013) A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech Syst Signal Process 35(1–2):108–126
67. Shi X, Wang R, Chen Q, Shao H (2014) Cutting sound signal processing for tool breakage detection in face milling based on empirical mode decomposition and independent component analysis. J Vib Control 21(16):3348–3358
68. Babouri M, Ouelaa N, Djebala A (2016) Experimental study of tool life transition and wear monitoring in turning operation using a hybrid method based on wavelet multi-resolution analysis and empirical mode decomposition. Int J Adv Manuf Technol 82(9–12):2017–2028
69. Wu Z, Huang N (2009) Ensemble empirical mode decomposition: a noise assisted data analysis method. Adv Adapt Data Anal 1:1–41
70. Wang Y, Yeh C, Young H, Hu K, Lo M (2014) On the computational complexity of the empirical mode decomposition algorithm. Phys A Stat Mech Appl 400(4):159–167
71. Zhao L, Yu W, Yan R (2016) Gearbox fault diagnosis using complementary ensemble empirical mode decomposition and permutation entropy. Shock Vib. https://doi.org/10.1155/2016/3891429
72. Wang J, Xie J, Zhao R, Zhang L, Duan L (2017) Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing. Robot Comput Integr Manuf 45:47–58
73. Szecsi T (1999) Cutting force modeling using artificial neural networks. J Mater Process Technol 92-93(3):344–349
74. Pimenov D (2013) The effect of the rate flank wear teeth face mills on the processing. J Frict Wear 34(2):156–159
75. Kalidass S, Palanisamy P, Muthukumaran V (2013) Prediction and optimization of tool wear for end milling operation using artificial neural networks and simulated annealing algorithm. Int J Mach Mach Mater 14(2):142–164
76. Wang G, Cui Y (2013a) On line tool wear monitoring based on auto associative neural network. J Intell Manuf 24:1085–1094

77. Wang G, Yang Y, Guo Z (2013b) Hybrid learning based gaussian ARTMAP network for tool condition monitoring using selected force harmonic features. Sensors Actuators A Phys 203(12):394–404
78. Torabi A, Meng J, Li X, Lim B, Zhai L, Oentaryo RJ (2015) A survey on artificial intelligence-based modeling techniques for high speed milling processes. IEEE Syst J 9(3):1069–1080
79. Salimiasl A, Özdemir A (2016) Analyzing the performance of artificial neural network (ANN)-, fuzzy logic (FL)-, and least square (LS)-based models for online tool condition monitoring. Int J Adv Manuf Technol 87(1–4):1–14
80. Owsley L, Atlas L, Bernard G (1997) Self-organizing feature maps and hidden Markov models for machine-tool monitoring. IEEE Trans Signal Process 45(11):2787–2798
81. Lu M, Wan B (2013) Study of high-frequency sound signals for tool wear monitoring in micromilling. Int J Adv Manuf Technol 66(9–12):1785–1792
82. Zhu K, Wong Y, Hong G (2009) Multi-category micro-milling tool wear monitoring with continuous hidden markov models. Mech Syst Signal Process 23(2):547–560
83. Geramifard O, Xu J, Zhou J, Li X (2014) Multimodal hidden markov model-based approach for tool wear monitoring. IEEE Trans Ind Electron 61(6):2900–2911
84. Siddhpura A, Paurobally R (2013) A review of flank wear prediction methods for tool condition monitoring in a turning process. Int J Adv Manuf Technol 65(1–4):371–393
85. Shawe T, Sun S (2011) A review of optimization methodologies in support vector machines. Neurocomputing 74(17):3609–3618
86. Hsueh Y, Yang C (2009) Tool breakage diagnosis in face milling by support vector machine. J Mater Process Technol 209(1):145–152
87. Wang G, Yang Y, Xie Q, Zhang Y (2014a) Force based tool wear monitoring system for milling process based on relevance vector machine. Adv Eng Softw 71(3):46–51
88. Abbasnejad M, Ramachandram D, Mandava R (2012) A survey of the state of the art in learning the kernels. Knowl Inf Syst 31(2):193–221
89. Douha L, Benoudjit N, Douak F, Melgani F (2012) Support vector regression in spectrophoto-metry: an experimental study. Crit Rev Anal Chem 42(3):214–219
90. Lei Y, He Z, Zi Y, Hu Q (2008) Fault diagnosis of rotating machinery based on a new hybrid clustering algorithm. Int J Adv Manuf Technol 35(9–10):968–977
91. Zhang T, Ye W, Shan Y (2015) Application of sliced inverse regression with fuzzy clustering for thermal error modeling of CNC machine tool. Int J Adv Manuf Technol 85(9–12):1–11
92. Fu S, Liu K, Xu Y, Liu Y (2016) Rolling bearing diagnosing method based on time domain analysis and adaptive fuzzy c - means clustering. Shock Vib (1): 1–8
93. Torabi A, Meng J, Xiang L, Lim B, Gan O (2016) Application of clustering methods for online tool condition monitoring and fault diagnosis in high-speed milling processes. IEEE Syst J 10(2):721–732
94. Grasso M, Albertelli P, Colosimo B (2013) An adaptive SPC approach for multi-sensor fusion and monitoring of time-varying processes. Procedia CIRP 12:61–66
95. Wang G, Liu C, Cui Y, Feng X (2014b) Tool wear monitoring based on cointegration modeling of multisensory information. Int J Comput Integr Manuf 27(5):479–487
96. Marksberry P, Jawahir I (2008) A comprehensive tool-wear/tool-life performance model in the evaluation of ndm (near dry machining) for sustainable manufacturing. Int J Mach Tool Manu 48(7):878–886
97. Zhang G, To S, Xiao G (2014) Novel tool wear monitoring method in ultra-precision raster milling using cutting chips. Precis Eng 38(3):555–560
98. Zhou J, Pang C, Lewis F, Zhong Z (2009) Intelligent diagnosis and prognosis of tool wear using dominant feature identification. IEEE Trans Ind Inf 5(4):454–464
99. Wu Y, Hong G, Wong W (2015) Prognosis of the probability of failure in tool condition monitoring application-a time series based approach. Int J Adv Manuf Technol 76(1–4):513–521
100. Girardin F, Rémond D, Rigal J (2010) Tool wear detection in milling - an original approach with a non-dedicated sensor. Mech Syst Signal Pr 24(6):1907–1920
101. Wang G, Guo Z, Qian L (2014) Online incremental learning for tool condition classification using modified fuzzy ARTMAP network. J Intell Manuf 25(6):1403–1411
102. Wang G, Yang Y, Li Z (2014) Force sensor based tool condition monitoring using a heterogeneous ensemble learning model. Sensors 14(11):21588–21602
103. Salehinejad H, Barfett J, Valaee S, Dowdell T (2017) Training neural networks with very little data—a draft. Arxiv, https://arxiv.org/pdf/1708.04347.pdf