

# Experimental Investigation of Tool Wear in End-Milling Operations using RSM and ML Approaches

Le Xuan Loi<sup>1\*</sup>, Nguyen Van Toan<sup>1</sup>, Nguyen Tai Hoai Thanh<sup>1</sup>

<sup>1</sup>Department of Machine Manufacturing, Faculty of Mechanical Engineering, Le Quy Don Technical University

\*Email address: mr.loimta02@gmail.com

**Abstract**— Accurate prediction of tool wear is critical for improving productivity and maintaining product quality in modern manufacturing systems. Flank wear is commonly adopted as a key tool life indicator due to its significant impact on machining accuracy, process stability, and operational reliability. This study presents and compares two predictive approaches, namely a regression based response surface methodology (RSM) and machine learning (ML) models for estimating tool wear in milling operations. In this work, flank wear is defined as the response variable, while cutting speed, feed rate, and depth of cut are considered as input parameters. A Design of Experiments (DOE) methodology with three factors at three levels is employed to systematically design the experiments. Machining trials are conducted on 40Cr metal alloy using a five-axis milling machine and a TiAlN-coated single-insert ball-end milling cutter with two cutting edges and a helix angle of 30°. The experimental data are used to develop the regression model. A comparative analysis between predicted and measured values indicates that ML model provides higher prediction accuracy within the trained range.

**Keywords**— Response Surface Methodology (RSM); machine learning (ML); Tool wear.

## I. INTRODUCTION

Tool wear and tool life are widely recognized as critical determinants of machining system productivity, manufacturing cost, product quality, and surface integrity. Premature or catastrophic tool failure can result in significant productivity losses, part rejection, and substantial economic penalties. Among various wear mechanisms, flank wear occurring on the relief face due to continuous rubbing against the machined surface is generally regarded as the dominant mode in most cutting operations and is frequently adopted as a primary criterion for tool life evaluation [1,2]. In contrast, crater wear becomes significant only under specific machining conditions, particularly at elevated cutting speeds and feed rates, where intensified thermal and chemical interactions at the tool chip interface accelerate material removal from the rake face [3,4].

The progression of flank wear has a direct impact on surface quality, as it alters the effective geometry of the cutting edge. Specifically, an increase in flank wear reduces the tool nose radius, thereby degrading the surface finish of the machined workpiece [5,6]. Moreover, excessive flank wear contributes to higher cutting forces and increased vibration levels, which further deteriorate machining stability and dimensional accuracy [7,8]. Consequently, controlling tool wear within acceptable limits is essential for ensuring process reliability and achieving cost-effective manufacturing. Maximizing tool utilization without compromising product quality remains a key objective in modern machining industries.

Tool wear behavior is strongly influenced by machining parameters, including cutting speed, feed rate, and depth of cut. Therefore, the development of accurate predictive models that relate these input variables to tool wear is crucial for process optimization. In this context, statistical and data-driven modeling approaches have been extensively employed. The Design of Experiments (DOE) methodology provides a systematic framework for planning and conducting experiments, enabling efficient data acquisition and robust analysis of factor interactions [4,9]. Regression-based models are commonly used to quantify the relationship between process parameters and tool wear, offering insights into the relative significance of control variables. The adequacy of such models is typically validated through analysis of variance (ANOVA), ensuring statistical reliability.

Furthermore, advanced computational techniques, such as artificial neural networks (ANNs) [10], have been increasingly utilized to enhance prediction accuracy by capturing nonlinear relationships inherent in machining processes [4,9,11]. Comparative analysis between experimental observations and model predictions serves to validate the robustness of these approaches. Ultimately, a comprehensive understanding of tool wear mechanisms and predictive modeling enables informed selection of machining parameters, thereby minimizing wear progression, improving surface quality, and enhancing overall process efficiency.

## II. EXPERIMENTAL AND METHODS

### 2.1. Description of the experimental setup



Fig. 1. Schematic of the experimental setup

Figure 1 presents a schematic representation of the experimental setup. All experiments were conducted on a five-axis CNC milling system (SPINNER U5-620, Germany),

comprising a multi-axis motion platform, workpiece, cutting tool, and an electron microscopy-based measurement system for tool wear evaluation. The workpiece material was 40Cr alloy steel. Prior to machining, the workpiece underwent volume hardening in a molten salt bath to ensure uniform hardness distribution and enhanced hardenability. Subsequently, both the top and bottom surfaces were ground to establish precise reference planes, thereby ensuring accurate positioning and machining angles. Surface hardness was measured at five uniformly distributed locations, and the average value was reported, with deviations at each point limited to within  $\pm 5\%$ . The cutting tool employed was a TiAlN-coated ball end mill (QBLMx1010-R5-75-10) featuring two cutting edges with a helix angle of  $30^\circ$ . Tool condition monitoring and flank wear measurement were carried out in situ using an electron microscopy-based system. This setup incorporated a digital microscope (AD7013MT) installed directly within the machining zone and integrated with dedicated computer software for real-time data acquisition and analysis, as illustrated in Figure 2a. Owing to its compact design, the device can be conveniently mounted within the working space without interfering with the machining process, while ensuring efficient data transfer and processing [102]. As the observed tool wear approached the predefined threshold, high-precision measurements of flank wear were subsequently performed using an AXIO-A2M microscope (Zeiss), as shown in Figure 2b.

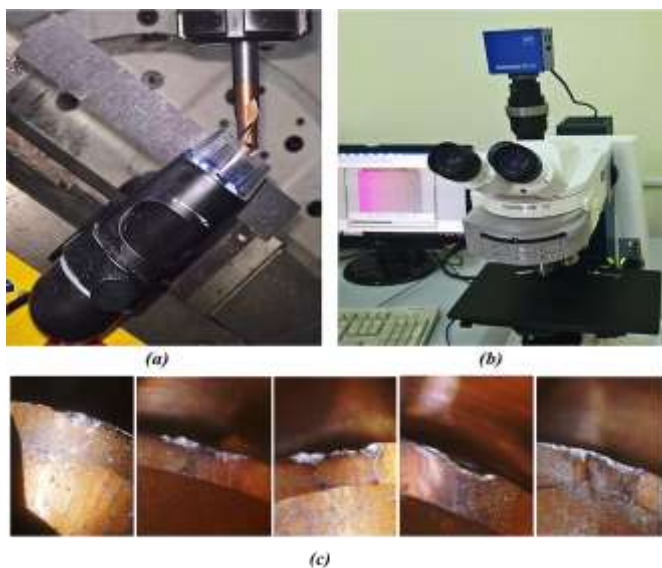


Figure 2. Microscopic system for flank wear measurement: (a) in-process measurement microscope; (b) verification microscope; and (c) Morphology of wear on the cutting edge of the tool.

The morphology of wear on the cutting edge, as illustrated in Fig. 2c, reveals a pronounced irregular degradation pattern along the tool edge. The worn region exhibits a combination of micro-chipping and localized material adhesion, indicating the coexistence of abrasive and adhesive wear mechanisms. The presence of adhered material layers on the rake face implies material transfer from the workpiece to the tool, which is characteristic of high-temperature machining conditions.

Machining conditions are governed by various process parameters that significantly influence tool wear. The proper identification and selection of these parameters are essential for achieving improved surface finish while minimizing tool degradation. Desired tool life can be attained through the appropriate adjustment of independently controllable variables that directly affect surface quality. Among the numerous controllable parameters, cutting speed, feed rate, and depth of cut were selected as the primary factors for conducting the experiments and developing the corresponding mathematical model.

## 2.2. Methodology

### 1. Response surface methodology (RSM)

In this study, the experimental design was established using Response Surface Methodology (RSM). The data obtained from the designed experiments (DoE) were employed to develop a mathematical response surface model and to train artificial neural network (ANN) models. The input variables considered were cutting speed ( $x_1$ ), feed rate ( $x_2$ ), and depth of cut ( $x_3$ ), while tool wear ( $Y$ ) was defined as the response variable. To ensure the robustness and generalizability of both the regression and ANN models, the process parameters were selected to span a wide and representative machining range, as presented in Table 1.

TABLE 1. The experimental factors and their corresponding levels

Independent variables	Units	Factor Level		
		Level 1(-1)	Level 2(0)	Level 3(+1)
Cutting speed	m/min	120	150	180
Feed rate	mm/tooth	0.05	0.1	0.15
Depth of cut	mm	0.1	0.15	0.2

For the experimental design, a three-level full factorial Design of Experiments (DoE) was initially adopted. Table 1 presents the experimental factors along with their corresponding coded and uncoded levels. A Central Composite Design (CCD) combined with a quadratic model involving three variables was subsequently employed to investigate the effects of the independent variables and to determine their optimal combination [12]. Based on the CCD framework, a total of 20 experimental runs were conducted, including six axial points, eight factorial points, and six replicates at the center point. The experimental runs were performed in a randomized order to minimize the influence of uncontrolled factors.

A quadratic polynomial model was employed to establish the relationship between the predicted response (surface roughness) and the independent variables. The mathematical expression of the proposed response model is presented in Eq. (1):

$$Y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k \beta_{ij} x_i x_j \quad (1)$$

where  $Y$  denotes the response variable;  $x_i, x_j, \dots, x_k$  represent the input variables;  $x_i^2, x_j^2, \dots, x_k^2$  correspond to the quadratic terms; and  $x_i x_j, x_i x_k, \dots, x_j x_k$  denote the interaction terms. The coefficient  $\beta_0$  is the intercept, while  $\beta_i, \beta_{ii}$  and  $\beta_{ij}$  represent the

linear, quadratic, and interaction coefficients, respectively, which are estimated from the experimental data [13].

The adequacy of the regression models was evaluated using the coefficient of determination ( $R^2$ ) and the adjusted coefficient of determination  $R^2_{adj}$ . Analysis of variance (ANOVA) was employed to assess the statistical significance and overall suitability of the developed models. The significance of the regression coefficients in the mathematical model was determined based on the F-value and corresponding p-value. The RSM analysis was conducted using specialized computational tools to establish the mathematical relationships between input and response variables. In this study, Design-Expert software (Version 12, Stat-Ease Inc., USA) was utilized for experimental design, data analysis, development of nonlinear models, and optimization of the response variables.

2. Machine Learning Model Development

In this study, widely adopted ML models (including Backpropagation Neural Network (BPNN)), were employed to predict tool wear. Before model training, the dataset was normalized using the MinMaxScaler technique to rescale all features to the range of 0 to 1. The dataset was partitioned into a training set (70%) and a testing set (30%) using the train\_test\_split function from the Scikit-learn’s model selection module in Python 3.9.7. The trained models were serialized using Python’s pickle module for subsequent multi-objective optimization. The performance of the developed models was evaluated using standard statistical metrics, as defined by the equations presented below.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2)$$

where  $y_i$  represents the actual value,  $\hat{y}_i$  denotes the predicted value,  $\bar{y}_i$  is the average value, and n is the number of samples.

TABLE 2. Design matrix values and responses

Run order	Codes values (x)			Tool wear (mm)		
				Measured values	Predicted value (RSM)	Predicted value (ML)
	$x_1$ (m/min)	$x_2$ (mm/rev)	$x_3$ (mm)			
1	150	0.1	0.15	0.183	0.1794	0.181
2	120	0.05	0.1	0.138	0.1252	0.137
3	150	0.1	0.2	0.209	0.2164	0.207
4	150	0.1	0.15	0.175	0.1794	0.176
5	150	0.05	0.15	0.141	0.1525	0.144
6	120	0.1	0.15	0.155	0.1658	0.153
7	180	0.05	0.1	0.168	0.1635	0.167
8	150	0.1	0.15	0.18	0.1794	0.181
9	180	0.15	0.2	0.235	0.2339	0.235
10	180	0.1	0.15	0.223	0.2319	0.225
11	150	0.1	0.15	0.181	0.1794	0.18
12	150	0.1	0.15	0.178	0.1794	0.177
13	150	0.15	0.15	0.235	0.2432	0.238
14	150	0.1	0.1	0.178	0.1903	0.181
15	120	0.15	0.2	0.203	0.1936	0.201
16	120	0.05	0.2	0.23	0.2267	0.231
17	180	0.15	0.1	0.315	0.3044	0.313
18	150	0.1	0.15	0.183	0.1794	0.181
19	180	0.05	0.2	0.192	0.182	0.185
20	120	0.15	0.1	0.185	0.1811	0.183

III. RESULTS AND DISCUSSION

A. Regression mathematical modeling for tool wear

The design matrix based on the CCD method, together with the experimental and predicted tool wear values, is presented in Table 2. The results indicate no significant discrepancy between the measured and predicted values, demonstrating good agreement and confirming the reliability of both the RSM and ML models. An empirical multivariate regression model describing the relationship between machining parameters and tool wear was developed using the quadratic polynomial form given in Eq. (1). The model coefficients were determined from the experimental data through analysis of variance (ANOVA). The resulting best-fit model for tool wear is presented in Eq. (3).

$$Y = 0.179 + 0.0196x_1 + 0.0269x_2 + 0.0077x_3 + 0.0212x_1x_2 - 0.0207x_1x_3 - 0.0222x_2x_3 + 0.0068x_1^2 + 0.0065x_2^2 + 0.0085x_3^2. \quad (3)$$

The developed equation can be employed to predict tool wear for arbitrary combinations of the input parameters within

the studied range. Analysis of variance (ANOVA) was conducted to evaluate the adequacy and significance of the regression model.

The results presented in Table 3 indicate that the experimental data are well represented by a quadratic polynomial model. The coefficient of determination ( $R^2$ ) for tool wear (Y) was found to be 0.961, implying that approximately 96.1 % of the variability in the response can be explained by the selected independent variables. Moreover, the close proximity of the  $R^2$  value to unity suggests that the influence of the cutting parameters on the response is effectively captured by the quadratic model, demonstrating a strong agreement between the predicted and experimental results [14].

B. Interaction effect of machining parameters on tool wear

Based on the analysis of variance (ANOVA) results (Table 3) and the response surface model presented in Eq. (3), the interactive effects of the independent variables namely cutting speed, feed rate, and depth of cut on tool wear were systematically evaluated. Equation (3) provides the predictive

model for tool wear derived from the experimental data. The analysis indicates that feed rate is the most influential parameter affecting tool wear, followed by cutting speed, while depth of cut exhibits a comparatively minor effect. Figure 3 further illustrates the influence of each machining parameter on tool

wear within the investigated range. The results consistently confirm that feed rate has the most pronounced impact on tool wear, whereas the effect of depth of cut is negligible. This limited influence may be attributed to the relatively narrow range of depth of cut selected in the present study.

TABLE 3. ANOVA analysis for the tool wear quadratic model.

Response	Model terms	Sum of square	Degree of freedom	Mean square	F-value	p-value (Prob. >F)		
Surface roughness	Mode	0.179	9	0.003	27.579	< 0.0001	significant	
	$x_1$ -cutting speed	0.125	1	0.005	45.163	< 0.0001		
	$x_2$ -Feed rate	0.216	1	0.009	84.966	< 0.0001		
	$x_3$ -Depth of cut	0.179	1	0.0008	7.0642	0.024		
	$x_1x_2$	0.152	1	0.003	30.938	0.0002		
	$x_1x_3$	0.165	1	0.003	29.499	0.0003		
	$x_2x_3$	0.163	1	0.003	33.919	0.0002		
	$x_1^2$	0.179	1	0.0006	5.8215	0.0365		
	$x_2^2$	0.233	1	0.0006	5.2375	0.0451		
	$x_3^2$	0.231	1	0.001	8.8311	0.014		
	Residual		0.179	10	0.00012			
	Total		0.03	19	0.00023			
	Lack of Fit		0.243	5	9.6E-06	23.326		0.018
Pure Error		0.19	5	0.00322				

$R^2 = 0.961$ , Adj.  $R^2 = 0.926$ , Adequate precision = 23.45, C.V. = 5.56 %

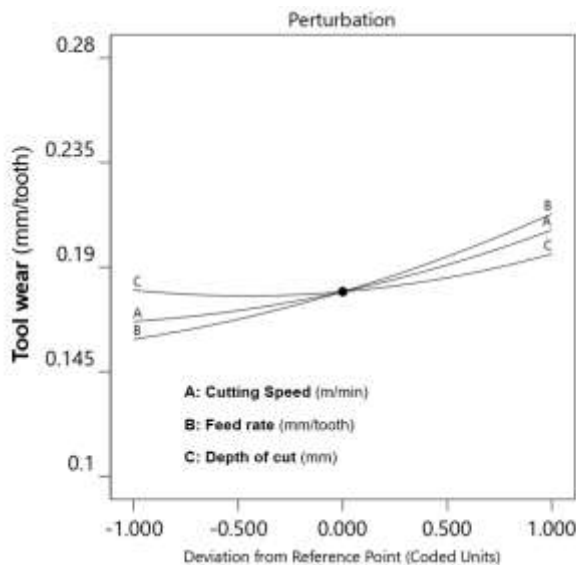


Fig. 3. The influence of each machining parameter on tool wear

The response surface plots derived from the quadratic polynomial regression model reveal the complex interactions among the independent variables, as illustrated in Fig. 4. These plots were generated by varying two variables within the experimental range while maintaining the third variable at its center level. Figure 4a depicts the combined effect of cutting speed and feed rate on tool wear. The results indicate that increasing the feed rate leads to a significant increase in tool wear, demonstrating its detrimental effect. Overall, both variables exhibit a strong influence; however, feed rate is identified as the dominant factor, followed by cutting speed. Notably, an increase in cutting speed tends to reduce tool wear within the investigated range. The interaction between cutting speed and depth of cut is presented in Fig. 4b. The results suggest that higher values of both cutting speed and depth of cut contribute to increased tool wear. Figure 4c illustrates the

interaction effect between feed rate and depth of cut, obtained by varying these two parameters while keeping cutting speed constant at its center value. It can be observed that tool wear exhibits only a slight variation as the depth of cut increases from 0.1 to 0.2 mm, indicating a relatively weak influence of this parameter. Figure 4d presents the contour plot, which estimates tool wear as a function of feed rate and cutting speed. In summary, the RSM analysis confirms that tool wear is predominantly influenced by feed rate, followed by cutting speed, while the effect of depth of cut remains comparatively insignificant within the investigated range.

### C. Machine learning-based predictive capability

Process parameters are intrinsically linked to tool wear in machining; therefore, establishing a robust mapping between input variables and tool wear is of critical importance. Figure 5 illustrates the predictive performance of both the RSM and ML models. As shown, the data points for both the training and testing sets are generally clustered along the diagonal line, indicating satisfactory overall predictive accuracy for tool wear. Nevertheless, a limited number of data points exhibit noticeable deviations between predicted and experimental values, which may be attributed to the models' limited ability to capture local variations or outliers, as well as potential data imbalance within the dataset.

In comparison, the RSM model (Fig. 5a) exhibits greater dispersion and occasional inconsistencies in prediction accuracy, with relatively larger errors observed in certain regions, resulting in an  $R^2$  value of 0.936. In contrast, the ML model (Fig. 5b) demonstrates a more stable predictive trend, characterized by smaller deviations between predicted and experimental values, and achieves a higher  $R^2$  value of 0.972. This reflects a stronger agreement with experimental observations and improved generalization capability. Overall, the selected input variables account for approximately 97.2% and 93.6% of the variance in tool wear for the ML and RSM models, respectively, thereby confirming the superior

predictive performance of the ML model in capturing the complex relationships between process parameters and tool wear. Furthermore, the comparative analysis indicates that the

ML model provides more reliable and consistent predictions, underscoring its effectiveness in modeling the nonlinear characteristics of the end-milling process.

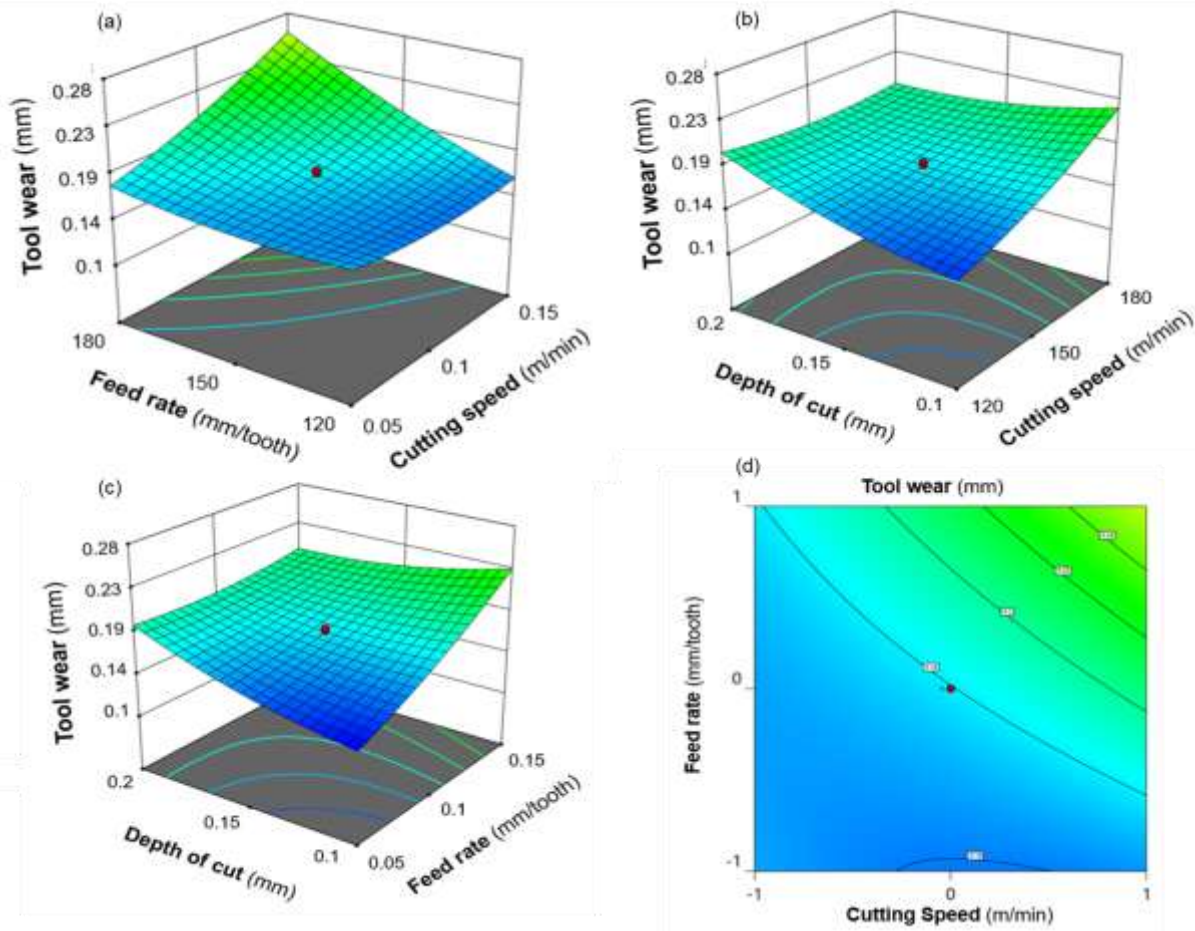


Fig. 4. Three-dimensional (3D) plot of the cutting parameters against the tool wear

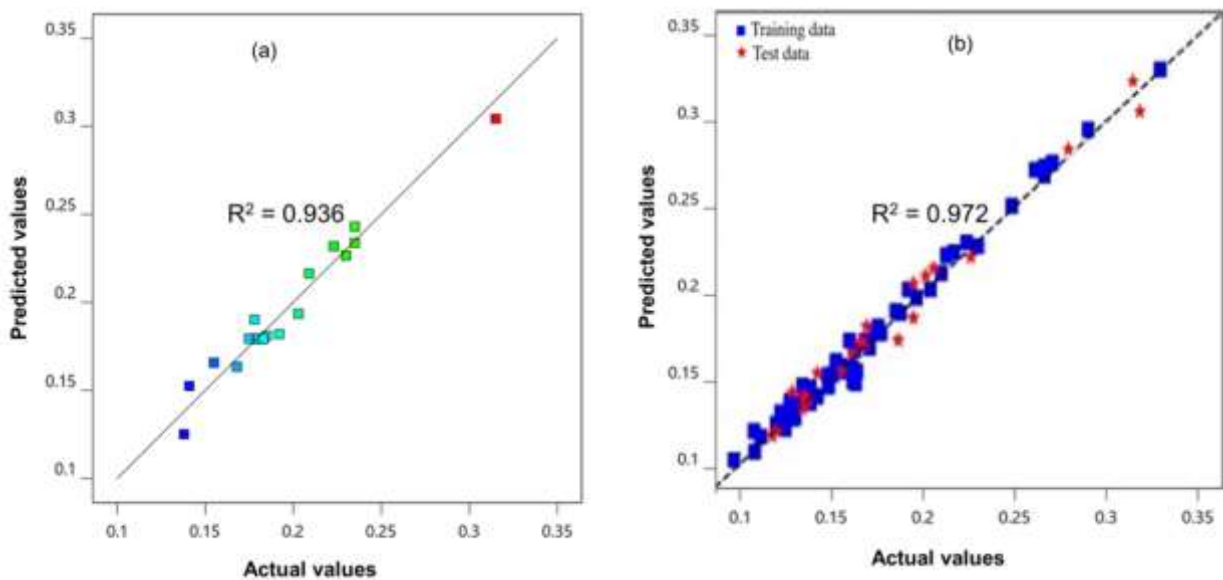


Fig. 5. Prediction accuracy of tool wear: (a) RSM model; (b) ML model.

## IV. CONCLUSIONS

This study employed the Design of Experiments (DOE) methodology to systematically investigate tool wear in end milling. Two predictive approaches, namely Response Surface Methodology (RSM) and a machine learning (ML) model, were developed and evaluated. Experimental trials were conducted on a five-axis CNC milling system to machine 40Cr steel alloy using a TiAlN-coated ball end mill, following a structured DOE framework. The collected experimental data were utilized to construct both regression-based and ML-based predictive models for tool wear estimation. A comparative analysis between the experimentally measured values and model predictions indicates that the ML model achieves superior predictive performance within the trained domain. In particular, the ML approach demonstrates higher robustness and accuracy in estimating flank wear compared to the regression model. These findings suggest that ML-based modeling offers enhanced capability for capturing complex relationships in machining processes and can be extended to other manufacturing applications. Furthermore, the proposed prediction framework provides a practical basis for optimizing machining parameters, thereby extending tool life under specified operating conditions. Overall, this study contributes to the advancement of tool condition monitoring systems and supports the development of automated and efficient machining processes in modern manufacturing environments.

## REFERENCES

- [1] R. Munaro, A. Attanasio, A. Del Prete, Tool Wear Monitoring with Artificial Intelligence Methods: A Review, (2023).
- [2] K. Zhu, X. Yu, The monitoring of micro milling tool wear conditions by wear area estimation, *Mech. Syst. Signal Process.* 93 (2017) 80–91. <https://doi.org/10.1016/j.ymssp.2017.02.004>.
- [3] J. Lars, M. Moreno, J.M. Andersson, R. Boyd, M.P. Johansson-j, S. Johnson, M. Od, L. Rogstr, Crater wear mechanism of TiAlN coatings during high-speed metal turning, 485 (2021). <https://doi.org/10.1016/j.wear.2021.204016>.
- [4] A.G.J.M. Manjaiah, Experimental Investigation on Tool Wear in AISI H13 Die Steel Turning Using RSM and ANN Methods, *Arab. J. Sci. Eng.* (2020). <https://doi.org/10.1007/s13369-020-05038-9>.
- [5] H. Pekşen, A. Kalyon, Optimization and measurement of flank wear and surface roughness via Taguchi based grey relational analysis, *Mater. Manuf. Process.* 00 (2021) 1–10. <https://doi.org/10.1080/10426914.2021.1926497>.
- [6] H. Hassanpour, A. Rasti, J. Hashemi, Effect of ball nose flank wear on surface integrity in high-speed hard milling of AISI 4340 steel using MQL, *Heliyon* 10 (2024) e37337. <https://doi.org/10.1016/j.heliyon.2024.e37337>.
- [7] L. Tan, Effects of tool wear on machined surface integrity during milling of Inconel 718, (2021).
- [8] J. Ma, Y. Li, D. Zhang, B. Zhao, X. Yan, X. Pang, Investigation of tool flank wear effect on system stability prediction in the milling of Ti - 6Al - 4 V thin - walled workpiece, *Int. J. Adv. Manuf. Technol.* (2022) 3937–3956. <https://doi.org/10.1007/s00170-022-10136-z>.
- [9] T. Alwarsamy, K. Sugapriya, A comparative analysis of tool wear prediction using response surface methodology and artificial neural networks, 12 (2014) 38–49.
- [10] U. Maheshwera, R. Paturi, S. Cheruku, V. Phani, K. Pasunuri, S. Salike, Modeling of tool wear in machining of AISI 52100 steel using artificial neural networks, *Mater. Today Proc.* (2020). <https://doi.org/10.1016/j.matpr.2020.06.581>.
- [11] R. Zaheer, A Study of the Optimization Algorithms in Deep Learning, 2019 Third Int. Conf. Inven. Syst. Control (2019) 536–539.
- [12] N.R. Draper, Response surface methodology: Process and product optimization using designed experiments, *J. Stat. Plan. Inference* 59 (1997) 185–186. [https://doi.org/10.1016/s0378-3758\(97\)81631-x](https://doi.org/10.1016/s0378-3758(97)81631-x).
- [13] S. Ghafari, H.A. Aziz, M.H. Isa, A.A. Zinatizadeh, Application of response surface methodology (RSM) to optimize coagulation-flocculation treatment of leachate using poly-aluminum chloride (PAC) and alum, *J. Hazard. Mater.* 163 (2009) 650–656. <https://doi.org/10.1016/j.jhazmat.2008.07.090>.
- [14] T. Zhou, L. He, J. Wu, F. Du, Z. Zou, Prediction of surface roughness of 304 stainless steel and multi-objective optimization of cutting parameters based on GA-GBRT, *Appl. Sci.* 9 (2019). <https://doi.org/10.3390/app9183684>.