



Article ID 1007-1202(2022)05-0439-14

DOI <https://doi.org/10.1051/wujns/2022275439>

# **Study on the Predictive of Dynamic Milling Force of Milling Process Based on Data Mining**

**JIN Lan, WU Lishuang<sup>†</sup>, ZHANG Xuefeng,  
XIE Liming**

School of Mechanical and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, Gansu, China

© Wuhan University 2022

**Abstract:** To improve surface accuracy of the work-piece and obtain potentially valuable information, a dynamic milling force prediction model was proposed based on data mining. In view of the current dynamic milling force obtained through finite element simulation and analytical calculation, in the finite element modeling, the model built is inevitably different from the actual working conditions, and the analytical calculation is slightly cumbersome and complex, and a dynamic milling force prediction model based on data mining is proposed. The model was established using a combination of regression analysis and Radial Basis Function (RBF) neural network. Using data mining as a means, the internal relationship between milling force, cutting parameters, temperature, vibration and surface quality is deeply analyzed, and the influence of dynamic milling force changes on different situations is extracted and summarized by the methods of cluster analysis and correlation analysis. The results show that the proposed dynamic milling force model has a good prediction effect, ensures the production quality, reduces the occurrence of flutter, improves the surface accuracy of the work-piece, and provides a more accurate basis for the selection of process parameters.

**Key words:** data mining; data modeling; dynamic milling force; prediction of milling force

**CLC number:** TG 506.9

**Received date:** 2022-05-11

**Foundation item:** Supported by Gansu Science and Technology Program (21YF5GA080)

**Biography:** JIN Lan, female, Associate professor, research direction: system dynamics. E-mail: lan\_jane@lut.edu.cn

<sup>†</sup> To whom correspondence should be addressed. E-mail: 18846431603@163.com

## **0 Introduction**

In recent years, the deep integration of information technology and manufacturing industry has become one of the ultimate goals of the manufacturing industry to promote "smart manufacturing"<sup>[1]</sup>. Data-driven predictive and proactive manufacturing is beginning to permeate all aspects of the machining process<sup>[2]</sup>.

Judging and predicting workpiece surface quality by monitoring signals related to vibration, cutting forces, temperature and etc. is an important aspect of intelligent cutting<sup>[3]</sup>. The quality, efficiency and competitiveness of milling can be greatly improved by converting the large amount of relevant data contained in the milling process, such as vibration, cutting forces and temperature, into useful information, exploring its intrinsic relevance to actual machining problems, and applying methods such as cluster analysis, correlation analysis, regression analysis, and artificial neural networks to generate new and unusual information on the basis of the available data sets to guide production<sup>[4]</sup>.

As early as 1985, Wang *et al*<sup>[5]</sup> used fuzzy cluster analysis to identify the cutting state, and this method has been widely used in data mining. Cai *et al*<sup>[6]</sup> used cluster analysis to classify and identify the energy efficiency status by applying image processing, and invented a method to locate and trace the source of low energy efficiency problems occurring in cutting and machining systems, providing a theoretical basis and a potential technical solution for online monitoring of the energy efficiency status of cutting and machining systems. Bracke

*et al*<sup>[7]</sup> outlined an integrated approach to detect grinding parameter settings which ensures repeatability on the one hand and economy on the other hand. Cluster analysis was used to identify groups of grinding parameters and grinding results with similar characteristics.

Based on the connection function, Pei *et al*<sup>[8]</sup> analyzed the correlation of force signal, vibration signal and roughness signal in milling process, and got a better result. Rodrigues *et al*<sup>[9]</sup> applied correlation analysis to explore the degree of correlation between roughness and cutting speed and feed rate when turning 45-gauge steel in order to find the optimal configuration of the parameters. Thasthakeer *et al*<sup>[10]</sup> investigated the correlation between the variation of different parameters and the rate of change of cutting forces and surface roughness after machining. They found that this correlation was higher under wet machining conditions. Based on the multiple regression analysis, Lin *et al*<sup>[11]</sup> established the prediction model of the cutting parameters and the machining vibration of the machined parts in face milling by the surface roughness, and got a good prediction result. In Ref. [12], to explore the relationship between machining parameters and cutting temperature, the analysis of variance was used to test the significance of the parameters and a regression equation was developed to demonstrate the accuracy of the predictive model by comparing predicted and experimental values of cutting temperature. Rajesh *et al*<sup>[13]</sup> used multivariate regression analysis, which included a variance analysis, to establish a predictive model for the relative effects of process parameters on incision width and surface roughness. Yeganeifar *et al*<sup>[14]</sup> used support vector machine, the neural network and regression analysis to predict and optimize the surface roughness and cutting forces during milling, and

compared them with each other in order to obtain a self-use and optimal selection scheme. Maher *et al*<sup>[15]</sup> used an adaptive neuro-fuzzy inference system model to study the degree of correlation between cutting parameters, cutting forces and surface roughness. Basak *et al*<sup>[16]</sup> applied Radial Basis Function (RBF) neural network to milling, and fitted RBF neural network model by 27 kinds of experimental data under different cutting conditions, which showed that this method can be applied to milling.

The advantage of data mining is the introduction of cluster analysis, correlation analysis and regression analysis which are efficient, convenient and accurate in solving practical machining problems. This method could start from the actual milling process and take cluster analysis, correlation analysis and regression analysis and RBF neural network as the main means to build a dynamic milling force prediction model. Results demonstrate that the proposed dynamic milling force model has a better prediction effect, which ensures production quality and reduces the occurrence of chattering to provide a more accurate process parameters selection basis.

## 1 Milling Data Gain with Multiple Sensors

In order to obtain the necessary experimental data, in this paper, the complex milling experiments including surface quality detection before milling, milling process, milling temperature detection, milling vibration detection, dynamic milling force detection and milling process are designed. The experimental scheme is shown in Fig. 1.

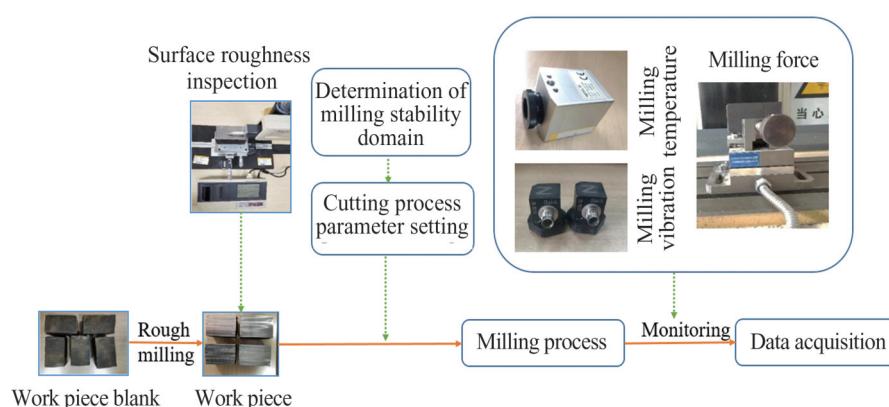


Fig. 1 Experimental scheme design idea

## 1.1 Design of the Milling Experimental Programme

In this paper, 12CrNi2 was selected as the experimental material (as shown in Fig. 2). This material belongs to alloy steel, which contains nickel and has certain sticky characteristics. Therefore, in order to make the cut surface have a high surface finish, a four-flute tungsten steel monolithic alloy end mill with the size of D12 mm×75 mm×4F is selected, as shown in Fig. 3.



Fig. 2 Work-piece after rough machining

Considering the influence of chatter on the actual cutting process before setting up the process parameters, the cutting stability region is found through the milling stability analysis in advance (Fig. 4). The range of process parameters is determined by combining the process manual.



Fig. 3 Integral alloy end mill

The technological parameters are as follows: cutting width  $a_e=0.5\text{-}6\text{ mm}$ , cutting depth  $a_p=0.1\text{-}2\text{ mm}$ , cutting speed  $v_c=72\text{-}140\text{ m/min}$ , feed per tooth  $f_z=0.025\text{-}0.063\text{ mm}$ , rotational speed  $n=1\ 000\text{-}4\ 000\text{ r/min}$ , and  $f$  is the feed speed. Then the response surface method is used to design the concrete parameter scheme. The specific cutting process parameters are shown in Table 1.

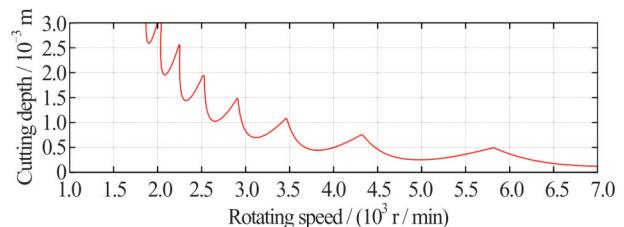


Fig. 4 Stability lobe diagram

Table 1 Process parameter scheme

Group	$a_p/\text{mm}$	$n/(\text{r}/\text{min})$	$a_e/\text{mm}$	$f_z/\text{mm}$	$f/(\text{mm}/\text{min})$	$v_c/(\text{m}/\text{min})$
1	2	1500	3.25	0.044	264	57
2	2	2750	0.5	0.044	484	104
3	2	2750	3.25	0.025	275	104
4	2	2750	3.25	0.044	484	104
5	2	2750	3.25	0.044	484	104
6	2	2750	3.25	0.063	693	104
7	2	2750	6	0.044	484	104
8	2	4000	3.25	0.044	704	151
9	1.05	1500	0.5	0.044	264	57
10	1.05	1500	3.25	0.025	150	57
11	1.05	1500	3.25	0.044	264	57
12	1.05	1500	3.25	0.044	264	57
13	1.05	1500	3.25	0.063	378	57
14	1.05	1500	6	0.044	264	57
15	1.05	2750	0.5	0.025	275	104
16	1.05	2750	0.5	0.044	484	104
17	1.05	2750	0.5	0.044	484	104

Continued on the next page

**Table 1(continued)**

Group	$a_p/\text{mm}$	$n/(r/\text{min})$	$a_e/\text{mm}$	$f_z/\text{mm}$	$f/(mm/\text{min})$	$v_c/(\text{m}/\text{min})$
18	1.05	2 750	0.5	0.063	693	104
19	1.05	2 750	3.25	0.025	275	104
20	1.05	2 750	3.25	0.025	275	104
21	1.05	2 750	3.25	0.044	484	104
22	1.05	2 750	3.25	0.044	484	104
23	1.05	2 750	3.25	0.044	484	104
24	1.05	2 750	3.25	0.044	484	104
25	1.05	2 750	3.25	0.044	484	104
26	1.05	2 750	3.25	0.044	484	104
27	1.05	2 750	3.25	0.063	693	104
28	1.05	2 750	3.25	0.063	693	104
29	1.05	2 750	6	0.025	275	104
30	1.05	2 750	6	0.044	484	104
31	1.05	2 750	6	0.044	484	104
32	1.05	2 750	6	0.063	693	104
33	1.05	4 000	0.5	0.044	704	151
34	1.05	4 000	3.25	0.025	400	151
35	1.05	4 000	3.25	0.044	704	151
36	1.05	4 000	3.25	0.044	704	151
37	1.05	4 000	3.25	0.063	1 008	151
38	1.05	4 000	6	0.044	704	151
39	0.1	1 500	3.25	0.044	264	57
40	0.1	2 750	0.5	0.044	484	104
41	0.1	2 750	3.25	0.025	275	104
42	0.1	2 750	3.25	0.044	484	104
43	0.1	2 750	3.25	0.044	484	104
44	0.1	2 750	3.25	0.063	693	104
45	0.1	2 750	6	0.044	484	104
46	0.1	4 000	3.25	0.044	704	151

## 1.2 Experimental Data Acquisition

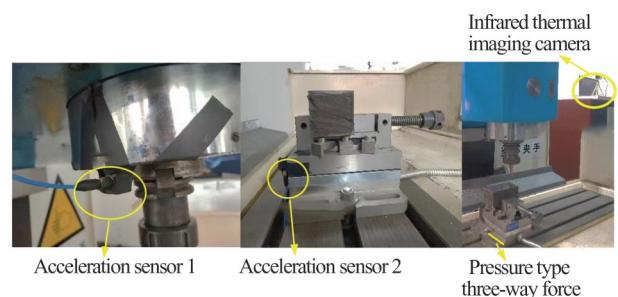
Before milling, one end of the three acceleration sensors is fixed at the spindle handle and fixture, respectively. The other end is connected to an LMS tester connected to a laptop computer, so as to collect the vibration of the spindle tool holder and the vibration of the work-piece itself during the milling process, and prepare for subsequent data processing. In order to measure the change of cutting temperature with time during the milling process, we focus the infrared thermal image on the tool tip, set a moderately sized temperature collection

frame, select the real-time highest temperature in the display collection frame, and use this temperature as the tool tip for milling temperature to realize real-time monitoring of milling temperature. The milling method is down milling. To avoid the influence of cutting fluids on cutting temperature variations, a dry cutting machining method is chosen. During cutting, the pre-treated work-piece is tightly fixed to the fixture of the pressure-type three-way force test sensor, the surface of the work-piece is gently tapped with a copper rod to keep its level, and the data on the real-time variation of the cutting

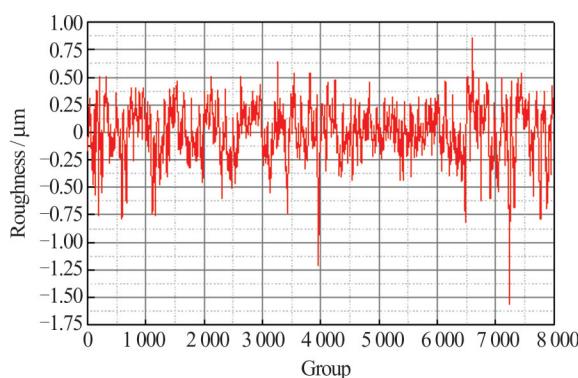
force is recorded by this sensor. The experimental conditions are shown in Fig. 5.

When the experiments were completed, roughly  $3.5 \times 10^8$  experimental data were gained. The resulting data were collated to obtain the experimental results. Taking the first set of experiments as an example, the data images are shown in Figs. 6-9.

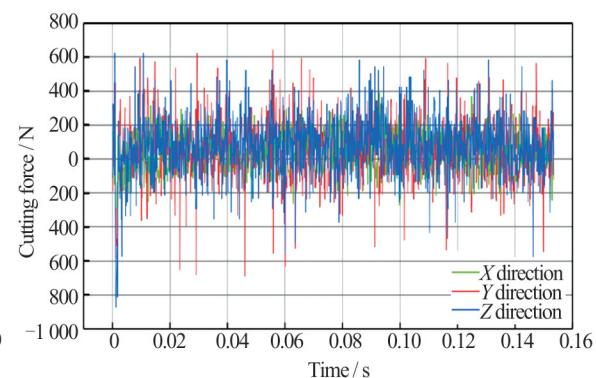
Figure 6 is the first set of roughness curves for the first piece of workpiece. According to the distribution of cutting traces, the surface of each workpiece is divided



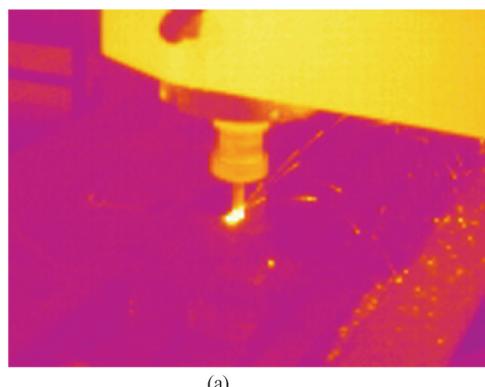
**Fig. 5 Experimental equipment installation**



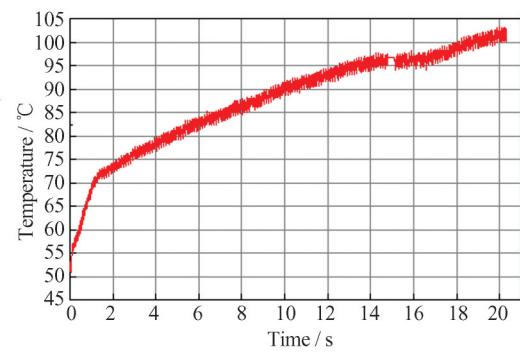
**Fig. 6 Roughness curve graph**



**Fig. 7 Dynamic cutting force**

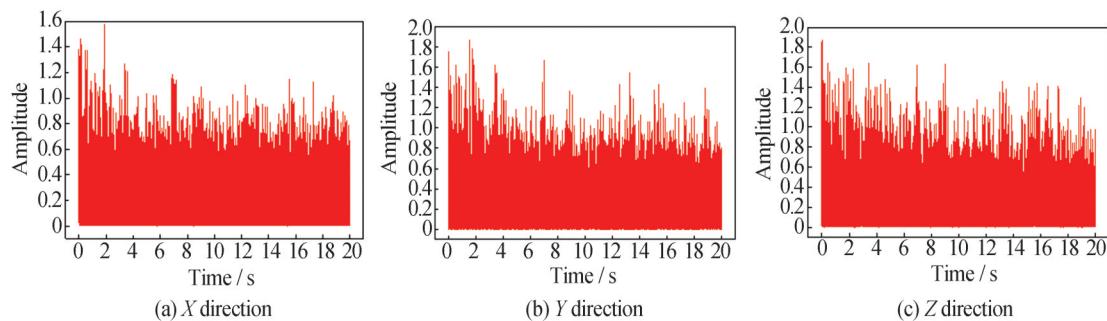


(a)



(b)

**Fig. 8 Infrared thermal imaging (a) and temperature change curve(b)**



**Fig. 9 Original vibration time domain graph**

"Amplitude" is acceleration amplitude measured by accelerometer

into 15 areas, and the roughness is measured separately, and the data records are saved. The SJ-400 measuring instrument under the Surftest brand is used to measure surface roughness as shown in Fig. 10. The dynamic cutting force curve of the first set of experiments is shown in Fig. 7. In the first set of experimental cutting processes, the maximum temperature is 103.7 °C, the minimum temperature is 50.6 °C, and the average temperature is 87.8 °C. Figure 8 shows the infrared thermal imaging and the temperature change curve. Figure 9 is a time-domain plot of the original vibration in the X, Y, Z directions measured by the accelerometer attached to the spindle.



Fig. 10 Surface roughness measuring instrument

### 1.3 Experimental Data Preprocessing

Before data mining can be performed on a sample of data, the data are usually needed to pre-process, i.e. clean, integrate, transform and normalize.

As the samples used for data mining in this paper are derived from experiments where the data acquisition sources are the respective test equipment, the acquired roughness, temperature, cutting force and vibration data have been pre-classified and placed at the time of acquisition. These data are imported into EXCLE tables separately and integrated to complete the integration work; as the data samples are collected under specific conditions and the collection process is supervised by personal, there is no missing data, mis-storage and duplication, but the data signal itself has trend terms and noise, so the data with numerical anomalies are needed to eliminate from the samples through the data screening function of EXCLE before analysis to obtain a more reasonable database, and then eliminate the trend items and reduce the noise. In order to complete the subsequent data mining work efficiently, workstations were chosen to be used for data processing where the data was within the computer's reach and, therefore, no data imputation was required.

As the data samples have various characteristics and their extends are different from each other, the data needs to be standardized. There are many types of experimental data in this paper, such as the temperature data, vibration data, the surface roughness data and cutting force data. The data itself has a large numerical span, so the Z-score standardization is used. After the Z-score normalization process, the mean value of the data could become 0, while the standard deviation becomes 1. The calculation is as follows:

$$x^* = \frac{x - \bar{x}}{\sigma} \quad (1)$$

where  $x$  is the original data of the sample,  $\bar{x}$  is the mean of the original data,  $\sigma$  is the standard deviation of the original data, and  $x^*$  is the new sample data after the standardization process. Taking the cutting force data of the first group of experiments as an example, its cutting force data after standardization is shown in Fig. 11.

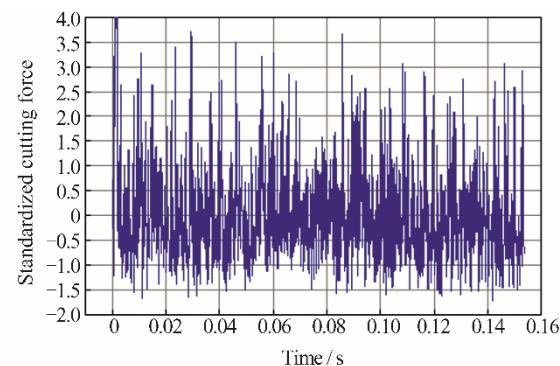


Fig. 11 Cutting force after standardized treatment

Considering that data preprocessing will make some features of the original data impossible to observe, the original data is qualitatively and quantitatively analyzed.

1) It is known from the temperature data that the influence of cutting width on cutting temperature in the cutting process parameters is more significant, and the cutting instability, severe roughness and machining surface hardening will make the overall temperature of the cutting process rise significantly.

2) By analyzing vibration data, it was found that higher amplitudes often mean higher temperatures. When the spindle speed reaches a certain value, the increase in cutting depth at this time will lead to increased vibration.

3) From the surface roughness data, it can be seen that if the spindle speed and feed speed can be reasonably increased, and the cutting depth can be reduced, it

will effectively improve the quality of the processing surface and improve the processing efficiency, otherwise it will be counterproductive.

## 2 Cluster Analysis and Correlation Analysis of Milling Data

### 2.1 Cluster Analysis

The cutting force data has three characteristics:  $X$ ,  $Y$  and  $Z$  direction components. In this paper, we propose to use the pre-processed dynamic milling force data as a classification, and use the  $k$ -means algorithm as the core to filter and cluster the dynamic milling force and its corresponding cutting process parameters and time according to the principle of "the highest degree of similarity".

Taking the first group of experiments as an example, after the sample data is preprocessed and normalized, its distribution is shown in Fig. 12. When performing the cluster analysis, the best number of classifications for the group of samples was first determined by drawing a line graph, and then the iteration value was set. The value of  $k$  is usually picked up by a large number of experiments,  $d$  is the sum of the distances under different  $k$  values from a large number of experiments. The appropriate  $k$  value can be determined by making a line chart with  $d$  as the ordinate and  $k$  as the abscissa. The corresponding  $k$  value at the inflection point is the best choice. Here, the number of clustering centers  $k=4$  and the number of iterations is 100. After 100 iterations, the clustering result suiting to the smallest selection of  $d$  is shown in Fig. 13, where different colors represent different categories.

The four data center coordinates are  $(0.692, -0.809, -0.210)$ ,  $(0.491, 0.142, 0.981)$ ,  $(-0.866, -0.253, -1.340)$  and  $(-0.730, 0.912, 0.115)$ , and the sums of the distances from each centre to the data in the corresponding class are 666.2035, 513.9014, 548.7038 and 635.1610, respectively. As can be seen from Fig. 14, the boundaries of the four data groups are obvious and the classification effect is good.

The remaining 45 groups of milling force data are clustered using the same method. Eventually it was found the majority (21 groups in total) of the experimental groups had milling force data that could be classified into 3 categories; 9 experimental groups had data that could be classified into 4 categories; 6 experimental groups had data that could be classified into 2 categories; few (2 groups in total) had data that could be classi-

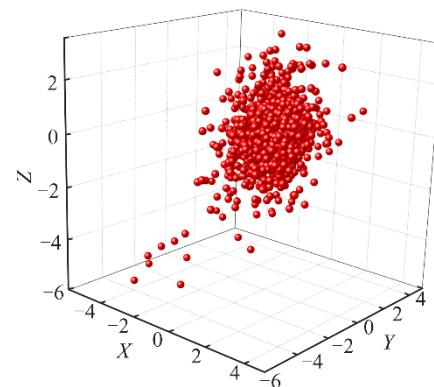


Fig. 12 Distribution of experimental samples in Group 1

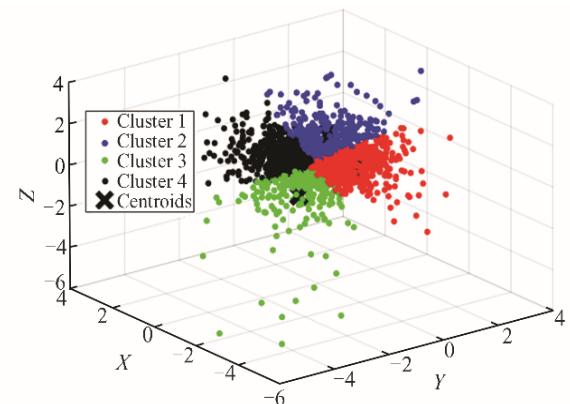


Fig. 13 Clustering of Group 1 samples

fied into 5 categories; and the rest of the data was unclassifiable data that each belonged to a category. Cluster analysis of data, the sample with a large amount of data, is divided into different sub-categories according to certain rules, so as to avoid problems such as large errors and insufficient accuracy when different types of data are analyzed together in the later stage.

### 2.2 Correlation Analysis

The clustering method described above classified the available 46 sets of milling force data into 126 categories, which shows that the main influencing factors for the variation of the milling forces are still different, even when they are generated under the same cutting process parameters. However, the clustering is of little significance for dynamic stiffness calculations at a later stage. In order to clarify the differences and characteristics of each type of data, further categorize them according to their main influencing factors, and improve the efficiency of data processing, the correlation analysis method was chosen to process each type of data and find their respective "labels".

From the above analysis it was found that temperature, roughness and vibration were the main influencing factors on dynamic milling forces, from which a preliminary judgement can be made that each group of experimental values of milling forces can be subdivided into multiple categories and their division may be related to these influencing factors. However, we know that dynamic milling forces are influenced by more factors than this. Therefore, in order to find a corresponding "label" for each group of data, a correlation analysis was chosen to correlate each category of data with temperature, roughness and vibration data, and to carry out a two-tailed test for significance, while data that were not significantly correlated with these three factors were grouped into a new category "Other".

Before conducting the correlation analysis, a scatter plot was drawn to make a preliminary judgement on the linearity or non-linearity of the relationship between the data, with a view to determining the type of correlation coefficient chosen. Taking the first six sets of data as an example, Fig. 14 shows the scatter plot of correlations corresponding to the first six sets of experiments.

As shown in Fig. 14, scatter plots of the correlation between the different classes of the first set of experi-

mental three-way milling forces and the vibration data in the Z1 direction, the vibration data in the Y2 direction, the roughness data and the temperature data were plotted, respectively, and it was found that all relationships were not linear and the data itself did not obey a normal distribution, so the Spearman coefficient was chosen for further analysis.

Table 2 is an excerpt from the results of the correlation analysis performed on the different classes of the 46 groups of experimental milling forces (since the original data were divided into 126 classes of more categories, they are not listed here).

It is important to note that the correlations are grouped in the same category, but in different directions, e.g. two categories are significantly correlated with vibration data, but the first category is positively correlated and the second is negatively correlated.

The correlation analysis revealed that 40 classes were significantly correlated with vibration data, 72 classes were significantly correlated with temperature data, 67 classes were significantly correlated with roughness data, and only 16 classes were classified as correlated with other factors, some of which were significantly correlated with two or three kinds of vibration,

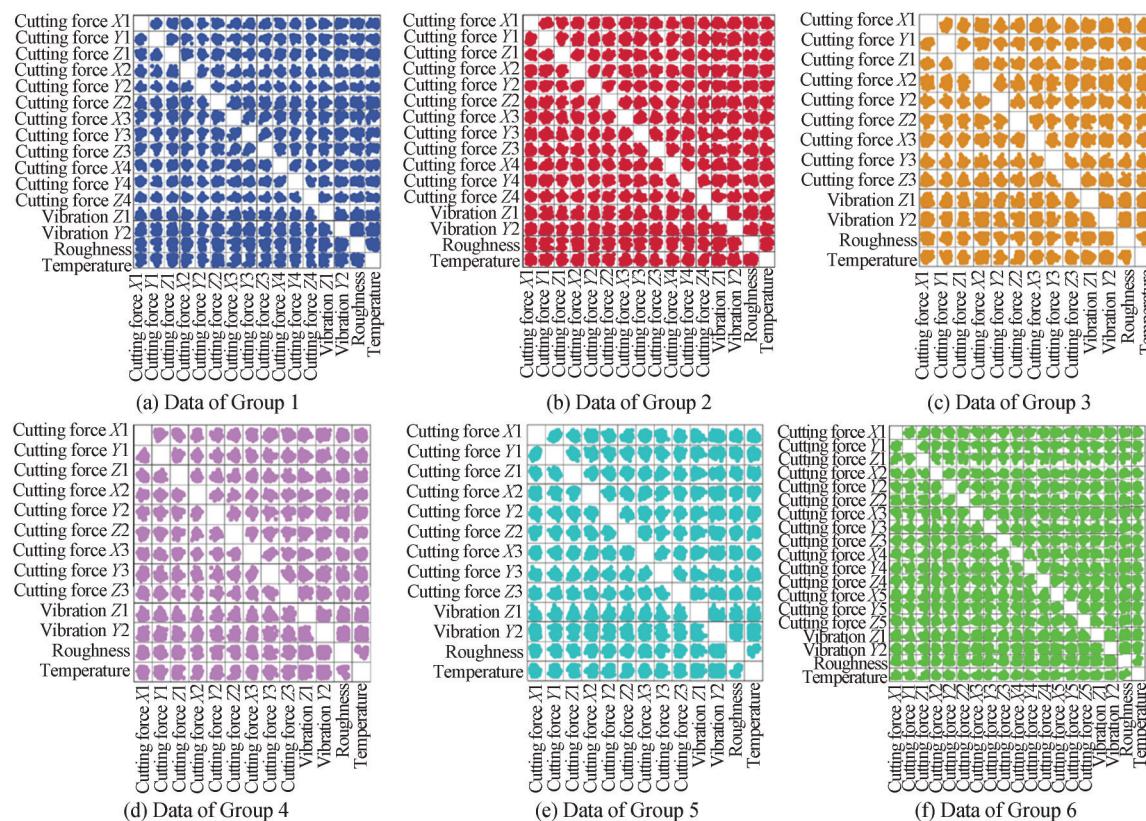


Fig. 14 Scatter plots of the correlation of the first set of experimental

**Table 2 Results of correlation analysis**

Group	Vibration	Roughness	Temperature	Other
1.1	—	—	—	✓
1.2	✓	✓	—	—
1.3	—	✓	✓	—
1.4	—	—	✓	—
2.1	✓	—	—	—
2.2	—	—	—	✓
2.3	—	✓	—	—
2.4	✓	—	—	—
3.1	✓	—	✓	—
3.2	—	✓	—	—
3.3	✓	—	—	—
⋮	⋮	⋮	⋮	⋮
41.1	—	—	—	✓
41.2	✓	✓	—	—
42.1	✓	✓	—	—
42.2	—	—	✓	—
42.3	✓	—	—	—
43	—	✓	✓	—
44	—	—	—	✓
45.1	—	—	✓	—
45.2	✓	—	✓	—
46	✓	—	—	—

"1.1, 1.2, 1.3, ..." refer to the different classes corresponding to the original data after clustering; "43, 44, 46, ..." are each in their own class; ✓ means with significant correlation

temperature and roughness data at the same time. Therefore, the milling force data obtained from all experimental groups can be classified into eight major classes as a whole: 1) vibration as the main factor; 2) temperature as the main factor; 3) roughness as the main factor; 4) vibration and temperature as the main factors; 5) vibration and roughness as the main factors; 6) temperature and roughness as the main factors; 7) vibration, temperature and roughness all as the main factors; 8) vibration, temperature and roughness are not major factors. The data modeling should be performed later to match with them and be handled separately.

### 3 Data Modeling

As mentioned above, the dynamics of milling forces are divided into eight categories. The relationship

between the dynamic milling force and corresponding independent variables was also different because of the different main categories of each type of data. According to the characteristics of different types of data, the dynamic milling force prediction model with high fitting accuracy is obtained by modeling the classified data with multi-element regression analysis and RBF neural network, respectively.

#### 3.1 Regression Analysis of Dynamic Cutting Force

Regression analysis is an important analysis tool for data mining, which has been widely used in modeling of various fields for its high efficiency, intuitionistic and accurate characteristics. In practical processing, the change of a certain index is caused by the coupling of several parameter variables, so multiple regression analysis is the most commonly used regression method in engineering application.

The current experimental formula of milling force is usually expressed in exponential form, which consists of correction factor, cutting process parameter and exponential factor<sup>[17]</sup>. The available process parameters are cutting width, cutting depth, cutting speed, feed per tooth, feed, spindle speed, milling cutter diameter and number of teeth. However, there is an absolute correlation between the feed speed, the number of teeth per tooth and the spindle speed, and between the cutting speed, spindle speed and the diameter of the milling cutter.

With consideration of degree of close correlation between dynamic cutting forces and time, the time parameter is introduced into the model at the same time. The regression model for the dynamic cutting forces was initially determined to contain the following six parameters: depth of cut  $a_p$ , feed per tooth  $f_z$ , cutting width  $a_e$ , cutting speed  $v_c$ , number of teeth of the milling tool  $z$  and time  $t$ . The model expression is:

$$F = A a_p^B f_z^C a_e^D t^E v_c^G z \quad (2)$$

Compared with the general model of multi-linear regression, the original model is in the form of polynomial summation, and the model is in the form of exponential class product, so the model is needed to deform, which are obtained from equations as follows:

$$\lg F = \lg A z + B \lg a_p + C \lg f_z + D \lg a_e + E \lg t + G \lg v_c \quad (3)$$

Let  $y = \lg F$ ,  $b = \lg A z$ ,  $x_1 = \lg a_p$ ,  $x_2 = \lg f_z$ ,  $x_3 = \lg a_e$ ,  $x_4 = \lg t$ ,  $x_5 = \lg v_c$ , then the regression model is trans-

formed into

$$y = b + Bx_1 + Cx_2 + Dx_3 + Ex_4 + Gx_5 \quad (4)$$

The independent variable and dependent variable data were processed using the least squares. To find the constant term  $B$  of a polynomial, add a column of 1 to the data of the independent variable, and the first coefficient is the constant term.

Through correlation analysis, take the data of "vibration as the main factor" as an example, carry on logarithmic analysis to the data sample, then carry on the regression analysis. The results show that the judgment parameters of the regression model are not ideal and from the residual analysis graph it is proved that the confidence intervals of many residuals do not contain zeros.

(i) When the first regression analysis of the data is not satisfactory, the reasonable expectation of goodness of fit  $R^2$  is given first, and then the iteration is repeated. In the process of iteration, the original data without zero in all residual confidence intervals are eliminated, and the remaining data are retained to form new independent and dependent data samples. And then regression analy-

sis is carried out again, until the goodness of fit is met.

(ii) At this point, according to the latest residual analysis image, adjust the fitting degree properly, and repeat the iterative process in step (1) until the evaluation index can judge the model more accurately.

Figure 15 shows the residual analysis of the first major category of data after the residuals are eliminated. The red lines in the two yellow circles in the graph indicate abnormal data, while the remaining green lines indicate the zero confidence interval from normal data to the residuals. The small circle in the centre of each line indicates the value of the residuals, while the whole line indicates the confidence interval of the residuals.

As shown in Fig. 15, after the exception data are culled, there are only 2 exception data left, which are negligible compared with the nearly 6 000 data samples. The analysis of the other scoring criteria showed that the  $R^2$  of fit was 0.776 6, the  $F$ -value was 3 989.293, and the probability of error  $P$  was 0. Therefore, it is clear that the constructed regression model is successful. The final regression analysis results for the eight categories of data are shown in Table 3.

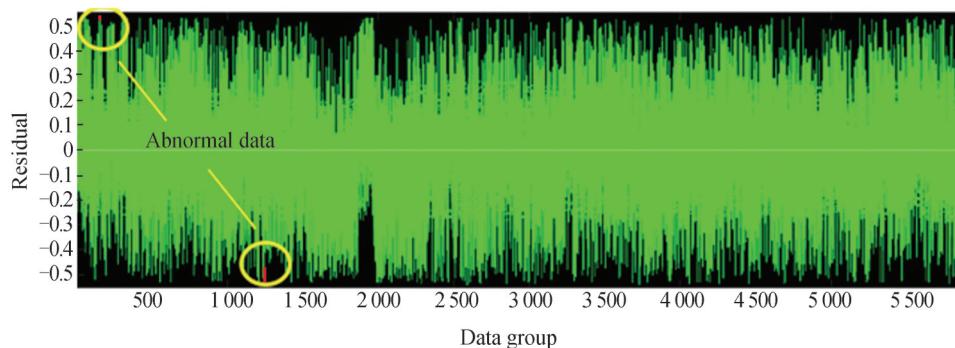


Fig. 15 Residual analysis graph

After judging and analyzing the final results of the above eight types of multiple linear regression models, it is found that the first five dynamic cutting force multiple linear regression models, "vibration was the main factor", "vibration and roughness were the main factors", "vibration and temperature were the main factors", "vibration, roughness and temperature were the main factors", and "vibration, temperature and roughness were not the main factors", fit better, all at 0.6  $F$  values are large, the probability of error is zero, and the number of confidence intervals for the residuals without zeros is less than 10, which can be directly applied. Their linear fit equations are:

1) Vibration is the main factor affecting the dynamic milling force variation:

$$F = 1.1594 \times 10^{23} a_p^{0.5397} f_z^{-4.9559} a_e^{-0.1141} t^{-0.0118} v_c^{2.9676} z \quad (5)$$

2) Vibration and roughness are the main factors affecting the dynamic milling force variation:

$$F = 0.0003 a_p^{-0.2574} f_z^{-1.1646} a_e^{-0.0949} t^{-0.0146} v_c^{-0.0751} z \quad (6)$$

3) Vibration and temperature are the main factors affecting the dynamic milling force variation:

$$F = 0.2500 a_p^{-0.2328} f_z^{-0.4130} a_e^{-0.0923} t^{-0.0199} v_c^{-0.1460} z \quad (7)$$

4) Vibration, roughness and temperature are the main factors affecting the dynamic milling force variation:

$$F = 3.6833 \times 10^6 a_p^{-0.1234} f_z^{0.5339} a_e^{0.1387} t^{-0.0663} v_c^{-1.2183} z \quad (8)$$

**Table 3 Results of regression analysis of milling force**

Group	<i>b</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>G</i>
Vibration	23.666 3	$1.1594 \times 10^{23}$	0.539 7	-4.955 9	-0.114 1	0.011 8	2.967 6
Vibration+roughness	-2.991 2	0.000 3	-0.257 4	-1.164 6	-0.094 9	-0.014 6	-0.075 1
Vibration+temperature	0	0.250 0	-0.232 8	-0.413 0	-0.092 3	-0.019 9	-0.146 0
Vibration+temperature+roughness	7.168 3	$3.6833 \times 10^6$	-0.123 4	0.533 9	0.138 7	-0.066 3	-1.218 3
Other	-4.001 4	$2.5 \times 10^{-5}$	-0.156 3	-1.295 2	0.586 3	0.011 9	0.934 0
Roughness	5.637 7	$1.0855 \times 10^5$	0.478 2	0.510 6	-0.060 8	-0.004 6	0.106 4
Temperature	1.682 2	12.026 5	-0.052 6	0.053 1	0.031 0	-0.077 0	0.460 1
Temperature+roughness	0.282 0	0.478 6	0.123 1	0.331 6	-0.040 2	-0.031 2	0.528 6

5) Vibration, temperature and roughness are not the main factors influencing the dynamic milling force variation (i. e. "other"):

$$F = 2.5 \times 10^{-5} a_p^{-0.1563} f_z^{-1.2952} a_e^{0.5863} t^{0.0119} v_c^{0.9340} z \quad (9)$$

However, multiple linear regression models were constructed for the main factors of roughness, temperature, and temperature and roughness. Although the *F*-values, *P*-values and residual analysis were good, the fit was low at 0.266 5, 0.184 1 and 0.272 1, respectively, indicating that the three models were poorly fitted and not very accurate. Therefore, there was no significant linear or fitted linear relationship between the logarithms of the independent and dependent variables for the latter three types of data. By combining their scatter plots with the number of independent variables, it was found that the scatter plots did not fit the common non-linear regression model. This result supports the conjecture that there may be a non-linear relationship between the independent and dependent variables of these three types of data that cannot be expressed directly, or that there is no direct correlation between the three types of data. Therefore, the latter three types of data could be further analyzed by using RBF neural network modeling.

### 3.2 RBF Neural Network Modeling of Dynamic Cutting Force

RBF neural network has a more intuitive structure, including the input layer, the implicit layer and the output layer. The input layer and the output layer respectively place the data samples of the independent variable and the dependent variable, and the number of input neurons is the number of independent variables. The im-

plicit layer can be seen as a "bridge". When building a network, there are two transformations: the nonlinear transformation from the input layer to the implicit layer, and the linear transformation from the implicit layer to the output layer. It can be seen that the implicit layer plays a "bearing" role in the construction of the RBF neural network. In fact, when the vector is at a low latitude, there will be a linear inseparability, and the mapping of the vector from low latitude to high latitude can be realized through the implicit layer, so as to achieve the purpose of linear separability.

Assuming that there is a non-linear relationship between the independent and dependent variables of the three types of data mentioned above and that the form of the relationship is complex and the deterministic relational equation can be obtained, we choose to use RBF neural networks for prediction and error analysis modelling. To build the model, a total of six neurons were set up in the input layer and  $lga_p$ ,  $lgf_z$ ,  $lga_e$ ,  $lgv_c$ ,  $lgz$  and  $lgf_t$  were entered into each neuron. The output layer has only one neuron and stores the milling force data in it. The three categories of data were divided into two separate parts. The last 200 data sets of each category were used as test samples and the rest of the data were used as training samples for the construction of the neural network.

To improve the accuracy of the neural network, relatively reasonable expansion coefficients were set for each of the three types of data. The schematic diagram of the neural network model is shown in Fig. 16. Finally, RBF neural network models were constructed for the three main factors of roughness (expansion coefficient

of 11), temperature (expansion coefficient of 11) and temperature and roughness (expansion coefficient of 11) respectively. The fitted curves and error curves for the experimental values of the test samples are shown in Figs. 17 and 18 for the network model with roughness and temperature as the main factors. The red curve in

Fig. 17 indicates the fitted value of the neural network, and the blue line curve indicates the experimental value, both of which have similar trends and relatively small differences. Combined with the error plot, it can be tentatively judged that the constructed model fits well and the error is relatively small and within the acceptable

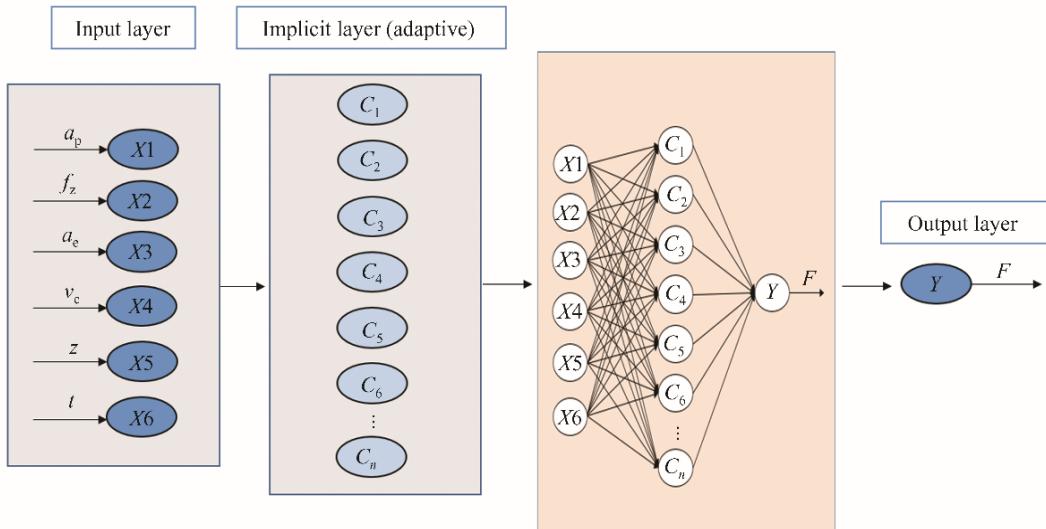


Fig. 16 Schematic diagram of RBF neural network model

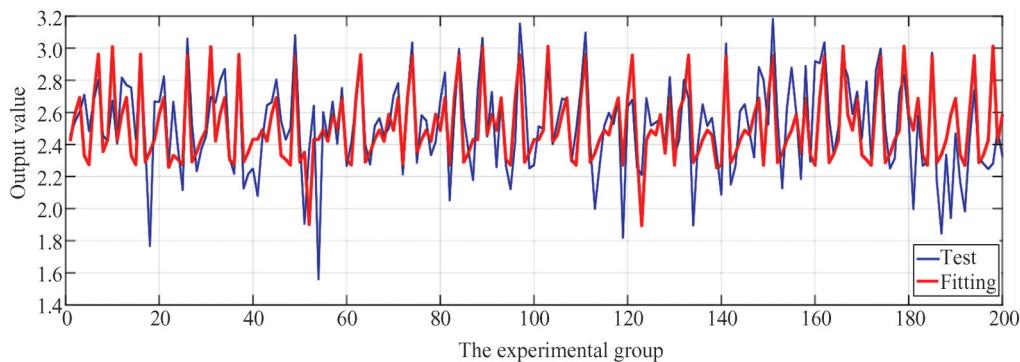


Fig. 17 Fitting curve of RBF neural network

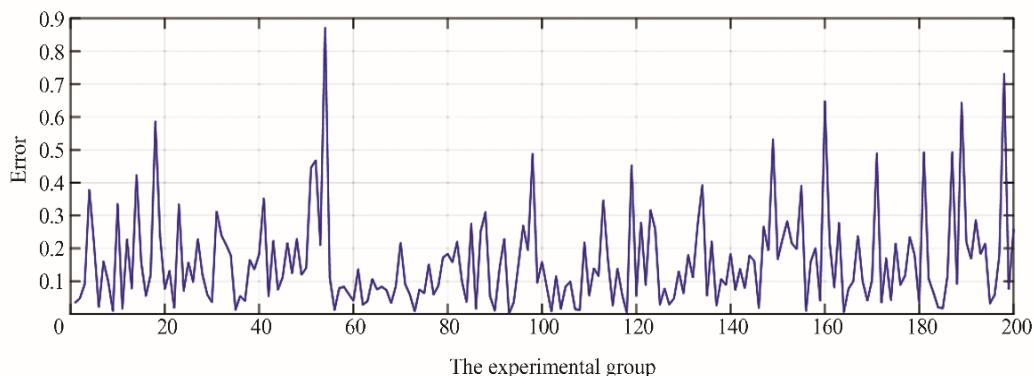


Fig. 18 RBF neural network error profile

range. To further judge the reasonableness of the model, mean squared error (known as MSE) is introduced, which was calculated as:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (T_i - \bar{T}_i)^2 \quad (10)$$

The MSE of the neural network model is 0.045 9 with temperature and roughness as influencing factors, 0.040 3 with temperature as influencing factor, and 0.040 4 with roughness as influencing factor. The MSE of the three models is very small, and the RNF neural network model has good prediction accuracy and can be applied in practice.

Comparing the modeling of the above regression analysis, it can be found that not all the relationships between dynamic milling force data and cutting process parameters can be obtained by multiple linear regression analysis, especially when considering the dynamic milling force variation at different times.

For the dynamic milling force data, where the vibration factor is not one of the main influencing factors, it is more meaningful to construct a model using RBF neural network. At present, the constructed model can directly predict the dynamic milling force by cutting process parameters and time, and then combine with the dynamic displacement change to obtain the dynamic stiffness of the milling tool system, but the specific expression of the dynamic stiffness of the milling tool system cannot be obtained.

## 4 Conclusion

A combination of data mining and cluster analysis with correlation analysis studies is conducted in this paper. The results show that the dynamic cutting force data is divided into eight main categories based on correlation as follows. "Vibration is the main factor", "vibration and temperature are the main factors", "vibration and roughness are the main factors", "vibration, temperature and roughness are the main factors", "vibration, temperature and roughness are not the main factors", "temperature is the main factor", "roughness is the main factor", "both temperature and roughness are the main factors".

Multi-linear regression analysis was applied to model the data of "vibration as the main factor", "vibration and temperature as the main factor", "vibration and roughness as the main factor", "vibration, temperature and roughness as the main factor", "vibration, temperature and roughness are not the main factors", and fit the

expressions to the five types of data. The three types of data, "temperature is the main factor", "roughness is the main factor", and "temperature and roughness are the main factors", were modeled by RBF neural network, and achieved high prediction accuracy.

Finally, with two prediction models, better predictions were obtained, providing a more accurate basis for selecting process parameters, ensuring production quality and reducing the occurrence of chatter.

## References

- [1] Lv L S, Deng Z H, Liu T, et al. Intelligent technology in grinding process driven by data: A review [J]. *Journal of Manufacturing Processes*, 2020, **58**: 1039-1051.
- [2] Kang H S, Lee J Y, Choi S, et al. Smart manufacturing: Past research, present findings, and future directions [J]. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 2016, **3**(1): 111-128.
- [3] Liu X L, Liu Q, Yue C X. Intelligent machining technology in cutting process [J]. *Journal of Mechanical Engineering*, 2018, **54**(16): 45-61(Ch).
- [4] Kantardzic M. *Data Mining: Concepts, Models, Methods, and Algorithms* [M]. Hoboken: John Wiley & Sons, Inc., 2011.
- [5] Wang M, Zhu J Y, Zhang Y Z. Fuzzy pattern recognition of metal cutting state (II)—Fuzzy ISODATA clustering analysis method [J]. *Journal of Aeronautics*, 1985, **6**(2): 173-179(Ch).
- [6] Cai Y, Shao H, Yuan J. *A Clustering Analysis Method of Energy Efficiency in Milling Process* [P]. CN: 110428000A, 2019-11-08.
- [7] Bracke S, Radetzky M, Rosebrock C, et al. Efficiency and effectiveness of high precision grinding manufacturing processes: An approach based on combined DEA and cluster analyses [J]. *Procedia CIRP*, 2019, **79**: 292-297.
- [8] Pei H J, Chen Y Y, Li G A, et al. Correlation between milling force, vibration and surface roughness based on copula function [J]. *Aeronautical Manufacturing Technology*, 2019, **62**(9): 59-67(Ch).
- [9] Rodrigues J R P, Gonçalves R S, Araújo A S Jr, et al. Correlation analysis of the roughness as function of cutting speed and feed rate (the roughness as function of machining parameters) [J]. *Acta Scientiarum Technology*, 2019, **41**(1): 37228.
- [10] Thasthakeer A T, Farid A A, Seng C T, et al. Analysis of the correlation between fractal structure of cutting force signal and surface roughness of machined workpiece in end milling operation [J]. *Fractals*, 2019, **27**(2): 1950013.

- [11] Lin Y C, Wu K D, Shih W C, et al. Prediction of surface roughness based on cutting parameters and machining vibration in end milling using regression method and artificial neural network[J]. *Applied Sciences*, 2020, **10**(11): 3941.
- [12] Shihab S K, Khan Z A, Mohammad A, et al. RSM based study of cutting temperature during hard turning with multi-layer coated carbide insert [J]. *Procedia Materials Science*, 2014, **6**: 1233-1242.
- [13] Rajesh K, Murali Krishnam Raju V V, Rajesh S, et al. Effect of process parameters on machinability characteristics of CO<sub>2</sub> laser process used for cutting SS-304 stainless steels [J]. *Materials Today: Proceedings*, 2019, **18**: 2065-2072.
- [14] Yeganeifar A, Niknam S A, Asadi R. The use of support vector machine, neural network, and regression analysis to predict and optimize surface roughness and cutting forces in milling [J]. *The International Journal of Advanced Manufacturing Technology*, 2019, **105**(1): 951-965.
- [15] Maher I, Eltaib M E H, Sarhan A A D, et al. Cutting force-based adaptive neuro-fuzzy approach for accurate surface roughness prediction in end milling operation for intelligent machining [J]. *The International Journal of Advanced Manufacturing Technology*, 2015, **76**(5): 1459-1467.
- [16] Basak S, Dixit U S, Davim J P. Application of radial basis function neural networks in optimization of hard turning of AISI D2 cold-worked tool steel with a ceramic tool [J]. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 2007, **221**(6): 987-998.
- [17] Wang P B. Radial Basis Function (RBF) neural network in gearbox mechanical fault diagnosis [C] // *Abstracts of the 4th Advanced Materials, Mechanics and Structural Engineering Conference (AMMSE 2017)*. Beijing: IOP Publishing, 2017: 321-325.

□