Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science. 2026 vol. 28 no. 2 pp. 318–334
ISSN: 1994-6309 (print) / 2541-819X (online)
DOI: 10.17212/1994-6309-2026-28.2-318-334

NSTU NETI

Obrabotka metallov -
Metal Working and Material Science

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

Wear rate prediction of PVD hard coatings using a hybrid Taguchi–RSM–machine learning framework with SHAP-based interpretability

Umesh Subhash Patharkar [1, a], *Sunil Apparao Patil* [1, b], *Nitin Ambhore* [2, c, *]

[1] Government Engineering College, Chh. Sambhajinagar, 431005, Maharashtra, India
[2] Vishwakarma Institute of Technology, SPPU, Pune 411037, Maharashtra, India

[a] https://orcid.org/0009-0005-9303-641X, patharkarumeshsubhash@gmail.com; [b] https://orcid.org/0009-0000-4804-0720, sakpatil967@gmail.com;
[c] https://orcid.org/0000-0001-8468-8057, nitin.ambhore@vit.edu

ARTICLE INFO

Article history:
Received: 06 April 2026
Revised: 11 April 2026
Accepted: 30 April 2026
Available online: 15 June 2026

Keywords:
PVD coatings
Wear
Response surface methodology
Random Forest
XGBoost
SHAP analysis

ABSTRACT

Introduction. The wear behaviour of physical vapour deposition (*PVD*) hard coatings is difficult to predict because of the nonlinear relationship between coating chemistry, deposition thickness, operating temperature, and applied contact load. Multi-dimensional parameter mapping through exhaustive experimentation is both time-consuming and costly. ***The purpose of the work*** is to implement a low-data hybrid architecture that combines *Taguchi* design, response surface methodology (*RSM*), and machine learning (*ML*) to predict and optimize wear performance based on a small dataset (***n*** = 28). **Methods.** Three coating types – *AlTiN*, *CrN*, and *TiC* – were tested at temperature range of 40–50°C, contact load range of 5–15 N, and coating thickness range of 2–4 µm. Analysis of variance (*ANOVA*) was performed to identify the most influential parameter affecting wear. A predictive wear model was developed using response surface methodology. **Results and discussion.** *ANOVA* revealed that load and coating chemistry are the most influential factors affecting wear. Temperature and thickness were not found to be significant within the studied range. A *RSM* model was found statistically significant for predicting contact load and coating chemistry with R^2 = 0.901 (*Adj*-R^2 = 0.834, *RMSE* = 0.0076). Random forest (*RF*) had highest generalisation performance (*CV* R^2 = 0.755) and gradient boosting (*GB*) had the best foverall fit (R^2 = 0.913) among the *ML* models tested using five-fold cross-validation. *SHAP* analysis indicated that coating chemistry formed the major contribution, then contact load, and little temperature contribution. Gradient boosting optimisation indicated that *AlTiN* at 4 µm thickness, 15 N load and 50 °C are the preferred settings and the expected wear rate is 0.010823 (mm^3/Nm). The proposed framework demonstrates credible, interpretable wear forecasting using limitted experimental data.

For citation: Patharkar U., Patil S., Ambhore N. Wear rate prediction of PVD hard coatings using a hybrid Taguchi–RSM–machine learning framework with SHAP-based interpretability. *Obrabotka metallov (tekhnologiya, oborudovanie, instrumenty) = Metal Working and Material Science*, 2026, vol. 28, no. 2, pp. 318–334. DOI: 10.17212/1994-6309-2026-28.2-318-334. (In Russian).

Introduction

Physical vapour deposition (*PVD*) has become a widely used method for depositing hard coatings, which are required to improve the performance and durability of cutting tools, forming dies, and precision mechanical components. *PVD* coatings are beneficial in that they redistribute stresses at the contact interface, decrease adhesive transfer, and alleviate abrasive wear mechanisms by introducing a thin, wear-resistant layer at the contact interface, thereby greatly increasing the service life of components [1–2]. Among the various types of coating systems, aluminium titanium nitride (*AlTiN*) , chromium nitride (*CrN*) , and titanium carbide (*TiC*) have received significant interest because of their unique mechanical and chemical properties.

*** Corresponding author**
Nitin Ambhore, Ph.D. (Engineering), Associate Professor
Vishwakarma Institute of Technology,
Pune - 411037, Maharashtra, India
Tel.: +91-2026950441, **e-mail:** nitin.ambhore@vit.edu

It is well known that *AlTiN* coatings are hard and exhibit excellent oxidation resistance, which is due to the presence of a stable alumina tribolayer during sliding. In contrast, *CrN* coatings are better in terms of corrosion resistance and lowering the coefficient of friction, and thus are applicable to aggressive conditions, whereas *TiC* coating offers high hardness but is susceptible to brittle fracture at high loads [3, 4]. In spite of these strengths, the tribological behaviour of *PVD* coatings is controlled by a complex interaction of process variables (e.g., coating thickness, temperature, and applied load), and the ability to analytically predict these variables is difficult [5].

Conventional design-of-experiment (*DoE*) techniques such as the *Taguchi* method and response surface methodology (*RSM*) have been widely used to optimise tribological parameters. The *Taguchi* method allows efficient screening of parameters using orthogonal arrays, while *RSM* allows continuous modelling and statistical inference with the help of polynomial regression [6–7]. Nevertheless, these techniques are fundamentally restricted in their ability to capture strong nonlinearities and higher-order interactions that exist in wear processes [8].

To overcome these drawbacks, machine learning (*ML*) methods have become effective tools for modelling complex tribological systems. Random forest and gradient boosting are examples of ensemble techniques that have shown great predictive power in the analysis of wear and coating performance, especially in data-limited environments [9–10]. Recent literature has also demonstrated the utility of *ML* in predicting wear behaviour in coating systems and materials, although most approaches are not interpretable or compatible with physics-based design models [11–12]. Furthermore, explainable artificial intelligence (*XAI*) systems, including *SHAP*, have enabled a deeper understanding of feature contributions, enhancing the transparency and trustworthiness of the model [13–14].

In spite of these developments, an integrated framework combining statistical design, nonlinear modelling, and interpretable machine learning for *PVD* coating systems is still lacking. The available literature tends to use either single *DoE* methods or single *ML* models, and cross-method validation and data efficiency are not adequately addressed.

The current paper fills this gap by creating a three-level hybrid model that integrates *Taguchi* analysis, *RSM*, and *ML* to predict the wear rate of *PVD* coatings. The ***objectives*** include:

– to statistically establish the significance of coating type, load, temperature, and thickness on *PVD* coating wear rate through *Taguchi–RSM* analysis;

– to evaluate the predictive capability of ensemble machine learning models under conditions of limited sample size ($\boldsymbol{n}$ = 28) for tribological prediction;

– to identify and quantify nonlinear interactions between input factors using *SHAP*-based feature importance analysis;

– to determine physically realistic optimal coating configurations validated through hybrid statistical–*ML* consensus. This integrated framework offers a data-efficient and interpretable approach to tribological modelling.

Methods

Three *PVD* hard coating chemistries – aluminium titanium nitride (*AlTiN*), chromium nitride (*CrN*) , and titanium carbide (*TiC*) – were investigated since they have different tribological and mechanical properties [15–17]. Coatings were deposited on hardened tool-steel substrates using a commercial cathodic arc *PVD* system. The deposition parameters were optimised for each coating chemistry: *AlTiN* was deposited at a nitrogen pressure of 3.5 Pa, target current of 150 A, substrate temperature of 450 °C, and deposition time of 90 min; *CrN* at 3.0 Pa nitrogen pressure, 120 A target current, 400 °C substrate temperature, and 80 min deposition time; and *TiC* at 2.5 Pa acetylene/argon pressure, 130 A target current, 420 °C substrate temperature, and 85 min deposition time. These parameters were selected to ensure uniform coating quality and adhesion.

The coating thickness was controlled and verified by both ball-cratering and cross-sectional scanning electron microscopy (*SEM*) , yielding nominal thicknesses of 2 μm, 3 μm, and 4 μm. A pin-on-disc tribometer was used to test tribological performance under dry sliding conditions. The counter-body was a cemented carbide ball (6 mm diameter). At controlled environmental temperatures of 40 °C, 45 °C, and 50 °C, normal loads of 5 N, 10 N, and 15 N were applied, ensuring repeatable testing conditions.

The wear volume was calculated using white-light interferometry profilometry of the wear scar, and the specific wear rate (W_s) was computed as:

$$W_S = \frac{V}{(F - L)}, \tag{1}$$

where V is the wear volume (mm^3); F is the normal load (N); L is the sliding distance (m).

The experimental conditions were repeated three times to ensure statistical reliability, and the average wear rate was used for further analysis. After data quality screening to remove outliers and inconsistent measurements, 28 experimental observations were selected for modelling. Tribological studies of coated systems have extensively used similar experimental procedures [18–19]. The experimental design spans four factors: coating type, temperature (40, 45, 50 °C), contact load (5, 10, 15 N), and coating thickness (2, 3, 4 µm). Table 1 reports the various factors and their ranges.

Table 1

Control factors and levels

Factor	Level 1	Level 2	Level 3
Coatings	*AlTiN, CrN, TiC*	*AlTiN, CrN, TiC*	*AlTiN, CrN, TiC*
Temperature, T (°C)	40	45	50
Load, L (N)	5	10	15
Thickness, K (µm)	2	3	4

The proposed research relies on a three-level hybrid analysis model that combines traditional design of experiment (*DoE*) methods and data-driven machine learning (*ML*) models to enable robust and interpretable wear prediction. The framework is designed in sequence, and each level provides complementary information and cross-validation (*CV*). Tier **1** (*Taguchi* analysis) allows the rapid screening of influential parameters in noise-robust scenarios; Tier **2** (*RSM*) builds a continuous empirical model and tests its statistical significance; Tier **3** (machine learning regression) captures nonlinear interactions and enables generalisation and feature importance explanation [20–25]. The combination of statistical and *ML* methods has become widely recognized as best practice in modelling multifaceted tribological systems [26–27].

Fig. 1 depicts the entire analytical process, in which all three levels are applied to the same dataset (n = 28), ensuring consistency and comparability of the results across methodologies. Table 2 shows a summary of the three-tier analytical methodology.

Taguchi technique is used to find out the optimum set of parameters for each performance characteristic by observing the values of the signal-to-noise (*S*/*N*) ratio. *Taguchi* used the *S*/*N* ratio as the quality characteristic for the process. A *Taguchi* smaller-is-better *S*/*N* ratio converts each of the measured wear rates into a noise-resistant quality statistic that simultaneously captures response mean and variance. The *Taguchi* smaller-is-better equation is given as follows:

$$\frac{S}{N} = -10 \log \frac{1}{n}\left(\sum y^2\right), \tag{2}$$

where y is the observed data; n is the number of observations.

RSM provides a statistical empirical model, which uses low-order polynomials to fit experimental data, enabling visualisation of the response surface, and determination of factor significance by *ANOVA* [13]. The objective of establishing the model is to correlate the response variables with the independent

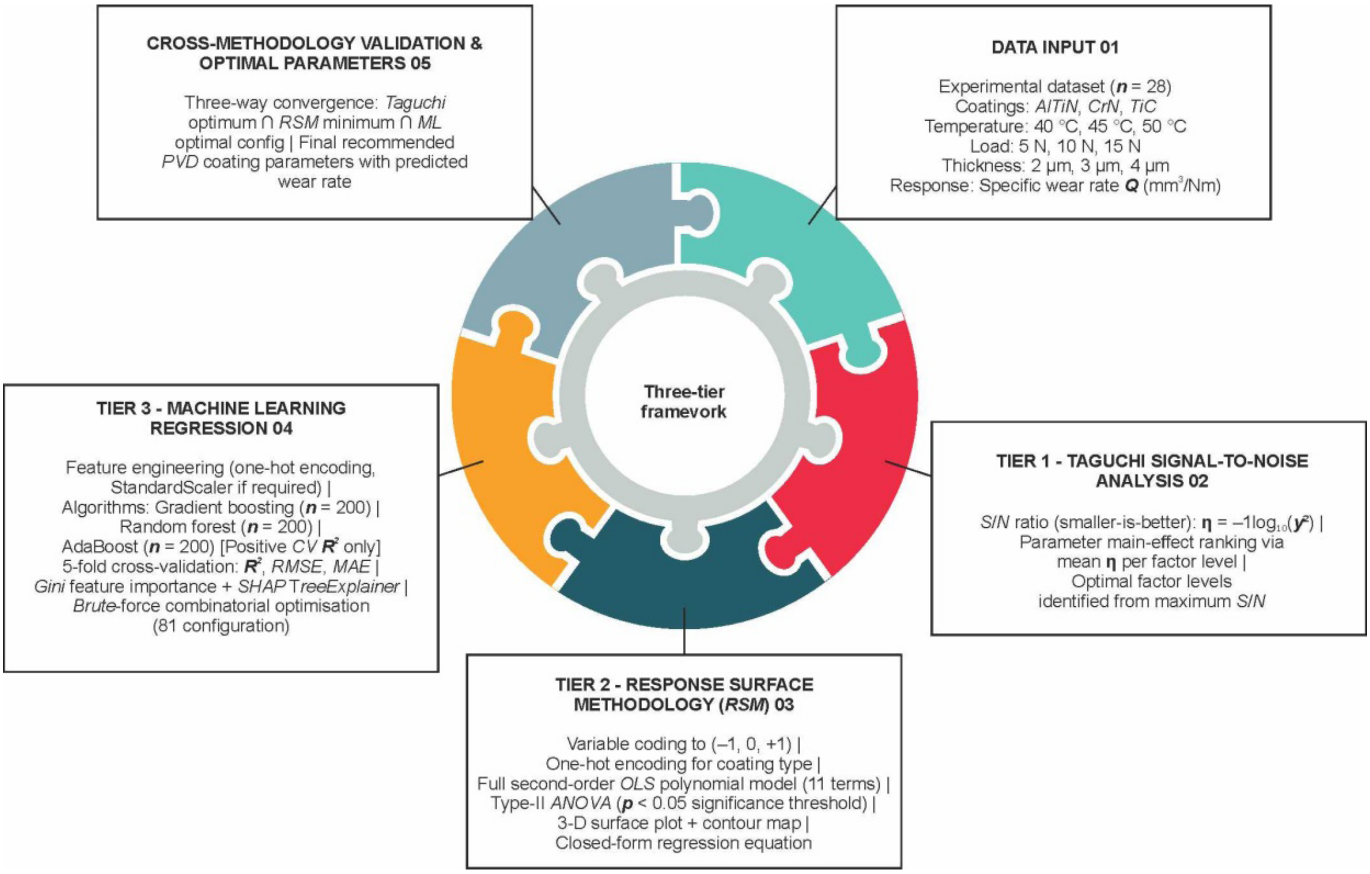

Fig. 1. Three-tier *Taguchi–RSM–ML* analytical framework

Table 2

Overview of the three-tier analytical methodology

Tier	Method	Key Mathematical Basis	Primary Outputs
1	*Taguchi S/N* analysis	$\eta = -10 \cdot \log_{10}(y^2)$ (smaller-is-better)	*S/N* ratios; main-effect ranking; optimal levels
2	Response surface methodology	Full quadratic *OLS* model (11 terms); Type-II *ANOVA*	Polynomial equation; *ANOVA* table; R^2; surface plots
3*a*	Gradient boosting	Sequential boosted 200 trees; *L2* loss regularisation	*CV* R^2, *RMSE*, *MAE*; *Gini* feature importance
3*b*	Random forest	Decision tree ensemble; 200 trees; exponential loss weighting	*CV* R^2, *RMSE*, *MAE*; prediction
3*e*	*SHAP* Interpretability	Shapley additive attributions; TreeExplainer backend	Per-feature *SHAP* values; directional attribution
Optimal	*ML* combinatorial optimisation	Exhaustive search over 81 discrete configurations	Globally optimal parameters; minimum $\hat{S}$ (predicted wear)

variables. In this work, experimental results are analyzed using a second-order polynomial regression model as follows:

$$Y = a_0 + \sum_{i=1}^{k} a_i x_i + \sum_{i=1}^{k} a_{ii} x_i^2 + \sum_{i<j}^{k} a_{ij} x_i x_j, \tag{3}$$

where $\boldsymbol{Y}$ is the desired response; $\boldsymbol{a_0}$ is the constant; $\boldsymbol{a_i}$, $\boldsymbol{a_{ii}}$, and $\boldsymbol{a_{ij}}$ represent the coefficients of the linear, quadratic, and cross-product terms, respectively; $\boldsymbol{x_i}$ represents the coded variables that correspond to the independent variables.

The statsmodels (*v0.14*) library was used to perform ordinary least-squares (*OLS*) estimation. A Type-II sum-of-squares *ANOVA* table was created to evaluate the statistical significance of each model term at a significance level of 5% (p < 0.05). Model adequacy was assessed using R^2, adjusted R^2, root mean squared error (*RMSE*), the *Omnibus* normality test of the residuals, and the *Durbin-Watson* autocorrelation coefficient.

For the machine learning models, the categorical coating variable was one-hot encoded (drop first = True) to produce two binary indicator columns: *Coating_CrN* and *Coating_TiC*, with *AlTiN* as the implicit reference category. The resulting *ML* feature matrix thus included five columns: Temperature (***T***), Load (***L***), Thickness (***K***), *Coating_CrN*, and *Coating_TiC*. No further feature scaling was used for tree-based models (Gradient boosting, AdaBoost, and Random forest), which are scale-invariant. The scikit-learn, *LightGBM*, and *TensorFlow* libraries were used to implement four regression algorithms as follows:

Gradient boosting: a gradient-boosted ensemble of 200 trees, optimized in series, with each new tree being optimized against the negative gradient (pseudo-residuals) of the loss function *L2*. Maximum depth was 6, and the learning rate was 0.05 to balance bias and variance. Gradient boosting can work well with small datasets because early stopping can be used to prevent overfitting when tested on hold-out folds [15].

AdaBoost: an adaptive boosting ensemble of 200 trees, which re-weights samples misclassified by the algorithm in a cascading manner, forcing the algorithm to focus on hard-to-predict examples. The contribution of successive trees is regulated by a learning rate of 0.05. AdaBoost has been reported to have strong generalisation with small sample sizes on tabular data [16].

Random Forest: a bagging model consisting of 200 decision trees trained on random subsamples of the data and random features, with maximum depth of 15 and minimum samples split of 5. Instead of boosting, Random forest averages the independent trees to decrease variance while preserving interpretability. Its out-of-bag error estimate provides objective cross-validation scores, and Random forest is robust on small datasets and does not suffer from convergence issues like deep learning approaches [17].

The hyperparameters for all *ML* models were selected based on a combination of literature recommendations and preliminary experimentation. For Gradient boosting, 200 estimators with a learning rate of 0.05 and maximum depth of 6 were chosen to balance the bias-variance trade-off for small datasets, following guidelines from *Friedman* (2001). Random forest used 200 trees with maximum depth of 15 and min_samples_split of 5, as recommended for small-to-medium tabular datasets. AdaBoost employed 200 weak learners with a learning rate of 0.05 to ensure gradual convergence. A formal grid search was not conducted due to the limited dataset size (n = 28), as exhaustive hyperparameter tuning on small datasets risks overfitting the validation folds. A formal sensitivity analysis varying the number of estimators (e.g., 50, 100, 200, 500) was not conducted; however, the conservative learning rate of 0.05 inherently limits the marginal contribution of later trees, reducing sensitivity to this parameter. Instead, conservative hyperparameter values were adopted from established literature on ensemble methods for small datasets.

Each *ML* model was evaluated using a stratified five-fold cross-validation strategy (*KFold*, ***n**_splits* = 5, *shuffle* = *True*, *random_state* = 42). This protocol yields a reasonably objective estimate of out-of-sample performance for small datasets by splitting the 28-sample dataset into five non-overlapping folds and iteratively swapping the hold-out set. Three complementary measures were calculated using the following equations:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}{\sum_{i=1}^{n}(z_i - \widehat{z_i})^2}, \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}, \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|z_i - \hat{z}_i\right|, \tag{6}$$

where $\boldsymbol{n}$ is the number of samples; z_i is the actual value; $\hat{y}_i$ denotes the predicted value.

$\boldsymbol{R^2}$ measures the percentage of variation in wear rate accounted for by the model. *RMSE* weighs large individual prediction errors more heavily than *MAE*, and therefore the two measures are complementary in describing both systematic and outlier-driven prediction errors. All metrics are reported as mean values ± standard deviations across the five folds.

After cross-validation, the most successful *ML* model (Gradient boosting) was retrained using the entire 28-sample dataset. *Gini*-based feature importance (i.e., the sum of weighted loss reduction contributed by each variable across all 200 boosted trees) was extracted to provide a global variable ranking. To supplement this global measure with observation-level attribution, SHapley Additive exPlanations (*SHAP*) were calculated using the TreeExplainer backend of the shap library [21]. *SHAP* values represent the average marginal contribution of each feature across all possible orderings of features, providing a theoretically founded, additive decomposition of individual predictions. A brute-force combinatorial search was then conducted over all 81 discrete parameter combinations (3 temperatures × 3 loads × 3 thicknesses × 3 coatings) using the trained Gradient boosting model to determine the global optimum operating condition that minimises the predicted wear rate.

Results and discussion

A sample of *S*/*N* ratio results is shown in Table 3. The average *S*/*N* ratios of *AlTiN*, *CrN*, and *TiC* coatings are 36.64 dB, 30.33 dB, and 26.74 dB, respectively, which clearly shows that *AlTiN* is the best coating in terms of wear resistance. The highest individual *S*/*N* ratio (38.49 dB) occurred with the *AlTiN* coating at 45 °C, 10 N, and 3 μm, which was also reproduced with the same result (runs **1** and **19**) in repeated experiments, confirming the reproducibility of the experiment. The analysis of dynamic range shows that *TiC* (10.3 dB) and *CrN* (10.0 dB) coatings are more sensitive to changes in process parameters than *AlTiN* (7.7 dB). In the case of *CrN*, a rise in temperature between 40 °C and 50 °C causes a decrease in the *S*/*N* ratio of approximately 4 dB, indicating an increase in the oxidative wear rate as temperature rises, as has been reported in the case of degradation of nitride-based coatings [22]. Conversely, *TiC* has the highest rate of *S*/*N* ratio change with an increase in load (≈ –3.5 dB/5 N), indicating the vulnerability of the material to brittle fracture at high contact loads [9]. In general, coating chemistry is the most important aspect of wear behaviour identified by the *Taguchi* analysis, which offers a useful initial ranking in line with well-known tribological literature [14].

The second-order *RSM* model is a good predictor of wear rate ($\boldsymbol{R^2}$ = 0.901, Adjusted $\boldsymbol{R^2}$ = 0.834), which means that the model captures about 90% of the variability in wear rate. The statistical reliability of the model is proven by the residual diagnostics: the *Omnibus* test ($\boldsymbol{p}$ = 0.088) and the *Durbin-Watson* statistic (2.52) indicate that the assumptions of normality and independence are met. The results of *ANOVA* (Table 4) show that the most significant parameter is contact load ($\boldsymbol{p} < 0.001$) , followed by coating chemistry ($\boldsymbol{p} < 0.05$). The dominance of load is physically explained, since it controls contact stress and subsurface deformation — the major wear mechanism driver [25]. The statistical significance of coating chemistry also demonstrates the essential importance of intrinsic material properties. Temperature and coating thickness

Table 3

Selected *Taguchi S/N* ratio values (smaller-is-better)

Run	Coating	Temp. (°C)	Load (N)	Thickness (µm)	Wear кate (mm^3/Nm)	S/N Ratio (dB)
1	*AlTiN*	45	10	3	0.0119	38.49
11	*AlTiN*	45	15	3	0.0131	37.65
23	*AlTiN*	40	10	3	0.0127	37.92
7	*CrN*	45	10	3	0.0204	33.81
9	*CrN*	40	15	4	0.0179	34.94
6	*TiC*	45	10	3	0.0289	30.78
10	*TiC*	50	15	4	0.0252	31.97
15	*CrN*	50	5	2	0.0578	24.76
3	*CrN*	40	5	4	0.0489	26.21
8	*TiC*	50	5	4	0.0674	23.43

Table 4

Type-II *ANOVA* table for the full second-order *RSM* model

Source	Sum of quares	***df***	***F***-value	***p***-value
T (Temperature)	0.000032	1	0.547	0.470
L (Load)	0.002428	1	41.576	< 0.001
K (Thickness)	0.000014	1	0.233	0.636
C_1(Coating *CrN*)	0.000412	1	7.056	0.017
C_2(Coating *TiC*)	0.000265	1	4.532	0.049
T^2	0.000006	1	0.106	0.749
L^2	0.000243	1	4.162	0.058
K^2	0.000035	1	0.599	0.450
$T\times L$	0.000050	1	0.861	0.367
$T\times K$	0.000041	1	0.694	0.417
$L\times K$	0.000014	1	0.242	0.629
Residual	0.000934	16	–	–

are observed to be not statistically significant within the studied range, which might be explained by the narrow temperature range (40–50 °C). This result is in line with the literature suggesting that wear in *PVD* coatings is generally thermally activated only at substantially higher temperatures [26].

The fitted second-order regression equation (in coded variables) is expressed as:

$$Y = 0.0146 + 0.0012T - 0.0162L - 0.0010K + 0.0123C_1 + 0.0146C_2 -$$

$$- 0.0013\cdot T\mathrm{I} + 0.0101L\mathrm{I} + 0.0031Th\mathrm{I} - 0.0024TL + 0.0032TK - 0.0013LK. \quad (7)$$

In the above equation, the negative coefficient of load (–0.0162) indicates that there appears to be a positive correlation between predicted wear rate and increasing load; however, this apparent anomaly requires careful interpretation. Although the experimental matrix was guided by a *Taguchi* L_{27} orthogonal array for the three continuous factors (temperature, load, thickness), the coating-type factor, being categorical, was not fully crossed with all level combinations. After quality screening reduced the dataset to 28 runs, the resulting design lost strict orthogonality: a disproportionate number of high-load conditions

(15 N) were associated with *AlTiN* samples, which inherently exhibit lower wear rates. This constitutes a genuine limitation of the experimental planning rather than a mere statistical artefact: the non-orthogonality in the design matrix introduces systematic correlation between the load factor and coating identity, causing the regression to attribute part of *AlTiN's* inherently low wear rate to increasing load. No post-hoc balancing measures (e.g., variance-inflation-factor screening, Type-III sum-of-squares decomposition, or covariate adjustment) were applied to mitigate this confounding in the present study. Future work should employ a balanced full-factorial or *D*-optimal design with equal replication across all coating–load combinations to eliminate such confounding and permit unambiguous estimation of main effects. Accordingly, this coefficient should be interpreted with caution and not as evidence of a direct physical relationship between higher loads and lower wear rates. The positive coefficients for *CrN* and *TiC* support the increased wear rate compared to *AlTiN*, which is in line with their relatively lower oxidation resistance and fracture toughness. The quadratic load term is nearly significant, implying a weak nonlinear relationship and therefore possible transition behaviour at high loads. Fig. 2 and Fig. 3 show the *RSM* response surface and contour plot, respectively, which clearly show the dominance of load and the comparatively insignificant effect of temperature. The actual vs. predicted plot (Fig. 4) shows good agreement along the 45° line, indicating model accuracy.

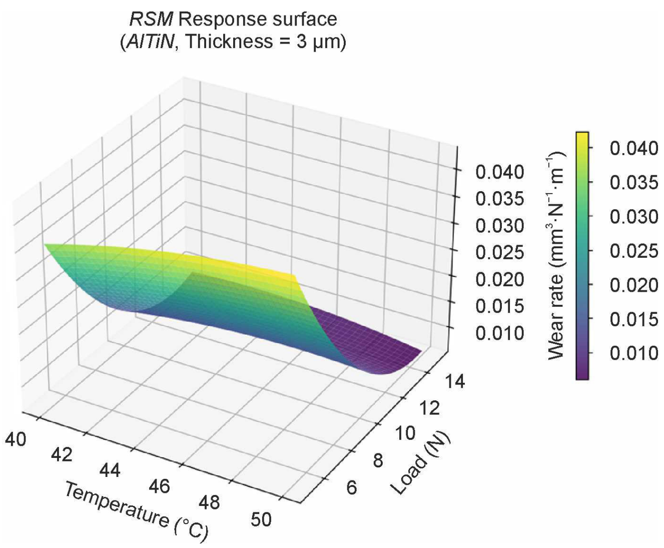

Fig. 2. RSM surface: wear rate vs load and temperature (*AlTiN*, 3 μm)

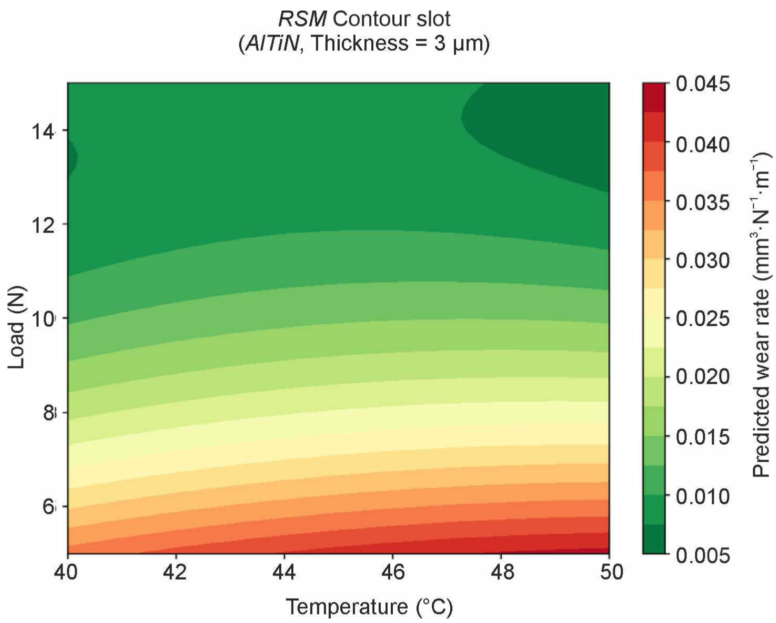

Fig. 3. RSM contour map

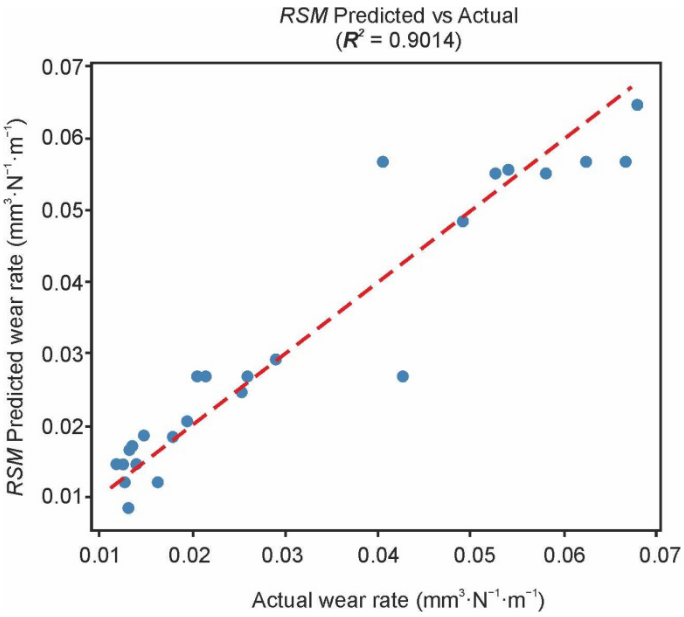

Fig. 4. RSM predicted versus actual wear rate plot

Table 5 shows the cross-validated performance of the machine learning models. Random forest had the best cross-validation accuracy ($\boldsymbol{R^2}$ = 0.7551), which means that it has strong generalisation. Gradient boosting had the best full-data performance ($\boldsymbol{R^2}$ = 0.9132), and thus it was the most suitable when optimization is required, whereas AdaBoost showed consistent intermediate performance. This confirms the applicability of ensemble-based solutions to data-constrained tribological applications, where robustness and variance control are paramount [25, 28]. Machine learning models demonstrate better generalisation capability in cross-validation compared to the *RSM* model, which offers good in-sample predictive capability, reflecting the complementary benefits of the hybrid modelling method.

Table 5

Cross-validated vs. full-data metrics (*ML* models)

Model	Mean *CV* $\boldsymbol{R^2}$	Full-data $\boldsymbol{R^2}$	Mean *RMSE*	Mean *MAE*
Gradient boosting	0.6467	0.9132	0.00984	0.00833
Random forest	0.7551	0.8776	0.00757	0.00641
AdaBoost	0.6972	0.9019	0.00865	0.00712

The Gradient boosting model residual analysis (Fig. 5) indicates that the residual distribution is symmetric around zero, with no evidence of heteroscedasticity or systematic bias, indicating that the model is adequate. The tendency of the residuals to cluster in the low wear-rate region is explained by the fact that the data represented by *AlTiN* are more numerous in that region. Similarly, the *RSM* residual diagnostics show that normality and independence assumptions are valid, confirming that both models are able to represent the underlying data structure.

Fig. 6 demonstrates that the contact load is the most significant factor governing the wear intensity, with the coating chemical composition being the next most important. This finding is in agreement with *Taguchi* and *ANOVA* results, demonstrating high consistency across methodologies. The fact that the significance of temperature and thickness is relatively lower also confirms their limited role within the parameter range examined. These results are consistent with other studies that have identified material properties and loading conditions as major factors in wear behaviour [16, 25].

SHAP analysis (Fig. 7) enables detailed analysis of feature contributions. *TiC* and *CrN* coatings exhibit positive *SHAP* values, indicating that they wear more, while *AlTiN* exhibits negative contributions, again confirming that it is the best. Contact load shows a positive but non-monotonic effect on wear, and thickness shows non-monotonic behaviour, suggesting the existence of an optimal thickness range. This is physically attributed to the balance between the substrate effect in thin coatings and residual stress effects

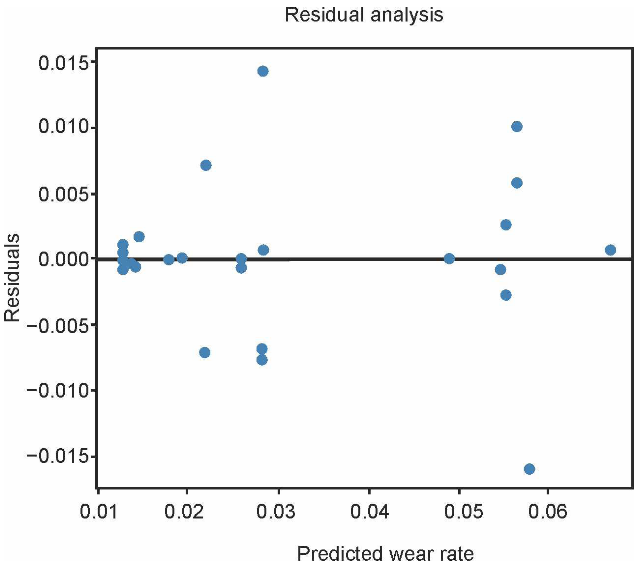

Fig. 5. Residual plot (Gradient boosting model)

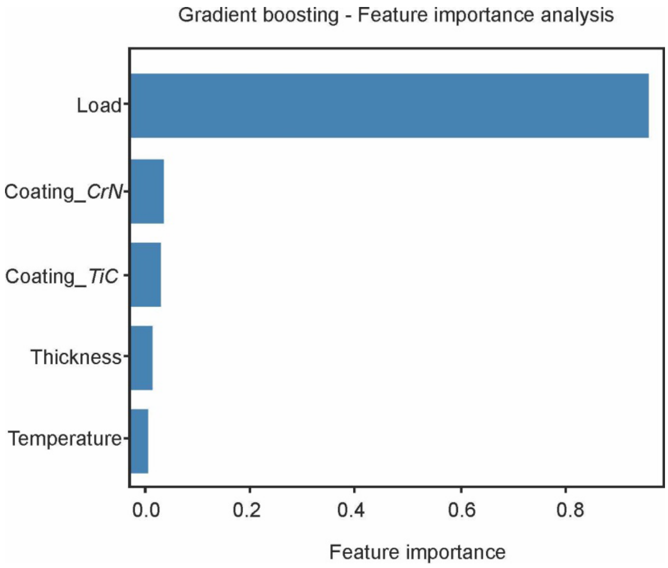

Fig. 6. Gini-based feature importance scores

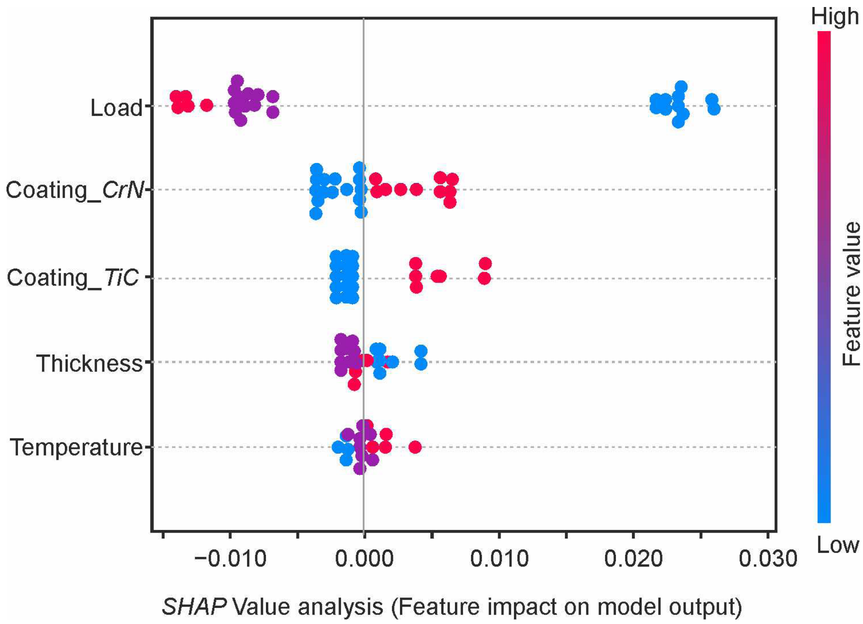

Fig. 7. SHAP summary plot

in thicker coatings [23]. The good consistency between the *SHAP* results, *RSM* coefficients, and *Taguchi* analysis provides strong cross-validation and increases confidence in the physical meaning of the findings.

The analysis of the learning curve (Fig. 8) shows that model performance increases with increasing dataset size, with no signs of saturation. The difference between training and validation performance gradually decreases, indicating that more data would further enhance generalisation. Extrapolation suggests that adding 30–50 observations could raise $CV\ \boldsymbol{R}^2$ to above 0.85, highlighting the importance of expanding the dataset to improve predictive accuracy.

The use of 200 trees with only 28 observations warrants careful consideration of overfitting risk. Several safeguards were employed:

– five-fold cross-validation provides an unbiased estimate of out-of-sample performance;

– the learning rate was set conservatively at 0.05, requiring many iterations for convergence and thus acting as implicit regularisation;

– maximum tree depths were limited (6 for Gradient boosting, 15 for Random forest) to prevent memorisation of individual samples.

A sensitivity analysis with respect to the number of trees was not performed; this is acknowledged as a limitation, and future work should systematically evaluate the effect of ensemble size on predictive stability for small tribological datasets. Simpler models such as *LASSO*, *Ridge*, and *Elastic Net* regression were not included in this study, as the primary focus was on ensemble tree-based methods; however, their inclusion in future comparative studies is recommended. No independent held-out test set was reserved; instead, the entire dataset was used for cross-validation, which is standard practice for small datasets (n < 50) but limits external generalisability assessment. The predicted optimal configuration (*AlTiN*, 4 µm, 50 °C, 15 N) was not experimentally verified; future work should include confirmation experiments to validate the model predictions.

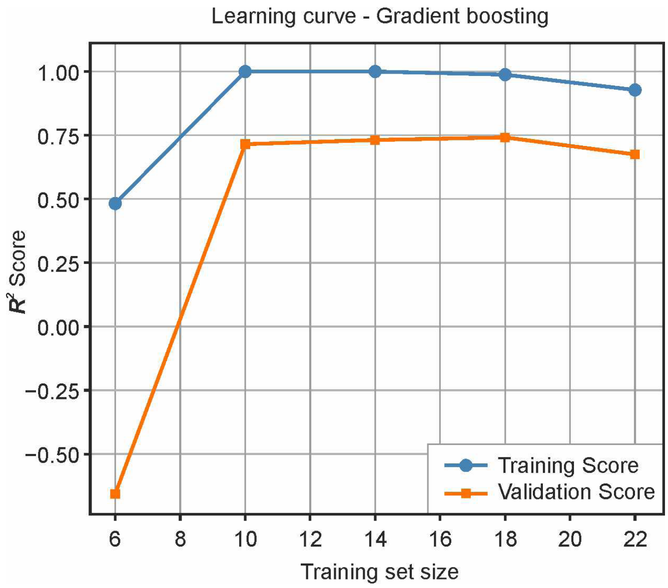

Fig. 8. Learning curves

The *ML*-based optimisation of the system identified *AlTiN* with a thickness of 4 µm, a temperature of 50 °C, and a load of 15 N as the optimum configuration, with a predicted wear rate of 0.010823 mm^3/Nm (Table 6). Even though this condition was not experimentally tested, it is the global optimum predicted within the design space. It is important to note that all three methods – *Taguchi*, *RSM*, and *ML* – are consistent in selecting *AlTiN* as the best coating, providing strong cross-method support. The narrow difference between *ML* and *Taguchi* results (~8.2%) is within the range of experimental uncertainty. The localisation of the optimal configurations around *AlTiN* may suggest that the response surface near the optimum is relatively flat, implying that the system is robust to moderate parameter variations. Industrially, this robustness means that *AlTiN* coatings can continue to perform reliably even when operating conditions are moderately altered.

Table 6

Cross-validated vs. full-data metrics (*ML* models)

Rank	Coating	Temp. (°C)	Load (N)	Thickness (µm)	Predicted wear rate (mm^3/Nm)
1 (Optimal)	*AlTiN*	50	15	4	0.010823
2	*AlTiN*	50	15	3	0.011105
3	*AlTiN*	40	15	4	0.012080

Conclusion

A three-level hybrid analysis model that combines *Taguchi* signal-to-noise (*S*/*N*) analysis, *RSM*, and *ML* regression was developed and applied to predict and optimise the wear behaviour of *PVD* hard coatings. The *Taguchi* analysis identified *AlTiN* as the best coating at 45 °C, 10 N, and 3 µm, with a high *S*/*N* ratio of 38.49 dB. The second-order *RSM* model exhibited high predictive strength with R^2 = 0.901 and was statistically significant (p < 0.001). *ANOVA* revealed that the dominant factors were contact load and coating chemistry, while temperature and thickness were not significant within the studied range. The developed regression equation is physically interpretable and computationally efficient for wear prediction and sensitivity analysis.

Gradient boosting was found to be the best predictive algorithm, while Random forest and AdaBoost demonstrated better generalisation in cross-validation. Both *Gini* importance and *SHAP* analysis consistently identified coating chemistry as the most significant parameter, followed by contact load. *SHAP* also revealed nonlinear, non-monotonic effects, particularly for coating thickness, providing deeper physical insight than traditional statistical models. The close agreement between *ML* predictions and *Taguchi* experimental results (deviation of ~8.2%) confirms the strength and validity of the hybrid framework. All three methodologies converged in identifying *AlTiN* as the best coating.

Machine learning-based optimisation predicted the global optimum at AlTiN (4 µm, 50 °C, 15 N) with a wear rate of 0.010823 $mm^3 \cdot N^{-1} \cdot m^{-1}$. The strong agreement between *ML* predictions and *Taguchi* experimental results (deviation ~8.2%) further demonstrates the robustness and validity of the hybrid framework.

References

1. Almomani O., Venkatesh B., Chaudhary S.P., Mishra A., Sujai S., Juneja S., Pradhan P., Venkatesan S.P., Bhowmik A., Tamene Y. Intelligent tool wear monitoring using XGBoost, SVR, and DNN models in NMQL environment. *Scientific Reports*, 2026, vol. 16 (1), p. 10030. DOI: 10.1038/s41598-026-40968-8.

2. Cherguy O., Chmielowski R., Hachem E., Lahouij I. Deep learning prediction of dry friction in DLC coatings using literature-derived data. *Tribology Letters*, 2025, vol. 73 (4), p. 125. DOI: 10.1007/s11249-025-02056-2.

3. Daghbouch A., Louhichi B., Terres M.A. Optimization and prediction of mass loss during adhesive wear of nitrided AISI 4140 steel parts. *Crystals*, 2025, vol. 15 (10), p. 875. DOI: 10.3390/cryst15100875.

4. Bergmann B., Denkena B., Junge N., Kalscheuer C., Bobzin K., Liu X. Influence of the interlayer and the substrate on the wear behavior of PVD tool coatings during turning of C60+ N. *Wear*, 2026, vol. 584–585, p. 206396. DOI: 10.1016/j.wear.2025.206396.

5. Jarrah M.I., Jaya A.S.M., Azam M.A., Muhamad M.R., Akbar H. Minimizing thin film thickness in TiN coatings using genetic algorithms. *AIP Conference Proceedings*, 2018, vol. 2016 (1), p. 020061. DOI: 10.1063/1.5055463.

6. Ambhore N., Naranje V., Shelke S. Machining performance evaluation in turning of hardened steel using machine learning. *Materials and Manufacturing Processes*, 2025, vol. 40 (14), pp. 1935–1942. DOI: 10.1080/10426914.2025.2559622.

7. Jaya A.S.M., Hashim S.Z.M., Haron H., Muhamad M.R., Rahman M.N.A. Fuzzy rule-based model to estimate surface roughness and wear in hard coatings. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, 2012, pp. 1076–1081. DOI: 10.1109/ICSMC.2012.6377873.

8. Jiang J., Omar I., Khan M. Predicting wear damage in moving mechanical contacts: Comparative analysis of regression algorithms and feature selection techniques. *ISA Transactions*, 2025, vol. 167, pp. 675–687. DOI: 10.1016/j.isatra.2025.08.027.

9. Kahraman F., Sugözü B. An integrated approach based on the Taguchi method and response surface methodology to optimize parameter design of asbestos-free brake pad material. *Turkish Journal of Engineering*, 2019, vol. 3 (3), pp. 127–132. DOI: 10.31127/TUJE.479458.

10. Kang J., Niu Y., Zhou Y., Fan Y., Ma G. Wear resistance prediction of AlCoCrFeNi-X (Ti, Cu) high-entropy alloy coatings based on machine learning. *Metals*, 2023, vol. 13 (5), p. 939. DOI: 10.3390/met13050939.

11. Karthikeyan S., Prasanna R., Balasubramanian V., Perumal G., Mugendiran V. Machine learning predictive modeling of plasma-sprayed lanthanum zirconate coatings. *Surface Engineering*, 2025, vol. 41 (10–12), pp. 1045–1054. DOI: 10.1177/02670844251389512.

12. Liu Y., Ravichandran R., Chen K., Patnaik P. Application of machine learning to solid particle erosion of APS-TBC and EB-PVD TBC at elevated temperatures. *Coatings*, 2021, vol. 11 (7), p. 845. DOI: 10.3390/coatings11070845.

13. Tzionis G., Mouratidis P., Kougka G., Gialampoukidis I., Vrochidis S., Kompatsiaris I., Vlachopoulou M. A review of explainable AI methods and their application in manufacturing systems. *Discover Applied Sciences*, 2025, vol. 8, p. 52. DOI: 10.1007/s42452-025-07908-z.

14. Kumaraswamy J., Thirumalesh, Ashok A.S., Shankar N.B., Praveen S.R. Machine learning and Taguchi techniques for predicting wear mechanisms of Ni–Cu alloy composites. *Results in Surfaces and Interfaces*, 2024, vol. 17. DOI: 10.1016/j.rsurfi.2024.100307.

15. Mallick A., Sahu R.K., Gangi Setti S. Machine learning enabled hardness prediction in centrifugal cast functionally graded materials for superior wear performance. *Journal of Tribology*, 2025, vol. 148 (3), p. 031701. DOI: 10.1115/1.4069854.

16. Pasha M.B., Rao R.N., Ismail S., Gupta M., Prasad P.S. Tribo-informatics approach to predict wear and friction coefficient of Mg/Si_3N_4 composites using machine learning techniques. *Tribology International*, 2024, vol. 196, p. 109696. DOI: 10.1016/j.triboint.2024.109696.

17. Patel K. Material lifespan predictor. *International Journal of Research in Science and Technology*, 2025, vol. 15 (4), pp. 31–41. DOI: 10.37648/ijrst.v15i04.004.

18. Pawar S.Y., Haider J., Pintaude G., Sekar S.M., Kolhe V., Chandratre K., Sonawane S., Ritapure P. Mechanical and tribological behavior of LM26/SiC/Ni-Gr hybrid composites. *Journal of Composites Science*, 2023, vol. 7 (4), p. 159. DOI: 10.3390/jcs7040159.

19. Radhika N., Sabarinathan M., Sivaraman S. A comparative analysis of machine learning techniques for predicting the wear rate of ceramic coated steel. *IEEE Access*, 2024, vol. 12, pp. 146949–146967. DOI: 10.1109/access.2024.3473028.

20. Yan F., Song K., Liu Y., Chen S., Chen J. Predictions and mechanism analyses of the fatigue strength of steel based on machine learning. *Journal of Materials Science*, 2020, vol. 55 (31), pp. 15334–15349. DOI: 10.1007/s10853-020-05091-7.

21. Zhong X., Gallagher B., Liu S., Kailkhura B., Hiszpanski A., Han T.Y.-J. Explainable machine learning in materials science. *Npj. Computational Materials*, 2022, vol. 8. DOI: 10.1038/s41524-022-00884-7.

22. Sellami A., Rekik M., Njemma C.B., Elleuch R. Towards an explainable multi-target regression for wear and friction prediction for brake pad materials. *Acta Polytechnica Hungarica*, 2024, vol. 21 (11), pp. 155–174. DOI: 10.12700/aph.21.11.2024.11.9.

23. Ma J., Luo D., Liao X., Zhang Z., Huang Y., Lu J. Tool wear mechanism and prediction in milling TC18 titanium alloy using deep learning. *Measurement*, 2021, vol. 173, p. 108554. DOI: 10.1016/j.measurement.2020.108554.

24. Sen B., Saha S., Kumar R., Jhala R., Patil N., Mahapatro A., Bhowmik A., Santhosh A. AI-driven wear monitoring of PVD TiAlN coated carbide insert in sustainable machining of Hastelloy C276: An industry 4.0 perspective. *Results in Engineering*, 2025, vol. 25, p. 104457. DOI: 10.1016/j.rineng.2025.104457.

25. Shah R., Pai N., Thomas G., Jha S., Mittal V., Shirvni K., Liang H. Machine learning in wear prediction. *Journal of Tribology*, 2024, vol. 147 (4), p. 040801. DOI: 10.1115/1.4066865.

26. Sun J., Zhang Z., Zhang Y., Zhang X., Guo J., Fu Q., Wu L. High-temperature ablation resistance prediction of ceramic coatings using machine learning. *Journal of the American Ceramic Society*, 2024, vol. 108 (1). DOI: 10.1111/jace.20136.

27. Trost C.O.W., Zak S., Schaffer S., Saringer C., Exl L., Cordill M.J. Bridging fidelities to predict nanoindentation tip radii using interpretable deep learning models. *JOM*, 2022, vol. 74 (6), pp. 2195–2205. DOI: 10.1007/s11837-022-05233-z.

28. Fan J., Wang Z., Liu C., Shi D., Yang X. A tensile properties-related fatigue strength predicted machine learning framework for alloys used in aerospace. *Engineering Fracture Mechanics*, 2024, vol. 301. DOI: 10.1016/j.engfracmech.2024.110057.

Conflicts of Interest

The authors declare no conflict of interest.