

Tool Wear Prediction Based on Multi-Information Fusion and Genetic Algorithm-Optimized Gaussian Process Regression in Milling

Zhiwen Huang¹, Member, IEEE, Jiajie Shao¹, Weicheng Guo¹, Weidong Li¹, Member, IEEE, Jianmin Zhu¹, Qichao He¹, and Dianjun Fang¹

Abstract—Tool wear condition monitoring plays a crucial role in intelligent manufacturing systems to enhance machining quality and efficiency. The indirect methods employ various sensor signals to monitor tool wear condition, attracting wide attention in industrial applications. Multi-information fusion technologies can promote tool wear monitoring results to be more accurate and reliable. For improving the prediction accuracy and ensuring the reliability of the indirect methods, this study proposes a tool wear prediction method based on multi-information fusion and genetic algorithm (GA)-optimized Gaussian process regression (GPR). First, wavelet packet denoising (WPD)-based signal processing is adopted to suppress the noise interference of multisensor signals. Then, kernel principal component analysis (KPCA)-based dimension reduction is employed to mine the most sensitive features to flank wear from candidate multidomain features. Next, a fusion model of GPR and GA optimization is designed to establish a nonlinear mapping relationship between sensitive characteristics and flank wear width. Finally, performance evaluations under three sets of milling tests are carried out to validate the effectiveness of the proposed method. Experimental results indicate that the proposed method can lower prediction error and uncertainty of flank wear width compared with other intelligent approaches, promoting a successful application of indirect monitoring methods in milling.

Index Terms—Gaussian process regression (GPR), genetic algorithm (GA), milling, multi-information fusion, tool wear monitoring.

Manuscript received 18 April 2023; accepted 15 May 2023. Date of publication 29 May 2023; date of current version 14 June 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 52105470, Grant 51975444, and Grant 51775323; in part by the National Key Research and Development Program of China under Grant 2022YFE0114300; and in part by the Postdoctoral Research Program of the University of Shanghai for Science and Technology under Grant 1022304302. The Associate Editor coordinating the review process was Dr. Weihua Li. (Zhiwen Huang and Jiajie Shao contributed equally to this work.) (Corresponding author: Jianmin Zhu.)

Zhiwen Huang, Weicheng Guo, Weidong Li, and Jianmin Zhu are with the School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China (e-mail: hzwedu@163.com; wcguo@usst.edu.cn; weidongli@usst.edu.cn; jmzhu_usst@163.com).

Jiajie Shao is with the School of Mechanical Engineering, Tongji University, Shanghai 200092, China (e-mail: sjj6901@163.com).

Qichao He is with Shanghai Aerospace Equipment Manufacturing General Company Ltd., Shanghai 200245, China (e-mail: vegetable_1989@126.com).

Dianjun Fang is with the School of Mechanical Engineering, Tongji University, Shanghai 200092, China, and also with the Qingdao Sino-German Institute of Intelligent Technologies, Qingdao 266000, China (e-mail: fang@tongji.edu.cn).

Digital Object Identifier 10.1109/TIM.2023.3280531

I. INTRODUCTION

NUMERICAL control milling is frequently used in various industrial fields, such as automobile [1] and aerospace [2]. The machining processes involve continuous material removal, which increases tool wear and may even result in cutting tool failure [3]. Tool wear condition directly affects the surface quality of finished workpieces [4]. Furthermore, according to relevant statistics [5], cutting tool failures account for about 20% of the machine downtime, and 3%–12% of machining time is lost on tool replacement. Therefore, tool wear monitoring has become increasingly critical to enhance the workpiece surface quality, reduce maintenance downtime, and improve production efficiency.

Many studies have been conducted throughout the years to promote the development of tool wear monitoring techniques, including direct and indirect measurement methods [6]. Measuring direct indicators (such as flank wear width or area [7]) by a microscope [8] or camera [9], the direct approach can reach high measurement precision. However, it is often performed offline or interfered with by cutting chips and coolant fluids, which are limited in industrial applications [10]. The indirect approach mines wear information from real-time sensor signals (such as cutting force, vibration, acoustic emission (AE), spindle motor current or power, and cutting temperature) [11], which can avoid interrupting continuous cutting processes. Therefore, the indirect method is more suitable for a practical application and has attracted wide attention and research.

A. Limits of Prior Arts

Recently, due to the complementarity and fault tolerance of different signals, multisensor fusion approaches enabled by artificial intelligence technologies have been widely applied in indirect tool wear monitoring [12], [13]. Among them, artificial neural network (ANN)-based methods (especially deep learning-based) have achieved outstanding performance in tool wear monitoring, benefiting from their powerful feature mining and data mapping abilities. However, their performances heavily rely on massively available samples with wear labels and a similar data distribution between testing and training samples [14]. However, because of data collected cost

and machining parameter variation, these two requirements are difficult to be met in practical applications, resulting in their performances deteriorating sharply. Moreover, these networks are generally regarded as black boxes and trained based on an empirical risk minimization principle, leading them to fall into a local minimum easily, especially under small samples [15].

As another of the most frequently used algorithms in tool wear monitoring, support vector regression (SVR) has clear interpretability and takes a structural risk minimization principle as an optimizing goal [16]. It holds excellent fitting and generalization abilities under small samples. However, like ANN-based approaches, SVR-based methods cannot provide an uncertainty estimation of predicted results directly [17]. Besides, other intelligent algorithms, such as least squares SVR (LSSVR) and random forest (RF), are also inconvenient to give uncertainties of their predicted results in indirect tool wear monitoring.

B. Research Motivation

The uncertainty analysis of predicted results is crucial to improve the stability and reliability of indirect methods. As an intelligent decision-making algorithm solved by Bayesian inference, Gaussian process regression (GPR) can provide uncertainty quantification of predicted results compared with other algorithms [17], such as ANN, SVR, and LSSVR. In addition, GPR holds better adaptability to deal with complex problems of high dimension and small samples, benefiting from its powerful nonlinear fitting ability and flexible parameter-solving approach [18]. Thus, GPR can monitor tool wear status more effectively and reliably under small samples. However, GPR is generally trained by a conjugate gradient algorithm, which makes its performance rely on the initial value and is prone to falling into a local optimum [19].

As an effective global optimization algorithm, the genetic algorithm (GA) is inspired by the natural selection and evolution mechanism [20]. Since the optimizing process does not depend on gradient computations, it has strong robustness and global search ability [21], which can effectively optimize the hyperparameters of GPR. However, the GA-optimized GPR model is still developing in tool wear monitoring, especially integrating into multisensor information fusion methods under small samples and multiple milling parameters.

C. Main Contribution

To further improve the accuracy and reliability of indirect methods under small samples, this study presents a tool wear prediction method based on multi-information fusion and GA-optimized GPR in milling operations. The major contributions are summarized as follows.

- 1) An indirect tool wear prediction method is developed to estimate flank wear width using multisensor signals, which increases the monitoring accuracy of tool wear conditions and reduces the prediction uncertainty of the indirect approach.
- 2) A fusion model based on GPR and GA optimization is designed to establish a nonlinear mapping between flank wear width and its sensitive multidomain features. It can

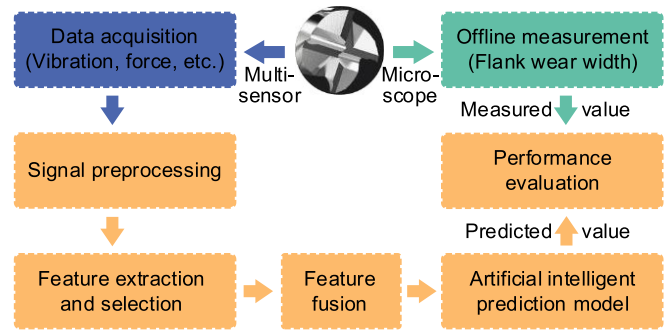


Fig. 1. Main steps of the indirect measurement approach.

achieve lower prediction error and provide uncertainty quantification of the predicted results.

- 3) Milling tests under multiple operating parameters are performed to verify the effectiveness of the proposed method, which can obtain better performance than other advanced approaches, promoting a successful practice of the indirect measurement technique under small samples.

The rest of this article is organized as follows. Section II reviews related works about indirect tool wear monitoring. The experimental setup, data acquisition, and the proposed methodology are described in Section III. Section IV elaborates on and discusses model performances and comparative results. The conclusion is drawn in Section V.

II. LITERATURE REVIEW

The indirect approach builds an intelligent prediction model for tool wear monitoring based on massive data historically collected from various sensors and then makes decisions upon the real-time signals online [22]. As shown in Fig. 1, it mainly consists of data acquisition, signal preprocessing, feature extraction and selection, feature fusion, and predictive model construction. It has received a lot of attention, and many achievements have emerged over the years.

In terms of data acquisition, cutting force [23], vibration [24], AE [25], motor current [26] or power [27], and cutting temperature [28] are the most frequent signals used to monitor tool wear. Recently, taking advantage of the complementarity and fault tolerance of various signals, multisensor signal fusion can comprehensively monitor tool wear information. It makes monitoring results more reliable and accurate and has been widely used in tool wear condition monitoring [13], [29]. Some relevant studies about tool wear prediction using multisensor signals are listed in Table I. It can be found that a combination of cutting force and vibration signals is one of the most common multisensor fusion strategies used for tool wear monitoring in machining processes. In addition, the sensitivity of cutting force and vibration signal to tool wear is different in the x -, y -, and z -directions [30]. Therefore, this study collects 3-D cutting forces and vibration signals to predict the flank wear width of milling cutters.

In terms of signal preprocessing, Kalman filtering [51], median filtering [52], bandpass filtering [53], and Wiener filtering [54] are usually used to alleviate the environmental

TABLE I
RELEVANT RESEARCH ABOUT TOOL WEAR PREDICTION
USING MULTISENSOR SIGNALS

Multi-sensor signals	Literatures
Cutting force and vibration	[30], [31], [32], [33], [34], [35], [36]
Cutting force and spindle power	[37]
Cutting force and AE	[38]
Cutting force, vibration, and AE	[39], [40], [41], [42], [43]
Cutting force, vibration, and current	[44]
Cutting force, vibration, and temperature	[45]
Cutting force, vibration, AE, temperature, and spindle current	[46], [47]
Other fusion signals	[48], [49], [50]

noise involved in multisensor signals. However, cutting force and vibration signals collected during machining operations are generally nonstationary and time-varying. The filtering schemes cannot effectively remove the environmental noise. Utilizing various basis functions and subdivided frequency ranges, wavelet packet denoising (WPD) has an advantage in effectively suppressing time-varying noise, thereby enhancing the signal-to-noise ratio of collected signals [55], [56]. Therefore, WPD-based signal preprocessing is employed in this study to reduce the interference of environmental noise.

In terms of feature extraction and selection, time, frequency, and wavelet domain (WD) features are first extracted to decrease the redundant information of multisensor signals. Generally, time domain (TD) features include maximum, peak-to-peak, variance, root mean square, skewness, kurtosis, and so on [34]. The frequency spectrum and its statistics are often utilized in the frequency domain (FD) [35]. Also, wavelet energy is frequently adopted in the time–frequency domain [36]. In this study, considering the complementarity of different domain features, multidomain features are extracted from multisensor signals. To further reduce the redundant information contained in initial features, various dimension reduction techniques have been applied for feature selection and fusion, such as kernel principal component analysis (KPCA) [30], [34], [57], minimum redundancy maximum relevance (MRMR) [58], and locally linear embedding (LLE) [59]. Among them, KPCA can effectively fuse principal components from numerous features, and it is a popular technique for nonlinear dimensionality reduction in tool wear monitoring [30], [34], [57]. Thus, in this study, the KPCA algorithm is adopted to fuse candidate features and obtain the most sensitive characteristics related to tool wear.

In terms of intelligent prediction models, ANN, SVR, LSSVR, multiple linear regression (MLR), RF, and GPR are commonly adopted to predict the flank wear of cutting tools, which are listed in Table II. Among them, ANN is one of the most frequently used intelligent models. Gao et al. [39] employed a gated recurrent unit for multisensor fusion and lowered the tool wear predicting error. He et al. [41] used stacked sparse autoencoders to improve tool wear prediction performance. Liu et al. [42] and Xu et al. [43] developed

TABLE II
INTELLIGENT MODELS APPLIED FOR TOOL WEAR
PREDICTION USING MULTISENSOR SIGNALS

Intelligent prediction models	Literatures
Artificial neural network (ANN)	[32], [35], [39], [40], [41], [42], [43], [44], [45]
Support vector regression (SVR)	[31], [34], [36], [50], [62], [63]
Least squares support vector regression (LSSVR)	[30], [64], [65], [66]
Multiple linear regression (MLR)	[37], [67], [68]
Random forest (RF)	[50], [69], [70]
Gaussian process regression (GPR)	[17], [19], [71]
Other models	[38], [47], [48], [72]

deep learning networks to fuse multisensor features and lowered prediction errors of flank wear. Paul and Varadarajan [45] proposed an ANN-based multisensor fusion model to predict tool wear. Feng et al. [49] introduced an attention mechanism to the convolutional neural network and improved the monitoring accuracy of tool wear using multisensor features. Although ANN-based methods (including deep learning methods) have achieved excellent performance in tool wear monitoring, monitoring results heavily depend on the following requirements. One is that mass training samples with wear labels are available [60]. Another is that testing samples obey the same data distribution as training samples [61]. However, these requirements are hard to be met in industrial applications due to the acquired and labeled costs of massive data and parameter variations of machining operations [14]. Moreover, ANN-based models are generally trained by an empirical risk minimization principle and easily fall into a local minimum under small samples [15].

In addition, taking advantage of clear interpretability and structural risk minimization principle, SVR and its variants are frequently used algorithms in tool wear monitoring [30], [34]. For example, Benkedjough et al. [62] used SVR to predict life prediction of cutting tools. Kong et al. [63] applied ν -SVR to learn the correlation between flank wear width and fused features. Taking the squared error instead of tube error as the objective function, LSSVR can reduce the calculative burden and increase modeling accuracy. Zhang et al. [66] applied LSSVR to predict the flank wear width of milling cutters and improved the predicting accuracy. However, they cannot give uncertainty quantification of tool wear monitoring results. Wu et al. [69] and Bustillo et al. [70] utilized RF for wear-modeling processes and obtained higher prediction accuracy than ANN and SVR. Although these methods hold strong fitting and generalization abilities for tool wear monitoring under small samples, they still cannot provide an uncertainty estimation of predicted results easily [17]. However, uncertainty analysis is vital to ensure reliability and enhance the applicability of indirect monitoring methods.

As a Bayesian inference-based intelligent algorithm, GPR can offer an uncertainty quantification to analyze the reliability of predicted results [71], [73]. Kong et al. [17] used GPR to obtain better accuracy than ANN and SVR in tool wear

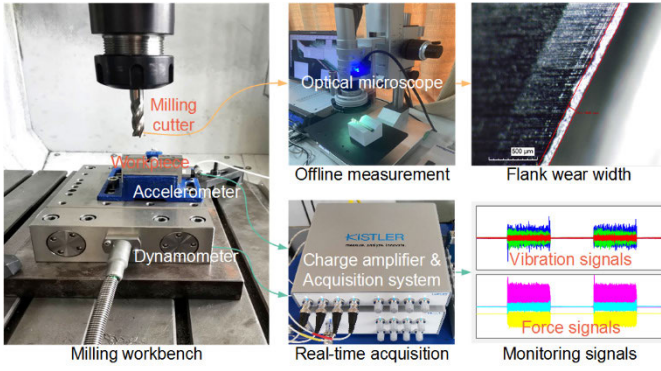


Fig. 2. Illustration of the data acquisition process.

prediction and analyzed a confidence interval (CI) under intrinsic parameters. Zhang et al. [19] and Li et al. [71] employed GPR to predict tool wear and improved the prediction accuracy and reliability. Although these methods have better adaptability to deal with tool wear monitoring problems of high dimensions and small samples, their performances are easily affected by the intrinsic parameters. In addition, GPR is generally trained by a conjugate gradient algorithm, making it hard to obtain the global optimum [19], [74].

As a natural selection and evolution mechanism-inspired probabilistic search algorithm, GA has strong robustness and global search ability [20], [21], which provides a potential solution to improve the performance of GPR. However, the GA-enabled GPR model is rarely investigated for monitoring tool flank wear in milling operations. Furthermore, the current studies scarcely combined them with the above-mentioned WPD-based multisensor signal preprocessing and KPCA-based multidomain feature fusion strategies. Therefore, this study presents an indirect tool wear prediction method based on multi-information fusion and GA-optimized GPR in milling operations, which synthetically utilize these methods to improve the predicting performance and quantitatively analyze the reliability.

III. MATERIALS AND METHODS

A. Experimental Setup and Data Acquisition

The milling experiments of die steel with high-speed steel cutters are implemented to validate the presented tool wear prediction method in this study. The experimental setup and data acquisition process are shown in Fig. 2. In the real-time data acquisition process, a dynamometer of Kistler 9347C is applied to monitor cutting forces, and an accelerometer of Kistler 8763B is used to monitor cutting vibrations. Meanwhile, an NI-DAQ system is adopted to collect multisensor signals at a sampling frequency of 2500 Hz. After finishing each milling operation, an optical microscope of the Hirox KH-7700 is adopted to offline measure the flank wear width of milling cutters.

In addition, three sets of milling tests under different machining conditions are carried out to verify the generalization ability of the proposed tool wear prediction method. The milling parameters are shown in Table III. Each set of

TABLE III
MACHINING PARAMETERS OF THREE SETS OF MILLING EXPERIMENTS

Group	Cutter	Spindle speed (r/min)	Feed speed (mm/min)	Feed engage-ment (mm/r)	Axial cutting depth (mm)	Radial cutting depth (mm)	Material removal rate
1	01-15	2000	500	0.0625	6	0.10	5.0
2	16-30	2000	500	0.0625	6	0.15	7.5
3	31-45	2000	500	0.0625	6	0.20	10.0

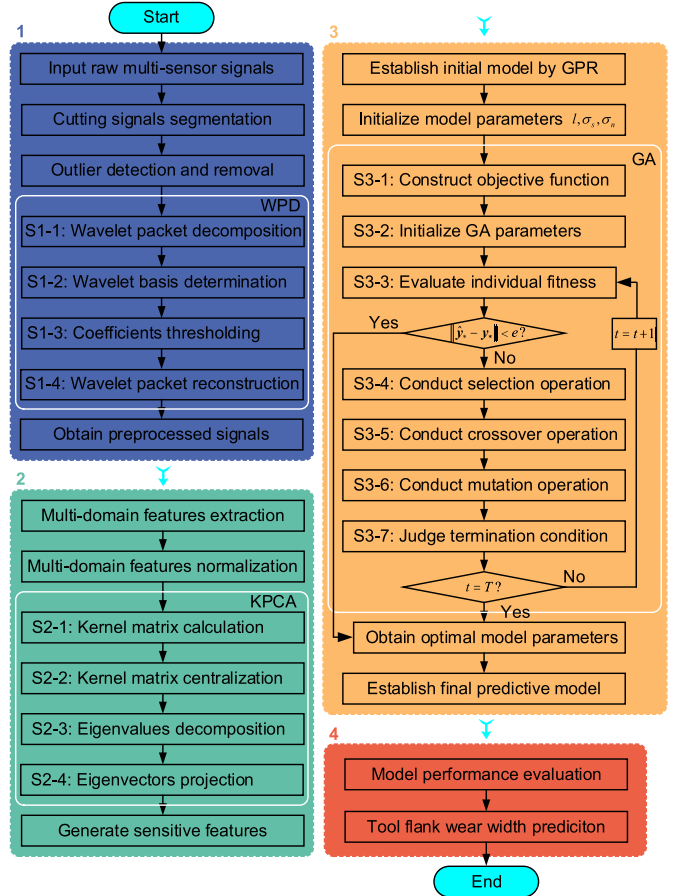


Fig. 3. Overall flowchart of the proposed tool wear prediction method.

milling tests applies 15 milling cutters. Further details of the experimental setup and data acquisition are available in the previous study [30].

B. Proposed Tool Wear Prediction Method

To further improve the prediction accuracy of tool wear and provide an uncertainty estimation of the indirect tool wear monitoring methods in milling operations, this study presents a tool wear prediction method based on multi-information fusion and GA-optimized GPR. Its overall flowchart is shown in Fig. 3.

As shown in Fig. 3, the proposed tool wear prediction method consists of four modules: 1) WPD-based sensor signal preprocessing; 2) KPCA-based multidomain feature fusion; 3) GA-GPR-based predictive model building; and 4) performance

evaluation and wear prediction. Each module is elaborated on in the following.

C. WPD-Based Sensor Signal Preprocessing

During the milling operations, the raw collected signals usually include noise information, and thus, signal preprocessing is crucial to enhance the signal-to-noise ratio of the monitoring signals. As one of the most effective time–frequency analysis approaches, WPD can be adopted to suppress the noise portion and enhance the characteristic component of the monitoring signals [75]. Thus, the WPD-based sensor signal preprocessing is performed on the collected signals.

As shown in Fig. 3, the cutting signals are first segmented from the raw collected signals to eliminate the interference of the noncutting phase. Then, for the segmented cutting signals, the outliers are detected and removed to control the unexpected disturbance from the data acquisition system. The outliers are defined as the sampling points above the mean plus or minus three times the standard deviation during a sampling period. Also, each outlier is replaced by the mean of 100 sampling points adjacent to it, considering the time-varying characteristics of the collected signals. In addition, the WPD approach is adopted to suppress the environmental noise contained in the cutting signals. Also, its main steps are summarized as follows.

Step 1-1: The four-level wavelet packet decomposition is carried out on the input time-varying and nonstationary cutting signals after selecting a “db8” wavelet with orthogonality and discreteness.

Step 1-2: The optimal wavelet tree is calculated to determine the optimal wavelet packet basis according to the given entropy criterion.

Step 1-3: The soft thresholding is selected to quantize each decomposed coefficient, considering the continuity and smoothness of the output signals.

Step 1-4: The denoising signal is obtained based on wavelet packet reconstruction using the processed coefficients.

Finally, after performing sensor signal preprocessing based on WPD, the preprocessed signals are utilized to predict the flank wear width of the milling cutter. Correspondingly, taking vibration signals in the z -direction collected from Group 2 as an example, its WPD-based signal preprocessing procedure is given in Fig. 4.

D. KPCA-Based Multidomain Feature Fusion

For the preprocessed signals, TD, FD, and WD features are first extracted to reduce the information redundancy [34]. As listed in Table IV, maximum, peak-to-peak, variance, root mean square, skewness, and kurtosis are adopted in the TD. Spectral skewness and spectral kurtosis are picked in the FD. A wavelet energy feature is chosen in the wavelet (time–frequency) domain. Thus, nine features are extracted from each channel signal within one second, and 54 candidate features are obtained for each milling feed.

Then, the candidate features are normalized to eliminate the effects of the physical unit in each domain, which is calculated

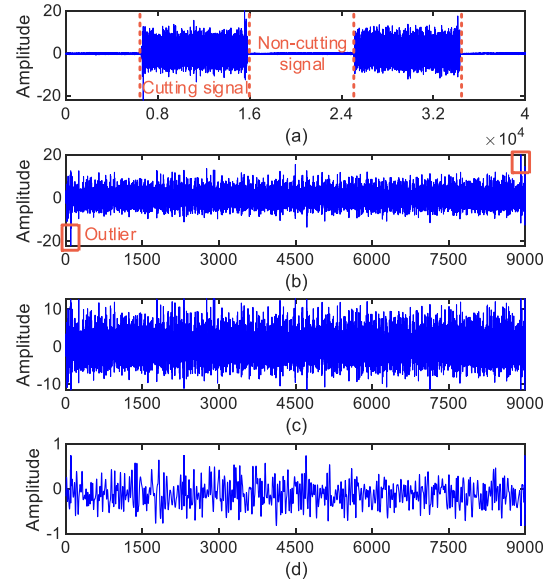


Fig. 4. Example of WPD-based signal preprocessing step using vibration signal from Group 2. (a) Online collected signals. (b) Segmented cutting signals. (c) Outlier removed signals. (d) Wavelet packet denoised signals.

TABLE IV
NINE MULTIDOMAIN FEATURES EXTRACTED
FROM THE PREPROCESSED SIGNALS

Feature	Expression
Maximum	$F_1 = \max(\mathbf{x})$
Peak-to-peak	$F_2 = \max(\mathbf{x}) - \min(\mathbf{x})$
Variance	$F_3 = \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})^2 / N$
Root mean square	$F_4 = \sqrt{\sum_{i=1}^N (\mathbf{x}_i)^2 / N}$
Skewness	$F_5 = E\left[\frac{(\mathbf{x} - \mu)}{\sigma}\right]^3$
Kurtosis	$F_6 = E\left[\frac{(\mathbf{x} - \mu)}{\sigma}\right]^4$
Spectral skewness	$F_7 = \sum_{i=1}^N \left(\frac{f_i - \bar{f}}{\sigma}\right)^3 \cdot S(f_i)$
Spectral kurtosis	$F_8 = \sum_{i=1}^N \left(\frac{f_i - \bar{f}}{\sigma}\right)^4 \cdot S(f_i)$
Wavelet energy	$F_9 = \sum_{i=1}^N wt_{\phi}^2(i) / N$

as follows:

$$\mathbf{x} = (\mathbf{x}^c - \mathbf{x}_{\min}^c) / (\mathbf{x}_{\max}^c - \mathbf{x}_{\min}^c) \quad (1)$$

where \mathbf{x}^c represents candidate features extracted from multi-sensor signals.

In addition, an example of multidomain feature extraction and normalization is shown in Fig. 5. It can be seen that not

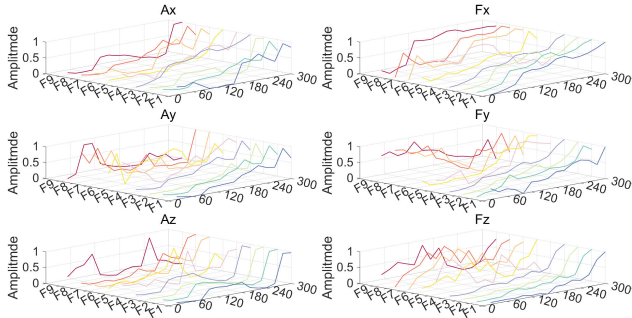


Fig. 5. Multidomain feature extraction and normalization of preprocessed signals from Group 2.

all candidate features are obviously sensitive to the tool wear process. Therefore, the feature selection is significant to obtain the most principal features [34].

As a widely used nonlinear dimensionality reduction approach, the KPCA can effectively solve the most principal components without performing the explicit calculation of the nonlinear mapping [76]. Thus, this study employs it to fuse the most sensitive features and further reduce the dimension of the candidate features. The algorithm of KPCA is briefly introduced in the following.

Step 2-1: The kernel matrix \mathbf{K} is computed according to the selected kernel function $K(\cdot)$, which is expressed as follows:

$$\mathbf{K} = \{K_{ij}\}_{n \times n}, \quad K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \quad (2)$$

where n represents the number of the input samples, \mathbf{x}_i and \mathbf{x}_j are the samples of input space, and $\phi(\cdot)$ denotes the nonlinear mapping from input space to feature space.

Step 2-2: Centralization of the kernel matrix \mathbf{K} is carried out as follows:

$$\widehat{\mathbf{K}} = \mathbf{K} - \mathbf{I}\mathbf{K} - \mathbf{K}\mathbf{I} + \mathbf{I}\mathbf{K}\mathbf{I} \quad (3)$$

where \mathbf{I} denotes a $n \times n$ matrix consisting of the element $1/n$.

Step 2-3: The eigenvalues of the centralized kernel matrix $\widehat{\mathbf{K}}$ is decomposed and computed by solving the following expression:

$$\widehat{\mathbf{K}}\boldsymbol{\alpha} = n\lambda\boldsymbol{\alpha} \quad (4)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ represents the eigenvectors and $n\lambda$ denotes the corresponding eigenvalues.

In particular, the above equation can be derived by

$$\begin{cases} \mathbf{C} = \frac{1}{n} \sum_{i=1}^n \phi(\mathbf{x}_i)\phi(\mathbf{x}_i)^T, & \mathbf{v} = \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i) \\ \lambda \langle \phi(\mathbf{x}_i), \mathbf{v} \rangle = \langle \phi(\mathbf{x}_i), \mathbf{C}\mathbf{v} \rangle, & i = 1, 2, \dots, n \end{cases} \quad (5)$$

where \mathbf{C} represents the covariance matrix of the mapping data and \mathbf{v} denotes its eigenvector.

Step 2-4: The projection of mapping data $\phi(\mathbf{x}_i)$ into the eigenvectors \mathbf{v} is implemented to obtain the principal component \mathbf{x}_i^p as follows:

$$\begin{aligned} \mathbf{x}_i^p &= \langle \phi(\mathbf{x}_i), \mathbf{v} \rangle = \frac{1}{n} \sum_{j=1}^n \alpha_j \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \\ &= \frac{1}{n} \sum_{j=1}^n \alpha_j K(\mathbf{x}_i, \mathbf{x}_j). \end{aligned} \quad (6)$$

Note that the number of the principal components is determined by the accumulation variance percentage, which is expressed as follows:

$$\sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i \geq \lambda. \quad (7)$$

In addition, the radial basis function (RBF) is selected as the kernel function, and the expression is given as follows:

$$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma_k^2). \quad (8)$$

After finishing multidomain feature fusion based on KPCA, the principal features related to tool wear are obtained, which are utilized to establish a predictive model for the flank wear width of the milling cutter.

E. GA-GPR-Based Predictive Model Building

1) GPR-Based Initial Model Building: As one of the most efficient regression methods, GPR has good adaptability and generalization ability to deal with complex nonlinear problems of high dimensions and small samples [17], [19]. Moreover, as a Bayesian inference-based algorithm, GPR can offer an uncertainty quantification to analyze the reliability of predicted results. Thus, this study adopts GPR to build the initial model of tool wear prediction. The details of its modeling are introduced as follows.

In general, GPR is the kernel-based nonparametric modeling based on Gaussian distribution. For a given training dataset $D = \{\mathbf{x}_i, y_i\}_{i=1}^n$, where \mathbf{x}_i represents the i th input sample and y_i denotes the corresponding observation value, the Gaussian process can be defined as follows:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (9)$$

where $m(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$ are the mean function and kernel function, respectively, which are expressed as follows:

$$\begin{cases} m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \\ k(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]. \end{cases} \quad (10)$$

Considering the noise in the data acquisition, a general GPR model can be given as follows:

$$y = f(\mathbf{x}) + \varepsilon \quad (11)$$

where $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$ represents the independent and identically distributed Gaussian noise with a mean of 0 and a standard variance of σ_n .

For any finite number of the observation values, y is also an individual Gaussian process, which can be expressed as

$$y \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}') + \sigma_n^2 \delta_{ij}) \quad (12)$$

where δ_{ij} is the Kronecker delta function.

In general, GPR first learns a prior function using the training dataset \mathbf{X} and then predicts a posterior function value \hat{y}_* of the test dataset \mathbf{X}_* . Also, the joint distribution of the observed values y and predicted values \hat{y}_* can be expressed as

$$\begin{bmatrix} y \\ \hat{y}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n & K(\mathbf{X}, \mathbf{X}_*) \\ K(\mathbf{X}_*, \mathbf{X}) & K(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix}\right) \quad (13)$$

where $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, and \mathbf{I}_n represents an $n \times n$ identity matrix.

According to the Bayesian theory, the posterior distribution is calculated as

$$\hat{\mathbf{y}}_* | \mathbf{X}, \mathbf{y}, \mathbf{X}_* \sim \mathcal{N}(m(\hat{\mathbf{y}}_*), \text{cov}(\hat{\mathbf{y}}_*)) \quad (14)$$

where $m(\hat{\mathbf{y}}_*)$ and $\text{cov}(\hat{\mathbf{y}}_*)$ are the mean and variance of $\hat{\mathbf{y}}_*$ respectively, their mathematical expression is given as follows:

$$\begin{cases} m(\hat{\mathbf{y}}_*) = m(\mathbf{X}_*) + K(\mathbf{X}_*, \mathbf{X}) \\ \quad \times [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n]^{-1} (\mathbf{y} - m(\mathbf{X})) \\ \text{cov}(\hat{\mathbf{y}}_*) = K(\mathbf{X}_*, \mathbf{X}_*) - K(\mathbf{X}_*, \mathbf{X}) \\ \quad \times [K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}_n]^{-1} K(\mathbf{X}, \mathbf{X}_*). \end{cases} \quad (15)$$

In practice, the performance of GPR model is affected by the kernel function. The squared exponential function has been widely used in GPR modeling and obtained more satisfactory prediction results [17], [19]. Therefore, it is selected as the kernel function of GPR model in this study. Also, its expression is shown as follows:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / 2l^2) \quad (16)$$

where σ_s represents the standard deviation of the observation signal and l denotes the characteristic length scale.

It can be seen that $\boldsymbol{\theta} = \{l, \sigma_s, \sigma_n\}$ is a key hyperparameter in the GPR model. Generally, the GPR model is trained using the conjugate gradient algorithm, but it is susceptible to the initial value and is difficult to converge to the global optimum [74]. In order to address these limitations, this study adopts an optimization algorithm to solve the hyperparameters.

2) *GA-Based Parameter Optimization*: GA is a global optimization probabilistic search algorithm inspired by the natural selection and evolution mechanism [20]. Since the optimization process does not depend on the gradient, it has strong robustness and global search ability [21]. Thus, the hyperparameters of the GPR model are optimized using GA. As shown in Fig. 3, the main steps of the GA optimization are briefly explained as follows.

Step 3-1 (Objective Function Construction): In this study, the expression is defined as follows:

$$\min \|\hat{\mathbf{y}}_* - \mathbf{y}_*\| \quad \text{s.t.} \quad \boldsymbol{\theta}^d \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^u \quad (17)$$

where $\boldsymbol{\theta}^d$ and $\boldsymbol{\theta}^u$ are the upper and lower bounds of the search interval for $\boldsymbol{\theta}$, respectively.

Step 3-2 (Parameters Initialization): There are the maximum evolutionary epoch T , maximum population size M , individual length L , search interval $[\boldsymbol{\theta}^d, \boldsymbol{\theta}^u]$, and initial population $P(0)$. The t th epoch population $P(t)$ is defined as

$$P(t) = \{X | X(q) = a_1 a_2 \cdots a_L, q = 1, 2, \dots, M\} \quad (18)$$

where $X(q)$ represents the q th individual of the t th epoch population $P(t)$ and $a_1 a_2 \cdots a_L$ denotes the coded value of the individual $X(q)$.

Step 3-3 (Individual Evaluation): Calculate the fitness value of each individual according to the fitness function, which is the same as the objective function.

Step 3-4 (Selection Operation): Based on the fitness evaluation of individuals, the selection operation is to inherit the optimized individuals to the next generation.

Step 3-5 (Crossover Operation): The crossover operator reorganizes the optimized individuals and gradually abandons the relatively inferior ones.

Step 3-6 (Mutation Operation): The individual value is mutated by changing some gene values of individual coding in the population. After implementing selection, crossover, and mutation operations, the updated population $P(t+1)$ can be obtained.

Step 3-7 (Termination Condition Judgment): After completing T iterations, the individual with the best fitness value obtained in the process of evolution is taken as the optimal solution of hyperparameters in the GPR model.

Finally, the optimal parameters of the GPR model are obtained using GA optimization to establish a predictive model for the flank wear width of milling cutters.

F. Performance Evaluation and Wear Prediction

To validate the performance and effectiveness of the proposed tool wear prediction method, four evaluation criteria are utilized in this study, which includes Pearson correlation coefficient (PCC), mean absolute percentage error (MAPE), mean absolute error (MAE), and root-mean-squared error (RMSE) [34]. The expressions are given as follows:

$$\begin{cases} \text{PCC} = \frac{\sum_{i=1}^N (\mathbf{y}_{*i} - \bar{\mathbf{y}}_*) (\hat{\mathbf{y}}_{*i} - \bar{\hat{\mathbf{y}}}_*)}{\sqrt{\sum_{i=1}^N (\mathbf{y}_{*i} - \bar{\mathbf{y}}_*)^2 \sum_{i=1}^N (\hat{\mathbf{y}}_{*i} - \bar{\hat{\mathbf{y}}}_*)^2}} \\ \text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{\mathbf{y}}_{*i} - \mathbf{y}_{*i}|}{\mathbf{y}_{*i}} \\ \text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{\mathbf{y}}_{*i} - \mathbf{y}_{*i}| \\ \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{y}}_{*i} - \mathbf{y}_{*i})^2} \end{cases} \quad (19)$$

where \mathbf{y}_{*i} represents the flank wear width measured by an optical microscope and $\hat{\mathbf{y}}_{*i}$ denotes the predicted flank wear width using the proposed tool wear prediction method.

In addition, a 95% CI of predicted results from the GA-GPR model is given by $\bar{\hat{\mathbf{y}}}_* \pm 2 \times (\text{cov}(\hat{\mathbf{y}}_*))^{1/2}$. To quantitatively analyze the reliability and stability of the proposed tool wear prediction method, CI average width (CIAW) and CI standard deviation (CISD) are utilized [17], which are expressed as follows:

$$\begin{cases} \text{CIAW} = \frac{1}{N} \sum_{i=1}^N 4 \times \sqrt{\text{cov}(\hat{\mathbf{y}}_{*i})} \\ \text{CISD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(4 \times \sqrt{\text{cov}(\hat{\mathbf{y}}_{*i})} - \text{CIAW} \right)^2} \end{cases} \quad (20)$$

where $4 \times (\text{cov}(\hat{\mathbf{y}}_{*i}))^{1/2}$ is the width of the 95% CI at the testing sample \mathbf{x}_{*i} .

Among these evaluation criteria, the larger the PPC value and the less the MAE/RMSE/MAPE value, the higher the predicted accuracy. Besides, the lower the CIAW/CISD value,

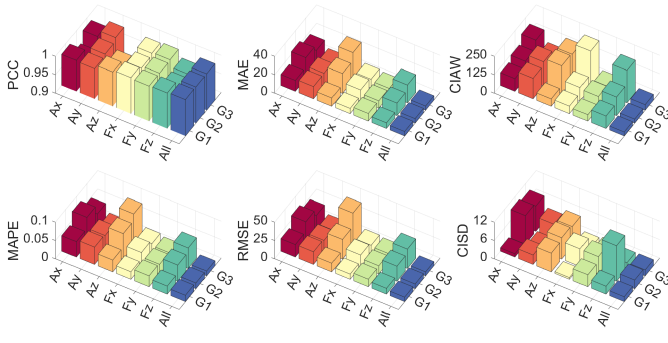


Fig. 6. Performance evaluation of tool wear prediction using different input signals.

TABLE V

OVERALL PERFORMANCE OF TOOL WEAR PREDICTION USING DIFFERENT SIGNALS

Signal	PCC	MAPE	MAE (um)	RMSE (um)	CIAW (um)	CISD (um)
Ax	0.9710±0.0541±	0.0151	16.8931±	22.9615±	124.7838±	6.2023±
Ay	0.9880±0.0390±	0.0059	11.9336±	14.5173±	133.0251±	3.0530±
Az	0.9550±0.0630±	0.0408	20.2237±	27.8209±	138.3177±	6.5005±
Fx	0.9935±0.0285±	0.0036	8.9421±	11.3360±	130.4226±	2.7420±
Fy	0.9955±0.0258±	0.0008	7.9504±	10.3227±	60.3874±	4.0131±
Fz	0.9830±0.0406±	0.0096	13.0284±	18.2628±	143.7790±	5.3649±
All*	0.9976±0.0179±	0.0007	5.4832±	6.8212±	37.7894±	2.7339±
All	0.9986±0.0003	0.0016	4.7295±	5.9631±	30.5832±	2.1118±

the more reliable and stable of prediction wear results. Finally, after implementing performance evaluation in prediction accuracy and reliability, the predictive tool wear prediction model with excellent performance is obtained, which is utilized to predict the flank wear width using in-process multisensor signals.

IV. RESULTS AND DISCUSSION

A. Multi-Information Fusion Analysis

1) *Multisensor Signal Fusion Analysis*: First, comparative experiments of the proposed tool wear prediction method using different signals are conducted to analyze the necessity of multisensor signal fusion. Also, evaluation results under different machining parameters are shown in Fig. 6, where “All” denotes a combination of vibration signals and cutting forces in x -, y -, and z -directions. In addition, overall performances of tool wear prediction using different signals are summarized in Table V, where “All*” represents no WPD-based signal denoising for multisensor signals. Note that quantitative results are given by “mean value \pm standard deviation.”

As shown in Fig. 6 and Table V, different signals are discrepant in the predicted results of tool flank wear. The evaluation criteria (including MAPE, MAE, RMSE, and CISD) of predicted wear adopting cutting forces are lower than that using vibration signals, which indicates that cutting forces are

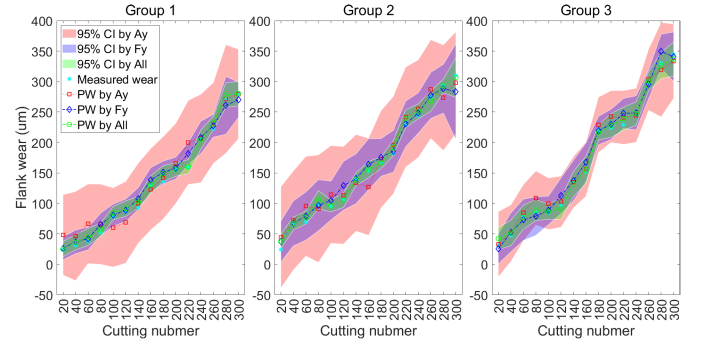


Fig. 7. Predicted results of tool flank wear using different input signals (95% CI).

more accurate in tool wear prediction than vibration signals. Besides, cutting forces or vibration signals in the y -direction obtain the lowest prediction error compared to other directions. However, all the single signals are significantly inferior to multisensor signal fusion in the performance of tool wear prediction. Finally, benefitting from the complementarity and fault tolerance of different signals, multisensor signal fusion lowers by 60.78%, 60.37%, 58.92%, 77.01%, and 30.83% of the mean of MAPE, MAE, RMSE, CIAW, and CISD compared to the best performance obtained by the vibration signals, respectively. Therefore, multisensor signal fusion is necessary and effective for tool wear prediction. In addition, predicted results under 95% CI of tool flank wear are shown in Fig. 7.

It can be seen from Fig. 7 that a 95% CI width of predicted results using vibration signals in the y -direction is significantly larger than that using cutting forces in the y -direction. Also, a 95% CI width of predicted results using multisensor signals is less than that using cutting forces in the y -direction. Besides, predicted wear using multisensor signals is closer to the measured wear than the single signals. Concretely, as shown in Table V, compared with the best performance of the force signals, predicted results of multisensor signals lower by 40.61%, 40.51%, and 42.23% in the mean value of MAPE, MAE, and RMSE. It also reduces by 49.36% and 47.38% in the mean value of CIAW and CISD, respectively. Therefore, these results display that multisensor signals are more accurate and reliable than the single signal in tool wear prediction.

In addition, as shown in Fig. 8, the WPD-based signal processing can overall lower the predicted error of tool flank wear when using multisensor signals. Correspondingly, predicted results under 95% CI of tool flank wear are drawn in Fig. 9. It can be found from Fig. 9 that a 95% CI width of predicted results without WPD is larger than that with WPD when using multisensor signals. Quantitatively, as shown in Table V, compared with no wavelet packet denoising (NWPD) for multisensor signals, predicted results based on WPD lower by 14.26%, 13.75%, 12.58%, 19.07%, and 22.75% in the mean value of MAPE, MAE, RMSE, CIAW, and CISD, respectively. Therefore, WPD-based signal processing can improve the accuracy and reliability of tool wear prediction using multisensor signals.

2) *Multidomain Feature Fusion Analysis*: Next, comparative experiments of tool wear prediction based on hybrid features extracted from different domains are carried out to

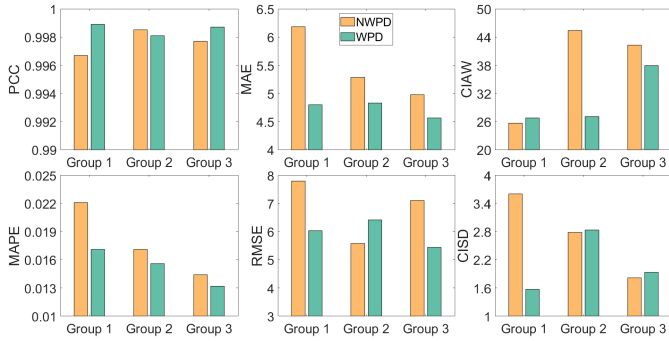


Fig. 8. Performance evaluation of tool wear prediction using WPD-based signal processing.

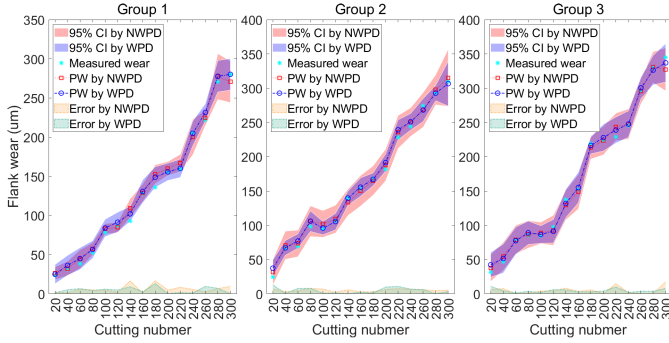


Fig. 9. Predicted results of tool flank wear using WPD-based signal processing (95% CI).

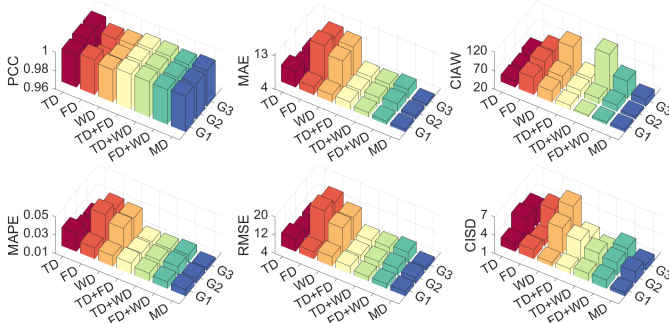


Fig. 10. Performance evaluation of tool wear prediction using different domain features.

analyze the necessity of multidomain feature fusion. In the experiments, all candidate features used in different domains are listed in Table IV. The evaluation results of these experiments are shown in Fig. 10. Besides, quantitative performances of tool wear prediction are given under different criteria in Table VI.

As shown in Fig. 10 and Table VI, a combination of two domains (i.e. TD + FD, TD + WD, and FD + WD) can lower the MAPE, MAE, and RMSE of predicted wear compared with the single domain (i.e., TD, FD, and WD), improving the prediction accuracy of tool flank wear. Also, it can lessen the CIAW and CISD of predicted wear, which enhances the reliability of tool wear prediction compared with the single domain feature. Moreover, fusion features from MD (i.e., TD + FD + WD) can further improve the accuracy and stability of predicted wear compared with combined features from the two domains. Also, a 95% CI width of predicted wear

TABLE VI
OVERALL PERFORMANCE OF TOOL WEAR PREDICTION
USING DIFFERENT DOMAIN FEATURES

Feature	PCC	MAPE	MAE (um)	RMSE (um)	CIAW (um)	CISD (um)
TD	0.9969±0.0010	0.0222±0.0056	6.7897±1.1488	7.9120±1.1682	38.2646±2.5558	3.4384±0.8333
	0.9910±0.0059	0.0342±0.0093	10.8032±3.3438	14.6135±5.3565	69.7716±3.0064	3.0977±1.2533
FD	0.9936±0.0013	0.0312±0.0070	9.8686±2.7186	12.4547±3.0202	68.1515±21.9956	4.4498±2.0855
	0.9982±0.0005	0.0190±0.0019	5.8854±0.1162	6.9984±0.3978	34.7776±2.3176	2.9525±1.0289
WD	0.9981±0.0005	0.0182±0.0007	5.6431±0.3001	6.8262±0.6635	59.8532±42.3778	2.5159±0.6900
	0.9972±0.0005	0.0187±0.0010	5.8389±0.7032	7.3115±1.0504	43.8706±17.6451	2.9478±0.3995
TD+FD	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292
	0.9982±0.0005	0.0190±0.0019	5.8854±0.1162	6.9984±0.3978	34.7776±2.3176	2.9525±1.0289
TD+WD	0.9981±0.0005	0.0182±0.0007	5.6431±0.3001	6.8262±0.6635	59.8532±42.3778	2.5159±0.6900
	0.9972±0.0005	0.0187±0.0010	5.8389±0.7032	7.3115±1.0504	43.8706±17.6451	2.9478±0.3995
FD+WD	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292
	0.9982±0.0005	0.0190±0.0019	5.8854±0.1162	6.9984±0.3978	34.7776±2.3176	2.9525±1.0289
MD	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292
	0.9982±0.0005	0.0190±0.0019	5.8854±0.1162	6.9984±0.3978	34.7776±2.3176	2.9525±1.0289

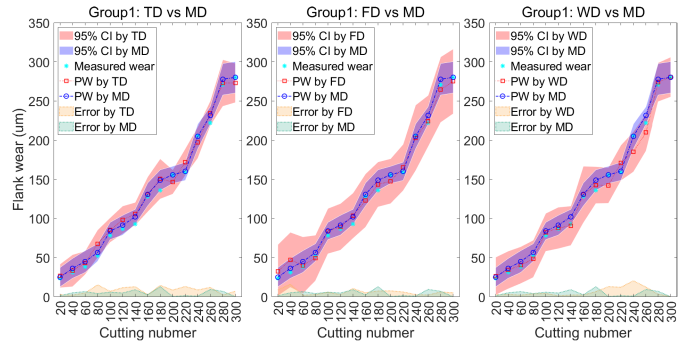


Fig. 11. Predicted results of tool flank wear using different domain features (95% CI).

using MD features is obviously less than the single domain features, which is shown in Fig. 11.

Specifically, compared with TD features that obtain the best performance in the single domain, predicted results using MD features lower by 31.06%, 30.34%, 24.63%, 20.07%, and 38.58% in the mean value of MAPE, MAE, RMSE, CIAW, and CISD, respectively. In addition, compared with combined features from two domains (i.e., TD + WD) that obtain the lowest error in two domains, predicted results by MD features lessen by 15.69%, 16.19%, and 12.64% in the mean value of MAPE, MAE, and RMSE, respectively. Also, it reduces by 12.06% and 28.47% in the mean value of CIAW and CISD, respectively, compared with combined features from two domains (i.e., TD + FD) that obtain the best reliability in the two domains. In conclusion, multidomain feature fusion not only can reduce the predicted wear of tool flank wear but also enhance the reliability and stability of tool wear prediction.

B. Comparisons With Other Methods

1) *Comparisons With Other Signal Fusion Methods:* In order to evaluate the effectiveness and advancement, the proposed tool wear prediction method is compared with other schemes in each critical process. First, performance

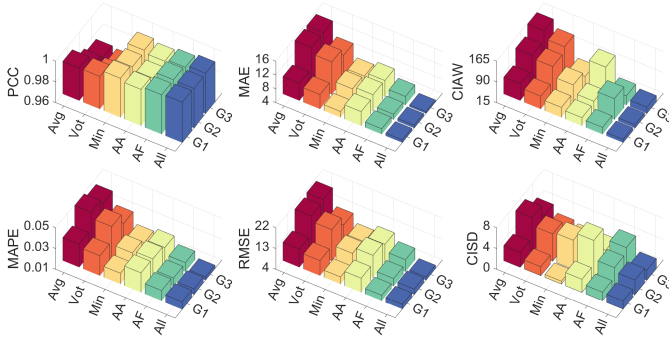


Fig. 12. Performance comparison of tool wear prediction between different signal fusion methods.

TABLE VII

OVERALL PERFORMANCE OF TOOL WEAR PREDICTION USING DIFFERENT SIGNAL FUSION METHODS

Method	PCC	MAPE	MAE (um)	RMSE (um)	CIAW (um)	CIRD (um)
Avg	0.9810±0.0076	0.0418±0.0076	13.1619±2.9206	17.5369±4.0204	121.7859±35.5963	4.6460±1.6070
Vot	0.9805±0.0080	0.0350±0.0060	10.9279±2.1405	13.8491±2.9206	104.1957±31.1963	3.2900±1.4647
Min	0.9964±0.0012	0.0227±0.0024	7.0966±1.0029	9.2764±1.6450	62.9299±12.2771	2.8952±2.0965
AA	0.9937±0.0013	0.0275±0.0017	8.5211±0.2993	12.0589±1.4734	72.5487±34.9949	4.1006±2.2055
AF	0.9976±0.0009	0.0189±0.0008	5.8850±0.4355	7.6557±1.1732	59.7143±22.3072	3.3027±1.2430
All	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292

comparisons between different signal fusion strategies are implemented to validate the effectiveness of the multisensor signal fusion scheme employed in the proposed method. The comparative results of these experiments are drawn in Fig. 12, where the Avg, Vot, and Min represent the average, voting, and minimum schemes of all the single input signals, respectively. Among them, the average scheme takes the same weight for different single signals, the voting scheme adopts different weights according to predicted error, and the minimum scheme utilizes the lowest predicted error of the single signal under each machining parameter. Besides, all accelerations (AA) denotes a fusion of all vibration acceleration signals in the x -, y -, and z -directions, and AF represents a fusion of all cutting forces in three directions. Correspondingly, the quantitative results are summarized in Table VII.

It can be found from Fig. 12 and Table VII that the predicted error of the voting scheme is less than the average strategy due to considering the sensitivity of different single signals. The minimum scheme can lower predicted error using the most sensitive signal in different machining parameters. When combining cutting forces in the x -, y -, and z -directions, the AF strategy can reduce predicted error compared with the average, voting, and minimum schemes of a single signal to a certain extent. However, its predicted accuracy and reliability are limited to the information limitation of single-type signal,

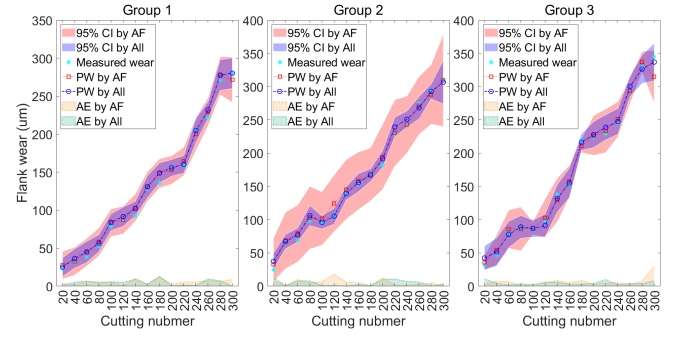


Fig. 13. Predicted results of tool flank wear using different signal fusion methods (95% CI).

TABLE VIII

OVERALL PERFORMANCE OF TOOL WEAR PREDICTION USING DIFFERENT FEATURE FUSION METHODS

Method	PCC	MAPE	MAE (um)	RMSE (um)	CIAW (um)	CIRD (um)
NFS	0.9985±0.0003	0.0177±0.0013	5.4740±0.2121	7.2224±0.8475	43.7768±16.5006	4.9008±2.0019
MRMR	0.9982±0.0004	0.0160±0.0014	4.9554±0.0344	6.0871±0.4093	50.6828±32.4517	3.6159±1.1806
LLE	0.9980±0.0003	0.0161±0.0017	4.9619±0.1980	6.3533±0.2185	51.0089±17.1694	2.9275±1.1368
KPCA	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292

and the predicted error is still less than the multisensor signal fusion used in the proposed method. Concretely, as shown in Table VII, the predicted wear of multisensor signal fusion is lower by 19.06%, 19.63%, and 22.11% in the mean value of MAPE, MAE, and RMSE, respectively. In addition, as shown in Fig. 13, a 95% CI width of predicted wear applying the all forces (AF) strategy is obviously larger than multisensor signal fusion. Overall, the predicted wear of multisensor signal fusion is lower by 48.78% and 36.06% in the mean value of CIAW and CIRD, respectively. Therefore, the multisensor signal fusion employed in the proposed method can improve the prediction accuracy and reliability compared with other signal fusion schemes.

2) *Comparisons With Other Feature Fusion Methods:* Then, to evaluate the effectiveness of multidomain feature fusion, the KPCA technique is compared with other feature fusion strategies, such as no feature selection (NFS) scheme, MRMR, and LLE. These methods are conducted on the presented GA-GPR prediction model using multisensor signals. Also, the performance comparison of these methods is shown in Fig. 14. Correspondingly, the overall results are quantitatively listed in Table VIII.

As shown in Fig. 14 and Table VIII, the feature selection strategies based on MRMR and LLE technique can achieve a lower predicted error (such as MPAE, MAE, and RMSE) than the NFS scheme that used all candidate multidomain features. However, these two techniques have a larger CIAW of the predicted wear compared with the NFS scheme, which reduces the reliability of tool wear prediction. Fortunately, the KPCA-based scheme can reduce a predicted error and improve the reliability of predicted results, which is shown

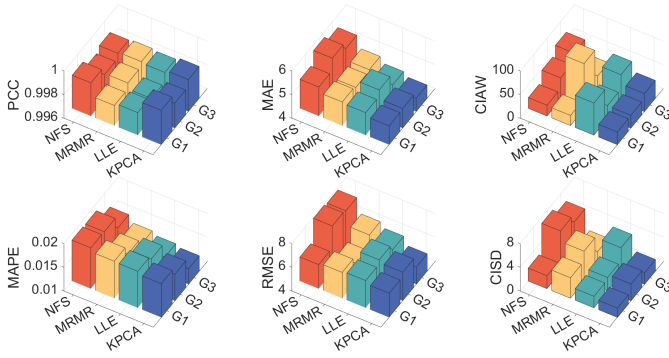


Fig. 14. Performance comparison of tool wear prediction between different feature fusion methods.

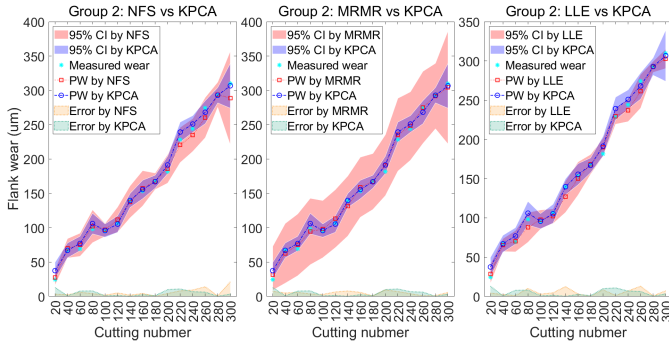


Fig. 15. Predicted results of tool flank wear using different feature fusion methods (95% CI).

in Table VIII and Fig. 15. Concretely, compared with NFS, MRMR, and LLE schemes, the KPCA-based strategy lowers by 13.60%, 4.56%, and 4.68% in the mean value of MAE, and it also reduces by 17.44%, 2.04%, and 6.14% in the mean value of RMSE. Furthermore, the KPCA-based scheme lowers by 30.14%, 39.66%, and 40.04% in the mean value of CIAW, and it reduces by 56.91%, 41.60%, and 27.86% in the mean value of CIRD compared with other feature selection schemes. Consequently, the KPCA-based multidomain feature fusion scheme can reduce a predicted error of tool flank wear and enhance the reliability of tool wear prediction.

3) *Comparisons With Other Prediction Methods:* Next, to demonstrate the effectiveness of the predictive model, the presented GA-GPR model is compared with other intelligent models, including MLR, BP, RBF, SVR, LSSVR, and RF. These models are also optimized by GA. Besides, all the compared models adopted the WPD-based signal preprocessing and KPCA-based feature fusion scheme same as the GA-GPR model. The performance evaluations under different machining parameters are shown in Fig. 16, and the overall performances under different evaluation criteria are quantitatively summarized in Table IX.

It can be found from Fig. 16 and Table IX that the GA-GPR model achieves the best performance of tool wear prediction among the compared methods. Compared with the GA-LSSVR model obtained the lowest error of other compared methods, the GA-GPR model lowers by 14.11%, 14.17%, and 11.38% in the mean value of MAPE, MAE, and RMSE, respectively, and it also reduces by 9.70%, 40.65%, and 18.17% in the standard

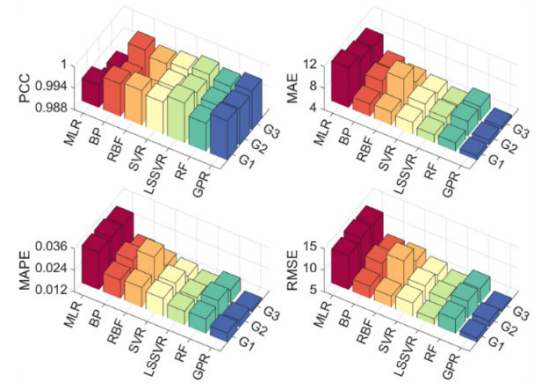


Fig. 16. Performance comparison of tool wear prediction between different prediction models.

TABLE IX
OVERALL PERFORMANCE OF TOOL WEAR PREDICTION
BETWEEN DIFFERENT PREDICTION MODELS

Model	PCC	MAPE	MAE (um)	RMSE (um)	CIAW (um)	CIRD (um)
MLR	0.9913±0.0020	0.0322±0.0007	10.3596±0.6964	13.1251±0.2508	226.4290±70.6067	3.5889±1.1321
BP	0.9963±0.0018	0.0225±0.0018	7.0078±0.5904	8.6730±0.7492	-	-
RBF	0.9954±0.0025	0.0242±0.0054	7.4980±1.5810	9.4679±1.9847	-	-
SVR	0.9969±0.0003	0.0202±0.0021	6.2448±0.3371	7.9576±0.4136	-	-
LSSVR	0.9984±0.0005	0.0178±0.0018	5.5106±0.1998	6.7287±0.4845	-	-
RF	0.9966±0.0007	0.0204±0.0010	6.3388±0.5185	8.5112±0.2856	-	-
GPR	0.9986±0.0003	0.0153±0.0016	4.7295±0.1186	5.9631±0.3965	30.5832±5.1792	2.1118±0.5292

variance of MAPE, MAE, and RMSE, respectively. Besides, the GA-GPR model can offer an uncertainty estimation to quantify the reliability of predicted results compared with BP, RBF, SVR, LSSVR, and RF models based on GA. Although the MLP model can provide a confidence interval of predicted results, the CIAW and CIRD values are significantly larger than the GPR model. Thus, the GA-GPR model can make tool wear prediction more accurate and reliable than other models, benefitting from a combination of the GA optimization and the GPR prediction model.

4) *Comparisons With Other Optimization Methods:* In addition, comparison experiments of predictive models under grid search (GS)- and GA-based optimizing schemes are conducted to validate the effectiveness of the model parameter optimization. Correspondingly, the overall performance of these optimized prediction models under different machining parameters is shown in Fig. 17.

It can be seen from Fig. 17 that the GA-optimized prediction models have a larger PCC and lower MAPE, MAE, and RMSE than GS-based models as a whole. Also, whether using GS- or GA-based model parameter optimization, GPR achieves higher correlation and lower error of tool wear prediction among other compared methods. Furthermore, as shown in Fig. 18,

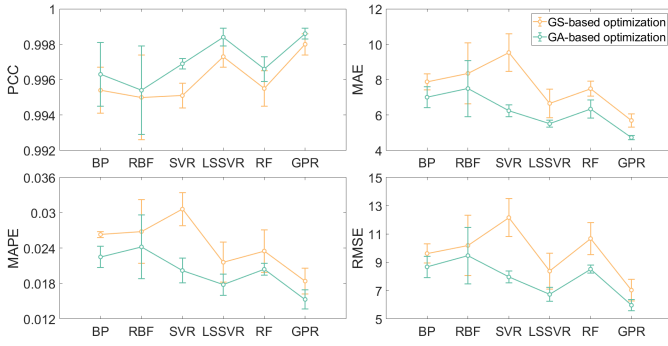


Fig. 17. Overall performance of tool wear prediction under different optimization methods.

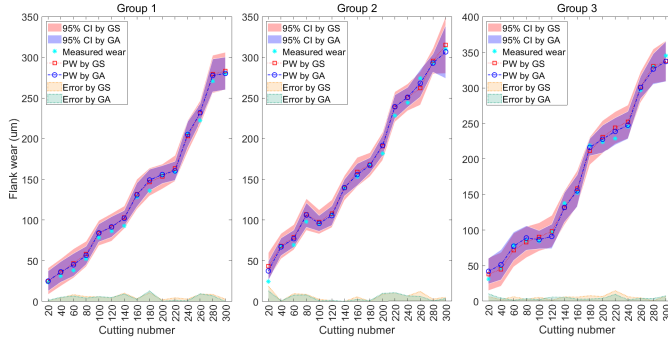


Fig. 18. Predicted results of tool flank wear using different optimization methods (95% CI).

a 95% CI width of predicted wear employing the GA-optimized GPR model is smaller than that using the GS-optimized GPR model.

The GA-GPR model reduces by 21.34% and 29.45% in the mean value of CIAW and CISD, respectively, which enhances the reliability of tool wear prediction compared with the GS-GPR model. Also, the GA-GPR model also lowers by 18.03%, 18.14%, and 14.54% in the mean value of MAPE, MAE, and RMSE, respectively, which improves the prediction accuracy of tool flank wear. Ultimately, the proposed tool wear prediction method achieves better performance than other methods in each critical process, including compared multi-sensor fusion schemes, multidomain feature fusion techniques, and intelligent prediction models.

C. Discussion

The indirect methods have been widely applied for tool wear monitoring integrated into intelligent manufacturing systems. When the training samples with labeled information are inadequate or the data distribution between training and testing samples is different, the monitoring accuracy of those data-driven indirect methods reduces significantly. In addition to the monitoring accuracy, the stability and reliability of indirect methods are also critical to integrated manufacturing systems. In order to improve the monitoring accuracy as well as provide the prediction reliability under small samples, this study investigates an indirect tool wear prediction method using multi-information fusion technologies and a GA-optimized GPR model synthetically.

TABLE X
PERFORMANCE COMPARISONS OF TOOL WEAR PREDICTION BETWEEN MAIN ABLATION METHODS

Methods	MAE (um)	CIAW (um)	CISD (um)	alpha-lambda	Optimizing time (s)
NWPD+KPCA	5.4832±	37.7894±	2.7339±	0.6889±	268.4948±
+GA-GPR	0.5128	8.6803	0.7318	0.0314	4.1025
WPD+NFS	5.4740±	43.7768±	4.9008±	0.6889±	280.3854±
+GA-GPR	0.2121	16.5006	2.0019	0.0629	3.9929
WPD+KPCA	5.7777±	38.8818±	2.6313±	0.7111±	968.7865±
+GS-GPR	0.2899	5.1291	0.6100	0.0629	59.0117
WPD+KPCA	4.7295±	30.5832±	2.1118±	0.8000±	269.6667±
+GA-GPR	0.1186	5.1792	0.5292	0.0544	1.5771

The main innovation of the developed tool wear monitoring method relies on the integration of the following technologies. First, the 3-D cutting forces and vibration signals provide multidimensional information for tool wear monitoring. Then, the WPD-based signal preprocessing can enhance the signal-to-noise ratio of the collected multisensor signals. Next, the multidomain features are extracted from the preprocessed signals to supply the comprehensive wear characteristics, and the KPCA technology is employed to fuse the principal features related to tool wear. Finally, the GPR-based prediction model can balance the monitoring accuracy and reliability, and the GA-optimized model can enhance the monitoring performance.

Intuitively, the main ablation methods are implemented to demonstrate the effectiveness of each technology integrated into the proposed method. In addition to the evaluation metrics, including MAE, CIAW, and CISD, the relevant research has shown that some advanced metrics such as alpha-lambda ($\alpha - \lambda$), prognostic horizon, and prediction distribution can be used to verify the accuracy and reliability of prognostic methodologies effectively [77]. Considering the consistency of these metrics displayed in the previous study [78], the $\alpha - \lambda$ metric is adopted to further highlight the contribution of each module. The $\alpha - \lambda$ metric is defined as follows:

$$\alpha - \lambda = \begin{cases} 1, & \text{if } (1 - \alpha)\lambda^* \leq \lambda_p \leq (1 + \alpha)\lambda^* \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where λ^* and λ_p denote the ground truth and the predicted value, respectively, and α represents an arbitrary error of the prediction model.

Also, the $\alpha - \lambda$ metric is averaged over the whole testing samples. The lower α and the larger $\alpha - \lambda$, the higher the accuracy and the better the reliability of the prediction method. To highlight the prediction performance of the proposed method, this study specifies 5% to be a chosen value for α . Besides, the optimizing time is also adopted to validate the ablation methods. The performance comparisons between the ablation methods are quantitatively summarized in Table X.

It can be found that compared with the ablation method (NWPD + KPCA + GA-GPR), the proposed method (WPD + KPCA + GA-GPR) can obviously reduce MAE, CIAW, and CISD and increase $\alpha - \lambda$ with almost no increment in optimizing time. The ablation method (WPD + NFS + GA-GPR) has higher MAE, CIAW, and CISD and lower

TABLE XI
COMPARISON FOR $\alpha - \lambda$ AND TIME CONSUMPTIONS
BETWEEN DIFFERENT PREDICTION MODELS

Methods	MAE (um)	alpha-lambda	Optimizing time (s)	Training time (s)	Testing time (s)
WPD+KPCA	10.3596±	0.2444±	-	0.2240±	0.0635±
+MLR	0.6964	0.1663	-	0.0655	0.0118
WPD+KPCA	7.0078±	0.6000±	533.3802±	0.5469±	0.0208±
+GA-BP	0.5904	0.0544	72.4053	0.0996	0.0074
WPD+KPCA	7.4980±	0.6222±	504.7083±	0.4427±	0.0365±
+GA-RBF	1.5810	0.1370	39.1026	0.1595	0.0074
WPD+KPCA	6.2448±	0.6444±	355.4948±	0.0781±	0.0208±
+GA-SVR	0.3371	0.1257	4.8246	0.0221	0.0074
WPD+KPCA	5.5106±	0.7556±	261.8594±	0.0990±	0.0208±
+GA-LSSVR	0.1998	0.0314	2.2186	0.0266	0.0074
WPD+KPCA	6.3388±	0.6889±	257.5052±	0.1583±	0.0260±
+GA-RF	0.5185	0.0831	2.7453	0.0829	0.0074
WPD+KPCA	4.7295±	0.8000±	269.6667±	0.1021±	0.0208±
+GA-GPR	0.1186	0.0544	1.5771	0.0517	0.0074

$\alpha - \lambda$, and increases the optimizing time than the proposed method. Moreover, compared with the ablation method (WPD + KPCA + GS-GPR), the proposed method can significantly reduce the optimizing time and prediction error and increase $\alpha - \lambda$. Therefore, the proposed method can balance the prediction accuracy and reliability of tool wear prediction, combined with the WPD-based multisensor signal processing, KPCA-based multidomain feature fusion, and GA-optimized GPR prediction model.

In addition, to further demonstrate the superiority of the proposed method, this study implements comprehensive comparisons of $\alpha - \lambda$ and time consumption with other advanced methods. Specifically, the comparative results are listed in Table XI. It is easily found that compared with other advanced methods, the proposed method obviously lowers the prediction error and improves the $\alpha - \lambda$ metric. Although the optimizing time of the proposed method is slightly more than the WPD + KPCA + GA-LSSVR method, their training and testing times are basically equal. Besides, the optimizing time of the WPD + KPCA + GA-RF method is slightly shorter than that of the proposed method, but the proposed method has an advantage in training and testing time. For the WPD + KPCA + GA-BP and WPD + KPCA + GA-RBF methods, the optimizing and training times are significantly longer than the proposed method. The WPD + KPCA + MLR method has a little time-consuming, but its prediction accuracy and reliability are not enough for effective tool wear prediction.

Consequently, when using multisensor signals for tool wear prediction, the proposed method (WPD + KPCA + GA-GPR) not only can prove higher prediction accuracy and reliability but also take less training and testing times, which has great utilization potentiality in online tool wear monitoring.

V. CONCLUSION

In this study, a tool wear prediction method based on multi-information fusion and GA-optimized GPR is proposed for indirectly measuring the flank wear width in milling. The multi-information fusion analysis and comprehensive comparisons with other advanced and ablation methods are carried out to verify the effectiveness of the proposed method under

different milling parameters. The main works are summarized as follows.

- 1) The proposed indirect method can improve the accuracy and reliability of tool wear prediction using multi-sensor signals, benefiting from the integration of the WPD-based signal preprocessing, the KPCA-based multidomain feature fusion, and the GA-optimized GPR prediction model.
- 2) The WPD-based signal preprocessing strategy lowers by 14.26%, 13.75%, 12.58%, 19.07%, and 22.75% of MAPE, MAE, RMSE, CIAW, and CISD than the NRPD-based approach without almost no increment in optimizing time, restraining the interference of environmental noise.
- 3) Compared with the NFS, MRM, and LLE schemes, the KPCA-based feature fusion method reduces MAE (13.60%, 4.56%, and 4.68%), RMSE (17.44%, 2.04%, and 6.14%), CIAW (30.14%, 39.66%, and 40.04%), and CISD (56.91%, 41.60%, and 27.86%), mining more sensitive characteristics of tool wear and thereby helping improve the prediction performance.
- 4) The GPR model surpasses the BP, RBF, SVR, LSSVR, and RF models in the accuracy of tool wear prediction and additionally provides an uncertainty estimation interval to quantify the reliability of the predicted results. Furthermore, the GA-GPR model obviously improves the prediction accuracy and reliability and significantly reduces the optimizing time compared with the GS-GPR model.

Comparison results demonstrate that the proposed tool wear prediction method achieves accurate and reliable monitoring of the flank wear width, promoting a successful application of indirect methods in milling operations. The study can also provide some beneficial suggestions for indirectly measuring tool flank wear in other machining operations (such as turning and drilling). For future research, the proposed method will be expanded to more processing parameters and machining operations, further enhancing its generalization ability. In addition, multi-information fusion methods based on other feature fusion technologies and intelligent prediction models will be developed to tool wear prediction, further improving the prediction accuracy and reliability.

REFERENCES

- [1] R. A. Mali, T. V. K. Gupta, and J. Ramkumar, "A comprehensive review of free-form surface milling—advances over a decade," *J. Manuf. Processes*, vol. 62, pp. 132–167, Feb. 2021, doi: 10.1016/j.jmapro.2020.12.014.
- [2] R. B. D. Pereira, L. C. Brandão, A. P. de Paiva, J. R. Ferreira, and J. P. Davim, "A review of helical milling process," *Int. J. Mach. Tools Manuf.*, vol. 120, pp. 27–48, Sep. 2017, doi: 10.1016/j.ijmactools.2017.05.002.
- [3] J. Ou, H. Li, G. Huang, B. Liu, and Z. Wang, "Tool wear recognition based on deep kernel autoencoder with multichannel signals fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021, doi: 10.1109/TIM.2021.3096283.
- [4] H. Guo, X. Lin, and K. Zhu, "Pyramid LSTM network for tool condition monitoring," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 2509511, doi: 10.1109/TIM.2022.3173278.
- [5] C. Liu, G. F. Wang, and Z. M. Li, "Incremental learning for online tool condition monitoring using ellipsoid ARTMAP network model," *Appl. Soft Comput.*, vol. 35, pp. 186–198, Oct. 2015, doi: 10.1016/j.asoc.2015.06.023.

- [6] D. Y. Pimenov, M. K. Gupta, L. R. R. da Silva, M. Kiran, N. Khanna, and G. M. Krolczyk, "Application of measurement systems in tool condition monitoring of milling: A review of measurement science approach," *Measurement*, vol. 199, Aug. 2022, Art. no. 111503, doi: [10.1016/j.measurement.2022.111503](https://doi.org/10.1016/j.measurement.2022.111503).
- [7] Y. Dai and K. Zhu, "A machine vision system for micro-milling tool condition monitoring," *Precis. Eng.*, vol. 52, pp. 183–191, Apr. 2018, doi: [10.1016/j.precisioneng.2017.12.006](https://doi.org/10.1016/j.precisioneng.2017.12.006).
- [8] X. Liang and Z. Liu, "Tool wear behaviors and corresponding machined surface topography during high-speed machining of Ti-6Al-4V with fine grain tools," *Tribol. Int.*, vol. 121, pp. 321–332, May 2018, doi: [10.1016/j.triboint.2018.01.057](https://doi.org/10.1016/j.triboint.2018.01.057).
- [9] K. Zhu, H. Guo, S. Li, and X. Lin, "Online tool wear monitoring by super-resolution based machine vision," *Comput. Ind.*, vol. 144, Jan. 2023, Art. no. 103782, doi: [10.1016/j.compind.2022.103782](https://doi.org/10.1016/j.compind.2022.103782).
- [10] D. Y. Pimenov, A. Bustillo, S. Wojciechowski, V. S. Sharma, M. K. Gupta, and M. Kuntoglu, "Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review," *J. Intell. Manuf.*, vol. 34, pp. 2079–2121, Mar. 2022, doi: [10.1007/s10845-022-01923-2](https://doi.org/10.1007/s10845-022-01923-2).
- [11] M. E. Korkmaz et al., "Indirect monitoring of machining characteristics via advanced sensor systems: A critical review," *Int. J. Adv. Manuf. Technol.*, vol. 120, nos. 11–12, pp. 7043–7078, Jun. 2022, doi: [10.1007/s00170-022-09286-x](https://doi.org/10.1007/s00170-022-09286-x).
- [12] K.-F. Zhang, H.-Q. Yuan, and P. Nie, "A method for tool condition monitoring based on sensor fusion," *J. Intell. Manuf.*, vol. 26, no. 5, pp. 1011–1026, Oct. 2015, doi: [10.1007/s10845-015-1112-y](https://doi.org/10.1007/s10845-015-1112-y).
- [13] J. A. Duro, J. A. Padget, C. R. Bowen, H. A. Kim, and A. Nassehi, "Multi-sensor data fusion framework for CNC machining monitoring," *Mech. Syst. Signal Process.*, vols. 66–67, pp. 505–520, Jan. 2016, doi: [10.1016/j.ymsp.2015.04.019](https://doi.org/10.1016/j.ymsp.2015.04.019).
- [14] Z. W. Huang, J. Shao, J. Zhu, W. Zhang, and X. Li, "Tool wear condition monitoring across machining processes based on feature transfer by deep adversarial domain confusion network," *J. Intell. Manuf.*, early access, doi: [10.1007/s10845-023-02088-2](https://doi.org/10.1007/s10845-023-02088-2).
- [15] M. S. Babu and T. B. Rao, "Multi-sensor heterogeneous data-based online tool health monitoring in milling of IN718 superalloy using OGM (1, N) model and SVM," *Measurement*, vol. 199, Aug. 2022, Art. no. 111501, doi: [10.1016/j.measurement.2022.111501](https://doi.org/10.1016/j.measurement.2022.111501).
- [16] Y. Cheng, X. Gai, Y. Jin, R. Guan, M. Lu, and Y. Ding, "A new method based on a WOA-optimized support vector machine to predict the tool wear," *Int. J. Adv. Manuf. Technol.*, vol. 121, nos. 9–10, pp. 6439–6452, Aug. 2022, doi: [10.1007/s00170-022-09746-4](https://doi.org/10.1007/s00170-022-09746-4).
- [17] D. Kong, Y. Chen, and N. Li, "Gaussian process regression for tool wear prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 556–574, May 2018, doi: [10.1016/j.ymsp.2017.11.021](https://doi.org/10.1016/j.ymsp.2017.11.021).
- [18] B. Qiang, K. Shi, N. Liu, P. Zhao, and J. Ren, "Application of cutting power consumption in tool condition monitoring and wear prediction based on Gaussian process regression under variable cutting parameters," *Int. J. Adv. Manuf. Technol.*, vol. 124, nos. 1–2, pp. 37–50, Jan. 2023, doi: [10.1007/s00170-022-10459-x](https://doi.org/10.1007/s00170-022-10459-x).
- [19] C. Zhang, W. Wang, and H. Li, "Tool wear prediction method based on symmetrized dot pattern and multi-covariance Gaussian process regression," *Measurement*, vol. 189, Feb. 2022, Art. no. 110466, doi: [10.1016/j.measurement.2021.110466](https://doi.org/10.1016/j.measurement.2021.110466).
- [20] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: Past, present, and future," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 8091–8126, Feb. 2021, doi: [10.1007/s11042-020-10139-6](https://doi.org/10.1007/s11042-020-10139-6).
- [21] K. Elbaz, S.-L. Shen, A. Zhou, Z.-Y. Yin, and H.-M. Lyu, "Prediction of disc cutter life during shield tunneling with AI via the incorporation of a genetic algorithm into a GMDH-type neural network," *Engineering*, vol. 7, no. 2, pp. 238–251, Feb. 2021, doi: [10.1016/j.eng.2020.02.016](https://doi.org/10.1016/j.eng.2020.02.016).
- [22] S. Y. Wong, J. H. Chuah, and H. J. Yap, "Technical data-driven tool condition monitoring challenges for CNC milling: A review," *Int. J. Adv. Manuf. Technol.*, vol. 107, nos. 11–12, pp. 4837–4857, Apr. 2020, doi: [10.1007/s00170-020-05303-z](https://doi.org/10.1007/s00170-020-05303-z).
- [23] Q. Zhu, B. Sun, Y. Zhou, W. Sun, and J. Xiang, "Sample augmentation for intelligent milling tool wear condition monitoring using numerical simulation and generative adversarial network," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021, doi: [10.1109/TIM.2021.3077995](https://doi.org/10.1109/TIM.2021.3077995).
- [24] S. Chiu et al., "Development of lightweight RBF-DRNN and automated framework for CNC tool-wear prediction," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022, doi: [10.1109/TIM.2022.3164063](https://doi.org/10.1109/TIM.2022.3164063).
- [25] C. H. R. Martins, P. R. Aguiar, A. Frech, and E. C. Bianchi, "Tool condition monitoring of single-point dresser using acoustic emission and neural networks models," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 3, pp. 667–679, Mar. 2014, doi: [10.1109/TIM.2013.2281576](https://doi.org/10.1109/TIM.2013.2281576).
- [26] X. Lin, B. Zhou, and L. Zhu, "Sequential spindle current-based tool condition monitoring with support vector classifier for milling process," *Int. J. Adv. Manuf. Technol.*, vol. 92, nos. 9–12, pp. 3319–3328, Oct. 2017, doi: [10.1007/s00170-017-0396-9](https://doi.org/10.1007/s00170-017-0396-9).
- [27] D. Axinte and N. Gindy, "Assessment of the effectiveness of a spindle power signal for tool condition monitoring in machining processes," *Int. J. Prod. Res.*, vol. 42, no. 13, pp. 2679–2691, Jul. 2004, doi: [10.1080/00207540410001671642](https://doi.org/10.1080/00207540410001671642).
- [28] Z. He, T. Shi, J. Xuan, and T. Li, "Research on tool wear prediction based on temperature signals and deep learning," *Wear*, vols. 478–479, Aug. 2021, Art. no. 203902, doi: [10.1016/j.wear.2021.203902](https://doi.org/10.1016/j.wear.2021.203902).
- [29] T. Mohanraj, S. Shankar, R. Rajasekar, N. R. Sakthivel, and A. Pramanik, "Tool condition monitoring techniques in milling process—A review," *J. Mater. Res. Technol.*, vol. 9, no. 1, pp. 1032–1042, Jan./Feb. 2020, doi: [10.1016/j.jmrt.2019.10.031](https://doi.org/10.1016/j.jmrt.2019.10.031).
- [30] Z. Huang, J. Shao, W. Guo, W. Li, J. Zhu, and D. Fang, "Hybrid machine learning-enabled multi-information fusion for indirect measurement of tool flank wear in milling," *Measurement*, vol. 206, Jan. 2023, Art. no. 112255, doi: [10.1016/j.measurement.2022.112255](https://doi.org/10.1016/j.measurement.2022.112255).
- [31] C. Liu, Y. Li, G. Zhou, and W. Shen, "A sensor fusion and support vector machine based approach for recognition of complex machining conditions," *J. Intell. Manuf.*, vol. 29, no. 8, pp. 1739–1752, Dec. 2018, doi: [10.1007/s10845-016-1209-y](https://doi.org/10.1007/s10845-016-1209-y).
- [32] M. Liu, X. Yao, J. Zhang, W. Chen, X. Jing, and K. Wang, "Multi-sensor data fusion for remaining useful life prediction of machining tools by IABC-BPNN in dry milling operations," *Sensors*, vol. 20, no. 17, p. 4657, Aug. 2020, doi: [10.3390/s20174657](https://doi.org/10.3390/s20174657).
- [33] W. Huang, X. Zhang, C. Wu, S. Cao, and Q. Zhou, "Tool wear prediction in ultrasonic vibration-assisted drilling of CFRP: A hybrid data-driven physics model-based framework," *Tribology Int.*, vol. 174, Oct. 2022, Art. no. 107755, doi: [10.1016/j.triboint.2022.107755](https://doi.org/10.1016/j.triboint.2022.107755).
- [34] 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](https://doi.org/10.1016/j.rcim.2016.05.010).
- [35] Z. Huang, J. Zhu, J. Lei, X. Li, and F. Tian, "Tool wear predicting based on multi-domain feature fusion by deep convolutional neural network in milling operations," *J. Intell. Manuf.*, vol. 31, no. 4, pp. 953–966, Apr. 2020, doi: [10.1007/s10845-019-01488-7](https://doi.org/10.1007/s10845-019-01488-7).
- [36] J. Wang, J. Xie, R. Zhao, K. Mao, and L. Zhang, "A new probabilistic kernel factor analysis for multisensory data fusion: Application to tool condition monitoring," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 11, pp. 2527–2537, Nov. 2016, doi: [10.1109/TIM.2016.2584238](https://doi.org/10.1109/TIM.2016.2584238).
- [37] P. Bhattacharyya and D. Sengupta, "Estimation of tool wear based on adaptive sensor fusion of force and power in face milling," *Int. J. Prod. Res.*, vol. 47, no. 3, pp. 817–833, Feb. 2009, doi: [10.1080/00207540701403376](https://doi.org/10.1080/00207540701403376).
- [38] I. Deiab, K. Assaleh, and F. Hammad, "On modeling of tool wear using sensor fusion and polynomial classifiers," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1719–1729, Jul. 2009, doi: [10.1016/j.ymsp.2009.02.001](https://doi.org/10.1016/j.ymsp.2009.02.001).
- [39] K. Gao, X. Xu, and S. Jiao, "Measurement and prediction of wear volume of the tool in nonlinear degradation process based on multi-sensor information fusion," *Eng. Failure Anal.*, vol. 136, Jun. 2022, Art. no. 106164, doi: [10.1016/j.engfailanal.2022.106164](https://doi.org/10.1016/j.engfailanal.2022.106164).
- [40] R. J. Kuo, "Multi-sensor integration for on-line tool wear estimation through artificial neural networks and fuzzy neural network," *Eng. Appl. Artif. Intell.*, vol. 13, no. 3, pp. 249–261, 2000, doi: [10.1016/S0952-1976\(00\)00008-7](https://doi.org/10.1016/S0952-1976(00)00008-7).
- [41] Z. He, T. Shi, and J. Xuan, "Milling tool wear prediction using multi-sensor feature fusion based on stacked sparse autoencoders," *Measurement*, vol. 190, Feb. 2022, Art. no. 110719, doi: [10.1016/j.measurement.2022.110719](https://doi.org/10.1016/j.measurement.2022.110719).
- [42] X. Liu, B. Zhang, X. Li, S. Liu, C. Yue, and S. Y. Liang, "An approach for tool wear prediction using customized DenseNet and GRU integrated model based on multi-sensor feature fusion," *J. Intell. Manuf.*, vol. 34, no. 9, pp. 885–902, 2022, doi: [10.1007/s10845-022-01954-9](https://doi.org/10.1007/s10845-022-01954-9).
- [43] X. Xu, Z. Tao, W. Ming, Q. An, and M. Chen, "Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion," *Measurement*, vol. 165, Dec. 2020, Art. no. 108086, doi: [10.1016/j.measurement.2020.108086](https://doi.org/10.1016/j.measurement.2020.108086).

- [44] X. Zhang, Y. Gao, Z. Guo, W. Zhang, J. Yin, and W. Zhao, "Physical model-based tool wear and breakage monitoring in milling process," *Mech. Syst. Signal Process.*, vol. 184, Feb. 2023, Art. no. 109641, doi: [10.1016/j.ymssp.2022.109641](https://doi.org/10.1016/j.ymssp.2022.109641).
- [45] P. S. Paul and A. Varadarajan, "A multi-sensor fusion model based on artificial neural network to predict tool wear during hard turning," *Proc. Inst. Mech. Eng. B, J. Eng. Manuf.*, vol. 226, no. 5, pp. 853–860, May 2012, doi: [10.1177/0954405411432381](https://doi.org/10.1177/0954405411432381).
- [46] M. Kuntoğlu and H. Sağlam, "Investigation of signal behaviors for sensor fusion with tool condition monitoring system in turning," *Measurement*, vol. 173, Mar. 2021, Art. no. 108582, doi: [10.1016/j.measurement.2020.108582](https://doi.org/10.1016/j.measurement.2020.108582).
- [47] N. Li et al., "Multi-sensor data-driven remaining useful life prediction of semi-observable systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 11, pp. 11482–11491, Nov. 2021, doi: [10.1109/TIE.2020.3038069](https://doi.org/10.1109/TIE.2020.3038069).
- [48] J. Wu, Y. Su, Y. Cheng, X. Shao, C. Deng, and C. Liu, "Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system," *Appl. Soft Comput.*, vol. 68, pp. 13–23, Jul. 2018, doi: [10.1016/j.asoc.2018.03.043](https://doi.org/10.1016/j.asoc.2018.03.043).
- [49] T. Feng, L. Guo, H. Gao, T. Chen, Y. Yu, and C. Li, "A new time-space attention mechanism driven multi-feature fusion method for tool wear monitoring," *Int. J. Adv. Manuf. Technol.*, vol. 120, nos. 7–8, pp. 5633–5648, Jun. 2022, doi: [10.1007/s00170-022-09032-3](https://doi.org/10.1007/s00170-022-09032-3).
- [50] V. Nasir, S. Dibaji, K. Alaswad, and J. Cool, "Tool wear monitoring by ensemble learning and sensor fusion using power, sound, vibration, and AE signals," *Manuf. Lett.*, vol. 30, pp. 32–38, Oct. 2021, doi: [10.1016/j.mfglet.2021.10.002](https://doi.org/10.1016/j.mfglet.2021.10.002).
- [51] M. Schwenzer, S. Stemmler, M. Ay, A. K. Rüppel, T. Bergs, and D. Abel, "Model predictive force control in milling based on an ensemble Kalman filter," *J. Intell. Manuf.*, vol. 33, no. 7, pp. 1907–1919, Oct. 2022, doi: [10.1007/s10845-022-01931-2](https://doi.org/10.1007/s10845-022-01931-2).
- [52] H. K. Tonshoff, X. Li, and C. Lapp, "Application of fast Haar transform and concurrent learning to tool-breakage detection in milling," *IEEE/ASME Trans. Mechatronics*, vol. 8, no. 3, pp. 414–418, Sep. 2003, doi: [10.1109/TMECH.2003.816830](https://doi.org/10.1109/TMECH.2003.816830).
- [53] B. Kilundu, P. Dehombreux, and X. Chiementin, "Tool wear monitoring by machine learning techniques and singular spectrum analysis," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 400–415, Jan. 2011, doi: [10.1016/j.ymssp.2010.07.014](https://doi.org/10.1016/j.ymssp.2010.07.014).
- [54] H. H. Shahabi and M. M. Ratnam, "On-line monitoring of tool wear in turning operation in the presence of tool misalignment," *Int. J. Adv. Manuf. Technol.*, vol. 38, nos. 7–8, pp. 718–727, Sep. 2008, doi: [10.1007/s00170-007-1119-4](https://doi.org/10.1007/s00170-007-1119-4).
- [55] C. Beale, C. Niezrecki, and M. Inalpolat, "An adaptive wavelet packet denoising algorithm for enhanced active acoustic damage detection from wind turbine blades," *Mech. Syst. Signal Process.*, vol. 142, Aug. 2020, Art. no. 106754, doi: [10.1016/j.ymssp.2020.106754](https://doi.org/10.1016/j.ymssp.2020.106754).
- [56] M. Cheng et al., "A novel multi-task learning model with PSAE network for simultaneous estimation of surface quality and tool wear in milling of nickel-based superalloy haynes 230," *Sensors*, vol. 22, no. 13, p. 4943, Jun. 2022, doi: [10.3390/s22134943](https://doi.org/10.3390/s22134943).
- [57] G. Song, J. Zhang, Y. Ge, K. Zhu, Z. Fu, and L. Yu, "Tool wear predicting based on weighted multi-kernel relevance vector machine and probabilistic kernel principal component analysis," *Int. J. Adv. Manuf. Technol.*, vol. 122, nos. 5–6, pp. 2625–2643, Sep. 2022, doi: [10.1007/s00170-022-09762-4](https://doi.org/10.1007/s00170-022-09762-4).
- [58] S. Ying, Y. Sun, C. Fu, L. Lin, and S. Zhang, "Grey wolf optimization based support vector machine model for tool wear recognition in fire-tree slot broaching of aircraft turbine discs," *J. Mech. Sci. Technol.*, vol. 36, no. 12, pp. 6261–6273, Dec. 2022, doi: [10.1007/s12206-022-1139-x](https://doi.org/10.1007/s12206-022-1139-x).
- [59] Z. Q. Li, P. Nie, and S. G. Zhao, "Identification method of tool wear based on locally linear embedding and support vector machine," *Appl. Mech. Mater.*, vols. 246–247, pp. 1289–1293, Dec. 2012, doi: [10.4028/www.scientific.net/AMM.246-247.1289](https://doi.org/10.4028/www.scientific.net/AMM.246-247.1289).
- [60] B. D. Deebak and F. Al-Turjman, "Digital-twin assisted: Fault diagnosis using deep transfer learning for machining tool condition," *Int. J. Intell. Syst.*, vol. 37, no. 12, pp. 10289–10316, Dec. 2022, doi: [10.1002/int.22493](https://doi.org/10.1002/int.22493).
- [61] C. Li, S. Zhang, Y. Qin, and E. Estupinan, "A systematic review of deep transfer learning for machinery fault diagnosis," *Neurocomputing*, vol. 407, pp. 121–135, Sep. 2020, doi: [10.1016/j.neucom.2020.04.045](https://doi.org/10.1016/j.neucom.2020.04.045).
- [62] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Health assessment and life prediction of cutting tools based on support vector regression," *J. Intell. Manuf.*, vol. 26, no. 2, pp. 213–223, Apr. 2015, doi: [10.1007/s10845-013-0774-6](https://doi.org/10.1007/s10845-013-0774-6).
- [63] D. Kong, Y. Chen, N. Li, and S. Tan, "Tool wear monitoring based on kernel principal component analysis and v -support vector regression," *Int. J. Adv. Manuf. Technol.*, vol. 89, nos. 1–4, pp. 175–190, Mar. 2017, doi: [10.1007/s00170-016-9070-x](https://doi.org/10.1007/s00170-016-9070-x).
- [64] D. Kong, Y. Chen, and N. Li, "Monitoring tool wear using wavelet package decomposition and a novel gravitational search algorithm–least square support vector machine model," *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, vol. 234, no. 3, pp. 822–836, Feb. 2020, doi: [10.1177/0954406219887318](https://doi.org/10.1177/0954406219887318).
- [65] Y. Xie, C. Zhang, and Q. Liu, "Tool wear status recognition and prediction model of milling cutter based on deep learning," *IEEE Access*, vol. 9, pp. 1616–1625, 2021, doi: [10.1109/ACCESS.2020.3047205](https://doi.org/10.1109/ACCESS.2020.3047205).
- [66] H. Zhang, C. Zhang, J. Zhang, and L. Zhou, "Tool wear model based on least squares support vector machines and Kalman filter," *Prod. Eng.*, vol. 8, nos. 1–2, pp. 101–109, Mar. 2014, doi: [10.1007/s11740-014-0527-1](https://doi.org/10.1007/s11740-014-0527-1).
- [67] P. Bhattacharyya, D. Sengupta, S. Mukhopadhyay, and A. B. Chattopadhyay, "On-line tool condition monitoring in face milling using current and power signals," *Int. J. Prod. Res.*, vol. 46, no. 4, pp. 1187–1201, Feb. 2008, doi: [10.1080/00207540600940288](https://doi.org/10.1080/00207540600940288).
- [68] T. Özel, Y. Karpat, L. Figueira, and J. P. Davim, "Modelling of surface finish and tool flank wear in turning of AISI D2 steel with ceramic wiper inserts," *J. Mater. Process. Technol.*, vol. 189, nos. 1–3, pp. 192–198, 2007, doi: [10.1016/j.jmatprotec.2007.01.021](https://doi.org/10.1016/j.jmatprotec.2007.01.021).
- [69] D. Wu, C. Jennings, J. Terpenney, R. X. Gao, and S. Kumara, "A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests," *J. Manuf. Sci. Eng.*, vol. 139, no. 7, Jul. 2017, Art. no. 071018, doi: [10.1115/1.4036350](https://doi.org/10.1115/1.4036350).
- [70] A. Bustillo, R. Reis, A. R. Machado, and D. Y. Pimenov, "Improving the accuracy of machine-learning models with data from machine test repetitions," *J. Intell. Manuf.*, vol. 33, no. 1, pp. 203–221, Jan. 2022, doi: [10.1007/s10845-020-01661-3](https://doi.org/10.1007/s10845-020-01661-3).
- [71] D. Li, Y. Li, and C. Liu, "Gaussian process regression model incorporated with tool wear mechanism," *Chin. J. Aeronaut.*, vol. 35, no. 10, pp. 393–400, Oct. 2022, doi: [10.1016/j.cja.2021.08.009](https://doi.org/10.1016/j.cja.2021.08.009).
- [72] K. Zhu and T. Liu, "Online tool wear monitoring via hidden semi-Markov model with dependent durations," *IEEE Trans. Ind. Informat.*, vol. 14, no. 1, pp. 69–78, Jan. 2018, doi: [10.1109/THI.2017.2723943](https://doi.org/10.1109/THI.2017.2723943).
- [73] G. Song, J. Zhang, K. Zhu, Y. Ge, L. Yu, and Z. Fu, "Tool wear monitoring based on multi-kernel Gaussian process regression and stacked multilayer denoising AutoEncoders," *Mech. Syst. Signal Process.*, vol. 186, Mar. 2023, Art. no. 109851, doi: [10.1016/j.ymssp.2022.109851](https://doi.org/10.1016/j.ymssp.2022.109851).
- [74] Z.-L. Ouyang and Z.-J. Zou, "Nonparametric modeling of ship maneuvering motion based on Gaussian process regression optimized by genetic algorithm," *Ocean Eng.*, vol. 238, Oct. 2021, Art. no. 109699, doi: [10.1016/j.oceaneng.2021.109699](https://doi.org/10.1016/j.oceaneng.2021.109699).
- [75] Z. Wang, Q. Zhang, J. Xiong, M. Xiao, G. Sun, and J. He, "Fault diagnosis of a rolling bearing using wavelet packet denoising and random forests," *IEEE Sensors J.*, vol. 17, no. 17, pp. 5581–5588, Sep. 2017, doi: [10.1109/JSEN.2017.2726011](https://doi.org/10.1109/JSEN.2017.2726011).
- [76] W. J. Lee, G. P. Mendis, M. J. Triebe, and J. W. Sutherland, "Monitoring of a machining process using kernel principal component analysis and kernel density estimation," *J. Intell. Manuf.*, vol. 31, no. 5, pp. 1175–1189, Jun. 2020, doi: [10.1007/s10845-019-01504-w](https://doi.org/10.1007/s10845-019-01504-w).
- [77] L. Biggio, A. Wieland, M. Arias Chao, I. Kastanis, and O. Fink, "Uncertainty-aware remaining useful life predictor," 2021, *arXiv:2104.03613*.
- [78] M. Bieber, V. Wim, and B. F. Santos, "The impact of metrics on the choice of prognostic methodologies," in *Proc. AIAA AVIATION Forum*, 2022, p. 3966, doi: [10.2514/6.2022-3966](https://doi.org/10.2514/6.2022-3966).



Zhiwen Huang (Member, IEEE) received the Ph.D. degree in mechanical engineering from the University of Shanghai for Science and Technology, Shanghai, China, in 2021.

He is currently an Assistant Research Fellow with the School of Mechanical Engineering, University of Shanghai for Science and Technology. His research interests include precision measurement techniques, condition monitoring, and intelligent control of the electromechanical system.



Jiajie Shao received the M.S. degree in mechanical manufacture and automation from the University of Shanghai for Science and Technology, Shanghai, China, in 2021. He is currently pursuing the Ph.D. degree in mechanical engineering with the School of Mechanical Engineering, Tongji University, Shanghai.

His research is currently focused on structural health monitoring, fault diagnosis, and transfer learning.



Weicheng Guo received the Ph.D. degree in mechanical engineering from Donghua University, Shanghai, China, in 2020.

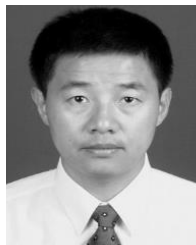
He was a Visiting Scholar with Purdue University, West Lafayette, IN, USA. He is currently a Lecturer with the School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai. His current research interests include precision machining of aerospace difficult-to-cut materials and intelligent manufacturing.



Weidong Li (Member, IEEE) received the Ph.D. degree in mechanical engineering from the National University of Singapore, Singapore, in 2002.

He is currently a Full Professor with Coventry University, Coventry, U.K., and the Dean of mechanical engineering with the University of Shanghai for Science and Technology, Shanghai, China. He has led a number of EU and U.K. projects in the areas of sustainable or digital manufacturing and cooperated with automotive, aeronautical, and manufacturing industries [e.g., Airbus, Jaguar Land Rover, Sandvik, and some manufacturing Small and Mid-sized Enterprise (SMEs)]. In the research areas, he has published over 120 research articles in peer-reviewed international journals and five books (Springer and World Scientific Publisher). His primary research interests include computer-aided manufacturing, sustainable manufacturing, and big data analytics for smart manufacturing.

Dr. Li is a fellow of Institution of Engineering and Technology (IET) and Institution of Mechanical Engineers (IMechE) U.K.



Jianmin Zhu received the Ph.D. degree in mechanical manufacture and automation from the Huazhong University of Science and Technology, Wuhan, Hubei, China, in 2000.

He is currently a Professor with the School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China. He has published over 120 articles, five books, and 30 patents. His research interests cover process monitoring and intelligent control of the electromechanical system, precision measurement techniques, and complex system modeling and control.

Prof. Zhu has won the Second Prize of National Science and Technology Progress Award, the Second Prize of the Ministry of Education Natural Science Award, and the Second Prize of Henan Province Science and Technology Progress Award.



Qichao He received the B.S. degree in mechanical engineering from the East China University of Science and Technology, Shanghai, China, in 2021.

He is currently working as an Engineer with Shanghai Aerospace Equipment Manufacturing General Company Ltd., Shanghai. His current research interests include precision machining and process optimization.



Dianjun Fang received the Ph.D. degree in mechanical engineering from the Technical University of Dortmund, Dortmund, Germany, in 1995.

He is currently a Professor with the School of Mechanical Engineering, Tongji University, Shanghai, China, and the Chinese Director of the Qingdao Sino-German Institute of Intelligent Technologies, Qingdao, China. He is the Honorary Chief Scientist in China with the Fraunhofer Institute for Material Flow and Logistics, Dortmund. His current research interests include structural health monitoring, digital twins, and intelligent logistics system.