JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Explainable Machine Learning Framework for Multi-Parameter Slurry Erosion Wear Prediction of Stainless Steel

[1]Gaurav Kumar, [2]Rajeev Kumar, [3]Gosiya Kaleem
[1, 2,3]Assistant Professor
[1, 2]Mechanical Engineering Department,[3]Computer Science and Engineering
[1,2,3] Dewan VS Institute of Engineering & Technology in Meerut, India

Abstract : **The slurry erosion is simultaneous with particle velocity, solid concentration and particle size and impingement angle, thus the single-parameter approach is not suitable for multi-variable engineering conditions. This study proposes an explainable machine-learning (XAI) approach to the prediction of slurry-erosion wear of SS304 stainless steel based on a nonlinear proof-of-concept dataset (N = 240) that was generated numerically instead of measured in the laboratory. R^2, RMSE, and MAE were used to compare the performance of five algorithms: Linear Regression, Decision Tree, Random Forest, Support Vector Regression, and Artificial Neural Network. The ANN outperformed SVR in 5-fold cross-validation (mean R^2 = 0.9714, RMSE = 9.68, MAE = 7.57) by a little margin. The dominant feature (41.61%) was particle velocity, followed by impingement angle (26.79%), particle size (19.71%) and slurry concentration (11.89%) as shown in the SHAP analysis of the ANN. The proposed methodology involves a parameter ranking with a quantitative measure and an interpretation of the framework, and is coupled with good predictive accuracy; it is suggested to serve as a methodology to be validated on future experimental data for SS304 slurries. Slurry erosion is a common phenomenon that occurs in the cutting processes of stainless steel.Slurry erosion is one of the phenomena that occur easily in cutting processes of stainless steel.**

Key Words: Slurry erosion; Stainless steel; Machine learning; Artificial neural network; Explainable AI; SHAP

I. INTRODUCTION

Slurry erosion is a material degradation process that occurs when solid particles are repeatedly abrading a solid subject by moving in a liquid medium, and it plays an important role in slurry pipeline, pump, valve, hydraulic machinery, mining machinery and power plant **[9,10]**. Repeated impacts produce micro-cutting, plowing, indentation and plastic deformation. Slurry-handling components are extensively used in stainless steel, especially in the case of SS304, because of its mechanical properties, corrosion resistance, manufacturability, and erosion resistance, which is strongly influenced by particle velocity, solid concentration, particle size and shape, and impingement angle, all of which interact in a nonlinear fashion. Kumar et al. **[1]** classified the erosion behaviour of SS304 with the help of slurry-pot rig, while More et al. **[2]** Adopted a Taguchi design and discovered that particle velocity is the most influential parameter, with a correlation coefficient of 0.987. The importance of using ML for multi-parameter erosion modelling has increased, as Singh et al. **[3]** predicted the erosion of Ni-WC-coated SS420 by using an ANN, and a recent work on 2205 duplex stainless steel coupled Taguchi analysis with ANN **[4].** However, conventional ML models are black boxes: while they may be highly accurate, it does not explain which parameters are the most important for erosion or how these parameters should be manipulated. Explainable AI (XAI) provides solutions for this, with SHAP **[5,6]** being one of the most promising techniques, and has been found useful in wear-prediction problems **[7,8].** The gap addressed here is not the previously reported use of ANN for erosion prediction, but rather the use of multi-parameter nonlinear dataset generation, comparative

benchmarking of five ML algorithms, and explainability using the SHAP method in a single and reproducible prediction-explanation framework for SS304 slurry erosion. An industrial application for such a predictive model could be in maintenance planning and material selection; the model could be used to predict the erosion rate for a proposed change in flow velocity, slurry loading, or pipe geometry, and it could also help a plant engineer understand, quantitatively, which operating variable is responsible for that change — something a black-box regression model could not do. This study aims to create a nonlinear multi-parameter computational dataset, compare five supervised learning algorithms based on R^2, RMSE, and MAE for both a held-out test split and 5-fold cross validation, select the most successful model, rank the four erosion parameters by global contribution using SHAP, and develop a computational framework for future experimental validation.

2. METHODOLOGY

2.1 A dataset of 240 observations was generated including four experimentally manageable erosion parameters **(Table 1)** based on the nonlinear response function $Ew = f(V,C,D,\theta) + \varepsilon$, where ε corresponded to controlled stochastic variation and a fixed random seed in R processing for reproducibility. The intent was not to replicate a previously published data set or to generate an artificial data set but to push forward with a physically plausible nonlinear regression benchmark for the framework being proposed.

Table 1: Ranges of computational parameters.

Parameter	Symbol	Range
Particle velocity	V	3.5–6.0 m/s
Slurry concentration	C	10–40 wt.%
Particle size	D	250–700 µm
Impingement angle	θ	15–90°
Observations	N	240

3. Data Partitioning and Modeling

A total of 192 samples (80%) were used for training, and 48 (20%) for independent testing, with an additional 5-fold cross-validation. A total of five regression algorithms were benchmarked (Table 2): linear regression (baseline), decision tree, random forest, support vector regression (RBF kernel), and an artificial neural network (ANN).

Table 2: Machine Learning Techniques Evaluations

Model	Key Characteristic
Linear Reg.	Baseline linear model
Decision Tree	Recursive space partitioning
Random Forest	Ensemble of trees, averaged output
SVR (RBF)	Kernel-based nonlinear regression
ANN (MLP)	2 hidden layers (20, 10 neurons)

Table 2b. ANN Design In Summary

Hyperparameter	Value
Hidden layers	2 (20, 10 neurons)
Input features	V, C, D, θ (standardized)
Training split	192 / 240 (80%)
Test split	48 / 240 (20%)
Validation scheme	5-fold cross-validation

2.3 Performance Evaluation Metrics

Prediction accuracy was measured using three statistical metrics: the coefficient of determination,

$$R^2 = 1 - [\Sigma(yi-\hat{y}i)^2/\Sigma(yi-\grave{y})^2]$$

which was used to quantify the degree of agreement between the actual and predicted response, with values closer to one corresponding to greater agreement; the root mean square error

$$RMSE = \sqrt{[(1/n)\Sigma(yi-\hat{y}i)^2]}$$

which was sensitive to whether the error was large or small; and the mean absolute error

$$MAE = (1/n)\Sigma|yi-\hat{y}i|$$

representing the average accuracy of prediction in the original response units: mg/h. Lower values of RMSE and MAE and an R^2 near unity are preferable for a better regression model.

2.4 SHAP-Based Explainability

The selected model by 5-fold cross-validation was interpreted with the framework of Shapely Additive explanations (SHAP), which breaks down the difference between the prediction and the baseline prediction into contributions to the model's output by the individual features, based on cooperative game-theoretic (Shapley) principles.

$$Ij = [(1/n)\Sigma|\varphi ij|| / \Sigma j[(1/n)\Sigma|\varphi ij|| \times 100$$

where $\Sigma|\varphi ij|$ is the absolute sum of local SHAP contribution from feature j for a given observation i, was used to compute the global importance of feature j, which is an interpretable, percentage-normalized ranking of the four erosion parameters across the entire dataset.

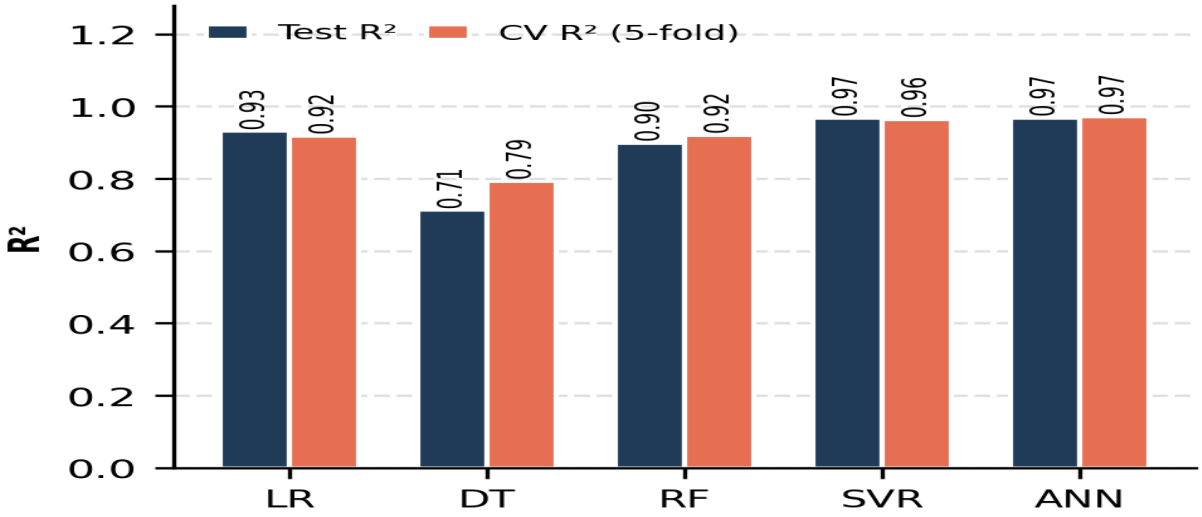
3. RESULTS AND DISCUSSION

3.1 Predictive Performance

The results of all five models are presented in Table 3, both for the test set and for five-fold cross-validation. The results of the test set of the models showed that the accuracy (R^2) of the models was very similar for SVR and ANN with $R^2 = 0.9680$, which is significantly higher than that obtained from Decision Tree with $R^2 = 0.7121$. Under 5-fold cross-validation — a more robust estimator — the ANN achieved the best overall performance (mean $R^2 = 0.9714$, RMSE = 9.68, MAE = 7.57, $\sigma(R^2) = 0.0077$), marginally ahead of SVR ($R^2 = 0.9639$). SHAP was therefore chosen for interpretation of the ANN.

Table 3. Test-Set with Five-fold CV Performance

Model	Test R^2	CV R^2	CV RMSE	CV MAE
Linear Reg.	0.930	0.917	16.85	12.62
Decision Tree	0.712	0.792	26.11	20.49
Random Forest	0.898	0.920	16.39	12.50
SVR	0.968	0.964	11.24	8.77
ANN	**0.968**	**0.971**	**9.68**	**7.57**



smooth transition params between cutting plough at low levels and indentation-dominated deformation at high levels avoidance Concentration — within the range of our sampling — had the modest global contribution, but this result was dependent upon computational choices and should not be construed as a general statement regarding its physical importance, even while particle size also made a moderate effect via mass and contact area.

Table 4. The importance of the global SHAP feature

Rank	Parameter	SHAP %
1	**Particle velocity**	**41.61**
2	Impingement angle	26.79
3	Particle size	19.71
4	Slurry concentration	11.89

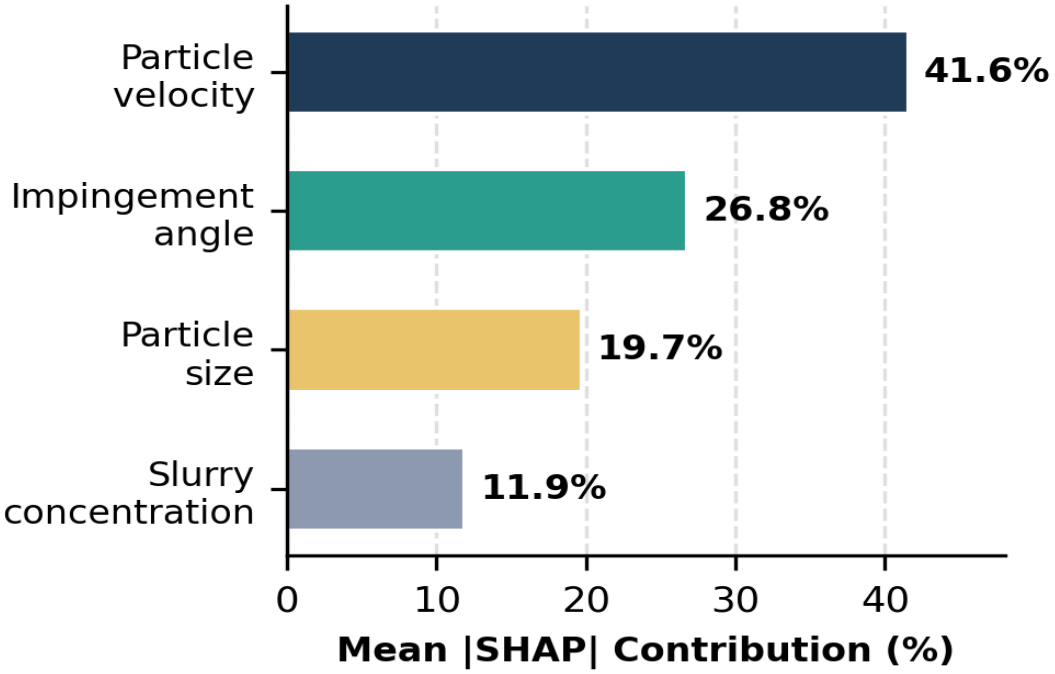


Fig. 2. SHAP's global feature priority ranking (ANN).

3.3 Prediction–Explanation Framework

The framework is comprised of two stages: **1)** ANN is applied to get EW, which: (V,C,D,θ)→ Ew; **2)** the second stage decomposes EW into SHAP contributions per instance per feature (Fig. 3). An estimate of the wear can be supplied to an engineer, as well as the operating parameter to be targeted for reduction in order to control erosion (particle velocity), which allows the engineer to be able to custom optimize the process, instead of adjusting parameters in general. Being a ranked prediction from some individual numerical output allows this two stage structure to identify an actionable ranking of diagnosis that would inform design or maintenance (unlike traditional regression).

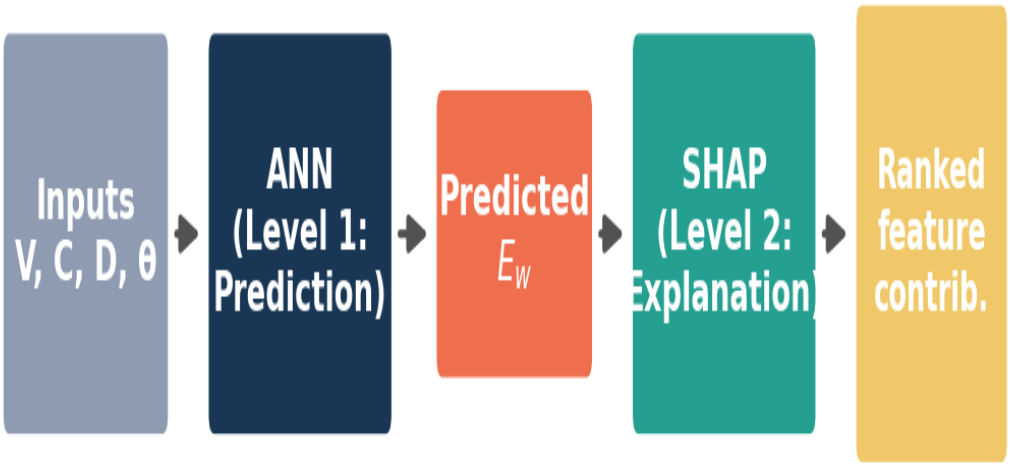


Fig. 3. A two-level prediction-explanation design.

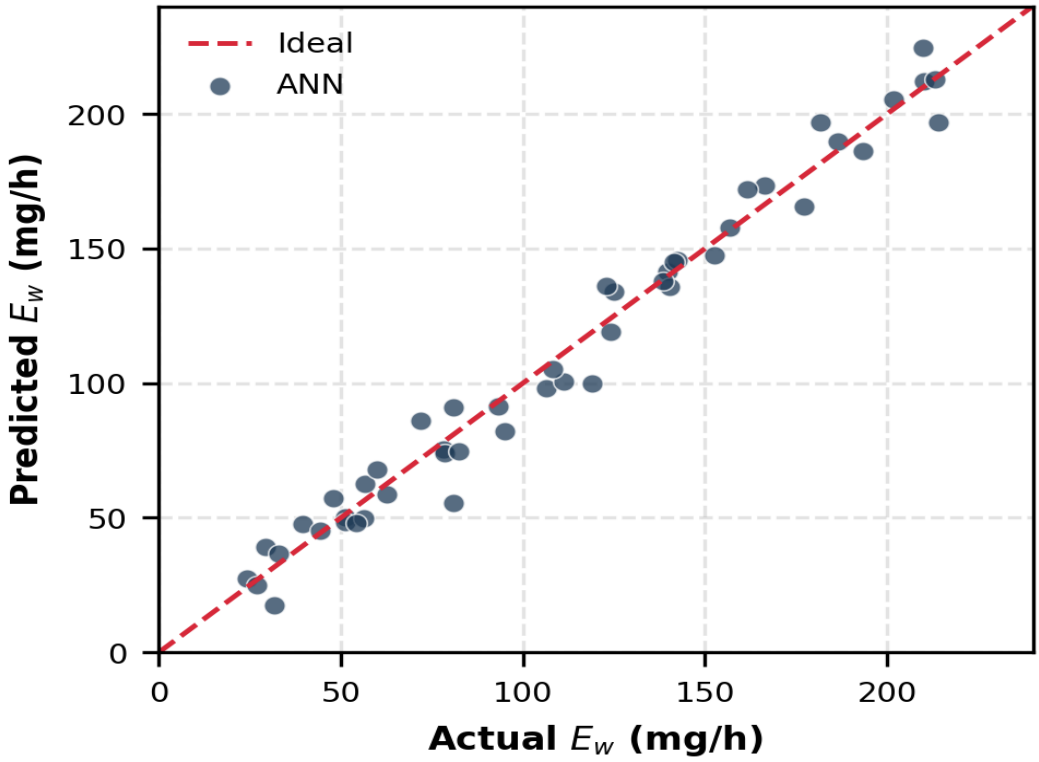


Fig. 4. ANN anticipated compared real erosion wear (test set, for illustration).

Annealing based model It should be noted, however, that given the combination of varying independent variables used in this study, the nonlinear relationship between velocity-concentration-size-angle exceeds even what simple axis-aligned partitioning can explain: ANN and SVR invariably outperform the limited Decision Tree from an engineering perspective. The low condition deviation of the ANN cross validation R^2 (0.0077) resulting from ANN training indicated selection based on the strong generalization capability within the parameter range studied for ANN output.

4. LIMITATIONS AND CONCLUSION

The primary limitation arises from the synthetically generated data used for computational reasons instead of experimental measurements, so the R^2, RMSE, MAE and SHAP percentages provided here are a proof-of-concept-derived result and should hopefully be representative of the framework methodology rather than experimentally- measured erosion behaviour as with SS304. More work should be done using the slurry-pot measurements (velocity, concentration, size, angle, exposure duration, mass loss, roughness hardness), instead of the synthetic data for comparison purposes with ensemble/ground-truth models and quantification of prediction-interval uncertainty and allowing for physics-informed models. Here, we developed a reproducible explainable-ML framework for slurry erosion prediction. The best, 5-fold cross validation performance was shown formulate: ANN (R^2 =0.9714, RMSE=9.68 and MAE=7.57) followed closely by SVR (R^2 =0.9639). According to the SHAP analysis, the particle velocity, impingement angle, particle size and slurry concentration were ranked as the four most significant parameters influencing wear prediction of erosion. The framework can also be employed to verify against experimental slurry-pot data of SS304, which could be easily expanded by including parameter sensitivity analysis, material selection and optimization of operation conditions for slurry-handling equipment.

REFERENCES

[1] G. Kumar, S. R. More, D. V. Bhatt, J. V. Menghani, SKIT Res. J. 7(2) (2017) 58–63.

[2] S. R. More, D. V. Bhatt, J. V. Menghani, Tribol. Ind. 39(4) (2017) 471–481.

[3] S. K. Singh et al., Metals 12(5) (2022) 706.

[4] Mater. Today Commun. 40 (2024) 109918.

[5] S. M. Lundberg, S.-I. Lee, Adv. Neural Inf. Process. Syst. (2017).

[6] S. M. Lundberg et al., Nat. Mach. Intell. 2 (2020) 56–67.

[7] S. V. Sotubadi, R. Liu, V. Nguyen, (2023).

[8] A. B. Arrieta et al., Inf. Fusion 58 (2020) 82–115.

[9] V. Javaheri, D. Porter, V.-T. Kuokkala, Wear 408–409 (2018) 248–273.

[10] G. R. Desale, B. K. Gandhi, S. C. Jain, Wear 259 (2005) 196–202.

[11] H. S. Grewal , Anupam Agrawal , H. Singh: Slurry Erosion Mechanism of Hydroturbine Steel: Effect of Operating Parameters

[12] Soderberg, S., Hogmark, S., Engmant, U., and Swahnt, H.(1981): Erosion classification of materials using a centrifugal erosion tester

[13] Bahadur, S., and Badruddin, R.(1990), "Erodent particle characterization and the effect of particle size and shape on erosion"

[14] Tilly, G. P.(1969): Sand erosion of metals and plastics: Abrief review

[15] Finnie, I.(1972): Some observations on the erosion of ductile metals

[16] Sunil Chandel, S. N. Singh, V. Seshadri, (2012): "Experimental study of erosion wear in a centrifugal slurry pump using coriolis wear test rig"

[17] Shobha Lata Sinha, Satish Kumar Dewangan, Anurag Sharma, (2015) "A Review on Particulate Slurry Erosive Wear of Industrial Materials: In Context with Pipeline Transportation of Mineral – Slurry"