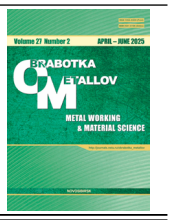
Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science. 2025 vol. 27 no. 2 pp. 126–141
ISSN: 1994-6309 (print) / 2541-819X (online)
DOI: 10.17212/1994-6309-2025-27.2-126-141

NSTU NETI

Obrabotka metallov - Metal Working and Material Science

Journal homepage: http://journals.nstu.ru/obrabotka_metallov



# Prediction of surface roughness in milling with a ball end tool using an artificial neural network

***Mikhail Gimadeev*** [a, *], ***Vadim Stelmakov*** [b], ***Aleksandr Nikitenko*** [c], ***Maksim Uliskov*** [d]

Pacific National University, 136 Tihookeanskaya St., Khabarovsk, 680035, Russian Federation

[a] https://orcid.org/0000-0001-6685-519X, 009063@togudv.ru; [b] https://orcid.org/0000-0003-2763-1956, 009062@togudv.ru;
[c] https://orcid.org/0000-0003-4729-5558, 005392@togudv.ru; [d] https://orcid.org/0009-0001-9858-423X, 2016104779@togudv.ru

**ARTICLE INFO**

*Article history*:
Received: 30 January 2025
Revised: 01 March 2025
Accepted: 27 March 2025
Available online: 15 June 2025

*Keywords*:
Milling
Roughness
Regression analysis
Artificial intelligence
Neural network
Standard error

*Funding*
This work has funded by the Ministry of science and higher education of Russian Federation (project № FEME–2024–0010).

**ABSTRACT**

**Introduction.** Milling stainless steel with a ball-end tool is a complex technological process that requires precise control of processing parameters to ensure high surface quality. In this regard, it is an urgent task to develop methods for predicting roughness parameters, such as ***Rz***. **The aim of this work** is to develop a predictive neural network model that can estimate surface roughness when milling stainless steel using a ball-end tool. **Method and methodology.** The main focus is on error backpropagation and gradient descent methods, as well as hyperparameter tuning, which are necessary to prevent overfitting and underfitting of the model. Experimental studies include the analysis of both controlled variables, such as feed per tooth, angle of inclination and diameter of the tool, and uncontrolled, including coolant supply and tool wear. **Results and discussions.** The use of coolant for milling austenitic steel has reduced the roughness parameters ***Rz*** by an average of 14%. A strong negative correlation has been established between the dimensional wear of the tool and the parameter ***Rz*** (−0.95). At the same time, wear in the range of 2…4 μm affects an increase in the ***Rz*** parameter by 21% compared to the minimum values. The data obtained were used to train eight configurations of artificial neural networks, which were used to predict roughness using the ***Rz*** parameter. The results show that the 3-16-16-1 network configuration showed the lowest *MSE* (0.0313), followed by 3-20-14-1 (0.0470) and 3-64-64-1 (0.0481), respectively. In addition, these configurations also demonstrated the lowest average absolute error values, which demonstrate the average of the absolute differences between predicted and observed values (0.1014; 0.1251 and 0.1155, respectively), and the coefficient of determination, which is a statistical measure indicating the proportion of data variability explained by the model (0.9944; 0.9916; 0.9904). A comparison of the experimental data with the predictions of various models allowed us to determine the average value of the absolute differences for the models according to the parameter **Ra** ≈ 0.074. The study suggests approaches to training neural network models for accurate prediction of roughness parameters, which makes a significant contribution to the methods of modeling machining processes.

**For citation:** Gimadeev M.R., Stelmakov V.A., Nikitenko A.V., Uliskov M.V. Prediction of surface roughness in milling with a ball end tool using an artificial neural network. *Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science*, 2025, vol. 27, no. 2, pp. 126–141. DOI: 10.17212/1994-6309-2025-27.2-126-141. (In Russian).

## Introduction

The quality of the machined surface plays a decisive role in ensuring the operational properties of machine parts [1]. Surface roughness (***Rz*** and ***Ra***) often serves as one of the main metrics for assessing the surface condition in the machining process [2]. Modeling methods for predicting ***Rz*** can be divided into three categories: experimental models, analytical models, and artificial intelligence (*AI*)-based models [3, 4]. In recent years, *AI*-driven models have become widely used among researchers to predict characteristics related to machining processes [5], and the use of artificial neural networks (*ANN*) is considered by the authors to predict surface roughness, tool wear, and other parameters in machining [6].

*** Corresponding author**
*Gimadeev Mikhail R.*, Ph.D. (Engineering), Associate Professor
Pacific National University,
136 Tihookeanskaya st.,
680035, Khabarovsk, Russian Federation
**Tel.:** +7 924 216-31-39, **e-mail:** 009063@togudv.ru

To function effectively, neural models use a vast network of simple computing processors known as "neurons." Neural networks are often used to solve complex problems in which the behavior of variables is not well known. One of their fundamental characteristics is the ability to learn from examples and apply this knowledge in a generalized way, which allows the creation of nonlinear models. This ability makes the use of *ANN* in multicriteria analysis very effective [7, 8].

The configuration of a neural network requires the definition of several important parameters: the number of nodes in the input layer, the number of hidden layers, the number of neurons in each hidden layer, and the number of neurons in the output layer. The state of neuron ***k*** is determined by the equation:

$$S_k = \sum_{i=1}^{n} (x_i w_{ki}) + b_k,$$

where $x_i$ is the output signal calculated by a neuron $i$; $w_{ki}$ is the synaptic weight between $i$ and $b_k$ neurons; $k$ is a weight associated with a constant, non–zero value known as the neuron's bias.

To use *ANN*, it is necessary to calculate synaptic weights and biases. The process of determining these parameters is called training and occurs iteratively, where the initial parameters are updated until the process reaches sufficient convergence.

The activation function ***f*** describes how the internal input and the current activation state influence the determination of the next state of the block. The most commonly used types of activation functions can be identified as follows:

Threshold function:

$$f(S_k) = \begin{cases} 1, & if\ S_k \geq 0 \\ 0, & if\ S_k < 0 \end{cases}.$$

A unit step function, or threshold function, is a mathematical function that takes the value **1** if its argument is greater than or equal to some threshold, and **0** otherwise.

Piecewise linear function, an example of which can be represented as:

$$f(S_k) = \begin{cases} aS_k + d, & if\ S_k \leq c; \\ eS_k + g, & if\ S_k > c, \end{cases}$$

where $a$, $c$, $d$, $e$, $g$ are constants.

A piecewise linear function consists of several linear sections, each defined on its own interval. The linear sections are connected to form a continuous function, although the derivative of such a function may be discontinuous at the junctions between sections.

Sigmoidal function:

$$f(S_k) = \left( \frac{1}{1 + \exp(a\mu)} \right),$$

where $a$ is the slope parameter of the sigmoidal function.

This function is the most commonly used and is characterized as an increasing function that balances linear and nonlinear behavior while maintaining its value within the range from 0 to 1.

The choice of activation function can significantly impact network performance. The rectified linear unit (*ReLU*), defined as $ReLU(X) = \max\{X, 0\}$, is currently the most widely used activation function and is popular in neural networks due to its non-saturation and non-linearity [9]. Compared to activation functions that exhibit saturation, such as the sigmoidal function, *ReLU* combined with gradient descent has superior performance. Gradient descent is a method used to minimize the loss function by adjusting the weights. In neural network training, the objective function is the output error of the network. The minima of the function form troughs, and the maxima form hills [10, 11].

Multilayer perceptrons (*MLP*) are recognized as the most widely used *ANN* models. A *MLP* consists of several layers: an input layer that receives raw data; one or more hidden layers that process the data by applying weights and activation functions; and an output layer that produces the final output or prediction based on the processed data. Neurons in each layer are connected only to neurons in the next layer, with no feedback or connections between neurons within the same layer. Additionally, a typical feature of *MLP* is the full connectivity between layers. An example of a network structure consisting of four layers: input layer, two hidden layers, and output layer, is shown in Fig. 1.

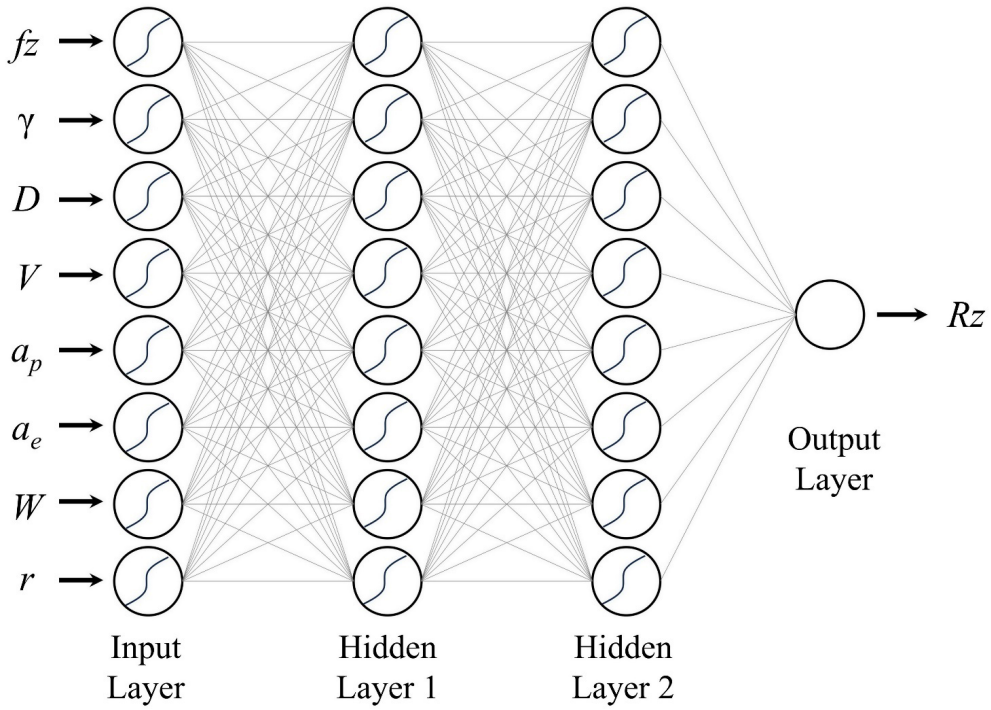



*Fig. 1.* Neural network structure for predicting the roughness parameter ***Rz***

In this structure (Fig. 1), the input layer has 8 nodes, each hidden layer has 8 nodes, and the output layer has 1 node. The nodes in the input layer represent the following factors: feed per tooth (***fz***, mm/tooth), angle of inclination (**γ**, °), tool diameter (***D***, mm), cutting speed (***V***, mm/min), cutting depth ($\boldsymbol{a_p}$, mm), side step ($\boldsymbol{a_e}$, mm), coolant supply (***W***, l/min), and tool wear (***r***, mm). The node in the output layer represents the predicted surface roughness parameter (***Rz***, μm).

The network shown is fully connected, meaning that each neuron in any layer is connected to all neurons in the previous layer. Signal flow through the network is from left to right, layer by layer. Considering a multilayer network with ***j*** and ***k*** nodes in each hidden layer, the example structure shown in Fig. 1 can be described by an **8**–***j***–***k***–**1** configuration. In general, the operation of this type of network is described by two main phases: forward propagation and back propagation.

The process of training *MLP* networks using the back propagation (*BP*) method follows the sequence: Forward Propagation → Loss Calculation → Back Propagation → Weight Update. An essential feature of *MLP* networks is the nonlinearity of neuron outputs, achieved by using the activation function.

Successfully building an *ANN* model based on the ***Rz*** response requires multi-factor experimentation and tuning. Although many researchers have applied *ANN* for modeling in various fields such as machine learning [12–14], there is still no clear guideline for building a predictive model. This study examines elements that can affect model performance and the ***Rz*** response using the capabilities of the *TensorFlow Python* library to reduce uncertainty and improve prediction quality.

Four performance indicators (metrics) were selected to evaluate the accuracy of models for predicting surface roughness [15]. These metrics include the coefficient of determination ($\boldsymbol{R^2}$), mean absolute error (*MAE*), mean squared error (*MSE*), and root mean squared error (*RMSE*). The coefficient of determination

($R^2$) shows what proportion of the variance of the dependent variable is explained by the independent variables of the model:

$$R^2 = \frac{\sum_{i=1}^{n} \left(Y_i - \widehat{Y_i}\right)^2}{\sum_{i=1}^{n} \left(Y_i - \bar{Y_i}\right)^2}.$$

where $n$ is the number of data; $Y_i$ is the observed values; $\hat{Y}$ is the predicted values; $\bar{Y}$ is the mean value of $Y$.

Despite its usefulness, $R^2$ has limitations: it does not account for the number of predictors and can be biased by outliers.

*MAE* is a measure of the absolute error $\left(| Y - \widehat{Y} |\right)$ between predicted and actual values:

$$MAE = \frac{1}{n} \sum_{i=0}^{n-1} \left|Y_i - \widehat{Y_i}\right|.$$

*MAE* is less sensitive to large errors than *MSE* and *RMSE* because it uses absolute error values.

*MSE* and *RMSE* are characterized by the mean square error and its square root, respectively:

$$MSE = \frac{1}{n} \sum_{i=0}^{n-1} \left|Y_i - \widehat{Y_i}\right|^2,$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} \left(Y_i - \widehat{Y_i}\right)^2}.$$

*MSE* is sensitive to large errors because squared differences increase with large deviations. Since *RMSE* is measured in the same units as the data itself, it is easier to interpret than *MSE*. However, like *MSE*, *RMSE* is also sensitive to large errors.

Analysis of these metrics is critical for a comprehensive assessment of predictive performance [15, 16]. When comparing these metrics, special attention is paid to *MSE*, which has the advantage of detecting and accounting large errors, making it useful in machine learning tasks where minimizing large deviations is important. Additionally, *MSE* is smooth and differentiable, simplifying gradient computation in optimization methods such as gradient descent. Therefore, *MSE* is often a more suitable choice for accuracy assessment.

Furthermore, the coefficient of determination $R^2$, with values close to **1**, is considered most favorable.

A preliminary data analysis is also performed before using machine learning models. This includes checking for normality and identifying/removing outliers that can significantly affect model accuracy.

Model optimization is an important step for effective solutions. Configuring hyperparameters ensures the best performance estimated from validation datasets within the selected algorithm. Hyperparameters play a significant role in controlling the learning process and significantly affect predictive accuracy. Proper setting of hyperparameters also helps reduce overfitting and underfitting, improving accuracy. Dropout is a method to prevent overfitting by randomly excluding neurons during training, preventing coadaptation [14].

***The aim of this work*** is to develop a predictive neural network model for assessing surface roughness when milling stainless steel with a ball-end tool.

To achieve this aim, the following ***tasks*** were addressed:

– study of predicting surface roughness parameter ***Rz*** when milling with a ball-end tool, including optimizing *ANN* architecture, selecting number of layers, and tuning model parameters to improve prediction accuracy.

– analysis of the influence of various input parameters, including tool tilt angle, on roughness prediction accuracy, and development of an approach to minimize input data without loss of model effectiveness, as well as study of model applicability with limited training sets.

– final testing of the developed model, assessment of accuracy using *MSE*, *RMSE*, *MAE*, and $R^2$ metrics, and evaluation of effectiveness through comparison of predicted and experimental data.

## Research methodology

Milling operations were performed on a *DMG MORI DMU 50* universal milling machine with 9 kW power and maximum rotation speed $n_{max}$ = 8,000 min$^{-1}$. The workpiece was made of austenitic stainless steel *AISI 321* with chemical composition,wt. %: $C \leq 0.12$; $Si \leq 0.8$; $Mn \leq 2.0$; $P \leq 0.035$; $S \leq 0.02$; *Ni* 9–11; *Cr* 17–19; $Ti < 0{,}8$; *Fe* – bal. The cutting tool was a hard alloy with multilayer (*TiN* and *TiNAl*) *PVD* coating and fine-grained base, with diameters of 6, 8, 10, and 12 mm from *Sandvik Coromant*. During the experiments, measurements of tool radius wear (***r***, mm) by levels were obtained using a *TT140* contact sensor from *Heidenhain*.

Surface roughness after milling was measured using a *SURFCOM 1800D* profilometer; for this device, the error according to the standard is 3 %. Filter – 50 % *Gaussian*. The parameter of the basic length (step cutoff) was chosen to be 0.8 mm (*ISO 4288:1996*) for all measurements, since the expected range should be $0.5 < Rz \leq 10$. Tracing was performed three times in the direction of tool feed.

Experimental design included controlled factors: $a_p$, $fz$, γ, ***D***, ***V***, and uncontrolled factors: ***W*** and ***r***. The response parameter was surface roughness ***Rz***.

After experiments, *ANN* models were built using *Python* with *TensorFlow* and *Keras* libraries for neural network creation, training, and regularization, *NumPy* for array operations, and *Scikit*-learn for data preprocessing. Experimental data were divided into training and test sets that went through a process of standardization and normalization, accounting for 70 % and 30 % of the total number of experiments performed, corresponding to 28 training attempts and 12 testing attempts.

The neural network training algorithm used is the back propagation (*BP*) method. This method calculates the gradient of the loss function with respect to the weights of the neural network. During forward propagation, the input data passes through the network, generating an output prediction. The error of this prediction is then calculated and propagated back through the network, from the output layer through all hidden layers to the input layer. At each layer, the gradient of the error with respect to the weights is calculated.

Hyperparameter values were tested and best values for the models are presented in Table 1.

The optimizer (*Adam*) updates weights according to computed gradients. These steps are repeated for each training epoch, allowing iterative improvement of predictions.

## Results and discussion

One of the most important distributions in statistics is the normal distribution, which describes typical behavior of many phenomena. To determine the distribution of ***Rz***, measured data after mechanical processing of 512 surfaces with coolant were analyzed. For all surfaces, technological parameters were $a_p$ = 0.2 mm, $a_e$ = 0.4 mm. Roughness results (***Rz***) are shown in Fig. 2.

Table 1

**Hyperparameters for the BPNN model**

| Model | Hyperparameters | Indicator |
|---|---|---|
| Sequential | Activation hidden layers | *Leaky ReLU* |
| | Kernel_regularizer | *l1*=0,0001, *l2*=0,0001 |
| | Dropout | 0.01 |
| | Optmizer | *Adam* |
| | Learning_rate | 0.001 |
| | Loss | mean_squared_error |
| | Metrics | *MSE*, *RMSE*, *MAE* |
| | Batch size | 16 |
| | Epochs | 500 |

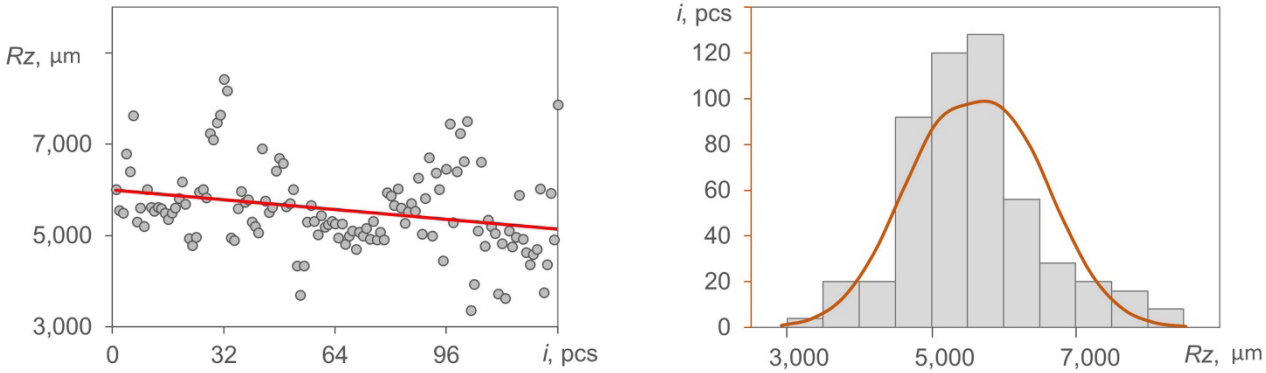



*Fig. 2*. Distribution of surface roughness parameter ***Rz***

*Pearson* goodness-of-fit test showed no basis for rejecting the hypothesis of normal distribution when conditions $f > 100$ and $p > 0.05$ are met. The sample ***Rz*** data deviate from the mathematical expectation of 5.357 μm by 0.389 μm on average. The data follow a normal distribution according to the **2σ** rule with a probability of 0.9873.

For neural network modeling, outliers were removed because they can distort results and reduce model ability to identify data patterns effectively [17]. For further experiments, depth and lateral pitch were kept constant at $a_p = 0.2$ mm and $a_e = 0.4$ mm [18], reducing the number of variable input parameters to three.

Variable ***W*** is often considered stochastic and uncontrollable, introducing unexplained variance independent of explanatory variables and the model. Typically, ***W*** and ***r*** are treated as integral components of variability; their influence on ***Rz*** was considered $fz = 0.4$ mm/tooth, $\gamma = 50°$, $D = 6$ mm, $z = 2$. Results are shown in Fig. 3.

Rational use of coolant is an important factor in increasing metal processing productivity. When using coolant, the roughness parameter ***Rz*** decreased by an average of 14 %.

The dissipation rate depends significantly on cutting speed ***Vc*** (m/min) and material removal volume ***Q*** (cm³/min):

$$Q = \frac{S_{zj} f z n_c z}{1000},$$

where: $n_c$ is the rotation frequency min$^{-1}$; $S_{zj}$ is the cross-sectional area during milling with a ball-end tool, mm$^2$:

$$S_{zj} = \left( 2 \cdot \sqrt{\left(\frac{a_p}{2}\right)^2 + \left( \frac{2 \cdot \sin\left( \frac{2 \cdot \cos^{-1}\left(1 - 2\frac{a_p}{2}\right) \cdot \frac{R}{2}}{2R} \right) R}{4} \right)^2} \right) \left( fz \cdot \sin\left( \tan^{-1}\left( \frac{a_p}{2 \cdot \sin\left( \frac{2 \cdot \cos^{-1}\left(1 - 2\frac{a_p}{2}\right)\frac{R}{2}}{2R} \right)} \right) 2 \right) \right).$$

The effect of coolant depends on rational choice of cutting conditions, tool wear, and tool/workpiece materials [19]. Coolant use and cutting speed of at least 75 m/min are necessary to achieve minimum roughness ***Rz***, beyond which ***W*** has little impact on model performance.

Surface roughness also depends on tool wear degree [20], with a strong negative correlation $R = -0.95$. As the number of machined workpiece surfaces (***i***, pcs) increases, dimensional tool wear occurs, typically within the range of 2 to 4 μm. This wear leads to an approximate 21 % increase in the ***Rz*** parameter relative to its minimum observed values.

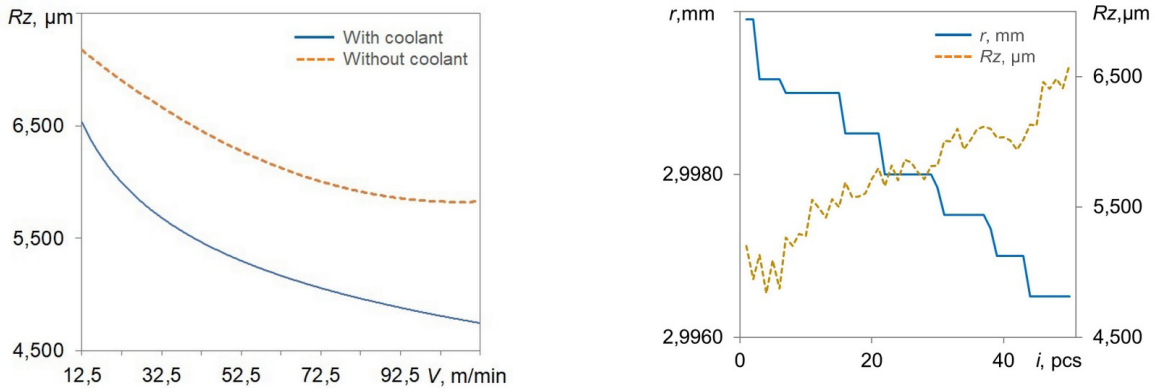



*Fig. 3.* Dependence of surface roughness parameter ***Rz*** on cutting speed, coolant and tool wear

An experiment with three factors at two levels each (Table 2) was conducted. Full factorial design yields 8 combinations ($k = 23 = 8$).

Based on theoretical factor influence on roughness, the model considered was: $Y(Rz) = a + b_{fz}X_{fz} + b_{\gamma}X_{\gamma} + b_{D}X_{D}$ Here are the calculated natural regression coefficients: $a = 3.00$; $b_{fz} = 2.77$; $b_{\gamma} = -0.55$; $b_{D} = -1.08$.

The parameters of the regression equation were estimated using the ordinary least squares method. Standardized **β**-coefficients: $\beta_{fz} = 0.17$; $\beta_{\gamma} = -3.33$; $\beta_{D} = -1.01$.

Comparison of the modules of the values of the standardized regression coefficients **β** allows us to conclude that **γ** (the angle of inclination of the ball-end tool) is the most influential factor in the formation of the roughness parameter *Rz*. When the ***fz*** factor is fixed, the remaining factors show negative standardized **β** coefficients ($\beta\gamma = -0{,}54$; $\beta_{D} = -1{,}09$), indicating a decrease in the response value ***Rz***. The value $R^2 = 0.14$ indicates that ***fz*** plays an important role in explaining ***Rz***, and fixing it significantly reduces the explanatory power of the model.

When **γ** is fixed, the standardized **β** coefficients ($\beta_{fz} = 2.77$; $\beta_{D} = -1.09$) show that ***fz*** has a strong positive contribution and ***D*** has a strong negative contribution. The high $R^2 = 0.84$ indicates that even when **γ** is fixed, ***fz*** and ***D*** remain important in explaining ***Rz***.

When ***D*** is fixed, the standardized coefficients $\beta_{fz} = 2.76$, $\beta_{D} = -0.54$. This means that when ***D*** is fixed, an increase in ***fz*** leads to an increase in ***Rz***, while an increase in **γ** leads to a decrease. The value $R^2 = 0.75$ also indicates that the model with fixed ***D*** explains the variation in ***Rz*** well, and the factors ***fz*** and **γ** remain important in explaining ***Rz***.

The regression model (*RM*), obtained after conducting an experiment on processing free form surface with a ball-end tool, characterizing the relationship between the roughness value, feed per tooth, diameter and angle of inclination of the tool, expressed by a normalized model has the following form:

$$Rz_{(RM)} = 3 + 2.77_{fz} - 0.55_{\gamma} - 1.08_{D} - 0.51_{fz\gamma} - 1_{fzD} + 0.22_{\gamma D} + 0.2_{fz\gamma D}.$$

Table 2

**Experimental data values**

| No. | *fz* | γ | *D* | *fz*·γ | *fz*·*D* | γ·*D* | *fz*·γ·*D* | *Rz* |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.5 | 50 | 12 | 1 | 1 | 1 | 1 | 3.05 |
| 2 | 0.1 | 50 | 12 | –1 | –1 | 1 | –1 | 0.13 |
| 3 | 0.1 | 10 | 12 | 1 | –1 | –1 | 1 | 0.18 |
| 4 | 0.1 | 10 | 6 | 1 | 1 | 1 | –1 | 0.37 |
| 5 | 0.5 | 10 | 12 | –1 | 1 | –1 | –1 | 4.33 |
| 6 | 0.5 | 50 | 6 | 1 | –1 | –1 | –1 | 6.38 |
| 7 | 0.1 | 50 | 6 | –1 | 1 | –1 | 1 | 0.26 |
| 8 | 0.5 | 10 | 6 | –1 | –1 | 1 | 1 | 9.32 |

In this study, we examine the effect of activation function on the performance of eight neural network models in predicting ***Rz*** (Fig. 4).

Loss function reflects model training efficiency. The *ReLU* activation speeds training but requires monitoring of both Train Loss and Validation Loss. Low training loss with high validation loss indicates

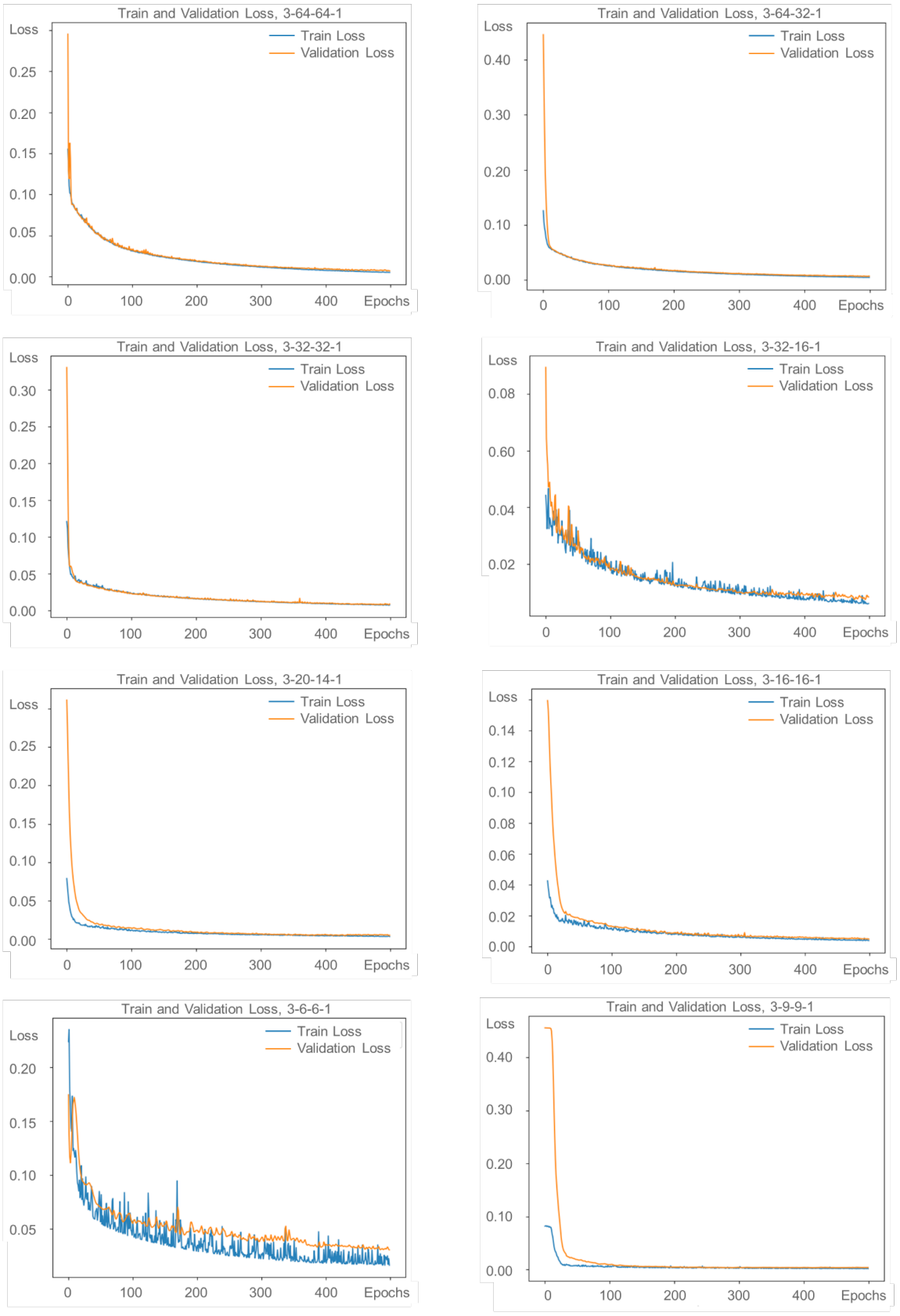



*Fig. 4.* Learning rates of various configurations

overfitting, meaning the model memorizes training data instead of generalizing. Fig. 3 shows 3–32–16–1 and 3–6–6–1 models remembered the training data too well, but cope poorly with new data.

According to results and based on *MSE* (Table 3), best network configurations are 3–64–64–1 (*MSE* = 0.0481), 3–20–14–1 (0.0470), and 3–16–16–1 (0.0313). These also have low *RMSE* (0.2174, 0.2135, 0.1770) and *MAE* (0.1155, 0.1251, 0.1014). *RMSE* is interpreted as error in the same units as data.

A design with $\boldsymbol{k}$ = 12 experiments was used to test the best models with factors fz ∈ {0.4; 0.5}, γ ∈ {10, 15, 20, 30, 40, 50}, D ∈ {6}, distributed randomly.

Table 4 and Fig. 5 present data showing the relationship between the values obtained during the experiment, calculated based on the developed regression model and the predicted *BPNN* responses.

Table 3

**Predictive performance of the neural network**

| Metrics | 3-64-64-1 | 3-64-32-1 | 3-32-32-1 | 3-32-16-1 | 3-20-14-1 | 3-16-16-1 | 3-6-6-1 | 3-9-9-1 |
|---|---|---|---|---|---|---|---|---|
| *MSE* | 0.0481 | 0.0621 | 0.0572 | 0.0685 | 0.0470 | 0.0313 | 0.0415 | 0.0603 |
| *RMSE* | 0.2174 | 0.2492 | 0.2391 | 0.2617 | 0.2135 | 0.1770 | 0.2037 | 0.2456 |
| *MAE* | 0.1155 | 0.1228 | 0.1656 | 0.1361 | 0.1251 | 0.1014 | 0.1306 | 0.1447 |
| $\boldsymbol{R^2}$ | 0.9904 | 0.9889 | 0.9898 | 0.9878 | 0.9916 | 0.9944 | 0.9926 | 0.9862 |

Table 4

**Predicted (*Y*(*Rz*)) and experimental (*Rz*) values for selected configurations, at $\boldsymbol{k}$ = 12**

| *fz* | γ | *D* | *Rz* | $Rz_{(RM)}$ | *Y*(*Rz*) | | |
|---|---|---|---|---|---|---|---|
| | | | | | 3-64-64-1 | 3-20-14-1 | 3-16-16-1 |
| 0.4 | 10 | 6 | 6.945 | 7.090 | 6.680 | 6.277 | 6.491 |
| 0.4 | 50 | | 4.610 | 4.840 | 4.938 | 5.025 | 5.116 |
| 0.4 | 20 | | 6.108 | 6.528 | 6.037 | 6.469 | 6.303 |
| 0.4 | 40 | | 5.400 | 5.403 | 5.357 | 5.503 | 5.261 |
| 0.5 | 20 | | 8.341 | 8.590 | 7.923 | 8.372 | 7.875 |
| 0.4 | 15 | | 6.614 | 6.809 | 6.426 | 6.495 | 6.330 |
| 0.5 | 10 | | 9.163 | 9.330 | 8.272 | 8.402 | 8.925 |
| 0.4 | 30 | | 5.826 | 5.965 | 6.067 | 5.987 | 6.307 |
| 0.5 | 15 | | 8.786 | 8.960 | 8.463 | 8.590 | 8.072 |
| 0.5 | 40 | | 6.992 | 7.110 | 7.035 | 6.977 | 7.407 |
| 0.5 | 30 | | 7.694 | 7.850 | 7.792 | 7.443 | 7.896 |
| 0.5 | 50 | | 6.024 | 6.370 | 6.548 | 6.745 | 6.541 |
| *MSE* | | | | 0.049 | 0.136 | 0.167 | 0.175 |
| *RMSE* | | | | 0.221 | 0.369 | 0.408 | 0.418 |
| *MAE*(***Rz***) | | | | 0.195 | 0.286 | 0.317 | 0.384 |
| *MAE*(***Ra***) | | | | 0.049 | 0.072 | 0.081 | 0.095 |
| $\boldsymbol{R^2}$ | | | | 0.973 | 0.924 | 0.907 | 0.903 |

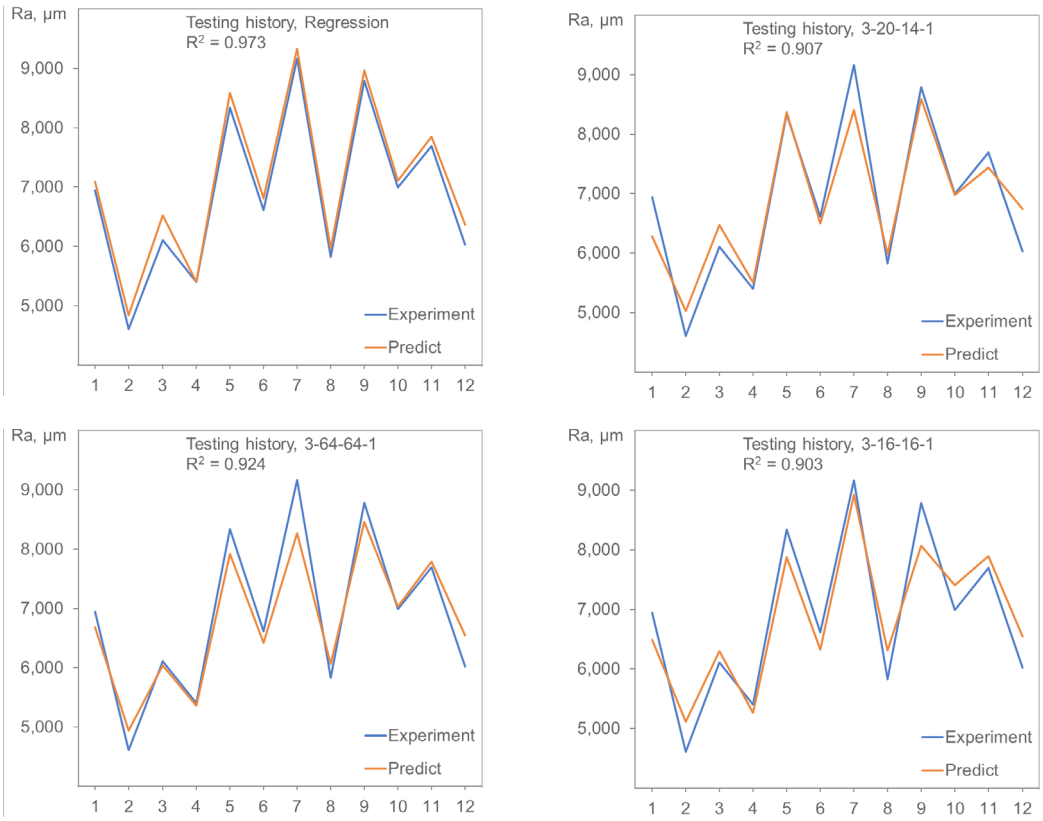



*Fig. 5*. Predicted and Experimental values for selected classifications

When evaluating Table 4 and Fig. 5, one can conclude that the models under consideration showed predicted values ($Y(Rz)$) close to the actual ones (***Rz***). The coefficient of determination $\boldsymbol{R^2}$ was 0.973 for the regression model and 0.924, 0.907, 0.903, respectively, for the considered configurations. For all configurations, $MAE \approx 0.2955$ μm means that, on average, model predictions deviate from actual ***Rz*** values by 0.2955 μm.

The ***Rz*** parameter correlates strongly with ***Ra*** (correlation coefficient 0.91) [21–23]. Statistical processing shows the relationship: $Ra = (Rz - 0.391) / 4.022$.

Comparison of experimental data with model predictions shows $MAE$ for $Ra \approx 0.049$ μm, which is negligible in surface roughness context, indicating close agreement between observations and true mean values. Thus, errors found do not significantly affect result accuracy, confirming data compliance.

## Conclusion

This paper discusses the development of an artificial neural network model to predict surface roughness when milling with a ball-end tool. The tuning process of *ANN* architecture, especially the selection of number of layers and neurons, is described to improve prediction accuracy.

The concept of parameter selection based on the significance of the contribution to the accuracy of surface roughness prediction ***Rz*** is considered to reduce the input factors to the minimum possible. The results show that it is possible to obtain accurate predictions of surface roughness even when taking into account a small number of input parameters with relatively small training sets.

Selecting the correct network configuration and input parameters is important to ensure accurate prediction. In addition, the study highlights the importance of considering the inclination angle of the ball-end tool (10° to 50°) in training the *ANN* models, and the increase in the angle affects the decrease in the magnitude of the roughness parameters.

The final tests conducted to check the adequacy of the proposed model showed that the model works well with reasonable accuracy under the given set of parameters. In conclusion, it can be said that this study contributes significantly to machining process modeling by milling.

## References


1. Suslov A.G., Medvedev D.M., Petreshin D.I., Fedonin O.N. Sistema avtomatizirovannogo tekhnologicheskogo upravleniya iznosostoikost'yu detalei mashin pri obrabotke rezaniem [System for automated wear-resistance technological control of machinery at cutting]. *Naukoemkie tekhnologii v mashinostroenii = Science intensive technologies in mechanical engineering*, 2018, no. 5 (83), pp. 40–44. DOI: 10.30987/article_5ad8d291cdd cd8.06334386.

2. Cai Y., Starly B., Cohen P., Lee Y.S. Sensor data and information fusion to construct digital-twins virtual machine tools for cyber-physical manufacturing. *Procedia Manufacturing*, 2017, vol. 10, pp. 1031–1042. DOI: 10.1016/j. promfg.2017.07.094.

3. Zain A.M., Haron H., Sharif S. Prediction of surface roughness in the end milling machining using artificial neural network. *Expert Systems with Applications*, 2010, vol. 37 (2), pp. 1755–1768. DOI: 10.1016/j.eswa.2009.07.033.

4. Cooper C., Wang P., Zhang J., Gao R.X., Roney T., Ragai I., Shaffer D. Convolutional neural network-based tool condition monitoring in vertical milling operations using acoustic signals. *Procedia Manufacturing*, 2020, vol. 49, pp. 105–111. DOI: 10.1016/j.promfg.2020.07.004.

5. Pimenov D.Y., Kumar Gupta M., da Silva L.R.R., Kiran M., Khanna N., Krolczyk G.M. Application of measurement systems in tool condition monitoring of milling: a review of measurement science approach. *Measurement*, 2022, vol. 199, p. 111503. DOI: 10.1016/j.measurement.2022.111503.

6. Guo M., Xia W., Wu C., Luo C., Lin Z. A surface quality prediction model considering the machine-tool-material interactions. *The International Journal of Advanced Manufacturing Technology*, 2024, vol. 131 (7–8), pp. 1–19. DOI: 10.1007/s00170-024-13072-2.

7. Chen C.H., Jeng S.Y., Lin C.J. Prediction and analysis of the Surface roughness in CNC end milling using neural networks. *Application Science*, 2022, vol. 12 (1), p. 393. DOI: 10.3390/app12010393.

8. Oktem H., Erzurumlu T., Erzincanli F. Prediction of minimum surface roughness in end milling mold parts using neural network and genetic algorithm. *Materials & Design*, 2006, vol. 27 (9), pp. 735–744. DOI: 10.1016/j. matdes.2005.01.010.

9. Bingham G., Miikkulainen R. Discovering parametric activation functions. *Neural Networks*, 2022, vol. 148, pp. 48–65. DOI: 10.1016/j.neunet.2022.01.001.

10. Galushkin A.I. *Neironnye seti: osnovy teorii* [Neural networks: fundamentals of theory]. Moscow, Goryachaya liniya – Telekom Publ., 2010. 496 p. ISBN 978-5-9912-0082-0. Available at: https://asu.tusur.ru/learning/010402/ d15a/010402-d15a-book1.pdf (accessed 02.04.2025).

11. Antsev A.V., Yanov E.S., Dang Ch.Kh. Optimizatsiya rezhimov rezaniya s uchetom urovnya vibratsii na osnove primeneniya metodov iskusstvennogo intellekta [Cutting modes optimization taking into account vibration level based on application of artificial intelligence methods]. *Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta = Bulletin of Voronezh state technical university*, 2020, vol. 16 (3), pp. 101–109. DOI: 10.25987/ VSTU.2020.16.3.014.

12. Erygin E.V., Duyun T.A. Prognozirovanie sherokhovatosti poverkhnosti pri chistovom frezerovanii s ispol'zovaniem neironnykh setei [Forecasting of the surface roughness in finishing milling using neural networks]. *Vestnik Belgorodskogo gosudarstvennogo tekhnologicheskogo universiteta im. V.G. Shukhova = Bulletin of Belgorod state technological university named after V.G. Shukhov*, 2019, no. 10, pp. 135–141. DOI: 10.34031/article_5db3f6 c82c1ea9.69942788.

13. Lapshin V.P., Turkin I.A., Gvindzhilia V.I., Dudinov I.O., Gamaleev D.O. Nekotorye voprosy sovmestnogo primeneniya determinirovannykh matematicheskikh modelei i neironnykh setei v tsifrovykh dvoinikakh protsessa obrabotki metallov rezaniem na metallorezhushchikh stankakh [Using deterministic models and neural networks in digital twins of metal cutting processes]. *STIN = Russian Engineering Research*, 2024, no. 10, pp. 6–11. (In Russian).

14. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 2014, vol. 15 (56), pp. 1929–1958.

15. Danish M., Gupta M.K., Irfan S.A., Ghazali S.M., Rathore M.F., Krolczyk G.M., Alsaady A. Machine learning models for prediction and classification of tool wear in sustainable milling of additively manufactured 316 stainless steel. *Results in Engineering*, 2024, vol. 22, p. 102015. DOI: 10.1016/j.rineng.2024.102015.

16. Chai T., Draxler R.R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 2014, vol. 7 (3), pp. 1247–1250. DOI: 10.5194/ gmd-7-1247-2014.

17. Ramírez-Gallego S., Krawczyk B., García S., Wozniak M., Herrera F. A survey on data preprocessing for data stream mining: current status and future directions. *Neurocomputing*, 2017, vol. 239, pp. 39–57. DOI: 10.1016/j.neucom.2017.01.078.

18. Wojciechowski S., Chwalczuk T., Twardowski P., Krolczyk G.M. Modeling of cutter displacements during ball end milling of inclined surfaces. *Archives of Civil and Mechanical Engineering*, 2015, vol. 15, pp. 798–805. DOI: 10.1016/j.acme.2015.06.008.

19. Kuscheva M.E., Klauch D.N., Kobelev O.A. Printsipy vybora smazochno-okhlazhdayushchikh tekhnologicheskikh sred dlya obrabotki metallov rezaniem [Principles of selection of cutting technological mediums for metal cutting]. *Izvestiya MGTU MAMI*, 2014, vol. 8, no. 1–2, pp. 73–76. DOI: 10.17816/2074-0530-67737. (In Russian).

20. Chehrehzad M., Kecibas G., Besirova C., Uresin U., Irican M., Lazoglu I. Tool wear prediction through AI-assisted digital shadow using industrial edge device. *Journal of Manufacturing Processes*, 2024, vol. 113, pp. 117–130. DOI: 10.1016/j.jmapro.2024.01.052.

21. Gimadeev M.R., Davydov V.M., Nikitenko A.V., Sarygin A.V. Formirovanie parametrov sherokhovatosti na osnove korrelyatsionnykh svyazei pri chistovom frezerovanii prostranstvenno-slozhnykh poverkhnostei [Formation of roughness parameters based on correlation relations during finishing milling of spatially complex surfaces]. *Uprochnyayushchie tekhnologii i pokrytiya = Strengthening Technologies and Coatings*, 2019, vol. 15, no. 6 (174), pp. 243–248.

22. Gimadeev M.R., Li A.A. Analysis of automated surface roughness parameter support systems based on dynamic monitoring. *Advanced Engineering Research (Rostov-on-Don)*, 2022, vol. 22 (2), pp. 116–129. DOI: 10.23947/2687-1653-2022-22-2-116-129.

23. Gimadeev M.R., Li A.A., Berkun V.O., Stelmakov V.A. Eksperimental'noe issledovanie dinamiki protsessa mekhanoobrabotki kontsevymi sferotsilindricheskimi frezami [Experimental study of the dynamics of the machining process by ball-end mills]. *Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science*, 2023, vol. 25, no. 1, pp. 44–56. DOI: 10.17212/1994-6309-2023-25.1-44-56.

## Conflicts of Interest

The authors declare no conflict of interest.