



Predictive model of surface roughness in milling of 7075Al based on chatter stability analysis and back propagation neural network

Long Bai¹ · Xin Cheng¹ · Qizhong Yang¹ · Jianfeng Xu^{1,2}

Received: 27 June 2022 / Accepted: 16 February 2023 / Published online: 10 March 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

Abstract

Accurate prediction of the machining quality such as the surface roughness is one of the main objectives of the intelligent manufacturing research. In this study, we investigate the feasibility of combining milling stability analysis and a back propagation (BP) neural network model to predict the surface roughness of aerospace aluminum alloy 7075Al in high-speed precision milling. The difference between the critical depth of cut obtained from the milling stability lobe diagram (SLD) and the actual depth of cut (termed the “chatter stability feature”) was used as a critical input variable of the neural network model, thereby improving its prediction performance. It is demonstrated that the proposed chatter stability feature has a strong correlation with the surface roughness, making it an information-rich input feature of the prediction model. The experimental results show that the prediction accuracy can be improved by 7.8% compared with a neural network model that only uses the cutting parameters (i.e., spindle speed, feed rate, depth of cut, and feed rate per tooth) as predictors. In addition, univariate and multivariate sensitivity analysis results suggest that the performance of the proposed approach is robust to errors in the SLD measurements. Compared to conventional methods which only consider the cutting parameters, an improvement in prediction accuracy can be expected with up to 10% errors in modal parameters.

Keywords Precision milling · Surface roughness · Chatter stability analysis · SLD · BP neural network

1 Introduction

High-speed milling and precision milling are widely used in manufacturing to satisfy the increasing requirements imposed on machined parts. Surface roughness is an important measure of the surface finish quality and has

a significant impact on different aspects of a machined product, including wearing, surface friction, and the ability to resist fatigue and heat transmission [1]. 7075 aluminum is widely used in the aerospace, transportation, and machinery manufacturing industries due to their low density, high specific strength, and good workability [2]. Due to the importance of the surface roughness of the machined parts, its prediction is an active research topic in the intelligent manufacturing community. There have been lots of analytical methods to evaluate the surface roughness of milling. For example, Quintana et al. [3] established a theoretical surface roughness model for ball end milling. They studied the influence of the geometric characteristics of ball end mill cutter on the surface roughness generation and analyzed the effect of the machining parameters on material removal rate. Bachrathy et al. [4] predicted surface errors based on the motion of the helical tool, in which they considered the influence of the tool parameters. However, since it is difficult to consider all the relevant factors affecting the surface roughness in a single physical model, researchers also explored the use of deep learning and neural network models for improving the prediction accuracy. Karayel [5] utilized the cutting speed, feed rate,

✉ Jianfeng Xu
jfxu@hust.edu.cn

Long Bai
bailong@hust.edu.cn

Xin Cheng
m202070537@hust.edu.cn

Qizhong Yang
m202070502@hust.edu.cn

¹ State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China

² Wuhan Digital Design and Manufacturing Innovation Center Co., Ltd, Wuhan, 430074, China

and depth of cut to train an artificial neural network (ANN) model and used it to control the surface roughness of the workpiece. Asiltürk et al. [6] combined the ANN and multiple regression methods to predict the surface roughness of AISI1040 under different combinations of cutting parameters. Bajić et al. [7] established a surface roughness model and optimized the cutting parameters for face milling. Groove [8] studied the effects of three factors—feed rate, tool nose radius, and cutting edge angle—on surface roughness. Lela et al. [9] adopted regression analysis, support vector machine, and Bayesian neural network to analyze the effects of cutting parameters on the surface roughness in end milling. As observed, the above research mainly focused on the use of static features, such as cutting parameters and cutter information, for predicting the roughness of the machined surface.

The use of features extracted from dynamic process data for predicting the machining quality of products can also be found in the literature. Abu-Mahfouz et al. [10] used the summary statistics of the vibration signal, such as the mean, standard deviation, kurtosis, and skewness, to predict the surface roughness of a workpiece in end milling. Tangjitsitcharoen et al. [11] developed a practical model to realize in-process prediction of surface roughness in ball-end milling by utilizing the dynamic cutting force ratio and processing parameters. Lou et al. [12] proposed a statistical model for predicting surface roughness using a neuro-fuzzy system and adopted an accelerometer to measure vibrations induced by cutting forces during milling operations. As a classic deep learning architecture, BP neural networks were shown previously to achieve satisfactory performance on surface texture prediction. Pan et al. [13] adopted BP neural networks to predict the surface roughness during grinding and studied the effect of different activation functions on the prediction accuracy. Markopoulos et al. [14] utilized BP and RBF neural networks with different training algorithms to study the prediction accuracy of the surface roughness in end-milling. Kao et al. [15] adopted multivariate linear regression and generalized regression neural networks to establish the relationship between the cutting forces and surface roughness.

The roughness prediction accuracy can be improved by considering real-time process data such as force and vibration. However, in an actual milling process, installing many sensors for online monitoring can be problematic considering the practical limitations of the machine tool and workpiece. In addition, training a model using a large amount of data collected during the milling process is time-consuming and prone to overfitting. *On the other hand*, the completeness of information contained in the training data is critical to *the* prediction accuracy when building a predictive model. In the absence of dynamic sensor signals, the commonly used cutting parameters and cutter information

do not account for the structural stiffness characteristics of the machine tool or the contact characteristics between the cutter and workpiece. *For this reason*, researchers have also shown significant interest in studying milling dynamics and milling stability lobe diagram (SLD). Quintana et al. [16] divided vibrations in machining into free, forced, and self-excited vibrations and proved that strong vibrations that occur in machining processes can seriously affect the machining accuracy and surface quality. Özşahin et al. [17] proposed a new method using revised bearing parameters to determine the tool point frequency response function (FRF) and used the accurate stability diagrams to avoid chatter vibrations. Wang et al. [18] investigated the harmful effects of self-excited regenerative chatter on workpieces and machine tool structures by studying the interaction between machine tool and thin-walled workpieces. Altintas et al. [19] established a classic linear milling dynamics model based on the regeneration effect and solved the milling SLD using a zero-order approximation method. Modal impact testing is a typical approach for obtaining the FRF of a tool-nose point [20], and Nguyen et al. [21] proposed a statistical approach to predict the pose-dependent FRFs for robotic milling.

Although current research on milling dynamics has covered many aspects to meet different practical needs, the quantitative relationship between the chatter stability and milling quality has not been fully investigated (particularly, when chatter has not yet occurred). According to the theory of the SLD and milling dynamics, the critical depth of cut (for a given spindle speed) is related to the modal parameters and the milling force coefficients, which define the dynamic characteristics of the milling system. In this study, we improved the vibration measurement in a modal experiment for high-speed precision milling to obtain more accurate FRFs. Based on this result, we introduce a new feature calculated from the SLD that considers the physical properties of the entire milling system, and hypothesize that the roughness prediction accuracy can be improved by using this feature. Even for practical industrial settings in which process data such as cutting forces and vibration signals are unavailable, the newly introduced chatter stability feature can still be used in combination with a simple back propagation (BP) neural network for fast online prediction of the roughness. The remainder of this paper is organized as follows. Section 2 introduces SLD and BP neural networks. The experimental work performed to acquire the data needed for the SLD calculations and the data used as the training and test datasets of the BP neural network are described in detail in Section 3. In Section 4, we compare the performance of the different methods on experimental data and demonstrate that the prediction accuracy can be improved by adopting the proposed approach. Finally, we conclude the paper in Section 5.

2 Methodology

2.1 Milling chatter stability analysis

As shown in Fig. 1, the milling cutter can be modelled as a linear system with two orthogonal degrees of freedom. Based on this model, the average cutting force per tooth cycle can be obtained as [19]:

$$\bar{F}_x = -\frac{Na_p}{4}K_{rc}f_z - \frac{Na_p}{\pi}K_{re}, \tag{1}$$

$$\bar{F}_y = \frac{Na_p}{4}K_{tc}f_z + \frac{Na_p}{\pi}K_{te}, \tag{2}$$

where N is the number of teeth of the milling cutter, a_p is the depth of cut, and $f_z = f/nN$ (n is the spindle speed) is the feed rate per tooth. The tangential (K_{tc} , K_{te}) and radial (K_{rc} , K_{re}) milling force coefficients (denoted with the subscript “c”) and edge force coefficients (denoted with the subscript “e”) can be calculated from Eqs. (1) to (2) by performing milling experiments with different combinations of cutting parameters and measuring the average milling force in each experiment.

Modal impact tests are often conducted to analyze the structural behavior of machining systems under specific installation conditions. The FRF of the structure can be measured using a modal impact test from which the modal parameters are identified, including the natural frequency, damping ratio, stiffness, and modal shape. A schematic of the hammer vibration test is shown in Fig. 2.

According to the chatter stability theory proposed by Altintas et al. [19], the critical depth of cut and the corresponding spindle speed can be calculated using the

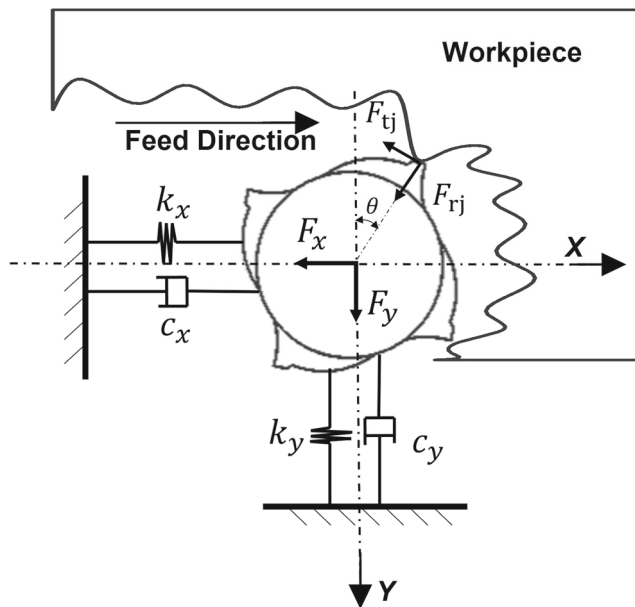


Fig. 1 Two-degree-of-freedom milling dynamics model

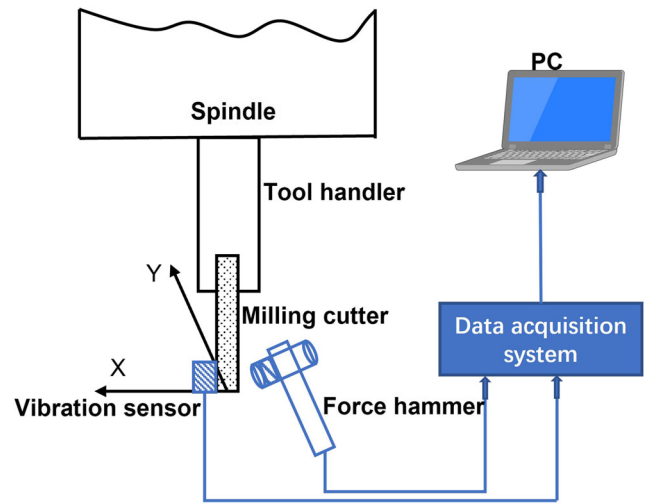


Fig. 2 Schematic of the modal impact test

structural parameters of the machine tool-cutter system and milling force coefficients as follows:

$$a_{plim} = -\frac{2\pi\Lambda_R}{NK_{tc}} \left[1 + \left(\frac{\Lambda_I}{\Lambda_R} \right)^2 \right], \tag{3}$$

where Λ_I and Λ_R are the imaginary and real parts, respectively, of the eigenvalues of the characteristic equation of the machine tool-cutter system. For a given chatter frequency f_c , the corresponding spindle speed n (in r/min) can be calculated by determining the tooth period T (in s), that is,

$$n = 60/(NT), \tag{4}$$

where

$$T = \frac{1}{f_c}(\pi - 2\theta + 2k\pi). \tag{5}$$

In Eq. (5), k is the number of lobes and θ is the rotation angle of the milling cutter.

The milling force model and dynamic milling model can be established given that the transfer function of the machine tool-cutter system, cutting parameters, and milling force coefficients are known. The critical depth of cut and the corresponding spindle speed can then be calculated following an analysis of the chatter stability, as discussed above.

2.2 BP neural network model

The BP neural network is a feedforward neural network that has a multilayer structure and works by an error backpropagation algorithm [22]. Owing to its simplicity, fast convergence speed, and good prediction performance when dealing with nonlinear problems, it is widely used in the field of intelligent manufacturing [23]. As shown in

Fig. 3, the BP neural network algorithm mainly includes two processes: forward propagation and error back-propagation. Considering the three-layer BP neural network in Fig. 3 as an example, the forward propagation procedure can be expressed as:

$$H_j = f_1 \left(\sum_{i=1}^n w_{ij} X_i - b_h^{(j)} \right), \quad j = 1, 2, \dots, l, \quad (6)$$

$$Y_k = f_2 \left(\sum_{j=1}^l w_{jk} H_j - b_o^{(m)} \right), \quad k = 1, 2, \dots, m. \quad (7)$$

where X_i represents the input data, and H_j and Y_k are the outputs of the hidden layer and the final predicted value, respectively. w_{ij} and w_{jk} are the weights between the input and hidden layers and the hidden and output layers, respectively, and $b_h^{(j)}$ and $b_o^{(m)}$ are the corresponding biases. f_1 and f_2 are the activation functions of the hidden and output layers, respectively.

The prediction error can be calculated as $E_m = \frac{1}{2} \sum_{k=1}^m (Y_k - T_k)^2$ (T_k is the true value), which can then be used to update the network weights and biases (i.e., back-propagation) as follows:

$$w_{ij}^+ = w_{ij} - \eta \frac{\partial E_m}{\partial w_{ij}}, \quad (8)$$

$$b_h^{(j)+} = b_h^{(j)} - \eta \frac{\partial E_m}{\partial b_h^{(j)}}. \quad (9)$$

In Eqs. (8)–(9), w_{ij}^+ and $b_h^{(j)+}$ are the weights and biases, respectively, after the update. $\frac{\partial E_m}{\partial w_{ij}}$ and $\frac{\partial E_m}{\partial b_h^{(j)}}$ are the error gradients against the weights and biases, respectively, and η is the learning rate.

Table 1 Specifications of the tungsten carbide milling cutter used in experiments

Parameters	Value
Radius of cutter blade	4 mm
Length of cutter blade	12 mm
Total length of cutter	50 mm
Radius of cutter handle	4 mm
Number of cutter blade	4
Helix angle	45°

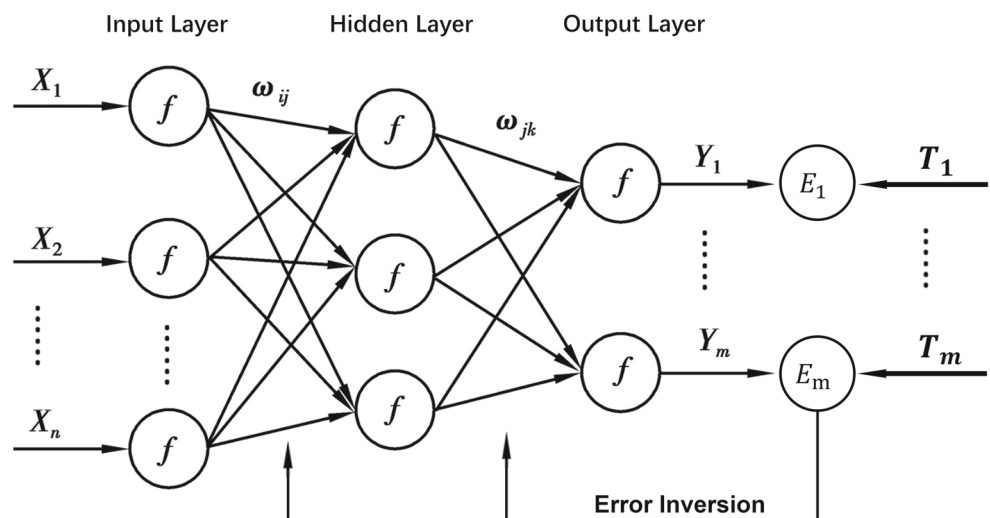
3 Experiments and data collection

3.1 Modal impact test

In the modal impact test performed in this study, a cylindrical flat-bottom milling cutter was installed on the tool holder, and the tool overhang length was set to 20 mm. The cutter parameters are listed in Table 1. Owing to the small size of the cutter used in the experiments, vibration sensors could not be installed at the cutter tip. Note that installing a vibration sensor at the tip of the cutter using an additional device would also affect the structural properties of the original machine tool-cutter system. Considering that the rigidity of the cutter-toolholder system is smaller than that of the spindle and the vibration is more intense, the vibration sensor that should have been installed at the cutter tip point is often installed on the plane where the tool holder and spindle are connected.

To address this issue, two vibration acceleration sensors (DYTRAN 3263A2, USA) were arranged in the same plane, where the cutter-toolholder combination and spindle were connected in an orthogonal direction. The input force and output vibration signals were measured and transmitted to the PC using a vibration and noise testing system

Fig. 3 The three-layer BP neural network structure



(LMS Test Lab & LMS SCADAS III, SIMENS, Germany). Figure 4 shows the experimental configuration of the modal impact test. Two contact accelerometers were mounted in the X- and Y-axis directions, respectively, and they were attached to the milling cutter shank with glue. In each hammering experiment, the hammer struck five times along the direction in which the vibration accelerometer was installed (X or Y), and the FRF of the corresponding direction was obtained.

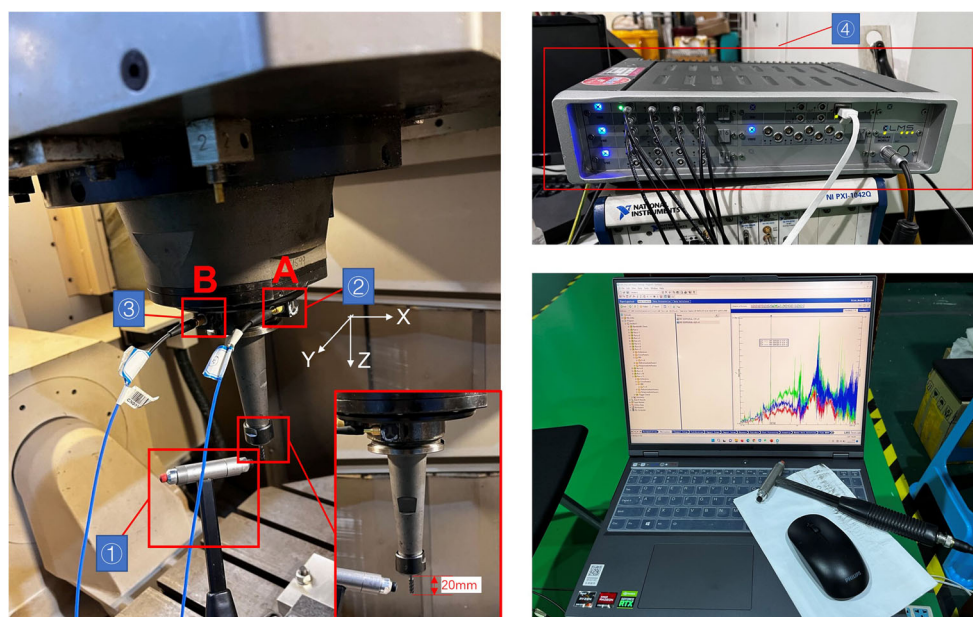
Figure 5a and 5b show the FRFs in the X- and Y-directions of the cutter tip point measured in one test. The modal parameters of the cutter-toolholder-spindle system were calculated by averaging the FRFs obtained from three separate experiments. Because the toolholder system is a revolving body, its structure in the X- and Y-directions is symmetrical, and the obtained FRFs shown in Fig. 5 can also reflect this. The main frequencies corresponding to the positions where the amplitude peaks appeared were also shown to be similar in these results. The natural frequency and damping ratio of the cutter-toolholder system in the main mode can be obtained based on the FRF results. The modal parameters obtained at two response points A and B (see Fig. 4) in each of the three hammering experiments in the X- and Y-directions are summarized in Table 2.

Considering that the above method adopted a compromised vibration sensor arrangement owing to the size limitation of the milling cutter, the obtained structural parameters are characteristics of the entire system, which consists of the tool nose point and the connection between the toolholder and the spindle. However, only the vibration response of the tool nose point is needed for chatter stability analysis, and in this study, a pair of Doppler optical vibrometers (PNV-RD-AVDI, HoloBright, Singapore) was used for its

measurement. As shown in Fig. 6, the contact vibration sensors were replaced by two non-contact Doppler vibrometers, which can avoid the size constraint of the milling cutter. The two Doppler vibrometers were placed in the X and Y directions, respectively. The distances between the vibrometers and the milling cutter were approximately 40 cm, and focus was achieved at the milling cutter tip by adjusting the focal length. Based on the experimental configuration shown in Fig. 6, the modal parameters can be obtained by following the procedure described in Section 2.

Because the laser vibrometer can only measure the vibration in one direction, the modal parameters of the cutter tip were obtained in separate experiments for the X- and Y-directions. As shown in Fig. 7, the FRFs in these two directions of the cutter tip are again similar, and the main *frequencies* of the *responses* can be *identified from the mode shapes*. By adopting this improved experimental design, the difficulty in installing the sensor on the small tool tip and the need to analyze the complex coupling behavior of the cutter nose point and vibration sensor can be addressed. The stiffness of the milling cutter was calculated based on the cantilever beam theory [24], and the obtained modal parameters are listed in Table 3. By comparing the results of Tables 2 and 3, it can be seen that there is a considerable difference between the modal parameters (especially the main frequencies) measured using contact accelerometers and non-contact vibrometers. Because the contact accelerometers were attached on the tool-holder (instead of the small tool tip) due to its size constraint, the results of the contact measurements are believed to be less accurate. As will be discussed in Section 4, the prediction accuracy can be improved by adopting SLDs that are more accurate.

Fig. 4 Modal impact test configuration. (① Hammer with force sensor; ② and ③: Vibration acceleration sensors in the X- and Y-directions; ④ LMS equipment.)



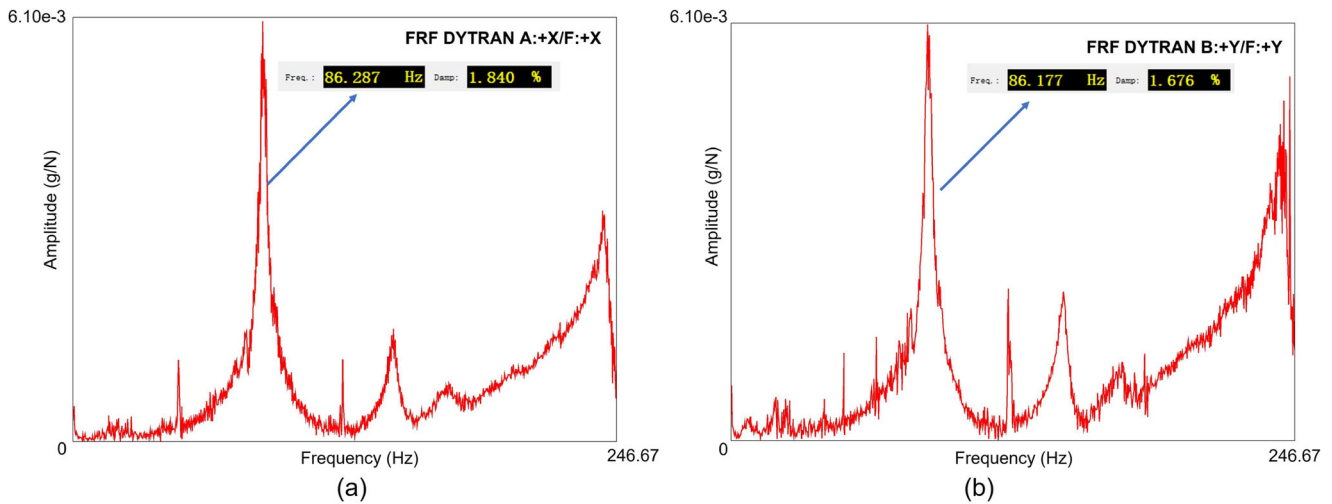


Fig. 5 FRFs of the cutter-toolholder system. (a X-direction; b Y-direction)

3.2 Milling dynamics analysis

It is necessary to collect milling force data to calculate the milling force coefficients using Eqs. (1)–(2). We performed trial cutting experiments under two sets of different cutting parameters, and the cutting forces shown in Fig. 8 were measured using a dynamometer (Kistler 9129AA, Switzerland) during these experiments. The cutting parameters can be arbitrarily selected in trial cutting experiments without considering the process conditions of the actual parts. The cutting parameters and milling force coefficient results are listed in Table 4. The milling SLDs of the cutter-toolholder and milling cutter tip point can be obtained by combining the milling force coefficients and modal parameters of the machine tool-cutter system based on the analysis in Section 2.1, and the results are shown in Fig. 9.

After obtaining the SLD, the difference between the depth of cut a_p adopted in an actual milling experiment and the critical depth of cut a_{plim} at the corresponding spindle

speed is calculated as a new input feature of the neural network model as follows:

$$\Delta a = a_{plim} - a_p. \quad (10)$$

Based on the physical interpretation of the milling SLD, it can be concluded that as the actual depth of cut a_p approaches the limit depth of cut a_{plim} , the milling status changes from “stable” to “chatter.” While it is well understood that chatter vibration qualitatively affects the machining quality, we hypothesized in this study that Δa also has a quantitative impact on roughness even for cases in which chatter has not actually occurred. As will be discussed below, this chatter stability feature can effectively improve the prediction accuracy of neural network models. The important implications of the newly introduced predictor of roughness can be summarized as follows. First, the performance of the predictive model can be improved compared to traditional approaches which only make use of the cutting parameters. Second, Δa can still be

Table 2 Cutter-toolholder modal parameters (main frequency f_c and damping ratio ξ) obtained in three hammering experiments in the X- and Y-directions

Direction	Test No.	f_c (Hz)		ξ		Mean of f_c (Hz)	Mean of ξ
		Point A	Point B	Point A	Point B		
X	1	86.2874	86.1849	1.84%	1.79%	86.3087	1.40%
	2	86.4192	85.9878	1.23%	1.55%		
	3	86.5354	86.4373	0.98%	1.02%		
Y	4	85.9953	86.0024	2.01%	1.98%	85.9641	1.94%
	5	86.1776	86.7343	1.68%	1.99%		
	3	84.9985	85.8767	2.12%	1.76%		



Fig. 6 Experimental configuration of the non-contact vibration measurement. (①: Milling cutter; ②: Vertically arranged Doppler vibrometers; ③: Hammer vibration test; ④: Data acquisition system.)

used as a chatter stability feature even if process data such as cutting force and vibration are unavailable in practical industrial settings. Third, the computational burden of the proposed method is significantly reduced compared to the neural network models that train directly on milling force and vibration data.

3.3 Milling experiments and data preparation

To prepare the training and test data for the predictive model, aluminum alloy 7075Al workpieces were machined on a Mikron vertical five-axis machining center. The size of the workpiece was $110 \times 90 \times 10 \text{ mm}^3$, and the size of the machined slots was $13 \times 4 \text{ mm}^2$ (see Fig. 10).

Table 3 Experimentally measured modal parameters (main frequency f_c , damping ratio ξ , and stiffness K) of the milling tool tip

Direction	f_c (Hz)	ξ (%)	K (N/m)
X	373.842	4.661	9.425×10^7
Y	392.373	3.751	9.425×10^7

The cutting parameters used in the milling experiments, including spindle speed, feed rate, depth of cut, and feed per tooth, are summarized in Table 5. Three workpieces (each containing 70 slots) were machined, which gave $70 \times 3 = 210$ experimental data samples in total. After the milling experiments were completed, a white light interferometer (Zygo Newview 9000, USA) was used to measure the roughness of the machined surface of the slots and calculate the ground-truth data for training the neural network model, as shown in Fig. 11. In this study, the arithmetic mean deviation of the contour (R_a) was used as a measure of machined surface roughness. The entire surface area of the slot was scanned, and R_a of the full-length topography of the machined feature was measured. The critical depth of cut corresponding to different spindle speeds can be calculated using the previously calculated milling SLD. We then calculated the difference between the critical depth of cut and actual depth of cut (i.e., the chatter stability feature). Examples of the experimental dataset containing both the input and target variables are listed in Table 6.

The roughness measurement results showed that 14 out of 210 machined slots had significantly higher R_a values than the rest, which are believed to be caused by chatter. As our main focus is to predict the surface roughness when the machining system is still in the “stable” state (it is desirable to avoid chattering in practical situations), these

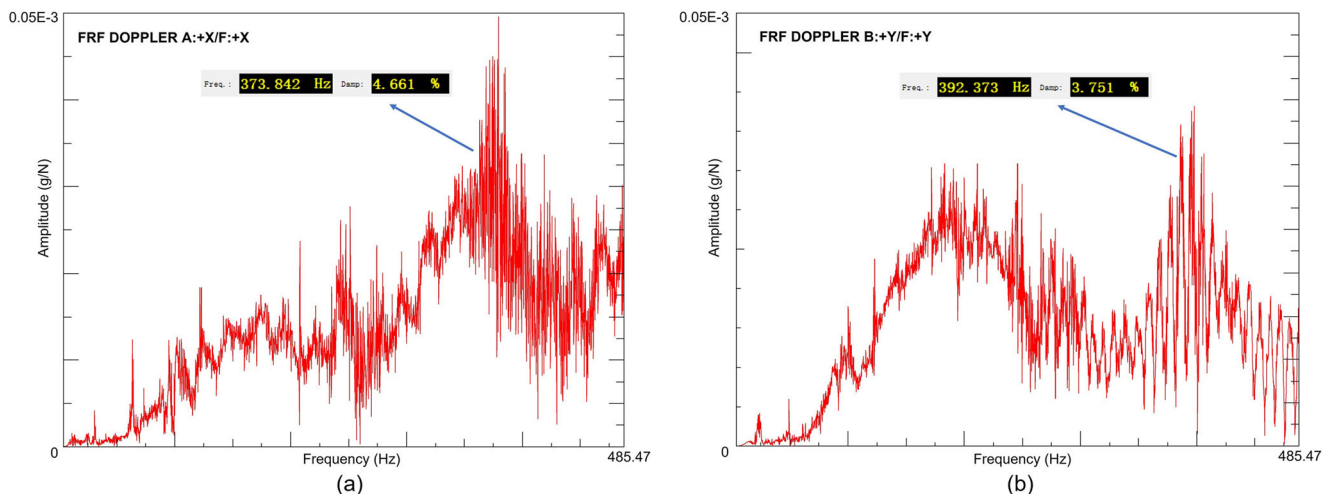


Fig. 7 FRFs of the milling cutter nose point. (a X-direction; b Y-direction)

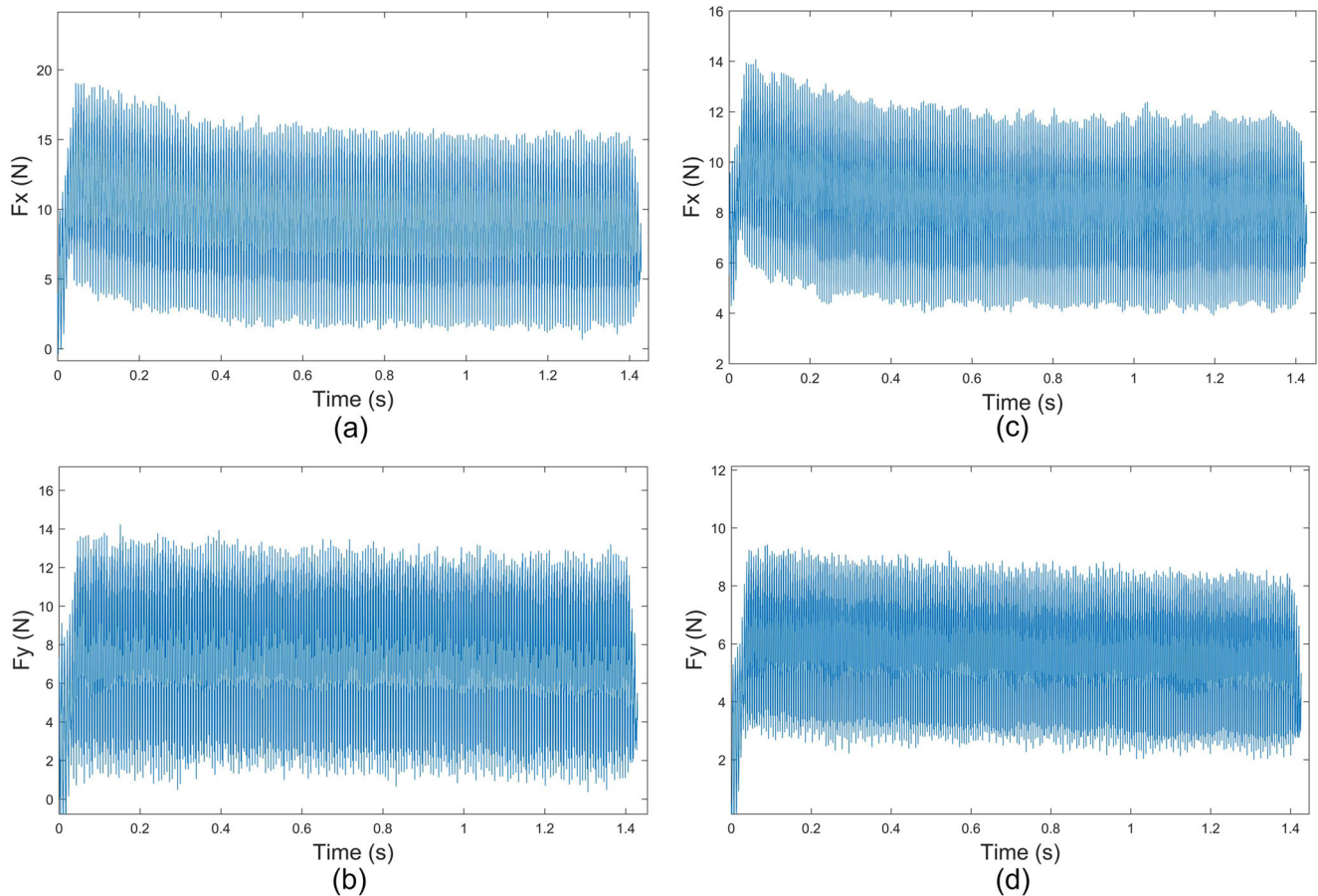


Fig. 8 Milling forces measured under specific cutting parameters, where **a** and **b** are milling forces of X- and Y-directions in test No.1; **c** and **d** are milling forces of X- and Y-directions in test No.2

outliers were removed from the experimental dataset. The remainder of the samples were then divided into training and test sets, and the numbers of samples used for training and performance evaluation (i.e., test) were 138 (70%) and 58 (30%), respectively.

4 Results

To demonstrate the performance improvement achieved by adopting the proposed approach, we trained three different BP neural network models and compared their prediction performances. More specifically, model 1 only

used traditional cutting parameters, including the spindle speed, feed rate, depth of cut, and feed per tooth. In addition to these static predictors, the input data of model 2 also contained the chatter stability feature Δa (Eq. (10)) obtained via modal impact tests with contact vibration sensors (Fig. 4). In comparison, model 3 was trained using the static features as well as the chatter stability feature obtained using the non-contact measurement method proposed in this study (Fig. 6). Detailed information regarding the BP neural network models adopted in this study, including the network structure, loss function, and optimizer, can be found in Table 7. The input and output variables of models 2 and 3 are also shown in Fig. 12. The

Table 4 Cutting parameters (spindle speed n , feed rate f , depth of cut a_p , and feed rate per tooth f_z) used in the trial cutting and the obtained milling force coefficients

Test No.	n (rpm)	f (mm/min)	a_p (mm)	f_z (mm/z)	K_{tc} (N/mm ²)	K_{rc} (N/mm ²)
1	7000	300	0.2	0.0107	1482	889
2	8000	300	0.2	0.0094		

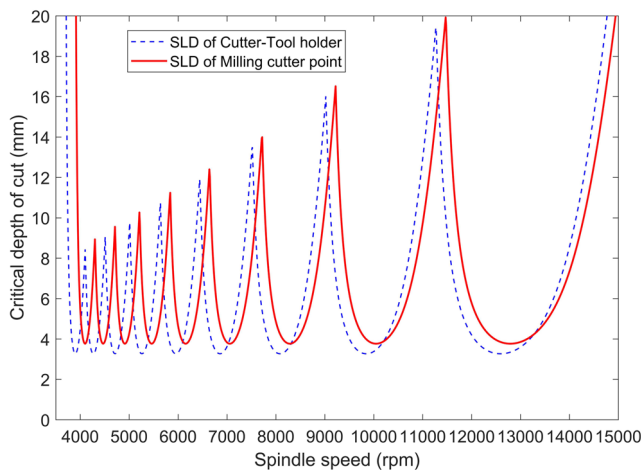


Fig. 9 Milling SLDs of the cutter-toolholder system and cutter tip

performance of the three models for roughness prediction was evaluated quantitatively in terms of the prediction accuracy, which is defined as:

$$\text{accuracy} = 1 - \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (11)$$

where N is the number of samples, y_i is the experimentally measured roughness value R_a , and \hat{y}_i is the prediction result given by the BP neural network model. The iteration diagrams of the accuracy and loss functions of the three models are shown in Fig. 13a–c, and Fig. 14 shows the



Fig. 10 Milling experiments (wet cutting) carried out on a Mikron vertical five-axis machining center. (①: Workpiece with the machined features.)

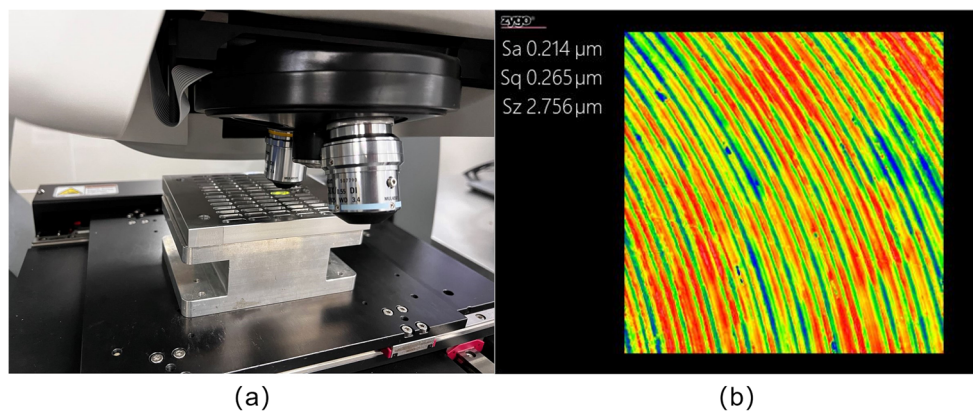
scatter plot (the predicted values versus the true values) of model 3 obtained on the test set. It can be observed from Fig. 13 that model 3 has significantly improved the accuracy of prediction with the addition of the chatter stability feature Δa introduced in this study, when other factors of the experiment remain the same.

Table 8 summarizes the prediction accuracy of the three models for the training, validation, and test sets (10% of the samples were used for validation during training). It can be observed from Table 8 that model 1 had an accuracy of 84.34% on the training set, and its accuracy was only 78.39% on the validation set. The neural network models built with the addition of the chatter stability feature improved the prediction accuracy for both the training and validation sets. The training accuracy of model 2 increased to 87.32% and its validation accuracy was 82.99%. Importantly, model 3 achieved the highest training accuracy of 93.16% and the validation accuracy reached 88.59%, confirming the effectiveness of the proposed approach. For all predictive models, the test set accuracy was close to the validation accuracy, showing that the trained neural network models were generalizable. As shown in Fig. 15, the prediction errors of the three models were mainly distributed between -0.06 and $0.06 \mu\text{m}$, and the prediction errors of model 3 had the smallest variance. In addition, the maximum prediction errors of models 1 and 2 exceeded $0.09 \mu\text{m}$, whereas the maximum prediction error of model 3 was within $0.09 \mu\text{m}$, suggesting that model 3 can effectively reduce large prediction errors.

The results in Table 8 suggest that an accurate SLD is desirable for obtaining satisfactory prediction performance. To further study the influence of the accuracy of the SLD, univariate and multivariate sensitivity analysis were performed on the modal parameters based on a uniform and Sobol sampling scheme [25], respectively. Considering that the spindle, tool, and the tool holder are all rotary structures, the mean values of the main frequency, damping ratio, and stiffness in the X- and Y-directions (see Table 3) were calculated and considered as the benchmark values. In the univariate sensitivity analysis, the three modal parameters were changed to 90%, 95%, 105%, and 110% of the benchmark values, obtaining 12 new sets of parameters as shown in Table 9. For each set of the modal parameters, a “degraded” version of the SLD was calculated, based on which the chatter stability feature values were extracted for the training and test samples. Table 9 summarizes the prediction accuracy obtained on these new datasets, from which it can be seen that the prediction performance of the BP model has deteriorated as a result of inaccurate SLDs. Among the three modal parameters, the prediction accuracy was shown to be least affected by the change of the damping ratio ξ , while the main frequency f_c and stiffness K exhibited similar influences. Importantly, all the accuracy

Table 5 Cutting parameter settings in the milling experiments

Cutting parameters	Values
Spindle speed (r/min)	5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000
Feed rate (mm/min)	200, 300, 400
Depth of cut (mm)	0.2, 0.5, 0.7, 1, 1.5, 2, 2.5
Feed rate per tooth (mm/z)	0.0054, 0.0058, 0.0063, 0.0069, 0.0075, 0.0083, 0.0094, 0.011, 0.0125, 0.015

Fig. 11 Roughness measurement (a white light interferometer; b 2D surface topography)**Table 6** Part of the experimental dataset including the chatter stability features

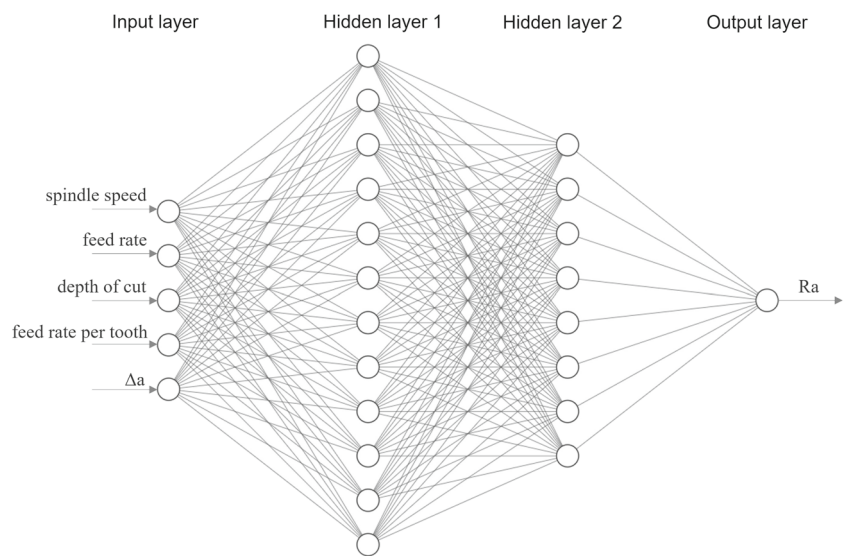
n (rpm)	f (mm/min)	a_p (mm)	f_z (mm/z)	Δa (mm)	R_a (μm)
5000	300	0.2	0.0150	3.1248	0.198
6000	300	0.2	0.0125	3.1405	0.211
7000	300	0.2	0.0107	3.1212	0.182
8000	300	0.2	0.0094	3.4212	0.171
9000	300	0.2	0.0083	3.7176	0.125
10000	300	0.2	0.0075	3.1776	0.132

Refer to Table 4 and Eq. (10) for explanations of the variables

Table 7 Details of the BP neural network models

Property	Value
Number of input features	4 (model 1), 5 (models 2 and 3)
Number of hidden layers	2
Number of nodes in hidden layer 1	12
Number of nodes in hidden layer 2	8
Loss function	Mean absolute error (MAE)
Activation function	Relu
Optimizer	Adam
Learning rate	0.0005

Fig. 12 The network structure of roughness prediction models 2–3



values listed in Table 9 are larger than that of model 1 (i.e., 78.86%—see Table 8), demonstrating the robustness of the proposed approach to errors of the modal parameters. For the purpose of multivariate sensitivity analysis, the Sobol sampling method, which is capable of covering multi-dimensional space with a limited number of samples [25], was adopted. In total, 50 sets of the modal parameters were obtained using this approach in three-dimensional parameter space, and the considered parameter ranges were the same as those in the univariate case. The training and test samples were also prepared following the same procedure as in the univariate case, and Fig. 16 shows the histogram plot of the prediction accuracy for these new sets of the modal parameters. It can be seen from Fig. 16 that when the modal parameters deviated from the benchmark values, the accuracy of the roughness prediction model was lower than that of the accurate model (i.e., 86.62% for model 3). However, in most cases (48 out of 50), the prediction accuracy was still higher than that of the conventional method which does not consider the chatter stability feature (i.e. model 1).

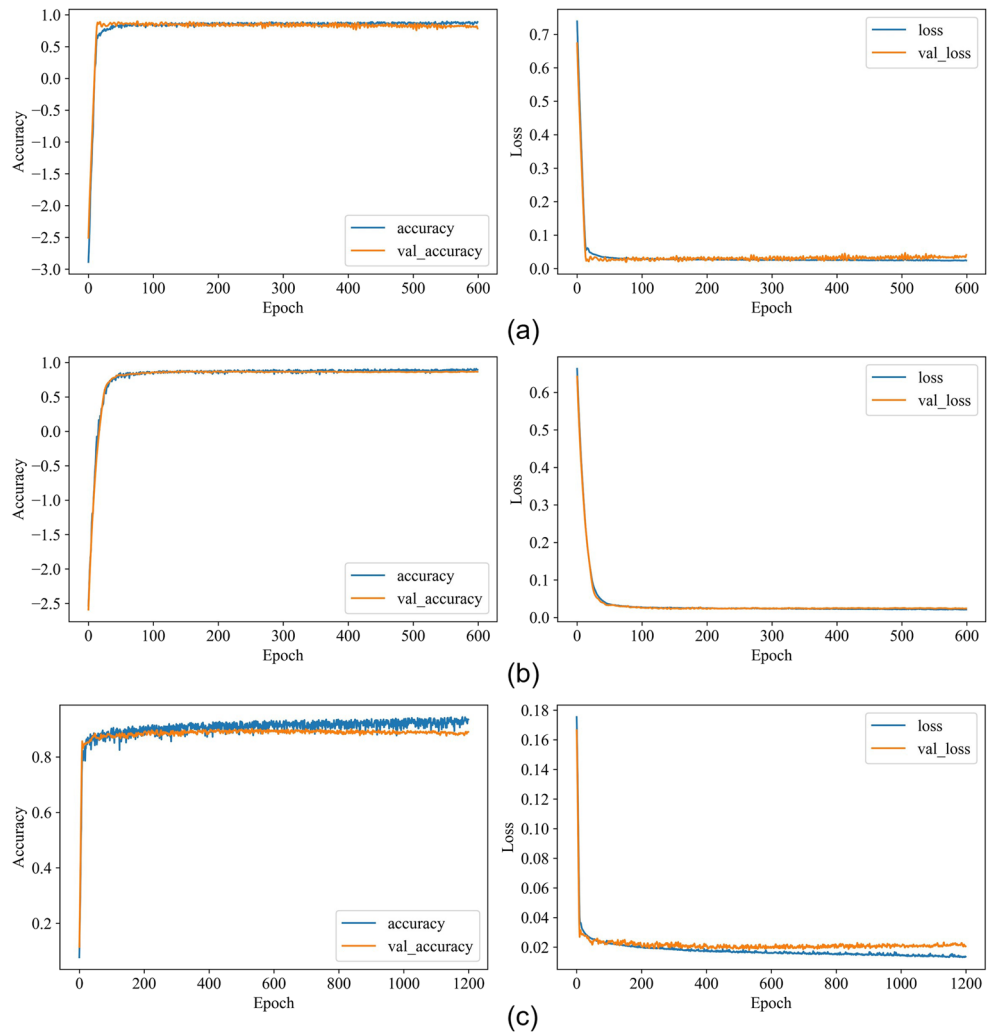
Finally, correlation analysis was performed to study the relationship between the input and target variables of the predictive model [26]. The Pearson correlation coefficient was selected for this purpose, which is defined as

$$r = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}}. \tag{12}$$

In Eq. (12), X_i, Y_i are the input and target variables, respectively, and \bar{X}, \bar{Y} denote their mean values.

Table 10 summarizes the Pearson correlation coefficients between the roughness value (i.e., the target variable) and different input variables of the predictive model. The results in Table 10 show that the spindle speed was negatively correlated with roughness. The feed rate, depth of cut, and feed rate per tooth were positively correlated with roughness. While these results could reveal the underlying relationships between the cutting parameters and the roughness to some extent, the absolute values of the Pearson correlation coefficient were shown to be small. Even for the depth of cut and feed rate per tooth, the Pearson correlation coefficient was below 0.3, which implies that the roughness value cannot be efficiently predicted using only these static cutting parameters. The chatter stability feature Δa (Eq. (10)) quantifies the difference between the limit depth of cut and the actual depth of cut. Based on the hypothesis of this paper, the machining process is more stable and the value of the machined surface roughness is smaller when Δa becomes larger. It can also be confirmed from Table 10 that Δa has a strong negative correlation with surface roughness. Among the input features of the predictive model studied in Table 10, the chatter stability feature Δa_2 (Eq. (10)) achieved the highest correlation coefficient of 0.3730 (absolute value) with the target variable, which explains why model 3 yielded the highest accuracy for the prediction of roughness. Note that the Pearson correlation coefficient of the chatter stability feature Δa_1 used in model 2 is only 0.0463 (absolute value). This confirms that the non-contact vibration measurement method adopted in this study can be used to obtain a more informative predictor of roughness, which is crucial for improving the accuracy of the BP neural network.

Fig. 13 Change of the accuracy and loss in the training process of: **a** model 1, **b** model 2, and **c** model 3



5 Conclusion

In this study, we analyzed the milling dynamics model of the machine tool-cutter system and adopted a non-

contact vibration sensor to obtain the FRF of the milling cutter tip point for the calculation of the milling SLD. To predict the roughness of the machined surface, we further quantified the milling SLD and introduced a chatter stability feature defined as the difference between the actual and critical depths of cut for a given spindle speed. Through milling experiments, it was demonstrated that the BP neural network model based on this chatter stability feature can improve the prediction accuracy by 7.8% compared with that using only conventional predictors, such as the static cutting parameters. The sensitivity analysis involving the

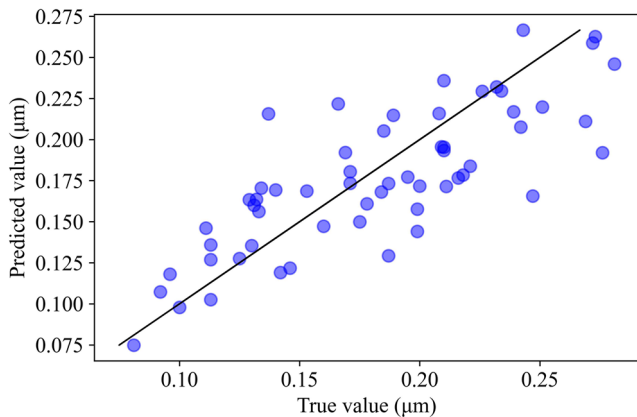


Fig. 14 Prediction results of model 3 on the test set

Table 8 Prediction accuracy of models 1–3 in training, validation, and test sets

Model No.	Training	Validation	Test
1	84.34%	78.39%	78.86%
2	87.32%	82.99%	81.23%
3	93.16%	88.59%	86.62%

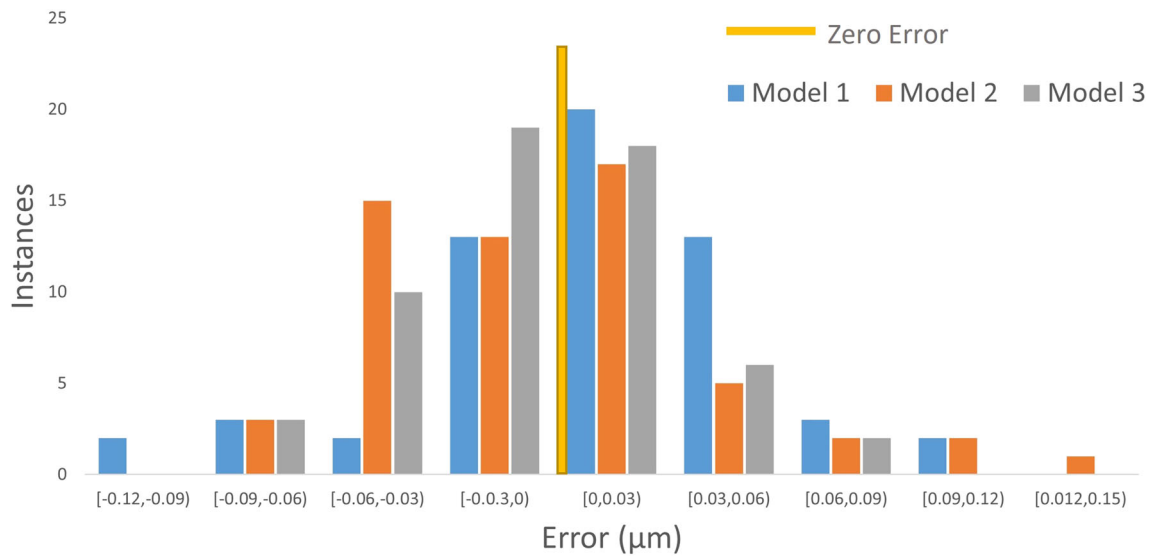


Fig. 15 Error histogram of models 1–3

Table 9 Prediction accuracy results of the univariate sensitivity analysis

No.	f_c (Hz)	ξ (%)	K (N/m)	Prediction accuracy (%)
1	344.797	4.206	9.425	79.64
2	363.953	4.206	9.425	80.57
3	402.263	4.206	9.425	81.15
4	421.419	4.206	9.425	81.50
5	383.108	3.785	9.425	82.13
6	383.108	3.996	9.425	83.81
7	383.108	4.416	9.425	84.29
8	383.108	4.627	9.425	84.44
9	383.108	4.206	8.483	80.65
10	383.108	4.206	8.954	81.88
11	383.108	4.206	9.896	81.19
12	383.108	4.206	10.368	80.33

Fig. 16 Histogram plot of the prediction accuracy obtained in the multivariate sensitivity analysis

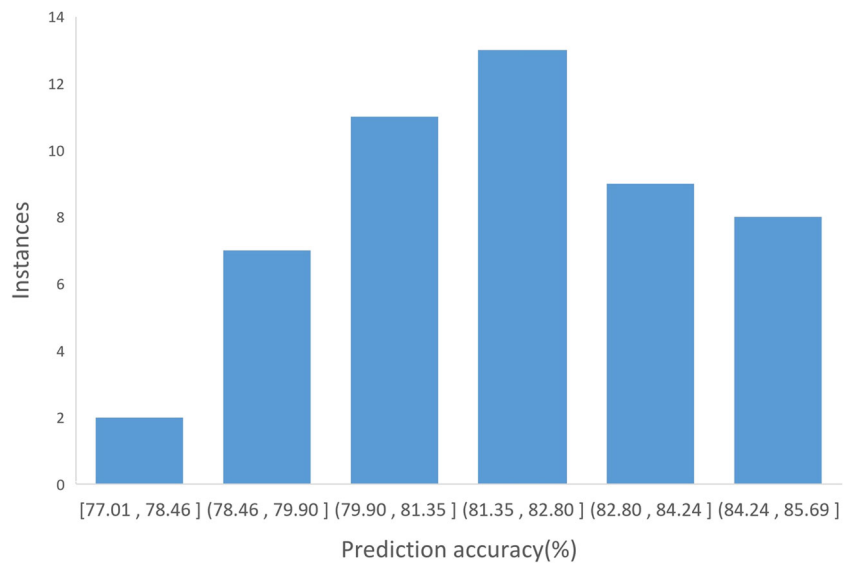


Table 10 Pearson correlation coefficients between different input features and the target variable

Input feature	Correlation coefficient
n (rpm)	-0.0528
f (mm/min)	0.0154
a_p (mm)	0.2080
f_z (mm/z)	0.1730
Δ_1 (mm)	-0.0463
Δ_2 (mm)	-0.3730

Refer to Table 4 and Eq. (10) for explanations of the variables

use of inaccurate SLDs demonstrated that performance improvement can still be achieved by adopting the proposed approach when the modal parameters are determined with (up to) 10% errors. Pearson correlation analysis further confirmed that the newly introduced feature had a strong correlation with the roughness.

Owing to the small size of the precision milling cutter, the vibration sensor could not be mounted at the cutter tip point during the modal test. An optical non-contact vibrometer was used to measure the vibration signals, and a more accurate critical depth of cut was calculated, which was shown to improve the predictive power of the chatter stability feature. By adopting the proposed approach, the prediction accuracy of the BP neural network can be improved without using a large amount of process data, and hence, it is readily applicable to existing machine tools (i.e., there is no need to install external sensors except for trial cutting). Future work will aim to explore the application of the proposed approach for quality prediction of different cutting methods, such as turning, grinding, and drilling.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Long Bai, Xin Cheng, and Qizhong Yang. The first draft and revised version of the manuscript were written by Long Bai and Xin Cheng.

Funding This work is supported by the National Key R&D Program of China (Grant No. 2020YFB1710400), the National Natural Science Foundation of China (Grant No. 52005205), and the National Science Fund for Distinguished Young Scholars (Grant No. 52225506).

Declarations

Conflict of interest The authors declare no competing interests.

References

- Çolak O, Kurbanoğlu C, Kayacan MC (2007) Milling surface roughness prediction using evolutionary programming methods. *Mater Des* 28(2):657–66. <https://doi.org/10.1016/j.matdes.2005.07.004>
- Muñoz-Escalona P, Maropoulos PG (2010) Artificial neural networks for surface roughness prediction when face milling. *AI 7075-T7351. J Mater Eng Perform* 19(2):185–93. <https://doi.org/10.1007/s11665-009-9452-4>
- Quintana G, Jd Ciurana, Ribatallada J (2010) Surface roughness generation and material removal rate in ball end milling operations. *Mater Manu Processes* 25(6):386–98. <https://doi.org/10.1080/15394450902996601>
- Bachrathy D, Insperger T, Stépán G (2009) Surface properties of the machined workpiece for helical mills. *Mach Sci Technol* 13(2):227–45. <https://doi.org/10.1080/10910340903012167>
- Karayel D (2009) Prediction and control of surface roughness in CNC lathe using artificial neural network. *J Mater Process Technol* 209:3125–3137. <https://doi.org/10.1016/j.jmatprotec.2008.07.023>
- Asiltürk I, Çunkaş M (2011) Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method. *Expert Syst Appl* 38:5826–5832. <https://doi.org/10.1016/j.eswa.2010.11.041>
- Bajić D, Lela B, Živković D (2008) Modeling of machined surface roughness and optimization of cutting parameters in face milling. *Metalurgija* 47:331–334
- Groove M (1996) Fundamentals of modern manufacturing, Prentice Hall, Upper Saddle River
- Lela B, Bajić D, Jozić S (2009) Regression analysis, support vector machines, and Bayesian neural network approaches to modeling surface roughness in face milling. *Int J Adv Manufact Technol* 42:1082–1088. <https://doi.org/10.1007/s00170-008-1678-z>
- Abu-Mahfouz I, El Ariss O, Esfakur Rahman AHM, Banerjee A (2017) Surface roughness prediction as a classification problem using support vector machine. *Int J Adv Manufact Technol* 92:803–815. <https://doi.org/10.1007/s00170-017-0165-9>
- Tangitsitcharoen S, Thesniyom P, Ratanakuakangwan S (2017) Prediction of surface roughness in ball-end milling process by utilizing dynamic cutting force ratio. *J Intell Manuf* 28(1):13–21. <https://doi.org/10.1007/s10845-014-0958-8>
- Lou SJ, Chen JC (1999) In-process surface roughness recognition (ISRR) system in end-milling operations. *Int J Adv Manufact Technol* 15:200–209. <https://doi.org/10.1007/s001700050057>
- Pan Y, Wang Y, Zhou P, Yan Y, Guo D (2020) Activation functions selection for BP neural network model of ground surface roughness. *J Intell Manuf* 31(8):1825–36. <https://doi.org/10.1007/s10845-020-01538-5>
- Markopoulos AP, Georgiopoulos S, Manolakos DE (2016) On the use of back propagation and radial basis function neural networks in surface roughness prediction. *J Ind Eng Int* 12(3):389–400. <https://doi.org/10.1007/s40092-016-0146-x>
- Kao YC, Chen SJ, Vi TK, Feng GH, Tsai SY (2021) Study of milling machining roughness prediction based on cutting force. *IOP Conf Ser: Mater Sci Eng* 1009(1):012027. <https://doi.org/10.1088/1757-899X/1009/1/012027>
- Quintana G, Ciurana J (2011) Chatter in machining processes: a review. *Int J Mach Tools Manuf* 51:363–376. <https://doi.org/10.1016/j.ijmactools.2011.01.001>
- Özşahin O, Budak E, Özgüven HN (2015) Identification of bearing dynamics under operational conditions for chatter stability prediction in high speed machining operations. *Precis Eng* 42:53–65. <https://doi.org/10.1016/j.precisioneng.2015.03.010>
- Wang D, Löser M, Ihlenfeldt S, Wang X, Liu Z (2019) Milling stability analysis with considering process damping and mode shapes of in-process thin-walled workpiece. *Int J Mech Sci* 159:382–397. <https://doi.org/10.1016/j.ijmecsci.2019.06.005>
- Altıntaş Y, Budak E (1995) Analytical prediction of stability lobes in milling. *CIRP Ann* 44:357–362. [https://doi.org/10.1016/S0007-8506\(07\)62342-7](https://doi.org/10.1016/S0007-8506(07)62342-7)

20. Pour M, Torabizadeh MA (2016) Improved prediction of stability lobes in milling process using time series analysis. *J Intell Manuf* 27:665–677. <https://doi.org/10.1007/s10845-014-0904-9>
21. Nguyen V, Melkote S (2021) Hybrid statistical modelling of the frequency response function of industrial robots. *Rob Comput Integr Manuf* 70:102134. <https://doi.org/10.1016/j.rcim.2021.102134>
22. Wu Y, Feng J (2018) Development and application of artificial neural network. *Wirel Pers Commun* 102:1645–1656. <https://doi.org/10.1007/s11277-017-5224-x>
23. Srikant RR, Krishna PV, Rao ND (2011) Online tool wear prediction in wet machining using modified back propagation neural network. *Proc Inst Mech Eng. Part B: J Eng Manuf* 225(7):1009–18. <https://doi.org/10.1177/0954405410395854>
24. Guo T, Meng L, Cao J, Bai C (2020) An identification method of the weak link of stiffness for cantilever beam structure. *Sci Prog* 103(3):0036850420952671. <https://doi.org/10.1177/0036850420952671>
25. Totis G, Sortino M (2020) Polynomial Chaos-Kriging approaches for an efficient probabilistic chatter prediction in milling. *Int J Mach Tool Manufact* 157:103610. <https://doi.org/10.1016/j.ijmactools.2020.103610>
26. Misaka T, Herwan J, Ryabov O, Kano S, Sawada H, Kasashima N et al (2020) Prediction of surface roughness in CNC turning by model-assisted response surface method. *Precis Eng* 62:196–203. <https://doi.org/10.1016/j.precisioneng.2019.12.004>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.