COMPARATIVE EVALUATION OF MACHINE LEARNING MODELS FOR SURFACE ROUGHNESS PREDICTION IN WEDM OF SKD61 UNDER LIMITED DATA CONDITIONS

Huynh Thanh Thuong[1]*, Hoang-Tien Cao[1,2], Dinh-Tu Nguyen[3]

[1]*Can Tho University, Vietnam*

[2]*CTU-AIMED Leading Research Team, Can Tho University, Vietnam*

[3]*Can Tho University of Technology, Vietnam*

*Corresponding author: thanhthuong@ctu.edu.vn

(Received: May 11, 2026; Revised: June 10, 2026; Accepted: June 22, 2026)

DOI: 10.31130/ud-jst.2026.24(6A).331E

Abstract - Surface roughness (Ra) is a key quality indicator in Wire Electrical Discharge Machining (WEDM) of SKD61 tool steel. This study compares three machine learning models, Artificial Neural Networks (ANN), Random Forest Regression (RFR), and Extreme Learning Machine (ELM), for Ra prediction under limited data conditions. Experimental data were obtained using a Taguchi L27 design with pulse-on time, pulse-off time, servo voltage, and wire feed rate as input variables. To assess model robustness, nine additional experiments outside the original design were conducted for independent validation. Model performance was evaluated using R^2, MAE, RMSE, and MAPE. Results showed that ELM provided favourable predictive performance, with a validation MAPE of 0.85%, compared with ANN and RFR under the investigated conditions. The proposed ELM-based framework provides a promising and computationally efficient approach for surface roughness prediction under limited-data WEDM conditions.

Key words - WEDM; Surface Roughness; Extreme Learning Machine; Artificial Neural Network; Random Forest; SKD61

1. Introduction

Wire Electrical Discharge Machining (WEDM) is widely recognised as one of the most effective non-traditional machining processes for producing intricate and high-precision components from hard-to-machine materials such as tool steels, superalloys, and heat-resistant alloys. The process relies on a series of controlled electrical discharges between a continuously moving wire electrode and the workpiece, enabling material removal without direct mechanical contact [1, 2]. Because of its ability to maintain dimensional accuracy and geometric stability, WEDM has become indispensable in die-making, aerospace, precision tooling, and mould manufacturing. Among the various tool steels, SKD61 is widely used for its high hardness, thermal fatigue resistance, and superior toughness; however, these same properties make it difficult to machine with conventional processes.

In WEDM, machining performance is often evaluated using two key responses: surface roughness (Ra) and material removal rate (MRR). Ra quantifies the functional quality, fatigue life, and surface integrity of the machined part, whereas MRR reflects productivity and cost efficiency. Prior studies consistently report that pulse-on time (T-on), pulse-off time (T-off), servo voltage (SV), and wire feed rate (WFR) are the predominant factors governing these responses [3, 4]. Longer pulse-on durations typically increase discharge energy, enhancing MRR but deteriorating surface finish due to larger molten craters [5, 6]. Conversely, higher servo voltage tends to stabilise the discharge gap, reducing arcing and improving surface quality [7, 8]. This inherent trade-off between Ra and MRR underscores the need for predictive modelling techniques that capture nonlinear interactions among parameters.

Traditional modelling approaches, including regression analysis, Taguchi methods, and response surface methodology (RSM), have been widely applied to analyse WEDM behaviour [9]. While these techniques provide valuable insight into factor significance, they generally assume linear or low-order polynomial relationships and may struggle to accurately represent the complex thermal–electrical interactions inherent to the WEDM process [10]. As a result, machine-learning models have received increasing attention for their ability to learn nonlinear patterns directly from experimental data. Artificial Neural Networks (ANNs), Random Forest Regression (RFR), and Extreme Learning Machines (ELMs) have been successfully applied in machining studies, demonstrating strong predictive performance for surface roughness, cutting forces, and process stability [11–18]. Among these methods, ANN has been the most extensively studied due to its flexibility in approximating nonlinear relationships [14–16]. However, ANN requires careful tuning of network architecture and may suffer from long training times and convergence issues when datasets are limited. RFR provides an ensemble-learning approach that handles nonlinear interactions and offers built-in variable-importance metrics [19].

More recently, ELM has emerged as a fast-learning neural network characterised by randomised hidden-layer weights and analytically determined output weights, enabling rapid training and promising generalisation performance, especially for small to medium-scale datasets [11, 20, 21]. In addition, in a related optimisation-oriented companion study by the same authors, the same WEDM experimental dataset was used to investigate multi-objective optimisation considering both surface roughness (Ra) and MRR [22]. That study systematically evaluated the effects of WEDM machining parameters and identified favourable parameter combinations using PSI and Pareto-front analysis, thereby providing valuable insights into parameter significance and optimisation behaviour. However, such optimisation-oriented analyses do not directly address the predictive capability of data-driven models, particularly their robustness when applied to previously unseen parameter combinations within the investigated parameter space. Consequently, there remains a need to develop and validate reliable predictive models for Ra under previously unseen parameter combinations within the investigated parameter space. Despite these advances, comparative studies that simultaneously evaluate ANN, RFR, and ELM for WEDM, particularly with independent validation, remain limited, especially for SKD61 tool steel.

Given these research gaps, the present study focuses on a generalisation-oriented evaluation of machine-learning models for predicting Ra in the WEDM of SKD61 tool steel. Three representative machine-learning techniques, ANN, RFR, and ELM, are developed using experimental data obtained from a Taguchi L27 design involving T-on, T-off, SV, and WFR. Unlike most existing WEDM studies, which assess predictive accuracy only within the designed experimental matrix, this work explicitly evaluates model robustness through additional machining experiments conducted outside the Taguchi L27 orthogonal array but within the same factor-level ranges. By combining internal testing with independent validation under identical data preprocessing and evaluation criteria, the study provides a fair and systematic comparison of model generalisation behaviour. The key contribution of this work lies not in proposing a new modelling framework, but in demonstrating how different learning paradigms perform under realistic small-sample conditions, thereby offering practical guidance for selecting reliable predictive models in industrial WEDM applications and providing new insight into model selection beyond conventional accuracy-based comparisons.

2. Experimental procedures

2.1. Material and experimental setup

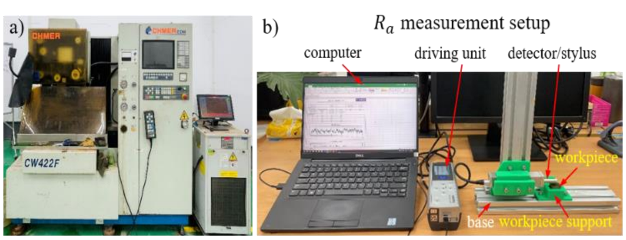

Figure 1. *a) WEDM machine, and b) Ra measurement setup*

The experiments were conducted using SKD61 tool steel [23], a hot-work die steel widely used in die-casting, forging, and extrusion applications owing to its high thermal fatigue resistance and toughness [24, 25]. These characteristics, however, make SKD61 difficult to machine using conventional cutting methods, thereby motivating the use of WEDM. All experiments were performed on a CNC wire electrical discharge machining system (CHMER CW422F), equipped with programmable control of pulse parameters and feed dynamics (Figure 1a). A brass wire electrode (Cu–Zn, 0.25 mm diameter) was selected for its good electrical conductivity and stable discharge characteristics when machining tool steels [1, 26].

Prior to machining, the workpiece thickness was maintained at 10 mm, and each trial was executed under identical dielectric flushing conditions to ensure consistency. A fresh wire was used for every experiment to eliminate the influence of wire wear on discharge stability. The general machining configuration and surface-roughness measurement setup followed standard industrial practice, consistent with prior WEDM studies [1, 7].

2.2. Machining parameters and surface roughness measurement

Table 1. *WEDM input parameters and their levels*

Parameter	Symbol	Level 1	Level 2	Level 3
Pulse-on time T-on, µs)	x_1	1	2	3
Pulse-off time (T-off, µs)	x_2	28	33	38
Servo voltage (SV, V)	x_3	45	50	55
Wire feed rate (WFR, mm/min)	x_4	6	7	8

Table 2*. Experimental matrix and measured responses.*

No.	Machining parameters				Ra (µm)
	T-on (µs)	T-off (µs)	SV (V)	WFR (mm/min)	
1	1	28	45	6	1.282
2	1	28	50	7	1.126
3	1	28	55	8	0.998
4	1	33	45	7	1.397
5	1	33	50	8	1.160
6	1	33	55	6	1.022
7	1	38	45	8	1.376
8	1	38	50	6	1.222
9	1	38	55	7	1.048
10	2	28	45	7	1.294
11	2	28	50	8	1.090
12	2	28	55	6	1.072
13	2	33	45	8	1.155
14	2	33	50	6	1.147
15	2	33	55	7	1.117
16	2	38	45	6	1.104
17	2	38	50	7	1.072
18	2	38	55	8	1.033
19	3	28	45	8	1.489
20	3	28	50	6	1.411
21	3	28	55	7	1.167
22	3	33	45	6	1.441
23	3	33	50	7	1.243
24	3	33	55	8	1.153
25	3	38	45	7	1.457
26	3	38	50	8	1.181
27	3	38	55	6	1.143

Four WEDM input parameters were selected for investigation, based on their documented influence on surface roughness (Ra): T-on, T-off, SV, and WFR. Each factor was assigned three levels, determined from preliminary screening tests to ensure stable sparking without wire rupture. Table 1 summarises the selected parameters and their levels.

A Taguchi L27 orthogonal array was employed, using four columns to represent the four three-level WEDM input parameters. This design substantially reduced the experimental effort compared with a full factorial design requiring 81 experiments. This design enables efficient exploration of factor effects and interactions using a relatively small number of experiments, aligning with established practices in WEDM research [3, 4]. The complete experimental matrix and measured outputs (Ra) are presented in Table 2. It should be noted that the experimental dataset presented in Table 2 was reused from a related optimisation-oriented companion study by the same authors [22]. In that companion study, the same WEDM experiments were used for multi-objective optimisation involving both Ra and MRR using PSI and Pareto-front analysis. In contrast, the present study uses the reused experimental dataset exclusively as a benchmark dataset for predictive modelling and comparative evaluation of Ra prediction performance using polynomial regression, ELM, ANN, and RFR. The overlap between the two studies is limited to the experimental dataset reported in Table 2. The modelling framework, machine-learning analysis, independent validation experiments, comparative results, discussion, and conclusions presented in the present work are original and were not included in the companion optimisation study.

Notably, Ra, as reported in Table 2, was measured using a HANDYSURF+35 (ACCRETECH) profilometer with a sampling length of 0.8 mm and a cut-off length of 4.0 mm, in accordance with JIS B 0601:2013 for surface texture profile evaluation. Five Ra measurements were taken at the four corners and the centre of each specimen, and the arithmetic mean of these measurements was used as the representative surface-roughness value for subsequent analysis. The standard deviation of repeated Ra measurements was consistently below 0.047 µm, indicating good measurement repeatability.

3. Development of regression models for predicting Ra

To predict surface roughness (Ra) from WEDM input parameters, regression models were developed using the experimental dataset. Four modelling approaches were considered: polynomial regression as a conventional statistical baseline, and three machine-learning models, namely ELM, ANN, and RFR, as illustrated in Figure 2. The dataset consisted of 27 experimental samples, each comprising four input features (T-on, T-off, SV, and WFR) and one output response (Ra), as presented in Table 2. Prior to model development, the dataset was divided into a training set (21 samples) and an internal test set (6 samples) using a fixed random seed to ensure reproducibility of the data partition. The training set was used for model development, whereas the internal test set was reserved exclusively for final model evaluation. All input features were normalised to the [0,1] range using min–max scaling based only on the training dataset to stabilise the training process and avoid data leakage. The adopted train–test framework was selected to maintain a consistent evaluation basis across all predictive models while enabling additional assessment using independent validation experiments. While repeated cross-validation may provide further insight into statistical stability under limited-data conditions, the present study primarily emphasises predictive behaviour on previously unseen parameter combinations through both internal testing and independent validation.

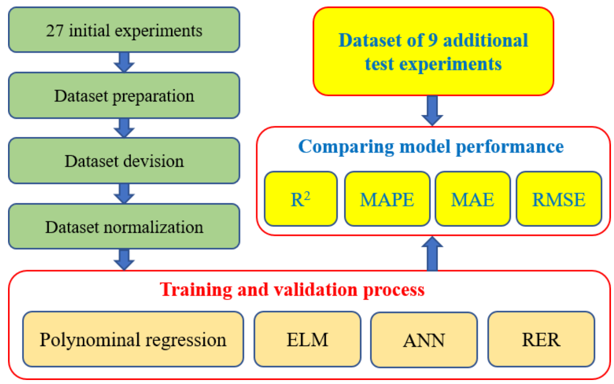

Figure 2. *Workflow of the model training and evaluation process*

Model performance was evaluated using standard statistical indicators, including the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE), following established practices in predictive modelling for manufacturing applications [11, 17, 19].

Polynomial regression was employed as a reference model to describe the relationship between WEDM input parameters and Ra. By incorporating polynomial terms and parameter interactions, polynomial regression can capture nonlinear behaviour commonly observed in machining processes. In this study, polynomial regression was implemented in Minitab (Minitab LLC, Pennsylvania, USA) to provide an empirical predictive baseline for comparison with machine-learning approaches [10]. To ensure a fair comparison with the machine-learning models, the polynomial regression model was developed using the same 21-sample training subset as ELM, ANN, and RFR. A stepwise regression procedure was applied to obtain a parsimonious polynomial formulation while maintaining predictive capability under limited-data conditions. Candidate linear, quadratic, and interaction terms were considered during model development, and the final formulation retained the statistically significant terms that balanced model simplicity and predictive performance.

ELM is a learning algorithm designed for single-hidden-layer feedforward neural networks (SLFNs) and is known for its high computational efficiency and good generalisation capability. In ELM, the input weights and hidden-layer biases are randomly assigned, while the output weights are analytically determined using a least-squares solution [11, 21, 27, 28]. This strategy significantly reduces training time while maintaining strong predictive accuracy. In the present study, the ELM model was configured with 12 hidden neurons and a hyperbolic tangent (tanh) activation function. The number of hidden neurons was selected empirically through preliminary trials to achieve stable prediction performance while avoiding underfitting and unnecessary model complexity under limited-data conditions.

ANNs are widely used for modelling complex nonlinear relationships between input and output variables. In this work, ANN models were developed to predict Ra using a two-hidden-layer architecture comprising 5 neurons in the first hidden layer and 12 neurons in the second hidden layer. A sigmoid activation function was employed, and network training was performed using the Adam optimisation algorithm for 2500 epochs with a batch size of 4. The adopted ANN architecture was selected empirically through preliminary trials to achieve satisfactory prediction accuracy while maintaining computational simplicity and avoiding excessive network complexity under limited-data conditions.

RFR is an ensemble learning method that combines predictions from multiple decision trees to improve robustness and predictive accuracy [19]. In this study, the RFR model was implemented in Python and trained using bootstrap sampling, square-root feature selection, parallel computation, and a fixed random seed to ensure reproducibility. To identify suitable model complexity under limited-data conditions, a manual grid search was conducted across several hyperparameter ranges. The number of trees varied from 8 to 100, the maximum tree depth from 3 to 10, the minimum number of samples required to split a node from 2 to 4, and the minimum number of samples per leaf from 1 to 4. Hyperparameter selection was performed during model development using the training dataset, while monitoring the balance between model complexity and overfitting tendency. The internal test dataset was reserved exclusively for final model evaluation and was not intended as a tuning criterion. The selected hyperparameter ranges were designed to provide sufficient model flexibility while reducing overfitting risk under small-sample conditions, rather than imposing fixed a priori settings that may not generalise across datasets.

Table 3. *Independent validation experiments within the investigated WEDM parameter ranges.*

No.	Machining parameters				Ra (µm)
	T-on (µs)	T-off (µs)	SV (V)	WFR (mm/min)	
1	1	28	45	7	1.263
2	1	28	45	8	1.249
3	1	28	50	6	1.123
4	2	33	45	6	1.156
5	2	33	50	7	1.146
6	2	33	55	8	1.118
7	3	38	50	6	1.177
8	3	38	55	7	1.145
9	3	38	55	8	1.148

In addition to internal testing using the Taguchi L27 dataset, nine independent validation experiments were conducted outside the L27 orthogonal array but within the same investigated factor-level ranges (Table 3). These experiments were excluded from both training and internal testing and were used to further evaluate model predictive capability using previously unseen parameter combinations. This validation strategy complements the internal testing procedure and enables comparative assessment of model behaviour within the investigated WEDM parameter space.

4. Results and discussion

4.1. Experimental results

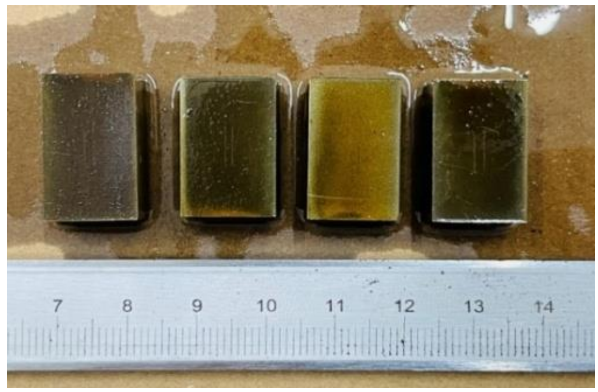

Figure 3. *Exemplified SKD61 steel samples after machining (from left to right, samples No. 7, 17, 22 and 24)*

Figure 3 illustrates representative machined samples obtained after the WEDM process, with the specimen numbers from left to right corresponding to experiments 7, 17, 22, and 24. Table 2 summarises the complete results of the 27 experiments carried out on SKD61 tool steel using the Taguchi L27 design. The measured surface roughness (Ra) values vary noticeably, ranging from 0.998 µm to 1.489 µm. This range reflects the strong influence of discharge parameters on surface quality and underscores the need to employ predictive modelling techniques to capture these variations accurately.

4.2. Performance of regression models for predicting Ra

The predictive performance of the developed regression models was evaluated using the testing dataset derived from the Taguchi L27 experiments. Figure 5 compares the predicted and experimentally measured Ra values obtained from polynomial regression, ELM, ANN, and RFR. Apparent differences in prediction accuracy are evident among the four modelling approaches.

The polynomial regression model was formulated without a constant term and employed uncoded process variables. Accordingly, Eq. (1) represents an empirical no-intercept polynomial model intended primarily for predictive representation of the experimental data rather than as a strict mechanistic or response-surface formulation. Consequently, the reported R^2, Adj-R^2, and Pred-R^2 values correspond to uncentered goodness-of-fit statistics and should be interpreted with appropriate caution. Therefore, these metrics were evaluated alongside residual analysis, internal testing performance, and independent validation results to provide a more comprehensive assessment of model fit and predictive capability. The resulting regression equation is presented in Eq. (1), and the corresponding ANOVA results are summarised in Table 4. The regression formulation exhibited statistically significant behaviour, yielding an F-value of 944.96 ($p < 0.0001$). In addition, the model achieved a coefficient of determination (R^2) of 99.56%, an adjusted coefficient of determination (Adj-R^2) of 99.45%, and a predicted coefficient of determination (Pred-R^2) of 99.35%. These statistical indicators suggest satisfactory fitting behaviour within the adopted 21-sample training dataset. The polynomial regression model was developed using the same 21-sample training subset as ELM, ANN, and RFR, thereby ensuring a fair comparison across all predictive approaches. Through stepwise regression and model simplification, a parsimonious formulation containing statistically significant terms was obtained. The resulting model exhibited satisfactory predictive performance within the adopted limited-data framework.

It should be noted that the final no-intercept formulation was retained during the model simplification procedure to obtain a parsimonious predictive model. Consequently, Eq. (1) should not be interpreted outside the investigated parameter ranges or as a physically based relationship. Instead, it is intended solely as an empirical predictive representation of the experimental observations within the studied WEDM conditions.

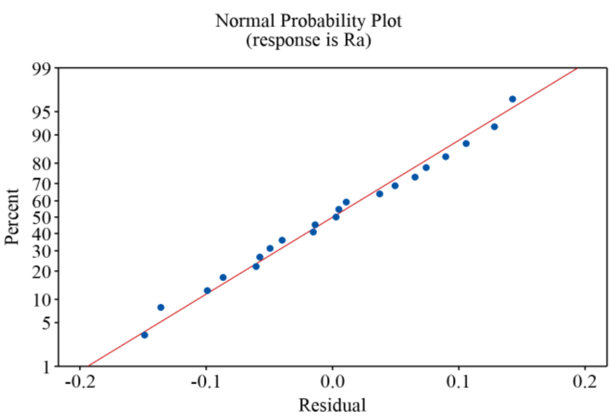

Figure 4. *Normal probability plot of residuals for the polynomial regression model developed using the 21-sample training dataset*

To further examine the polynomial regression model, a residual analysis was performed. Figure 4 shows the normal probability plot of the residuals obtained from the 21-sample training dataset. Most residual points lie close to the reference line, with no evident extreme outliers or pronounced departures from linearity. This observation indicates that no substantial violation of the normality assumption is apparent from the residual distribution. Accordingly, the uncentered R^2-based statistics were interpreted in conjunction with residual analysis, internal testing results, and independent validation performance to assess the predictive capability of the PR model.

The comparative predictive performance of the developed models on the internal testing dataset is summarised in Table 5, while the corresponding agreement between predicted and experimental Ra values is illustrated in Figure 5. Although the polynomial regression model demonstrated good fitting behaviour on the training dataset, its predictive capability on previously unseen data remained lower than that of ELM and ANN. The comparatively weaker performance observed during internal testing suggests that, despite improved statistical characteristics and simplified structure, parametric polynomial models may still exhibit limited generalisation capability under limited-data WEDM conditions. Accordingly, interpretation of the polynomial model was based not only on regression statistics but also on its behaviour under internal testing and independent validation.

Among the machine-learning models, both ELM and ANN demonstrated strong predictive performance compared with polynomial regression and RFR, as shown in Table 5 and Figure 5. The ANN model achieved an R^2 value of 0.99607, indicating close agreement between predicted and measured Ra values. Nevertheless, slight deviations were observed for several samples, reflecting the sensitivity of gradient-based learning to network architecture and training conditions. ELM achieved a comparable level of predictive performance ($R^2 = 0.99616$) and also showed close agreement with the experimental measurements across the test dataset. Considering both internal testing and independent validation results, ELM demonstrated promising predictive performance and predictive robustness under the investigated WEDM conditions.

Table 4. *ANOVA result for the regression model*

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	4	30.9036	99.55%	30.9036	7.72590	944.96	0.00000
x_2	1	30.3648	97.82%	1.0894	1.08938	133.24	0.00000
x_1^2	1	0.0822	0.26%	0.2406	0.24064	29.43	0.00005
x_2^2	1	0.2506	0.81%	0.4513	0.45128	55.20	0.00000
$x_1 \times x_3$	1	0.2060	0.66%	0.2060	0.20603	25.20	0.00011
Error	17	0.1390	0.45%	0.1390	0.00818		
Total	21	31.0426	100.00%				

$$Ra = 0.09546x_2 + 0.139x_1^2 - 0.001445x_2^2 - 0.0099x_1x_3 \quad (1)$$

The RFR model produces a noticeably lower prediction accuracy, with an R2 value of 0.88971. This reduced performance suggests that tree-based ensemble methods may be less effective at capturing the smooth, continuous relationship between WEDM parameters and surface roughness when the available dataset is limited. The piecewise-constant nature of decision-tree predictions can lead to discontinuities that are not well-suited to modelling gradual variations in Ra.

A quantitative comparison of model performance is summarised in Table 5 using multiple statistical indicators, including R2, MAE, MAPE, and RMSE. The ELM model consistently achieves the lowest error values (MAE = 0.0086, RMSE = 0.0093, MAPE = 0.68%), followed closely by the ANN model. Both machine-learning approaches significantly outperform polynomial regression and RFR, demonstrating a stronger capability to capture the nonlinear effects of WEDM process parameters on surface roughness.

In addition, Table 6 presents the point-wise prediction results and percentage errors for the internal testing dataset. Both ELM and ANN produced relatively small prediction errors, generally below 1.2%, indicating strong agreement between predicted and experimental Ra values. It is noted that the predicted values are reported to four decimal places to avoid masking small differences between model outputs. In contrast, RFR and PR exhibited larger deviations for several samples, reflecting comparatively lower predictive accuracy. These detailed prediction results are consistent with the statistical indicators reported in Table 5 and further confirm the high predictive capability of the ELM and ANN models under internal testing conditions.

Overall, the results confirm that machine-learning-based models generally provide improved predictive capability relative to polynomial regression within the investigated WEDM parameter range. Among the evaluated approaches, ELM exhibited a favourable balance between prediction accuracy and robustness. Because the internal testing subset consisted of only six samples, the associated performance metrics should be interpreted with appropriate caution. Therefore, model assessment in the present study was not based solely on the fixed train–test split but was complemented by independent validation using nine previously unseen experimental conditions within the investigated parameter domain.

While repeated cross-validation, repeated train–test splits, or leave-one-out validation may further improve statistical robustness under extremely limited datasets, the adopted fixed train–test framework was selected to maintain consistency across all evaluated models and to facilitate additional independent validation using previously unseen parameter combinations within the investigated parameter space. In the present work, this evaluation framework was further strengthened through nine independent validation experiments, providing additional verification of model generalisation under previously unseen parameter combinations within the investigated parameter space.

4.3. Validation using nine independent validation experiments

Figure 6 compares the predicted and experimentally measured Ra values obtained from nine independent validation experiments using different regression models. The results show that the ELM and ANN predictions agree more closely with the experimental measurements, whereas RFR exhibits larger deviations. Polynomial regression shows the most pronounced discrepancies, particularly in experimental cases 4 and 6. Overall, the ELM model demonstrated promising predictive behaviour under the investigated validation conditions.

To further assess predictive capability, nine independent validation experiments outside the L27 orthogonal array were conducted and used exclusively for independent validation. Table 7 summarises the predictive errors of the polynomial regression, ELM, ANN, and RFR models based on MAE, RMSE, and MAPE. The results clearly reveal substantial differences in predictive behaviour under previously unseen validation conditions within the investigated parameter ranges.

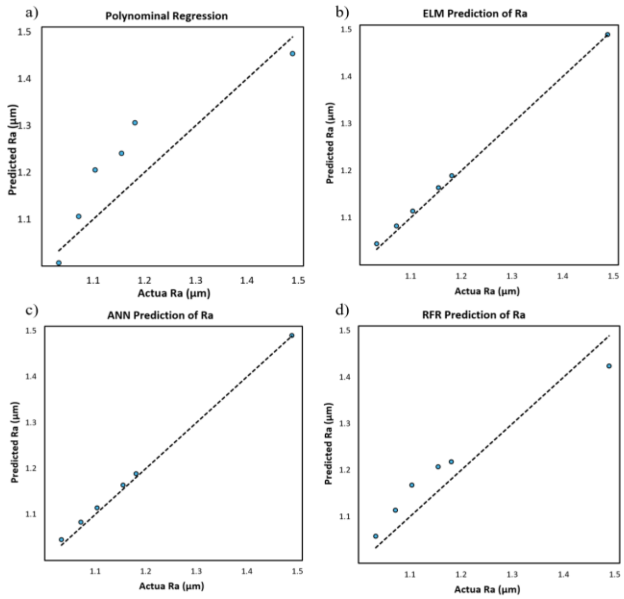

Figure 5. *Comparison of predicted and actual Ra values using different models: (a) Polynomial Regression, (b) ELM, (c) ANN, and (d) RFR*

Table 5. *Performance comparison of different regression models*

Model	R^2	MAPE (%)	MAE (μm)	RMSE (μm)
Polynomial regression	0.73149	5.86	0.0678	0.0777
ELM	0.99616	0.78	0.0086	0.0093
ANN	0.99607	0.78	0.0086	0.0094
RFR	0.88971	4.03	0.0476	0.0498

Table 6. *Experimental and predicted surface roughness values with percentage errors for the internal testing dataset*

No.	Actual Ra (μm)	ELM		ANN		RFR		PR	
		Predicted Ra (μm)	% Error	Predicted Ra (μm)	% Error	Predicted Ra (μm)	% Error	Predicted Ra (μm)	% Error
13	1.155	1.1640	0.78	1.1642	0.80	1.2075	4.55	1.2408	7.45
16	1.104	1.1142	0.92	1.1144	0.94	1.1680	5.80	1.2049	9.15
17	1.072	1.0830	1.03	1.0833	1.05	1.1139	3.91	1.1059	3.17
18	1.033	1.0449	1.15	1.0452	1.18	1.0578	2.40	1.0069	2.52
19	1.489	1.4900	0.07	1.4899	0.06	1.4238	4.38	1.4537	2.35
26	1.181	1.1894	0.71	1.1888	0.66	1.2184	3.17	1.3056	10.58

The ELM model achieved the lowest overall errors (MAE = 0.0099, RMSE = 0.0114, and MAPE = 0.85%), indicating a strong capability to capture the nonlinear relationship between WEDM parameters and surface roughness. In contrast, the ANN model exhibited moderate accuracy (MAPE = 1.98%), reflecting a reasonable but less stable generalisation performance. The RFR produced higher prediction errors (MAPE = 2.93%), suggesting that tree-based models are less suited for modelling smooth continuous responses such as Ra under limited data conditions.

Polynomial regression performed worst among all models, with an MAPE of 4.03%, indicating that low-order polynomial structures are insufficient to represent the complex discharge–material interaction dynamics inherent in WEDM. Overall, the independent validation results suggest that ELM demonstrated promising predictive performance under the investigated conditions.

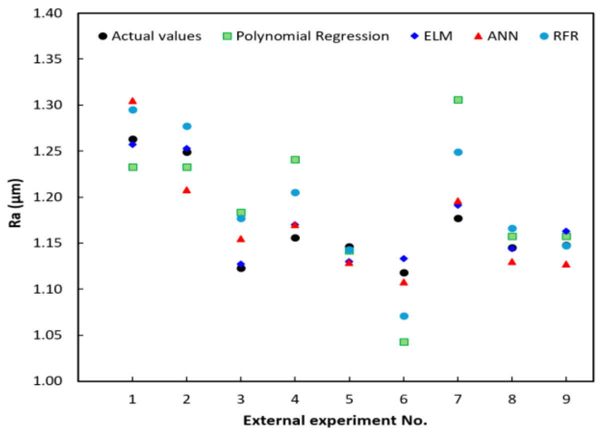

Figure 6. *Comparison of predicted and actual Ra values of nine independent validation experiments for different models*

Table 7. *Performance comparison of different models with nine independent validation experiments*

Model	MAPE (%)	MAE (µm)	RMSE (µm)
Polynomial regression	4.03	0.0467	0.0619
ELM	0.85	0.0099	0.0114
ANN	1.98	0.0234	0.0260
RFR	2.93	0.0341	0.0408

The observed differences in validation performance among ELM, ANN, and RFR may be attributed to their intrinsic modelling characteristics and the limited size of the available dataset. ANN models rely on iterative gradient-based optimisation, making them sensitive to local minima, weight initialisation, network depth, and hyperparameter selection [29, 30]. With only 21 training samples, ANN tends to approximate the training data rather than fully capture the underlying functional relationship, leading to reduced robustness under unseen machining conditions. This behaviour is consistent with prior reports indicating that neural networks may require careful architecture selection and sufficient training data to achieve stable generalisation [28–31].

For RFR, the reduced accuracy arises from the piecewise-constant approximation employed by decision trees, which limits their ability to model smooth nonlinear responses such as surface roughness in WEDM. Tree-based models partition the input space into discrete regions, potentially introducing discontinuities in the predicted response [19]. Moreover, random forests rely on an ensemble of decision trees and bootstrap aggregation; when trained on very small datasets, the resulting model may still be sensitive to data partitioning and local variations in the training samples [19].

Unlike ANN and RFR, ELM employs a single-pass analytical solution for output weight estimation, avoiding local minima and slow convergence [11]. The randomised hidden-layer mapping provides an effective nonlinear feature representation, enabling accurate approximation even with limited data. Previous studies have reported the suitability of ELM for small-sample regression problems due to its universal approximation capability and low model complexity [20, 21, 32], which is consistent with the stable performance observed in the present study.

A further point that merits consideration is the influence of model architecture on predictive performance. Although the adopted ELM and ANN configurations provided satisfactory results within the present training, testing, and independent validation framework, the available dataset remains relatively small for a comprehensive investigation of architecture sensitivity. Consequently, the effects of alternative hidden-layer structures and neuron numbers were not examined in detail in the present study. This limitation should be acknowledged when interpreting the reported results, and it represents a worthwhile direction for future research involving larger datasets and more extensive architectural analyses.

Table 8*. Experimental and predicted surface roughness values with percentage errors for nine independent validation experiments*

No.	Actual Ra (µm)	ELM		ANN		RFR		PR	
		Predicted Ra (µm)	% Error	Predicted Ra (µm)	% Error	Predicted Ra (µm)	% Error	Predicted Ra (µm)	% Error
1	1.263	1.257	0.48	1.305	3.33	1.295	2.53	1.233	2.38
2	1.249	1.253	0.32	1.208	3.28	1.277	2.24	1.233	1.28
3	1.123	1.127	0.36	1.155	2.85	1.177	4.81	1.184	5.43
4	1.156	1.170	1.21	1.170	1.21	1.205	4.24	1.241	7.35
5	1.146	1.130	1.40	1.129	1.48	1.143	0.26	1.142	0.35
6	1.118	1.133	1.34	1.108	0.89	1.071	4.20	1.043	6.71
7	1.177	1.191	1.19	1.196	1.61	1.249	6.12	1.306	10.96
8	1.145	1.144	0.09	1.130	1.31	1.166	1.83	1.157	1.05
9	1.148	1.163	1.31	1.127	1.83	1.147	0.09	1.157	0.78

Table 8 summarises the experimental and predicted Ra values, along with the corresponding percentage errors, for nine independent validation experiments. These detailed prediction results provide additional insight into model behaviour under previously unseen parameter combinations within the investigated parameter space and complement the statistical indicators reported earlier. The results indicate that ELM maintained consistently low prediction errors across the validation experiments, with percentage errors generally remaining below approximately 1.5%. ANN also demonstrated good predictive capability, although slightly larger deviations were observed in several cases. By contrast, RFR and PR exhibited greater variability and larger prediction errors for certain validation conditions. Overall, these findings support the independent validation results and suggest that ELM demonstrated promising predictive accuracy and generalisation capability under the investigated WEDM conditions.

In addition, prior research on surface roughness prediction in WEDM has primarily focused on individual modelling approaches and evaluated performance within designed experimental datasets. For example, Paturi et al.

developed an ANN model to predict surface roughness in WEDM of Inconel 718, demonstrating the ability of neural networks to capture nonlinear relationships between machining parameters and surface quality [14, 16, 33, 34]. Similarly, machine-learning techniques such as ELM, weighted ELM, support vector regression (SVR), and quantum SVR have been applied to predict surface roughness of machined aluminium alloy in WEDM, achieving high R^2 values within the design space [13]. However, these studies typically do not include systematic comparisons across multiple predictive models or independent validation using previously unseen experiments. In contrast, the present work incorporates both internal testing and independent validation on previously unseen parameter combinations within the investigated parameter space, thereby providing additional assessment of model generalisation behaviour and predictive reliability for WEDM process planning. This independent validation strategy was adopted to provide an additional assessment of predictive behaviour using genuinely unseen observations, thereby complementing the internal testing results obtained from the limited experimental dataset.

5. Conclusions

The present study provides a comparative evaluation of three machine-learning models, ELM, ANN, and RFR, for predicting Ra in the WEDM of SKD61 tool steel. By combining a Taguchi L27 dataset with additional independent validation experiments, the study emphasises model generalisation under limited data conditions, addressing a key limitation of many existing WEDM studies that assess performance only within the experimental design space.

The results reveal clear differences in predictive behaviour among the evaluated models. ELM consistently achieved the lowest MAE, RMSE, and MAPE values across both internal testing and independent validation experiments, indicating promising predictive performance under the investigated conditions. ANN also demonstrated strong predictive capability, although somewhat larger prediction errors were observed during the independent validation experiments, while RFR produced higher prediction errors, indicating limited suitability for modelling the smooth and continuous Ra response under small-sample constraints.

These findings highlight that evaluating predictive models solely based on fitting accuracy can be misleading and confirm that independent validation experiments are valuable for assessing practical reliability in WEDM applications. Overall, under the specific training, testing, and independent validation framework adopted in this study, ELM demonstrated promising predictive performance and may provide a practical framework for Ra prediction under the investigated WEDM conditions. Future work may focus on expanding the experimental dataset, incorporating real-time process signals, or integrating ELM with optimisation techniques to further enhance intelligent WEDM process planning.

Acknowledgement: This study is made possible thanks to the CAD/CAM/CNC Lab at the Faculty of Mechanical Engineering, Can Tho University, for their permission to use the WEDM machine for the experiment.

REFERENCES

[1] K. Ho, S. Newman, S. Rahimifard, and R. D. Allen, "State of the art in wire electrical discharge machining (WEDM)", *Int. J. Mach. Tools Manuf.* vol. 44, pp. 1247–1259, 2004. https://doi.org/10.1016/j.ijmachtools.2004.04.017

[2] K. P. Rajurkar, M. M. Sundaram, and A. P. Malshe, "Review of Electrochemical and Electrodischarge Machining", *Procedia CIRP,* vol. 6, pp. 13–26, 2013. https://doi.org/10.1016/j.procir.2013.03.002

[3] D. Scott, S. Boyina, and K. P. Rajurkar, "Analysis and optimization of parameter combinations in wire electrical discharge machining", *Int. J. Prod. Res*, vol. 29, pp. 2189–2207, 1991. https://doi.org/10.1080/00207549108948078

[4] T. A. Spedding and Z. Q. Wang, "Parametric optimization and surface characterization of wire electrical discharge machining process", *Precis. Eng,* vol. 20, pp. 5–15, 1997. https://doi.org/10.1016/S0141-6359(97)00003-2

[5] M. Rozenek, J. Kozak, L. Dabrowski, and K. Lubkowski, "Electrical discharge machining characteristics of metal matrix composites", *J. Mater. Process. Technol,* vol. 109, pp. 367–370, 2001. https://doi.org/10.1016/S0924-0136(00)00823-2

[6] S. S. Mahapatra and A. Patnaik, "Optimization of wire electrical discharge machining (WEDM) process parameters using Taguchi method", *The International Journal of Advanced Manufacturing Technology*, vol. 34, pp. 911–925, 2006. https://doi.org/10.1007/S00170-006-0672-6

[7] N. Naeim, M.A. AbouEleaz, and A. Elkaseer, "Experimental Investigation of Surface Roughness and Material Removal Rate in Wire EDM of Stainless Steel 304", *Materials,* vol. 16, pp. 1022-1038, 2023. https://doi.org/10.3390/ma16031022

[8] V. Aggarwal, S.S. Khangura, and R.K. Garg, "Parametric modeling and optimization for wire electrical discharge machining of Inconel 718 using response surface methodology", *The International Journal of Advanced Manufacturing Technology*, vol. 79, pp. 31–47, 2015. https://doi.org/10.1007/S00170-015-6797-8

[9] P.N.Q. Huy, T. Kha, T.B. Loc, N. Van Cuong, N.D. Tu, and H.T. Thuong, "Studying the effects of cutting parameters on surface roughness in milling of aluminium alloy using the Taguchi method", *The University of Danang - Journal of Science and Technology*, vol. 23, No. 11, pp. 54–59, 2025. https://doi.org/10.31130/UD-JST.2025.23(11).331.

[10] M.A. Abdullah, B.A. Ahmed, and S.K. Ghazi, "Enhancing of Material Removal Rate and Surface Roughness in Wire EDM Process using Grey Relational Analysis", *Engineering, Technology & Applied Science Research,* vol. 14, pp. 17422–17427, 2024. https://doi.org/10.48084/etasr.8450

[11] G. Bin Huang, Q.Y. Zhu, and C.K. Siew, "Extreme learning machine: Theory and applications*", Neurocomputing,* vol. 70, pp. 489–501, 2006. https://doi.org/10.1016/j.neucom.2005.12.126

[12] T. Thankachan *et al.*, "Prediction of surface roughness and material removal rate in wire electrical discharge machining on aluminum based alloys/composites using Taguchi coupled Grey Relational Analysis and Artificial Neural Networks", *Appl. Surf. Sci*, vol. 472, pp. 22–35, 2019. https://doi.org/10.1016/j.apsusc.2018.06.117

[13] M. Ulas, O. Aydur, T. Gurgenc, and C. Ozel, "Surface roughness prediction of machined aluminum alloy with wire electrical discharge machining by different machine learning algorithms", *Journal of Materials Research and Technology,* vol. 9, pp. 12512–12524, 2020. https://doi.org/10.1016/j.jmrt.2020.08.098

[14] S. Shakeri, A. Ghassemi, M. Hassani, and A. Hajian, "Investigation of material removal rate and surface roughness in wire electrical discharge machining process for cementation alloy steel using artificial neural network", *The International Journal of Advanced Manufacturing Technology*, vol. 82, pp. 549–557, 2015. https://doi.org/10.1007/S00170-015-7349-Y

[15] G. Zhang, Z. Zhang, J. Guo, W. Ming, M. Li, and Y. Huang,

"Modeling and Optimization of Medium-Speed WEDM Process Parameters for Machining SKD11", *Materials and Manufacturing Processes*, vol. 28, pp. 1124–1132, 2013. https://doi.org/10.1080/10426914.2013.773024

[16] U. M. R. Paturi, S. Cheruku, S. Salike, V. P. K. Pasunuri, and N.S. Reddy, "Estimation of machinability performance in wire-EDM on titanium alloy using neural networks", *Materials and Manufacturing Processes*, vol. 37, pp. 1073–1084, 2022. https://doi.org/10.1080/10426914.2022.2030875

[17] C. Naresh, P.S.C. Bose, and C.S.P. Rao, "Artificial neural networks and adaptive neuro-fuzzy models for predicting WEDM machining responses of Nitinol alloy: comparative study", *SN Applied Sciences,* vol. 2, pp. 314-337, 2020. https://doi.org/10.1007/S42452-020-2083-Y

[18] H.T. Thuong and N. Van Cuong, "Prediction of surface roughness in finish milling of 6061 aluminium alloy using artificial neural network and extreme learning machine", *The University of Danang - Journal of Science and Technology,* vol. 24, No. 3, pp. 87–93, 2026. https://doi.org/10.31130/UD-JST.2026.24(3).724E

[19] L. Breiman, "Random forests", *Mach. Learn,* vol. 45, pp. 5–32, 2001. https://doi.org/10.1023/A:1010933404324

[20] S. Ding, X. Xu, and R. Nie, "Extreme learning machine and its applications", *Neural Computing and Applications*, vol. 25, pp. 549–556, 2013. https://doi.org/10.1007/S00521-013-1522-8

[21] G. Bin Huang, D.H. Wang, and Y. Lan, "Extreme learning machines: a survey", *International Journal of Machine Learning and Cybernetics,* vol. 2, pp. 107–122, 2011. https://doi.org/10.1007/s13042-011-0019-y

[22] H. T. Thuong, C. H. Tien, N. D. Tu, "Integrated PSI–Pareto front-based multi-objective optimisation of WEDM parameters for SKD61 tool steel", *International Journal of Advanced Technology and Engineering Exploration*, 2026. Accepted 18/06/2026.

[23] S.H. Yeh, L.H. Chiu, Y.T. Pan, and S.C. Lin, "Relative Dimensional Change Evaluation of Vacuum Heat-Treated JIS SKD61 Hot-Work Tool Steels", *Journal of Materials Engineering and Performance*, vol. 23, pp. 2075–2082, 2014. https://doi.org/10.1007/S11665-014-0961-4

[24] M. Wang, Y. Wu, Q. Wei, and Y. Shi, "Thermal Fatigue Properties of H13 Hot-Work Tool Steels Processed by Selective Laser Melting", *Metals*, vol. 10, pp. 116-126, 2020. https://doi.org/10.3390/MET10010116

[25] H. Nguyen Le Dang, V.T. Nguyen, V.H. Hoang, X.T. Vo, and V.T.T. Nguyen, "Durability Comparison of SKD61 and FDAC Steel Mold Inserts in High-Pressure Die-Casting Process", *Machines*, vol. 13, pp. 352-365, 2025. https://doi.org/10.3390/MACHINES13050352

[26] A. Muttamara and P. Nakwong, "Enhancing Wire-EDM Performance with Zinc-Coated Brass Wire Electrode and Ultrasonic Vibration", *Micromachines,* vol. 14, pp. 862-876, 2023. https://doi.org/10.3390/MI14040862

[27] J. Wang, S. Lu, S.H. Wang, and Y.D. Zhang, "A review on extreme learning machine", *Multimedia Tools and Applications*, vol. 81, pp. 41611–41660, 2022. https://doi.org/10.1007/S11042-021-11007-7

[28] F.J. Pontes, J.R. Ferreira, M.B. Silva, A.P. Paiva, and P.P. Balestrassi, "Artificial neural networks for machining processes surface roughness modeling", *The International Journal of Advanced Manufacturing Technology*, vol. 49, pp. 879–902, 2009. https://doi.org/10.1007/S00170-009-2456-2

[29] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer New York, NY, 2006

[30] S. Kumar, A. Batish, R. Singh, and T.P. Singh, "A hybrid Taguchi-artificial neural network approach to predict surface roughness during electric discharge machining of titanium alloys", *Journal of Mechanical Science and Technology,* vol. 28, pp. 2831–2844, 2014. https://doi.org/10.1007/S12206-014-0637-X

[31] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, pp. 436–444, 2015. https://doi.org/10.1038/nature14539

[32] G.B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification", *IEEE Transactions on Systems, Man, and Cybernetics, Part B* (Cybernetics), vol. 42, pp. 513-529, 2012. https://doi.org/10.1109/TSMCB.2011.2168604

[33] U.M.R. Paturi, H. Devarasetti, N.S. Reddy, N. Kotkunde, and B.K. Patle, "Modeling of surface roughness in wire electrical discharge machining of Inconel 718 using artificial neural network", *Mater. Today Proc*, vol. 38, pp. 3142–3148, 2021. https://doi.org/10.1016/j.matpr.2020.09.503

[34] S. Sarkar, S. Mitra, and B. Bhattacharyya, "Parametric optimisation of wire electrical discharge machining of γ titanium aluminide alloy through an artificial neural network model", *The International Journal of Advanced Manufacturing Technology*, vol. 27, pp. 501–508, 2006. https://doi.org/10.1007/S00170-004-2203-7.