Materials Today Communications 50 (2026) 114562

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

Application of machine learning in predicting corrosion inhibition capacity of *Spinacia oleracea* leaf extract on copper

Omotayo Sanni [a,*], Oluwatobi Adeleke [b], Samuel Ayodele Iwarere [a], Tien-Chen Jen [b], Michael Olawale Daramola [a,*]

[a] Sustainable Energy & Environment Research Group (SEERG), Department of Chemical Engineering, Faculty of Engineering, Built Environment & Information Technology, University of Pretoria, Hatfield, Pretoria 0028, South Africa

[b] Department of Mechanical Engineering Science, University of Johannesburg, Cnr Kingsway and University Roads, Auckland Park, Johannesburg 2092, South Africa

ARTICLE INFO

Keywords:
Machine learning
Corrosion inhibition
Plant extract
Copper
Predictive analysis

ABSTRACT

In many different industries, material corrosion is a major problem and it has a big financial impact. In this case, plant extracts employed as corrosion inhibitors, provide affordable way to prevent copper from corroding in acidic media, providing an alternative to the dangerous chemicals now in use. This study presents an integrated experimental and machine learning approach for investigating the corrosion inhibition performance of *Spinacia oleracea* leaf extract on copper in nitric acid medium. Experimental procedure involving gravimetric analysis under different concentrations, temperatures, and exposure durations, inhibitory efficiency was conducted. Different machine learning (ML) models, namely Artificial Neural Network (ANN), Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (tree) were developed for predicting the corrosion rate. To overcome the black-box limitation of the ML models, an interpretable feature analysis was carried using Shapley Additive ExPlanations (SHAP).The accuracy and validity of the models were evaluated using statistical tests like Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Deviation (MAD), Mean Absolute Percentage Error (MAPE), and Variance Accounted For (VAF).The best prediction accuracy was obtained with SVM, giving an averaged validation-based RMSE, MAE, MAD, MAPE, and VAF values of 0.386, 0.287, 0.192,0.08773, and 99.093, respectively, across 5 folds. SHAP interpretability identified inhibitor concentration as the most influential variable controlling corrosion inhibition. The data-driven framework that combines experimental gravimetric analysis with SHAP-enhanced ML in this study contributes to the broader development of transparent, eco-friendly, and data-driven corrosion prediction models.

1. Introduction

The field of corrosion science seeks to avoid or at least regulate material deterioration with the ultimate goal of preserving infrastructure safety and so contributing significantly to the global economy. Usually, a redox electrochemical process with an anodic component that oxidizes the metal and a cathodic component that reduces oxygen and water molecules from the surrounding environment is what causes metallic corrosion [1–4]. Cl^-, SO_4^{2-}, and OH^- are examples of aggressive species that can speed up these chemical reactions and, consequently, surface deterioration. Metal protection against corrosion in aqueous HNO_3 solutions has been an important factor in the field of metal protection, because HNO_3 is widely utilized in removing the oxide layers and deposits from metallic surface [1]. Therefore, metallic materials can readily corrode and are susceptible to corrosion in basic, acidic, or neutral media. Researchers are developing a variety of corrosion control techniques to eliminate or reduce this detrimental effect. Coatings (metallic, inorganic and organic), changes in metal potential (cathodic and anodic protection), environmental changes (corrosion inhibitor use, etc.), metal nature (positional changes, surface compressive stress and tensile stress), and design advancements are some of the corrosion control techniques that are available [5–8]. Among these methods, one of the most practical, cheap and popular technique for reducing metallic corrosion is the utilization of corrosion inhibitors [9,10]. The great attraction of the hydrophobic interactions between their tails and polar groups for the metal surface are known to be the main factors governing the corrosion inhibitor moleculeadsorption at metal-water interface [11–13]. Researchers are very interested in green corrosion inhibitors

* Corresponding authors.
E-mail addresses: tayo.sanni@yahoo.com (O. Sanni), michael.daramola@up.ac.za (M.O. Daramola).

https://doi.org/10.1016/j.mtcomm.2025.114562
Received 18 September 2025; Received in revised form 3 November 2025; Accepted 22 December 2025
Available online 23 December 2025

because of theirsustainablenature, easy biodegradability, non-toxicity, low cost, simple preparation methods, and natural production, even though the majority of reported chemical inhibitors showed good protection and high inhibition efficiency when taking into account their effects on the environment [14–16]. Furthermore, there are no dangerous chemicals or heavy metals in green corrosion inhibitors. In order to reduce corrosion, it is critical to determine the ideal dosage and frequency of corrosion inhibitor dosages.

Hence, extensive experiment is conducted wherein the metallic corrosion rates are measured in various operational and environmental conditionsviadiverse dose-schedules and dosages of inhibitor. This experiment is costly andtime-consuming but without reliable theoretical model, they are nevertheless necessary. Computer guided research has been giving valuable insights into the role of potentially corrosion inhibitive molecules and corrosion mechanisms. While mechanistic corrosion prediction models without inhibitors have been implemented and developed for commercial purpose [17–24], little success in incorporating corrosion inhibitoreffect in this model is reported. Firstly, diverse inhibitor molecule is observed to perform optimally under diverse condition for unknownreasons. Often, the corrosion inhibitors composition formulation is not disclosed to the researcher/operator as it is proprietary knowledge belonging to the manufacturer, making ittricky to model their behavior. Secondly, one complexity is that the exact corrosion inhibitormode of action isnot clear. Thirdly, the inhibitors efficiency depends on large numbers of operational and environmental factor, for instance flow rate, water chemistry,temperature, etc [25,26].

According to the literature, extract of plant, flower, seed, fruit, and leaves contain active compound that are promising for corrosion inhibition in aggressive solutions. Moreover, these compounds are widely available, cheap, and renewable corrosion inhibitor alternatives [27, 28]. Various plant extracts have been tested experimentally for the corrosion resistance of metal. The corrosion inhibition properties of *Equisetum arvense* extract on copper substrate in seawater medium was studied by Esquivel-Lopez et al. [29]. The authors reported maximum inhibition efficiency of about 87.5 %. Similarly, the *Commiphoramyrrha* extract, was tested by [30] on copper in a 2 M HNO_3 media. The experimental results showed that the inhibition efficiency increased as the inhibitor concentration increased, up to 91.8 % at 300 ppm concentration. Also, green *Eucalyptus* leaf extract as corrosion inhibitor, on mild steel in 1 M HCl corrosive medium, was evaluated by [31]. The maximum corrosion inhibition achieved with this inhibitor is 88 % at 800 ppm concentration. The *Eucalyptus globulus* leaves extract tested by Haldhar and Prasad [32] on low carbon steel immersed in a 0.5 M H_2SO_4 solution reached a maximum corrosion inhibition value of 93.09 % at 600 mg/L concentration. About 95.8 % was obtained with a 200 mg/L concentration of *Ficus tikoua* leaf extract [33] on carbon steel in 1 M HCl solution. But experimental testing is resource-intensive, expensive, and time-consuming [34–36].To address this gap, recently, materials informatics has become popular because of the rapid development of technology and data science through machine learning (ML) [37]. Overall, this work's novelty is developing a ML-based predictive model for predicting corrosion inhibition capacity of *Spinacia oleracea* leaf extractas corrosion inhibitor for copper in acidic media.

ML has been used to solve several corrosion-related problems over the past ten years, including modelling CO_2 corrosion [38,39], detecting corrosion through automated image analysis [40], modelling the growth of corrosion defects in pipelines [41], material inspection [42], modelling the rate of corrosion in marine environments [43], determining the time at which corrosion begins for embedded steel in reinforced concretes [44,45], electrochemical impedance spectra prediction [46], modelling pipeline ageing [47,48], and the pitting corrosion spatial distribution characterization [49]. Wurgeret al. [50] estimated the inhibition efficiency of untested molecules for magnesium alloy corrosion using a multidisciplinary approach that included density functional theory calculations, data mining, unsupervised machine learning based on clustering techniques, and inhibitor efficiencies high-throughput experimental screening. The suggested workflow demonstrated the existence of a distinct relationship between the inhibitors molecular structure acting as magnesium corrosion inhibitor and the corrosion inhibition efficiency, despite the comparatively small amount of experimental input data available for benchmarking their technique.

Recent research has shown that bio- and swarm-optimized ANN models can improve prediction accuracy and feature sensitivity in complicated engineering systems [51].While several studies have adopted ML techniques in their black box nature in corrosion studies, a significant gap exists in the interpretability of the corrosion operational parameters governing the inhibition efficiency. This study fills this gap by developing an explainable AI (XAI) framework with SHapley Additive exPlanations (SHAP)-based interpretability analysis of corrosion predictors. In addition, most prior works have focused on ferrous metals, electrochemical datasets, or synthetic inhibitors, with limited interpretability of model outcomes. The present study integrates an experimental–computational workflow that combines gravimetric corrosion data of Spinacia oleracea (SLE) extract on copper with multi-model ML prediction and SHAP-based feature interpretability. The SHAP-augmented ML framework for a green inhibitor on a non-ferrous metal in nitric acid medium takes this study a step further from what has been previously done in corrosion research. This research aims at developing an experimental and interpretable machine learning framework forpredicting the corrosion inhibition efficiency of *Spinacia oleracea* (spinach) leaf extract on copper in nitric acid medium. This aim is achieved through the following objectives (i) experimental investigation of the corrosion behavior of copper at varying concentrations, temperatures, and exposure durations using *Spinacia oleracea* extract as a green inhibitor (ii)comparing the performance of different ML models namely developing Artificial Neural Network (ANN), Random Forest (RF), Support Vector Machine (SVM) and Decision Tree assessed and validated using a k-foldscross validated for predicting the corrosion rate (iii) apply SHAP-based interpretability analysis to identify and quantify the influence of key operational variables on model prediction and corrosion behavior.This study supports corrosion informatics and sustainable material protection strategies through an interpretable, data-driven methodology that incorporates an eco-friendly inhibitor.

2. Materials and method

2.1. Experimental data description

A set of tests on the inhibition of copper against corrosion in nitric acid solutions at various concentrations of*Spinacia oleracea* extract provided the experimental data that was employed. Gravimetric tests were used in the trials to estimate the weight loss, corrosion rate, and the efficiency of the inhibitor which varies with time. 99.9 % pure copper rods with a diameter of 5.8 mm were used to make the copper electrode. The electrode was ground with 600 grit and then 2000 grit emery paper prior to each experiment. Analytical grade 65 % HNO_3 was diluted with double distilled water to provide the aerated 1 M HNO_3 solution used for the studies. ASTM G1 and ASTM G31 are followed in the measurement of weight loss and the computation of corrosion rates. Eqs. 1 and 2 were used to compute corrosion rates and inhibition efficiency, respectively (ASTM International) [52].

$$CR(mm/year) = \frac{87.600 \quad \times WL}{\rho at} \tag{1}$$

Where WL represents the weight loss (mg), a is the specimen surface area, ρ is the copper sample density (g/cm^3) and t is the exposure period (hours). The inhibitive efficacy ($IE\%$) was determined according to Eq. 2:

$$\frac{IE(\%) = WL_0 - WL}{WL_0} \tag{2}$$

Where WL_0 refers to sample weight loss without inhibitor and $WL$is the samples weight loss with inhibitor in the corrosive environment. The concentrations in the tube, which was submerged in a temperature-controlled water bath, varied from 0 (blank) to a maximum of 100:500 ppm. The specimens were taken out of the test solution after a day and cleaned with distilled water. The lack of insoluble corrosion products oxide on the specimen show that copper corrodes to Cu^{2+} under these circumstances. To ensure accuracy and reproducibility, the experiments were carried out three times using copper coupons that had the same chemical composition.

2.2. Model data preparation and preprocessing

Relevant features for the model were extracted from the experimental procedures comprising exposure time (hours), temperature (°C), and inhibitor's concentration (ppm) as the input variable and corrosion rate (mm/year)derived from gravimetric analysis as the output variable. The experimental dataset used for model training and validation comprised 48 corrosion data samples. The statistical properties of the key variables are provided in Table 1. Temperature and inhibitor's concentration exhibit a higher standard deviation, depicting significant fluctuations in their values. This impact on the rate of corrosion. The mean corrosion rate of 5.40 mm/year suggests moderate corrosion activity, but the wide range (0.1242–16.99 mm/year) implies that different conditions led to highly variable outcomes.

The correlation heat map for the output and input variables is shown in Fig. 1, giving a preliminary assessment of variable interdependence [53]. The exposure time exhibits a slight negative correlation of −0.15 with the corrosion rate, suggesting that increased exposure time marginally reduces the corrosion rate. However, temperature has a positive correlation of 0.42 with the corrosion rate. This indicates that an increase in temperature is associated with a rise in corrosion rate and establishes the severity of corrosion at higher temperatures. A significant negative correlation (-0.80) indicates that increased inhibitor concentrations markedly diminish corrosion rates, thus establishing the inhibitor's effectiveness in reducing the rate of corrosion. The dataset was visually and statistically checked for outliers and data inconsistencies before model training. Extreme values of input and output variables were identified using a z-score analysis as in Eq. 3. Data points with $|z| > 3$ were considered potential outliers [54]. This study optimized data scaling to reduce overfitting and improve ML model robustness [55].Eq. 4 was used to scale the features by normalizing the training data to ensure a similar scale.

$$z = \frac{x - \mu}{\sigma} \tag{3}$$

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

Where x is the mean of the variables and $x_{norm}$is the normalized variable.$x_{\min}$ and $x_{\max}$ are the minimum variable and maximum variable, respectively.

2.3. Model development

2.3.1. Hyperparameter settings

The hyperparameters of the four machine learning models namely Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest (RF) and Decision Tree (DT) models, were optimized to achieve an optimal prediction performance. Model hyperparameters were tuned using a grid-search optimization strategy.The choice of these models was based on their performance in addressing complex non-linear relationships, which are inherent in the corrosion inhibition problem. A 2-hidden-layer architecture was selected for the ANN, with 10 neurons in each layer. The inherent complexity of the inhibition problems requires more than a single-layer architecture [56].Furthermore, a Levenberg Marquardt (LM) training algorithm and Rectified Linear Unit (ReLU) activation function was defined for the training process over 500 iterations. Control parameters specified for the SVM models includea RBF kernel, with the Gamma and epsilon values set at 0.1, respectively. A regularization parameter as high as 100 was set to reduce bias. A maximum depth of 5 was selected for the decision tree to prevent overfitting by limiting tree growth and ensuring generalization. Decision trees split data based on the most important features, allowing feature importance analysis for corrosion prediction. A 100 estimator was selected to ensure stability and accuracy, as it reduces variance and prevents individual trees from overfitting.

Table 1
Properties of key model variables.

Parameter	Maximum	Minimum	Mean
Time (hour)	50	30	40
Temperature (^{O} C)	504	24	264
Inhibitor concentration (ppm)	500	0	250
Corrosion rate (mm/year)	16.993	0.124	5.398

2.3.2. SHapley Additive exPlanations (SHAP)

To provide an interpretable explanation of the influences of corrosion parameters on the corrosion rate, a Shapley Additive exPlanations (SHAP)-based feature ranking framework was employed. SHAP offers a model-agnostic methodology for decomposing a machine learning model's prediction into the additive impacts of individual features, ensuring both global and local interpretability [57]. For a model with n input feature i, the SHAPley value ϕ is computed as in Eq. 5.

$$\phi_i = \sum_{S \subseteq N} \frac{(n - |S| - 1)!|S|!)}{n!} [v(S \cup \{i\}) - v(s)] \tag{5}$$

Where ϕ_i is the SHAPley values representing the importance of each feature, n denoted the number of features while N represents the group input in the dataset. S is a subset of N. The SHAP algorithm's basic principle is that the sum of all feature contributions is obtained by subtracting the baseline ϕ_o and the model's predicted value $f(x)$ as in Eq. 6 [58].

$$f(x) = \sum_{i=1}^{N} \phi_I(x) + \phi_o \tag{6}$$

The value of $\phi_i(x)$ signifies the extent to which the feature influences the model's prediction relative to the baseline ϕ_o for the data instance x. The predicted outcome is represented by the baseline value, ϕ_o

2.3.3. k-fold cross-validation and performance evaluation

In this study, a k-fold cross-validation was adopted for model training and testing to enhance the reliability and generalization of the developed machine learning models. The complete dataset was randomly partitioned into $k = 5$ equal folds. During each iteration, 4 folds were used to train the models, while the remaining fold was used for testing. At each iteration, the following statistical metrics, Variance Accounted For (VAF), Mean Absolute Error (MAE), Mean absolute deviation (MAD), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) and coefficient of determination (R^2) were computed (Eqs. 7–12), and the average values were reported to evaluate the model's stability and predictive consistency.

$$MAPE = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{y_k - \widehat{y_k}}{y_k} \right| \times 100\% \tag{7}$$

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N} [y_k - \widehat{y_k}]}{N}} \tag{8}$$

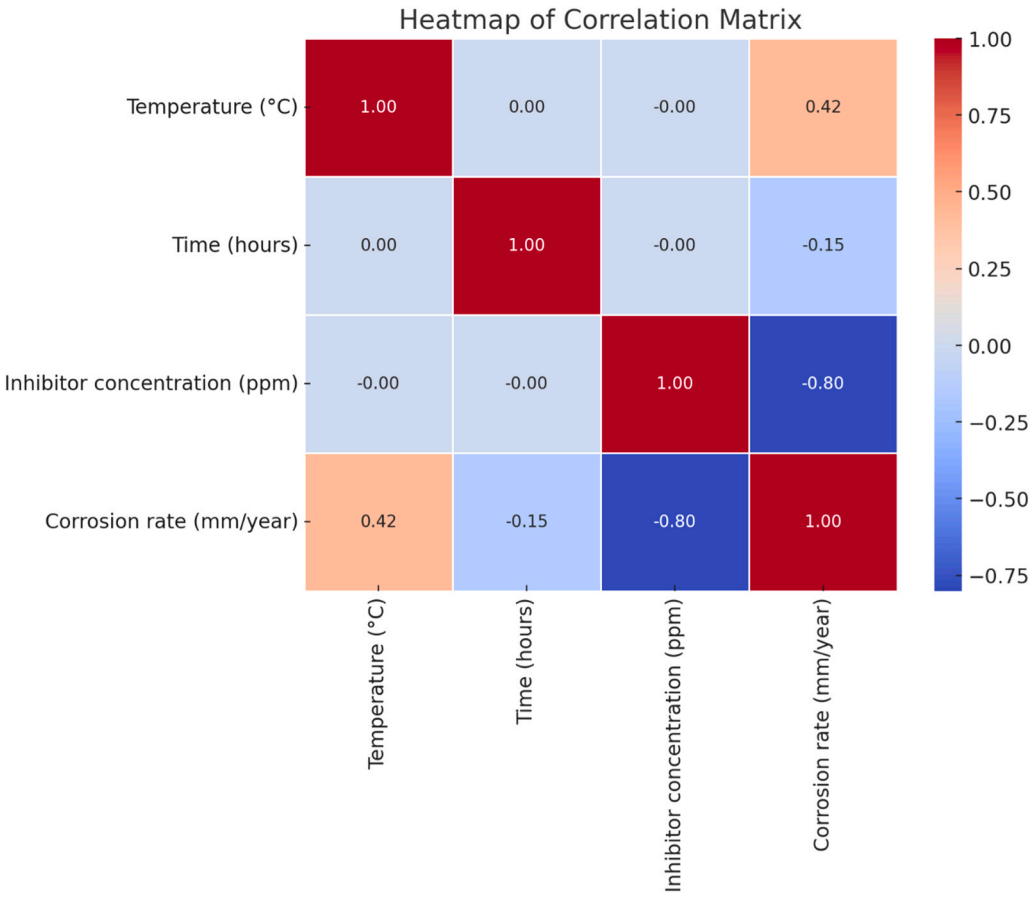

Fig. 1. correlation heat map of the input and output data.

$$MAE = \frac{\sum_{k=1}^{N} |\widehat{y_k} - y_k|}{N} \quad (9)$$

$$MAD = \frac{\sum_{k=1}^{N} |y_{k-}\overline{y}|}{N} \quad (10)$$

$$VAF = 1 - \left[\frac{\mathrm{var}(\widehat{y_k} - y_k)}{\mathrm{var}(y_k)}\right] \times 100 \quad (11)$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N} (y_{k-}\widehat{y_k})^2}{\sum_{k=1}^{N} (y_{k-}\overline{y})^2} \quad (12)$$

The experimental dataset used in this study is relatively meager and includes a narrow operating range of exposure conditions (time, temperature, and inhibitor concentration). To address the challenge of limited model variability and overfitting risk, systematic hyperparameter tuning, and k-fold (5-fold) cross-validation were carried out in addition to the normalization data processing process to provide more reliable estimates of model

Table 3
IE% and fraction of surface covered variation with SLE concentrations (exposure time = 96 h; T = 30°C).

Exposure time (Hours)	Concentration (ppm)	Inhibition efficiency (%)	Degree of surface coverage (θ)
24	100	68.565	0.686
	200	70.987	0.709
	300	87.957	0.879
	400	97.752	0.977
	500	98.996	0.989
48	100	68.429	0.684
	200	70.643	0.706
	300	81.818	0.818
	400	97.182	0.972
	500	98.915	0.989
72	100	68.222	0.682
	200	70.637	0.706
	300	80.357	0.804
	400	96.682	0.967
	500	98.868	0.988
96	100	68.153	0.682
	200	70.437	0.704
	300	80.218	0.802
	400	96.668	0.967
	500	98.852	0.988

3. Results and discussion

3.1. Evaluation of the corrosion inhibition performance of Spinacia oleracea extracts (SLE)

In this study, the SLE's effectiveness as corrosion inhibitor was assessed using a gravimetric test. According to Table 3, the inhibitory efficiency of SLE increases steadily with concentration, reaching a maximum efficiency at a "limiting concentration." Accordingly, inhibition efficiency rose in the following manner with concentration at the 100–500 ppm level: 68.57: 70.99: 87.96: 97.75: 98.99. Fig. 2 graphically depicts the trend, which is practically linear. It may be the result of organic compounds in the corrosion inhibitor molecule adhering to the surface of the copper, which tends to reduce surface roughness, which is

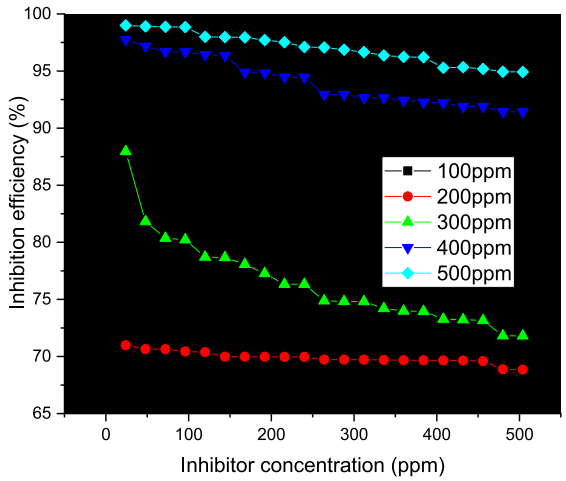

Fig. 2. Variation of percentage inhibition efficiency with concentration.

clearly advantageous [59]. Thus, it may be concluded that inhibitors molecules adsorbed onto the copper and/or corrosion products, forming protective layers that prevented the metal substrate from coming into touch with the acidic solution [60], hence slowing down the pace of corrosion. Therefore, it may be concluded that the geometric blocking effect was adhered to by the corrosion inhibition mechanism of SLE. Surface coverage (θ) increases as a result of SLE's adsorption on the metal surface or its interaction with corrosion product on the metal surface. The contact between corrosive ions in the solution and the metal matrix is successfully blocked by this improvement in physical shielding capability. Fig. 2 showed the relationship between the concentrations of SLE and the corrosion inhibition efficiency in a 1 mol/L HNO_3 solution at room temperature. As the concentration of SLE increased, so did the corrosion inhibition efficiency value derived from the weight loss method. It was determined that the ideal concentration for using SLE as a corrosion inhibitor was 500 ppm. Therefore, we have suggested that SLE be taken into account as a copper inhibitor in acidic settings.

3.2. *SHAP-based interpretability analysis of corrosion predictors*

The SHAP analysis provides insight into the understanding of the corrosion inhibition mechanism by quantifying how each factor contributes to the corrosion response, highlighting the dominant variables that drive inhibition efficiency. Shown in Fig. 3 is the SHAP summary plot, which visualizes the magnitude and direction of the influence of each corrosion and inhibition parameter on corrosion rates. It is shown that inhibitor concentration is the most dominant feature influencing the corrosion rate. High inhibitor concentrations correlate with significantly negative SHAP values, signifying that an increase in *Spinacia oleracea* extract concentration markedly decreases the predicted corrosion rate. This observation correlates with the experimental findings presented in Table 3, where inhibition efficiency increased steadily with concentration, reaching up to 98.99 % at 500 ppm. The negative SHAP contribution aligns with the physical mechanism of inhibitor adsorption, wherein increased molecule coating of the copper surface restricts metal-acid interaction and inhibits electrochemical processes.

In addition, temperature was observed to exhibit a predominantly positive SHAP contribution. This implies that a higher corrosion rate is recorded with increasing temperature. High-temperature points correlate with substantial positive SHAP values, signifying their significant corrosion-accelerating influence. This trend aligns with the positive correlation (r = 0.42) between temperature and corrosion rate reported earlier (Fig. 1), reflecting the thermally activated nature of the dissolution process, which enhances ion mobility and metal oxidation kinetics. The exposure time variable exhibits a mixed and relatively diminished effect on the model output, with SHAP values clustered near zero. Longer exposure durations slightly reduce corrosion rate predictions, indicating that extended immersion enables a more stable inhibitor film to form on the copper surface. However, the modest range of SHAP values suggests that time alone does not significantly alter corrosion dynamics compared to concentration and temperature.

The magnitude of contribution shown in the SHAP bar plot in Fig. 4 shows the average absolute contribution of each experimental parameter to the predicted corrosion rate, grading feature relevance quantitatively. The bar lengths show the mean absolute SHAP value, indicating each variable's overall impact on model output, independent of direction. The plot shows that inhibitor concentration has the highest mean SHAP value, making it the most important corrosion inhibition predictor. Its dominance shows that the corrosion rate is strongly correlated with *Spinacia oleracea* extract dose. The model correctly predicted that larger inhibitor concentrations increase surface coverage and decrease metal dissolution, highlighting the inverse link between concentration and corrosion rate. With a moderate mean SHAP score, temperature is the second most important. This shows that temperature accelerates corrosion despite its secondary role. Temperature increases ionic mobility and metal oxidation reactions in nitric acid environments. Exposure time has the lowest mean SHAP value, indicating little effect on corrosion prediction. After the inhibitor film stabilizes on the copper surface, immersion duration barely significantly affects the corrosion rate within the tested range.

Fig. 5 presents the SHAP decision plot showing how inhibitor concentration, temperature, and exposure time affect the model's corrosion rate prediction. From the base value (average model prediction) to the individual predicted corrosion rate, each line shows the variables' cumulative contribution to the model output. The plot shows that inhibitor concentration has the greatest detrimental effect on model output. However, temperature favorably affects model output, increasing corrosion rate forecasts, especially at high temperatures. The red-colored temperature segments in multiple samples rise, indicating that ion mobility and oxidation kinetics in the nitric-acid environment accelerate corrosion. The slight effect of exposure time moderates' corrosion rate over time. Time's blue segments show small negative shifts, indicating that extended immersion durations solidify and preserve the inhibitor layer's protective efficacy, though less so than

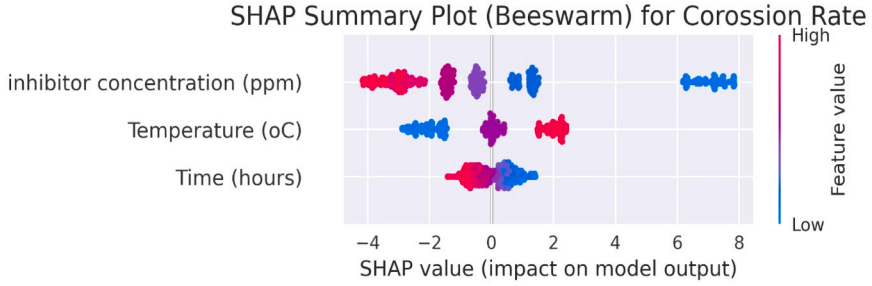

Fig. 3. SHAP summary plot of corrosion features' influence on corrosion rates.

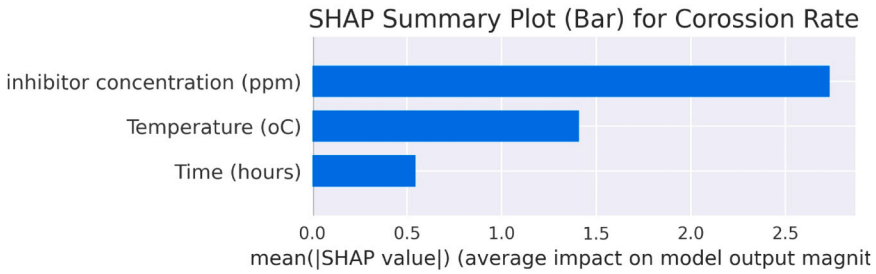

Fig. 4. SHAP bar plot of corrosion features' influence on corrosion rates.

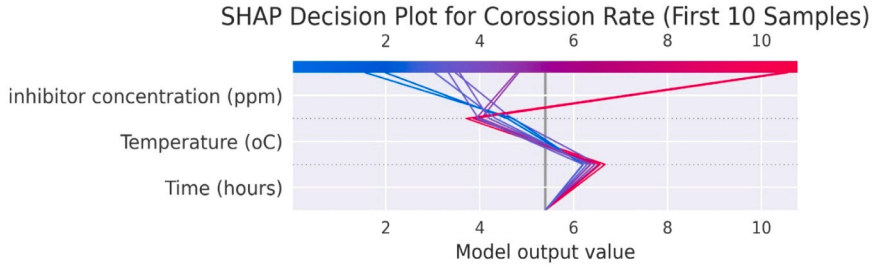

Fig. 5. SHAP decision plot of corrosion features' influence on corrosion rates.

concentration and temperature.

Shown in Fig. 6 is the SHAP force plot of corrosion features' influence on corrosion rates, which breaks down the corrosion rate prediction for a single sample into experimental variable components for local interpretability. It shows how unique feature values for the first data instance collectively change the model's prediction away from the baseline (mean prediction of all samples), unlike global feature-importance plots. The red arrows indicate features that push the prediction toward higher corrosion rates, whereas the blue arrows represent features that lower the predicted value. The inhibitor concentration and exposure time both have large red contributions (negative SHAP values), lowering the predicted corrosion rate relative to the baseline. Since higher inhibitor concentration improves surface coverage and film stability, *Spinacia oleracea* extract protects against corrosion. Temperature (blue segment) increases projected corrosion rate, slightly offsetting inhibitory effects. This supports the thermally triggered nature of corrosion processes; higher temperatures accelerate electrochemical reactions and undermine inhibitor film stability.

3.3. Performance evaluation of the developed model

The ML models were developed and assessed using a 5-fold cross-validation, with four out of the five equal folds used for training while the remaining fold was used for validation. The models' performance was evaluated using relevant statistical metrics. With 5 iterations, the process was repeated, and the performance metrics presented in Table 4 and 5represent the average values obtained across the five folds.Table 4 presents the average statistical metrics aggregated across all 5 folds at the training phase, revealing a trend that shows thatthe SVM model outperformed other models at the training phase, exhibiting the lowest error metrics (RMSE = 0.386, MAE = 0.287, MAD = 0.192, MAPE = 0.08773) alongside the highest VAF (99.09 %) and R^2 (0.9908). The results demonstrate that the model accounts for almost all the variance in the training data, rendering it exceptionally proficient in discerning the underlying patterns [61]. The Random Forest model exhibited robust predictive accuracy as the second-highest performing model, with statistical metrics comparable to those of the SVM model. The model has a slightly higher RMSE of 0.405 and MAE of 0.306,indicating a minor compromise in accuracy relative to SVM. However, the ensemble learning methodology of Random Forest frequently improves generalization, indicating it may outperform an overfitted SVM. The Decision Tree model exhibited satisfactory performance during training, although it recorded a higher RMSE, MAE, and MAPE values compared to SVM and Random Forest. The R^2 (0.9828) and VAF (98.28 %) demonstrate that the model accounts for the majority of the variation; nevertheless, its higher error metrics imply a diminished fit at training. The ANN model showed the largest variability across folds, implying limited generalization and relatively higher prediction errors.

Table 4
Average training performance of all developed models across 5-fold cross-validation.

Model	RMSE	MAE	MAD	MAPE	VAF	R^2
ANN	0.735	0.576	0.448	0.18345	96.664	0.967
SVM	0.386	0.287	0.192	0.08773	99.093	0.991
Decision Tree	0.528	0.361	0.210	0.13313	98.281	0.983
Random Forest	0.405	0.306	0.205	0.12977	98.985	0.989

Fig. 7 compares the performance of the ANN, SVM, decision tree, and random forest based on the averaged training statistical metrics values across folds. Random Forest and SVM demonstrate superior overall performance, evidenced by their minimal RMSE, MAE, and MAD,

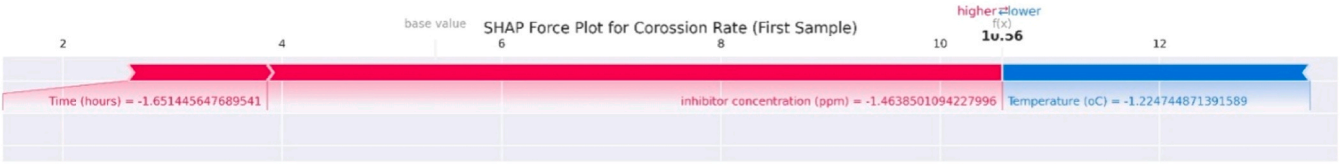

Fig. 6. SHAP Force plot of corrosion features' influence on corrosion rates.

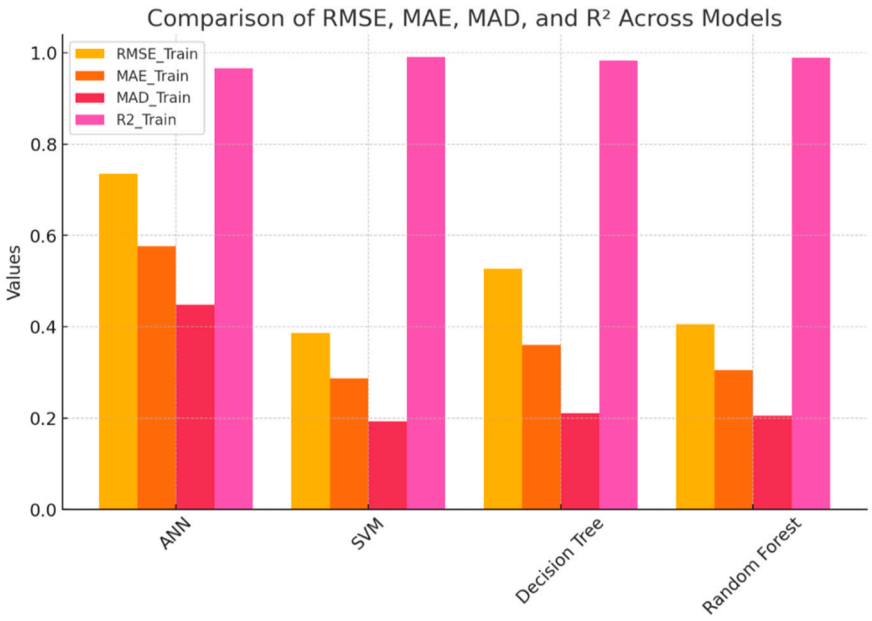

Fig. 7. Comparison of averaged model performance across 5-fold training subsets.

alongside elevated R^2 values. The ANN has the highest error metrics (RMSE, MAE, and MAD), and the widest deviation, suggesting overfitting and underfitting tendencies in some folds. The Decision Tree has moderate performance, with marginally increased RMSE and MAE relative to SVM and Random Forest. If the objective is to minimize errors, Random Forest or SVM would be the superior options. The VAF and MAPE values are depicted in the trend in Fig. 4.

Based on VAF and MAPE values, the stacked chart in Fig. 8 compares the performance of all models. All models offer robust performance based on the VAF values, indicating their efficacy in modeling data variance. The VAF values show negligible variation among models, indicating that each model well captures the fundamental data patterns. SVM and Random Forest have the lowest MAPE values, rendering them the most dependable models regarding predictive accuracy. ANN have satisfactory performance; nonetheless, they are marginally surpassed by SVM and Random Forest regarding MAPE.

The average regression plots obtained from all 5-folds for all models for predicting the corrosion rate are presented in Fig. 9. Random Forest demonstrates the highest accuracy based on the correlation coefficient, with predictions that closely correspond to experimental corrosion data.

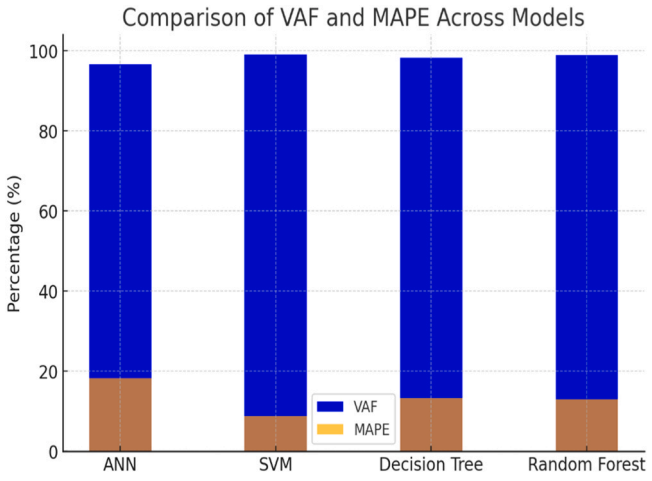

Fig. 8. Mean VAF and MAPE values across 5-fold cross-validation for all models.

SVM demonstrates outstanding performance, exhibiting little dispersion and divergence from the ideal fit. The Decision Tree exhibits accuracy but demonstrates indications of overfitting, presumably attributable to its architecture. ANN demonstrates the most variability, signifying its inferior reliability compared to the other models in this particular application (Fig. 10).

Thetrend comparison chart in Figure 10visualizestheexperimental vs. predicted corrosion rate averaged over the folds**.** The experimental corrosion rate displays noticeable peaks and variations, necessitating models capable of effective generalization. The Decision Tree exhibits localized mispredictions, suggesting it may lack generalizability across different parts of the dataset. Random Forest and SVM proficiently encapsulate these variations, rendering them formidable candidates for precise prediction. All models closely adhere to the experimental corrosion rate trend [62]. Random Forest and SVM demonstrate the most accurate fit, since their predictions nearly correspond with the experimental values. ANN display slight discrepancies, especially in regions of abrupt changes, potentially leading to increased percentage error. Moreover, ANN exhibits minor discrepancies, especially in areas with significant fluctuations. Decision Tree forecasts may occasionally deviate, potentially signifying overfitting to certain values. Both ANN and Decision Tree exhibit challenges in accurately identifying peaks, indicating possible difficulty in managing high fluctuations. Consistent underestimation or overestimation of corrosion rate by a model may result in elevated RMSE, MAE, or MAPE values. Random Forest and SVM exhibit negligible variance, indicating they probably produce reduced error metrics.

Table 5 presents the average validation-phase results (mean of 5 folds). The SVM exhibits the lowest error values with RMSE. MAE, MAD, and MAPE values of 0.408, 0.308, 0.227, and 9.806 % respectively, signifying it as the most precise model with little predictive error. Likewise, SVM exhibits the highest VAF and R^2 values of 98.89 % and 0.9889, signifying that it accounts for nearly 99 % of the variation in the data. The RMSE, MAE, and MAPE values of 0.419, 0.320, and 0.1190 % respectively, for RF demonstrate a strong performance, although slightly below SVM's. ANN exhibits the lowest VAF (96.63 %) and R^2 (0.9644), hence emphasizing its inferior prediction capability. Similarly, ANN exhibits the greatest error metrics (RMSE = 0.732, MAE = 0.587, MAPE = 0.15983 %), indicating that it is the least dependable model during the validation stage. The Decision Tree outperforms the ANN but underperforms compared to SVM and Random Forest. Random Forest

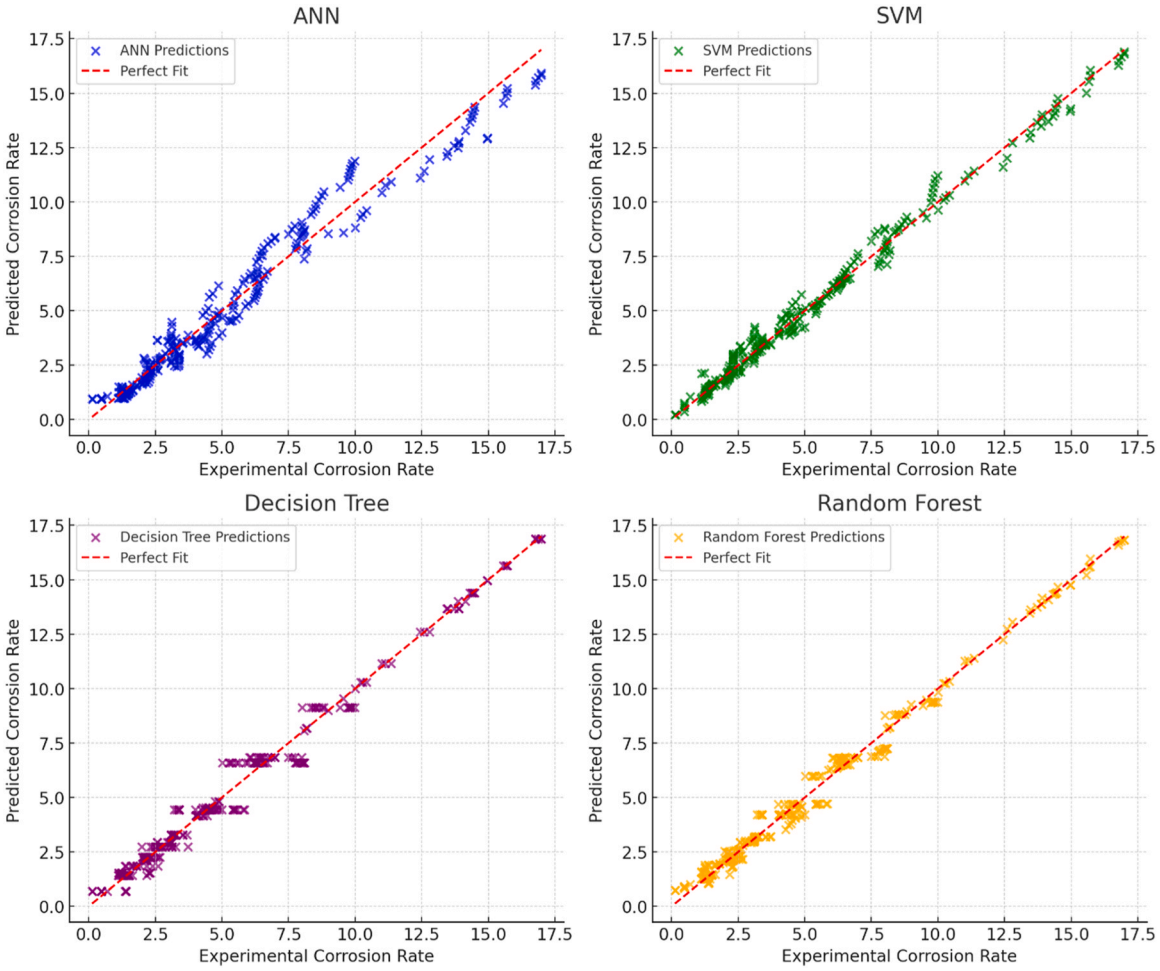

Fig. 9. Averaged regression plots of predicted vs experimental corrosion rate for the developed models.

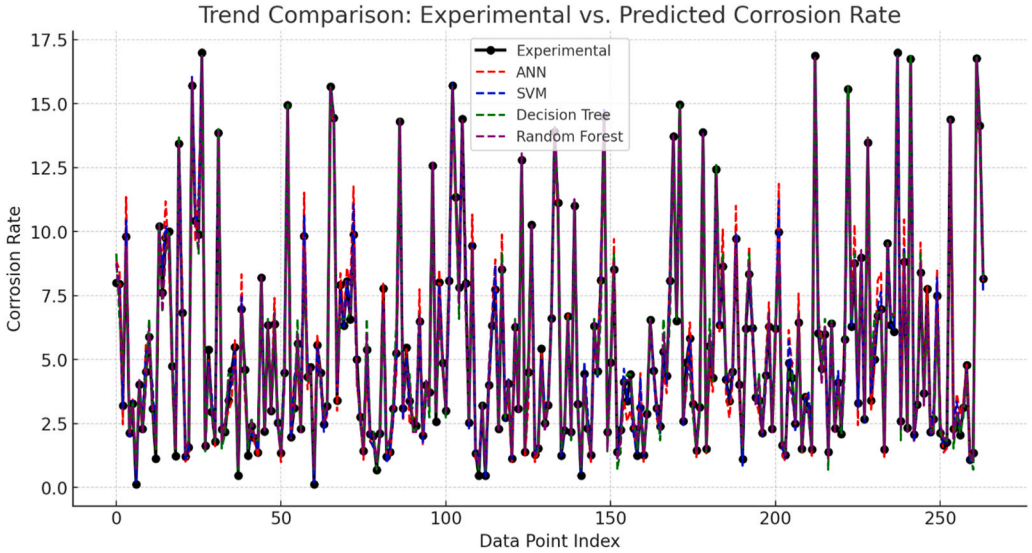

Fig. 10. Mean experimental–predicted trend across 5-folds.

exhibits remarkable performance, achieving a VAF of 98.84 % and a R^2 of 0.9884, rendering it nearly as robust as SVM. The Decision Tree attains a little lower R^2 (0.9811), indicating it accounts for less variance compared to SVM and Random Forest. The best corrosion rate prediction model is SVM based on lowest error values and higher accuracy. If computing performance and interpretability are insignificant, SVM is preferred for corrosion rate prediction. Random Forest is a reliable option that requires little adjustment, and it is a useful model for interpretability and stability.

The radar chart in Fig. 11 depicts the trend in the aggregated error

Table 5
Average validation-phase performance of all developed models across 5-fold cross-validation.

Model	RMSE	MAE	MAD	MAPE	VAF	R^2
ANN	0.732	0.587	0.497	0.15983	96.625	0.9644
SVM	0.408	0.308	0.227	0.09806	98.891	0.9889
Decision Tree	0.534	0.373	0.232	0.12651	98.106	0.9811
Random Forest	0.419	0.320	0.271	0.11905	98.839	0.9884

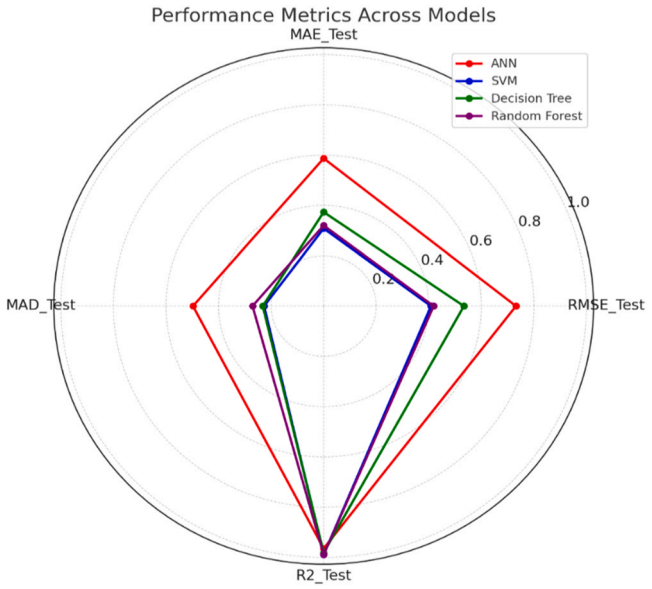

Fig. 11. Averaged error and accuracy metrics (mean of 5 folds).

profileobtained from the 5-fold cross-validation. It revealsthat SVM has the smallest polygon, indicating lower values for RMSE, MAE, and MAD across folds. It possesses one of the highest R^2 values, indicating it accounts for the majority of the variance in the data. This validates SVM as the superior model regarding prediction accuracy. The random forest exhibits slightly higher RMSE, MAE, and MAD compared with SVM; however, it significantly outperforms both ANN and Decision Tree. Its R^2 is almost equivalent to that of SVM, establishing it as a viable candidate for predicting corrosion rates. The Decision Tree has worse error metrics compared to ANN, although it is less effective than SVM or Random Forests. The R^2 value is somewhat worse to that of SVM and Random Forest, indicating it accounts for slightly less variation. ANN possesses the largest polygon, indicating higher error values. It has inferior reliability in predicting corrosion rates relative to alternative models.

The validation-phase VAF and MAPE averaged across folds are depicted in the trend in Fig. 12. SVM exhibits the highest VAF and lowest MAPE, establishing it as the most precise model. The random forest model has performance roughly comparable to that of the SVM. The decision tree exhibits commendable performance; yet, it demonstrates somewhat reduced VAF and elevated MAPE compared to SVM and Random Forest models.

Fig. 13 shows an error histogram depicting the residual error distribution for the developed models. This chart shows the frequency of different error values with a fitted probability density function curve. Residuals are more widely spread for ANN, indicating that it has higher errors compared to other models. The distribution is less symmetric, meaning ANN's predictions do not show a balance between over-predictions and under-predictions. For SVM, the error distribution is centered around zero, with less spread, indicating low error variance. For the decision tree, residuals are somewhat skewed, showing that it introduces systematic bias in certain predictions. The distribution for the random forest model is quite symmetric, with residuals closely

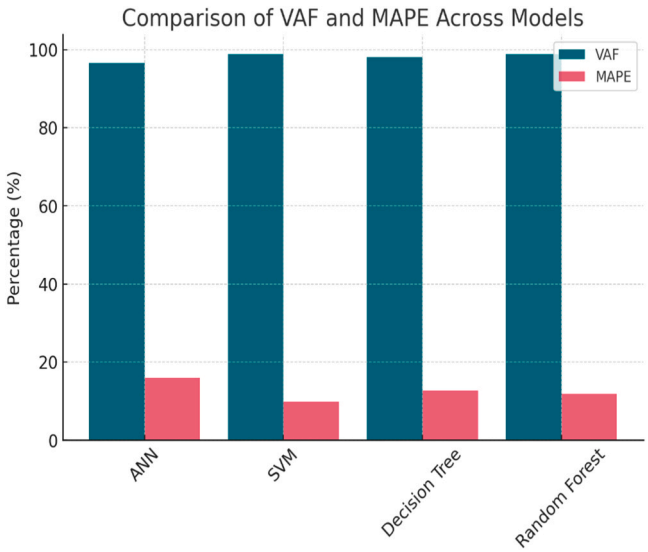

Fig. 12. VAF–MAPE trend averaged across 5 cross-validation folds.

concentrated around zero. It can be deduced from the error distribution diagram that SVM is the best model, showing the most normal error distribution and lowest error spread, while Random Forest is nearly as good, with slight skewness but still maintaining strong predictive ability. Decision Tree is moderately accurate but introduces some systematic bias. ANN has the highest error spread and the least symmetric distribution, confirming its poorer performance in corrosion rate prediction.

Fig. 14 consolidates the validation-phase experimental–predicted comparison averaged across all folds. Localized mispredictions made by the decision tree raise the possibility that it is not generalizable to other parts of the dataset. These variances are effectively captured by random forest and SVM, making them a viable option for accurate prediction. Given that their predictions almost match the experimental values, Random Forest and SVM show the best fit while ANN exhibits minor inconsistencies, particularly in areas with sudden changes, which could result in a higher rate of errors [63,64]. Predictions from decision trees may occasionally diverge, which could indicate overfitting to particular values. Increased RMSE, MAE, or MAPE values may arise from a model's consistent underestimation or overestimation of corrosion rate. Given their low variance, Random Forest and SVM most likely result in lower error metrics.

A comparative assessment of the developed ML models revealed that a significant variation exists in terms of their stability and generalization capacity. It was observed that the SVM and RF models exhibited a more consistent performance across the 5 folds, depicting a marginal discrepancy at the training and validation phase. This trend further establishes their greater reliability and robustness for a limited experimental dataset where the risk of overfitting is higher. However, the ANN model was observed to exhibit a substantially higher variability owing to its data size sensitivity.

Uncertainty analysis was carried out as an essential instrument for improving the model reliability of the SVM, as the most consistent models across folds. A split-conformal prediction interval (PIs) and bootstrap-based confidence bands were incorporated, aligning with recent methodological recommendations for quantifying uncertainty in ML models [65,66]. Shown in Table 6 are the key uncertainty indicators, namely Prediction Interval Coverage Probability (PICP), Mean Prediction Interval Width (MPIW), and the Normalized Prediction Interval Width (NPIW).Calibration residuals from each fold were used to determine non-parametric quantiles of absolute error for the 95 % PIs. The SVM model had a PICP of 0.958, demonstrating that 95.8 % of genuine corrosion-rate data fell within the 95 % confidence ranges (Table 5). The MPIW = 4.541 mm $year^{-1}$ and NPIW = 0.269 suggest excellent coverage

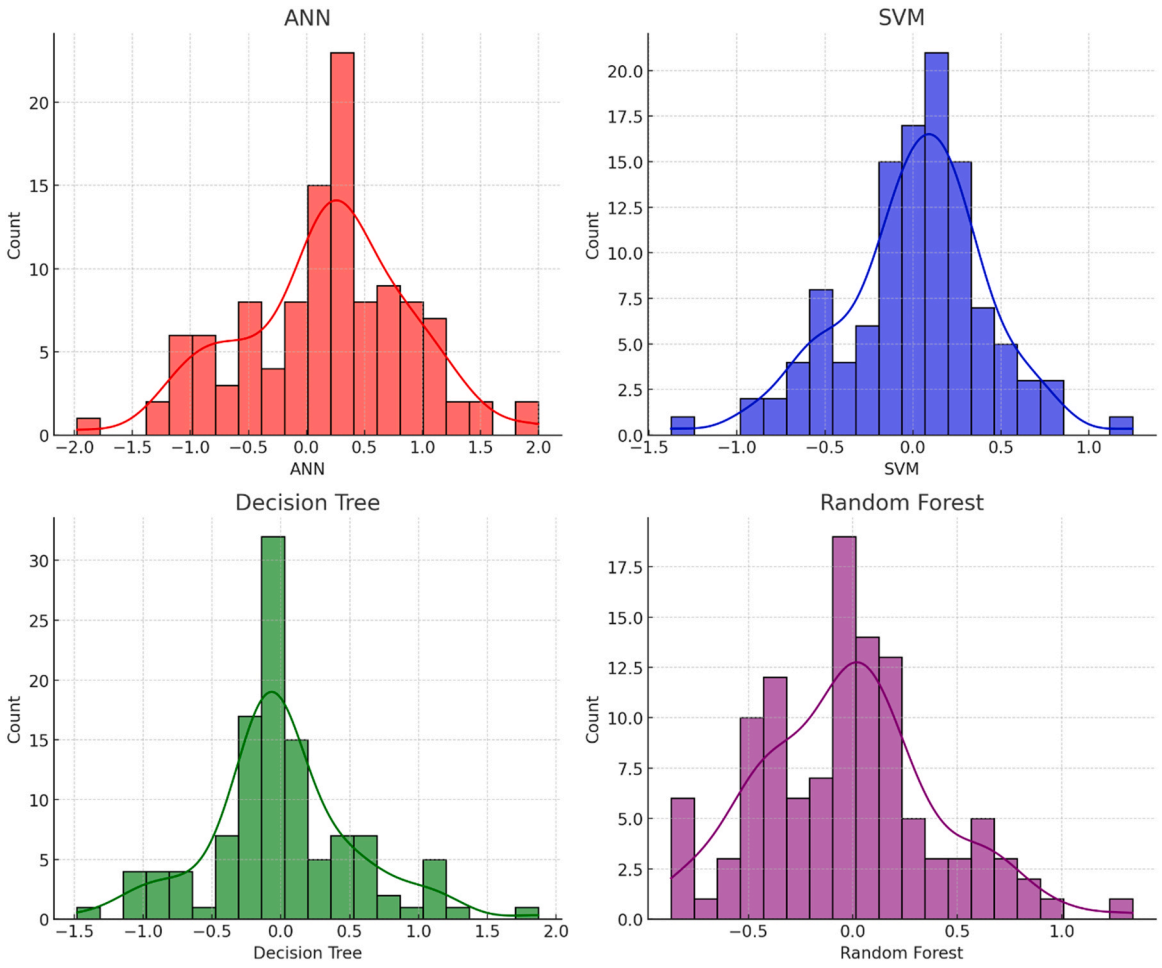

Fig. 13. Aggregated error histogram across 5 folds for the ANN, SVM, RF, and DT models.

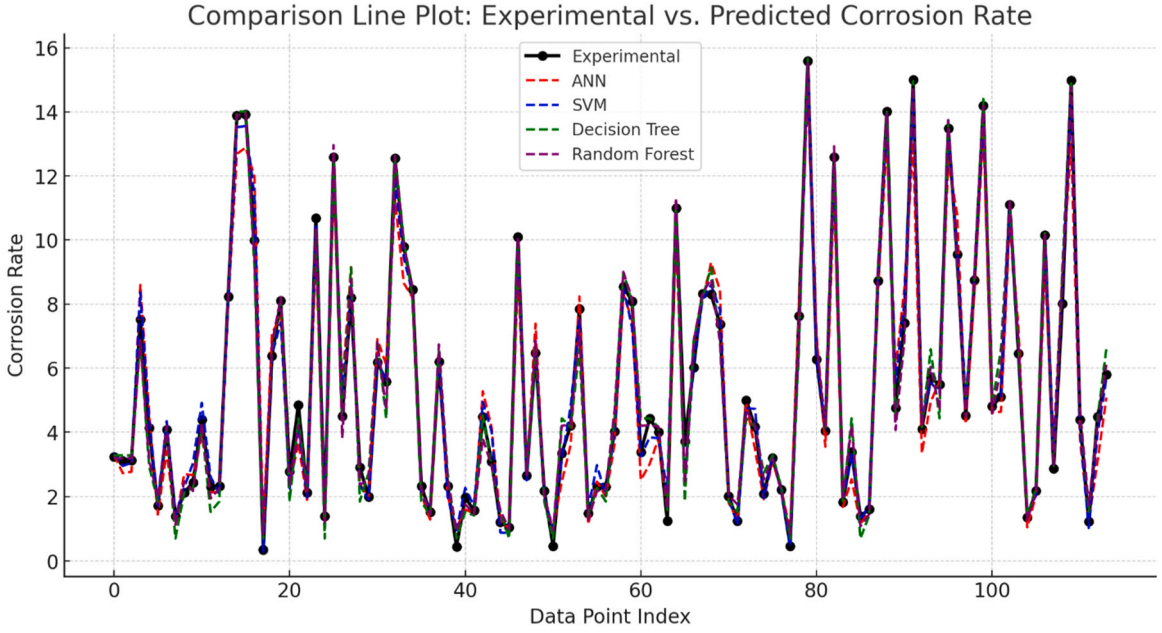

Fig. 14. Comparison of experimental and predicted corrosion rates averaged across 5-fold cross-validation.

reliability and crisp intervals. Fold-wise analysis showed consistent interval behavior, with PICP ranging from 0.880 to 1.000 and MPIW from 3.79 to 5.06 mm year^{-1} .

Table 6
Per-fold conformal PI summary.

Fold	PICP	MPIW	NPIW
1	1.000	4.960	0.294
2	0.921	4.075	0.242
3	0.986	4.814	0.285
4	0.88	3.794	0.225
5	1.00	5.062	0.300

4. Conclusions

This study develops an experimental and SHAP-enhanced ML framework for investigating the performance of SLE as a green corrosion inhibitor for copper in acidic media. The active components of SLE were extracted using a one-step water extraction process, and SLE demonstrated exceptional corrosion inhibition performance. The experimental investigation reveals that SLE displays a high inhibition efficiency of about 98.99 % at a concentration of 500 ppm, maintaining consistent performance over extended exposure time. The experimental result was further validated with data-driven insights using ML models namely ANN, SVM, Decision Tree, and Random Forest to predict the corrosion performance considering input variables such as exposure time, temperature, and inhibitor concentration. The SVM displayed the best performance in comparison to other models with averaged validation error metrics, with RMSE, MAE, MAD, MAPE, and VAF values of 0.386, 0.287, 0.192,0.08773, and 99.09, respectively. The SHAP-based feature interpretability revealed inhibitor concentration as the strongest negative influence on corrosion rate, while temperature exhibited a positive corrosion-enhancing effect, and exposure time had a marginal stabilizing impact. SLE molecules stick to the surface of the Cu and form a barrier that prevents the corrosive media from reaching the metal surface, which is the cause of the inhibitory mechanism. The data-driven framework that combines experimental gravimetric analysis with SHAP-enhanced ML in this study provides useful insight into the corrosion inhibition process and highlights the SLE extract potential as a sustainable, eco-friendly inhibitor that can be applied in industrial settings. Due to the small dataset and narrow experimental conditions, which may increase overfitting risk and restrict generalizability, this study recommends the expansion of the dataset across broader temperature, concentration, and exposure ranges, incorporating cross-laboratory validation, and exploring larger or ensemble-based models to enhance reliability and real-world applicability in future research.

CRediT authorship contribution statement

Omotayo Sanni: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Oluwatobi Adeleke**: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis. **Samuel Ayodele Iwarere**: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision. **Tien-Chen Jen**: Writing – review & editing, Writing – original draft, Validation. **Michael Olawale Daramola**: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] E. Karamnia, M. Yousefpour, A study of pomegranate peel extract effect on corrosion inhibition performance on aluminum in HNO3 solution, Colloids Surf. A Physicochem. Eng. Asp. 694 (2024 Aug 5) 134080.
[2] A. Hossen, R. Mahmud, A. Islam, S.K. Ahsan, M.I. Mondal, Minimization of corrosion in aquatic environment–a review, Int J. Hydro 7 (1) (2023) 9–16.
[3] O. Sanni, S.A. Iwarere, M.O. Daramola, Introduction: Corrosion basics and corrosion testing. InElectrochemical and Analytical Techniques for Sustainable Corrosion Monitoring, Elsevier, 2023 Jan 1, pp. 1–23.
[4] X. Li, P. Liu, C. Han, T. Cai, Y. Cui, W. Xing, C. Zhi, Corrosion of metallic anodes in aqueous batteries, Energy Environ. Sci. (2025).
[5] L. Mulky, P. Rao, Tribo-corrosion control with molecules of bio-origin: experimental studies and theoretical insights, J. Bio TriboCorros. 10 (3) (2024 Sep) 46.
[6] S. Rao, A. Arathi, S.S. Shree, G.K. Prashanth, H.S. Lalithamba, M.R. Kushalatha, S. Rao, L. Avinash, A sustainable approach for the corrosion control of mild steel using Cocous nucifera gum: an electrochemical investigation, Inorg. Chem. Commun. 170 (2024 Dec 1) 113423.
[7] C. Verma, A.H. Al-Moubaraki, A. Alfantazi, K.Y. Rhee, Heterocyclic amino acids-based green and sustainable corrosion inhibitors: adsorption, bonding and corrosion control, J. Clean. Prod. (2024 Feb 20) 141186.
[8] L. Afia, A. Ait Mansour, Z. Khadfy, M. Bazzaoui, R. Mamouni, R. Salghi, Valorizing the potential of saffron petals extract for aluminium corrosion control: an integrated approach involving extraction, experimental and computational analysis, Colloids Surf. A Physicochem. Eng. Asp. 696 (2024 Sep 5) 134240.
[9] O. Sanni, S.A. Iwarere, M.O. Daramola, Evaluation of corrosion inhibition of essential oil-based inhibitors on aluminum alloys, ACS Omega 7 (45) (2022 Nov 4) 40740–40749.
[10] D.I. Răuță, E. Matei, S.M. Avramescu, Recent development of corrosion inhibitors: types, mechanisms, electrochemical behavior, efficiency, and environmental impact, Technologies 13 (3) (2025 Mar 5) 103.
[11] X. Guo, B. Shi, Z. Fu, G. Yang, Y. Li, L. Wang, L. Lu, L. Ma, D. Zhang, Atom-scale insight into the adsorption behavior of imidazole corrosion inhibitors at defective copper/water interfaces, Corros. Sci. (2025 Jan 27) 112744.
[12] E.M. Saoudi Hassani, R. Salim, E. Ech-chihbi, H. Zejli, I. Mehdaoui, R. Mahmoud, Y.A. Younous, G. A. Shazly, A. Aqdas, A. Taleb, M. Taleb, Investigation of water treatment residues as corrosion inhibitors in acidic environment, Sci. Rep. 15 (1) (2025 Apr 9) 12164.
[13] H. Singh, S. Sharma, Understanding the Adsorption Behaviour of Corrosion Inhibitors on Metal–Water and Air–Water Interfaces from Molecular Simulations. Functional Materials for the Oil and Gas Industry, CRC Press, 2023 Sep 6, pp. 51–73.
[14] N. Okon Eddy, A.O. Odiongenyi, E.E. Ebenso, R. Garg, R. Garg, Plant Wastes as alternative sources of sustainable and green corrosion inhibitors in different environments, Corros. Eng. Sci. Technol. 58 (5) (2023 Aug) 521–533.
[15] R.A. El-Nagar, N.A. Khalil, Y. Atef, M.I. Nessim, A. Ghanem, Evaluation of ionic liquids based imidazolium salts as an environmentally friendly corrosion inhibitors for carbon steel in HCl solutions, Sci. Rep. 14 (1) (2024 Jan 22) 1889.
[16] H. Kumar, P. Yadav, R. Kumari, R. Sharma, S. Sharma, D. Singh, H. Dahiya, P. Kumar, S. Bhardwaj, P. Kaur, Highly efficient green corrosion inhibitor for mild steel in sulfuric acid: experimental and DFT approach, Colloids Surf. A Physicochem. Eng. Asp. 675 (2023 Oct 20) 132039.
[17] C.D. Taylor, Corrosion informatics: an integrated approach to modelling corrosion, Corros. Eng. Sci. Technol. 50 (7) (2015 Oct) 490–508.
[18] F. Vitse, S. Nesic, Y. Gunaltun, D. Larrey de Torreben, P. Duchet-Suchaux, Mechanistic model for the prediction of top-of-the-line corrosion risk, Corrosion 59 (12) (2003 Dec 1).
[19] B.F. Pots, Prediction of corrosion rates of the main corrosion mechanisms in upstream applications, NACE-05550. InNACE CORROSION, Nace, 2005 Apr 3.
[20] R.C. Woollam, S.E. Hernandez, Assessment and comparison of CO2 corrosion predicition models, SPE-100673. SPE International Oilfield Corrosion Conference and Exhibition, SPE, 2006 May 30.
[21] K. Nickless, R.A. Atadero, Mechanistic deterioration modeling for bridge design and management, J. Bridge Eng. 23 (5) (2018 May 1) 04018018.
[22] L.B. Coelho, D. Zhang, Y. Van Ingelgem, D. Steckelmacher, A. Nowé, H. Terryn, Reviewing machine learning of corrosion prediction in a data-oriented perspective, npj Mater. Degrad. 6 (1) (2022 Jan 26) 8.
[23] J.P. Cruz, E.G. Veruz, I.V. Aoki, A.M. Schleder, G.F. de Souza, G.L. Vaz, L.O. de Barros, R.T. Orlowski, M.R. Martins, Uniform corrosion assessment in oil and gas pipelines using corrosion prediction models–Part 1: models performance and limitations for operational field cases, Process Saf. Environ. Prot. 167 (2022 Nov 1) 500–515.
[24] S. Kovacevic, W. Ali, E. Martínez-Pañeda, J. LLorca, Phase-field modeling of pitting and mechanically-assisted corrosion of Mg alloys for biomedical applications, Acta Biomater. 164 (2023 Jul 1) 641–658.
[25] Shwetha KM, B.M. Praveen, B.K. Devendra, A review on corrosion inhibitors: types, mechanisms, electrochemical analysis, corrosion rate and efficiency of corrosion inhibitors on mild steel in an acidic environment, Results Surf. Interfaces (2024 Jul 20) 100258.
[26] R. Aslam, M. Mobin, S. Zehra, J. Aslam, A comprehensive review of corrosion inhibitors employed to mitigate stainless steel corrosion in different environments, J. Mol. Liq. 364 (2022 Oct 15) 119992.
[27] A. Dehghani, G. Bahlakeh, B. Ramezanzadeh, M. Ramezanzadeh, Potential of Borage flower aqueous extract as an environmentally sustainable corrosion

inhibitor for acid corrosion of mild steel: Electrochemical and theoretical studies, J. Mol. Liq. 277 (2019) 895–911.
[28] A. Sedik, D. Lerari, A. Salci, S. Athmani, K. Bachari, İ.H. Gecibesler, R. Solmaz, Dardagan Fruit extract as eco-friendly corrosion inhibitor for mild steel in 1 M HCl: Electrochemical and surface morphological studies, J. Taiwan Inst. Chem. Eng. 107 (2020) 189–200.
[29] A. Esquivel López, C. Cuevas-Arteaga, M.G. Valladares-Cisneros, Universidad Autónoma del Estado de Morelos Study of the corrosion inhibition of copper in synthetic seawater by *Equisetum arvense* as green corrosion inhibitor, Rev. Mex. Ing. QuíM. 19 (2019) 603–616.
[30] S. Al-Nami, Corrosion inhibition effect and adsorption activities of methanolic myrrh extract for Cu in 2 M HNO_3, Int. J. Electrochem. Sci. (2020) 1187–1205.
[31] A. Dehghani, G. Bahlakeh, B. Ramezanzadeh, Green Eucalyptus leaf extract: a potent source of bio-active corrosion inhibitors for mild steel, Bioelectrochemistry 130 (2019) 107339.
[32] R. Haldhar, D. Prasad, Corrosion resistance and surface protective performance of waste material of *Eucalyptus globulus* for low carbon steel, J. Bio TriboCorros. 6 (2020) 48.
[33] Q. Wang, B. Tan, H. Bao, Y. Xie, Y. Mou, P. Li, D. Chen, Y. Shi, X. Li, W. Yang, Evaluation of *Ficus tikoua* leaves extract as an eco-friendly corrosion inhibitor for carbon steel in HCl media, Bioelectrochemistry 128 (2019) 49–55.
[34] M. Akrom, S. Rustad, A.G. Saputro, H.K. Dipojono, Data-driven investigation to model the corrosion inhibition efficiency of Pyrimidine-Pyrazole hybrid corrosion inhibitors, Comput. Theor. Chem. 1229 (2023).
[35] M.E.A. Ben Seghier, D. Höche, M. Zheludkevich, Prediction of the internal corrosion rate for oil and gas pipeline: implementation of ensemble learning techniques, J. Nat. Gas. Sci. Eng. 99 (2022).
[36] A.H. Alamri, N. Alhazmi, Development of data driven machine learning models for the prediction and design of pyrimidine corrosion inhibitors, J. Saudi Chem. Soc. 26 (6) (2022).
[37] A. Agrawal, A. Choudhary, Deep materials informatics: applications of deep learning in materials science, MRS Commun. 9 (3) (2019) 779–792.
[38] J. Wang, Z. Zhang, X. Liu, Y. Shao, X. Liu, H. Wang, Prediction and interpretation of concrete corrosion induced by carbon dioxide using machine learning, Corros. Sci. 233 (2024 Jun 1) 112100.
[39] Z. Dong, M. Zhang, W. Li, F. Wen, G. Dong, L. Zou, Y. Zhang, Development of a predictive model for carbon dioxide corrosion rate and severity based on machine learning algorithms, Materials 17 (16) (2024 Aug 14) 4046.
[40] M. Khayatazad, L. De Pue, W. De Waele, Detection of corrosion on steel structures using automated image processing, Dev. Built Environ. 3 (2020 Aug 1) 100022.
[41] G. Qin, Y.F. Cheng, P. Zhang, Finite element modeling of corrosion defect growth and failure pressure prediction of pipelines, Int. J. Press. Vessels Pip. 194 (2021 Dec 1) 104509.
[42] D. Curiel, F. Veiga, A. Suarez, P. Villanueva, Advances in robotic welding for metallic materials: application of inspection, modeling, monitoring and automation techniques, Metals 13 (4) (2023 Apr 5) 711.
[43] R.E. Melchers, Modeling and prediction of long-term corrosion of steel in marine environments, Int. J. Offshore Polar Eng. 22 (04) (2012 Dec 1).
[44] K. Bhargava, A.K. Ghosh, Y. Mori, S. Ramanujam, Modeling of time to corrosion-induced cover cracking in reinforced concrete structures, Cem. Concr. Res. 35 (11) (2005 Nov 1) 2203–2218.
[45] U.M. Angst, Predicting the time to corrosion initiation in reinforced concrete structures exposed to chlorides, Cem. Concr. Res. 115 (2019 Jan 1) 559–567.
[46] B. Py, A. Maradesa, F. Ciucci, Gaussian processes for the analysis of electrochemical impedance spectroscopy data: prediction, filtering, and active learning, Electrochim. Acta 439 (2023 Jan 20) 141688.
[47] C.I. Ossai, B. Boswell, I.J. Davies, Predictive modelling of internal pitting corrosion of aged non-piggable pipelines, J. Electrochem. Soc. 162 (6) (2015 Mar 11) C251.
[48] C.I. Ossai, B. Boswell, I.J. Davies, Reliability Analysis and Performance Predictions of Aged Pipelines Subjected to Internal Corrosion: A Markov Modelling Technique, in: InInternational Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 57205, American Society of Mechanical Engineers, 2015 Aug 2.
[49] L. Organ, J.R. Scully, A.S. Mikhailov, J.L. Hudson, A spatiotemporal model of interactions among metastable pits and the transition to pitting corrosion, Electrochim. Acta 51 (2) (2005 Oct 10) 225–241.
[50] T. Würger, C. Feiler, F. Musil, G.B. Feldbauer, D. Höche, S.V. Lamaka, M. L. Zheludkevich, R.H. Meißner, Data science based Mg corrosion engineering, Front. Mater. 6 (2019 Apr 5) 53.
[51] J. Khatti, K.S. Grover, P. Samui, A comparative study between LSSVM, LSTM, and ANN in predicting the unconfined compressive strength of virgin fine-grained soil, Front. Built Environ. 11 (2025).
[52] ASTM International Standard Practice for Preparing, Cleaning, and Evaluating Corrosion Test, ASTM International (2011), pp. 1-9.
[53] F. Mwiti, A. Gitau, D. Mbuge, R. Njoroge, D.L. Antille, J. Khatti, Assessment of off-road agricultural traction in situ using large scale machine learning and neurocomputing models, Sci. Rep. 15 (1) (2025).
[54] O. Adeleke, T.-C. Jen, Explainable AI and machine learning-based analysis of municipal solid waste generation rate: a South African case study, Waste Manag. 206 (2025) 115036.
[55] J. Khatti, P.G. Asteris, A. Bardhan, Data proportionality and its impact on machine learning predictions of ground granulated blast furnace slag concrete strength, Front. Struct. Civ. Eng. 19 (8) (2025) 1305–1333.
[56] Adeleke, O., Akinlabi, S.A., Jen, T., &Dunmade, I. (2021).Application of artificial neural networks for predicting the physical composition of municipal solid waste: An assessment of the impact of seasonal variation.
[57] F. Tempel, E.A.F. Ihlen, L. Adde, I. Strümke, Explaining human activity recognition with SHAP: validating insights with perturbation and quantitative measures, Comput. Biol. Med. 188 (2025) 109838.
[58] O. Adeleke, K.O. Olatunji, D.M. Madyira, T.-C. Jen, Application of multimodal machine learning-based analysis for the biomethane yields of NaOH-pretreated biomass, Sci. Rep. 15 (1) (2025) 24372.
[59] B. Anandkumar, N.G. Krishna, R.V. Solomon, T. Nandakumar, J. Philip, Synergistic enhancement of corrosion protection of carbon steels using corrosion inhibitors and biocides: molecular adsorption studies, DFT calculations and long-term corrosion performance evaluation, J. Environ. Chem. Eng. 11 (3) (2023 Jun 1) 109842.
[60] Y.G. Avdeev, Y.I. Kuznetsov, Organic inhibitors of metal corrosion in acid solutions. I. Mechanism of protective action, Russ. J. Phys. Chem. A 97 (3) (2023 Mar) 413–427.
[61] P. Dutta, S. Paul, A. Kumar, Chapter 25 - Comparative analysis of various supervised machine learning techniques for diagnosis of COVID-19, in: In.S. L. Tripathi, V.E. Balas, S.K. Mohapatra, K.B. Prakash, J. Nayak (Eds.), Electronic Devices, Circuits, and Systems for Biomedical Applications, Academic Press, 2021, pp. 521–540.
[62] A.J. Kilani, O. Adeleke, C.A. Fapohunda, Application of machine learning models to investigate the performance of concrete reinforced with oil palm empty fruit brunch (OPEFB) fibers, Asian J. Civ. Eng. 23 (2) (2022) 299–320.
[63] W.S. Lee, V. Alchanatis, C. Yang, M. Hirafuji, D. Moshou, C. Li, Sensing technologies for precision specialty crop production, Comput. Electron. Agric. 74 (1) (2010) 2–33.
[64] J. Wang, Q. Chen, Y. Chen, RBF Kernel Based Support Vector Machine with Universal Approximation and Its Application, in: F.-L. Yin, J. Wang, C. Guo (Eds.), Advances in Neural Networks – ISNN 2004, Springer, Berlin Heidelberg, 2004, pp. 512–517.
[65] L. Chen, Y. Fissha, M. Hasanipanah, R. Ghodhbani, H. Dehghani, J. Khatti, Accurate prediction of blast-induced ground vibration intensity using optimized machine learning models, Def. Technol. 52 (2025) 32–46.
[66] K. null Jitendra null, N. null Denise-Penelope, K. null, Assessment of bearing capacity of concrete piles in alluvial soils using bio and swarm-optimized artificial neural network models, Bull. Comput. Intell. 1 (1) (2025) 53–75.