Latif *et al. Biochar* (2026) 8:88
https://doi.org/10.1007/s42773-026-00606-y

Biochar

ORIGINAL RESEARCH Open Access

Deep learning-aided prediction and mechanistic analysis of reaction kinetics in biochar-catalyzed antibiotic degradation

Junaid Latif[1,2], Na Chen[1,2*], Jia Xie[1,2], Zheng Ni[1,2], Lang Zhu[1,2], Azka Saleem[3], Kai Li[1,2] and Hanzhong Jia[1,2*]

Abstract

The proliferation of antibiotics in the environment poses a severe threat to public health and ecosystems. While biochar-based catalytic systems offer a promising remediation strategy, their design is complicated by nonlinear interactions between material properties and reaction conditions. To overcome this, we present an interpretable machine learning framework for the accurate prediction of reaction rate constants (*k*) for antibiotic degradation. A comprehensive dataset was compiled from the literature, encompassing 16 features across three domains including biochar catalyst properties, elemental composition, and reaction conditions. Among the six machine learning (ML) algorithms evaluated, the transformer-based deep learning model TabPFN (Tabular Prior-data Fitted Network) demonstrated superior performance (Test $R^2 = 0.91$, RMSE = 0.021), outperforming tree-based, kernel-based, and neural network models. Model interpretability analyses revealed that catalyst properties contributed the largest share of predictive capability (59.3%), followed by reaction conditions (25.9%) and elemental composition (14.8%). Persistent free radicals (formed at 450–550 °C) and high total pore volume (> 0.23 $cm^3\ g^{-1}$) were identified as key drivers of reaction kinetics, along with optimal oxidant (0.5–5.5 mg L^{-1}) and pollutant concentrations (< 22 mg L^{-1}). These insights were embedded into a user-friendly web-based GUI, enabling rapid *k* prediction for new biochar catalysts with prediction errors below 20% on external validation samples. This work provides both a robust predictive tool and a generalizable, data-driven methodology for understanding and optimizing complex environmental catalytic processes.

Highlights

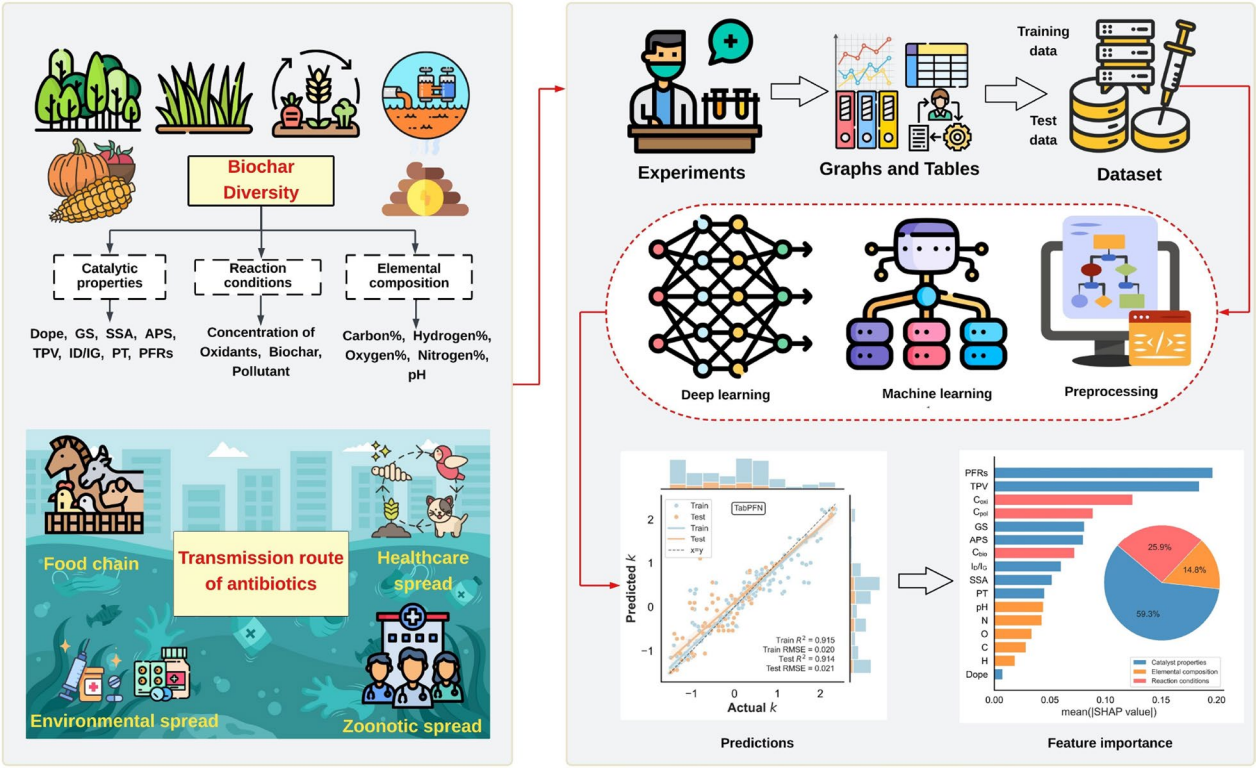
- Advanced AI predicts how fast biochar materials can break down antibiotic pollutants.
- The study identifies the key material traits that are most effective for improving pollutant removal.
- An interactive web tool enables fast screening and smarter design of high-performance catalysts.

Keywords Machine learning, TabPFN, Biochar, Antibiotics, Graphical user interface

*Correspondence:
Na Chen
nachen@nwafu.edu.cn
Hanzhong Jia
jiahz@nwafu.edu.cn
Full list of author information is available at the end of the article

Springer

Graphical Abstract



1 Introduction

Antibiotics have emerged as a global environmental concern due to their persistence and associated risks to human and ecological health (Qin et al. 2021). Often termed as "pseudo-persistent" pollutants, they are continually introduced into ecosystems despite various removal efforts. A significant proportion of antibiotics, either in their parent or metabolized forms, are discharged into sewage systems and eventually reach wastewater treatment plants (Luo et al. 2021). In China alone, annual antibiotic consumption reaches approximately 163,000 tons, accounting for nearly 50% of global usage (Yu et al. 2023). Reported concentrations range from 0.20 to 27 ng L^{-1} in coastal waters and sediments (Ohore et al. 2019), and from 0.4 to 0.7 μg L^{-1} in groundwater and surface water bodies (Yi et al. 2019). In these environments, antibiotics can inhibit the growth of primary producers and exert chronic toxicity on aquatic organisms. Human exposure through the food chain can trigger allergic reactions and disrupt physiological functions (Zhang et al. 2021b). More critically, the development of antibiotic resistance, where microorganisms survive and proliferate in the presence of antibiotics, has emerged as one of the most urgent public health threats (Balaban et al. 2019). Resistance to multiple major antibiotics is increasingly observed, leading to infections that are difficult or impossible to treat (Nathan and Cars 2014). Consequently, removing antibiotics and associated biological contaminants remains a critical challenge, requiring advanced treatment and prediction strategies.

Biochar, produced via biomass pyrolysis, has attracted increasing attention as an efficient and sustainable catalyst for antibiotic removal and degradation. Numerous studies have demonstrated that biochar-based catalytic systems can effectively degrade a wide range of antibiotics, including the fluoroquinolones group such as ciprofloxacin (Luo et al. 2019) and ofloxacin (Liu et al. 2020), sulfonamide such as sulfamethoxazole (Huang et al. 2019; Liang et al. 2019), and tetracyclines (Huang et al. 2019). During pyrolysis, persistent free radicals (PFRs) are generated and remain stabilized on the biochar surface, facilitating electron transfer to oxidants such as persulfate, hydrogen peroxide, and molecular oxygen, leading to the formation of reactive oxygen species (ROS) responsible for the degradation of emerging organic contaminants (Jia et al. 2019; Zhang et al. 2021a). Beyond PFRs,

biochar's high surface area, porous texture, and tunable surface chemistry play a critical role in activating oxidants (Brillas 2022) through both radical and non-radical pathways (Yu et al. 2020; Luo et al. 2021). However, pristine biochar often displays limited catalytic activity, motivating targeted structural modifications such as heteroatom doping (e.g., nitrogen and sulfur) to stabilize PFRs and enhance redox activity (Latif et al. 2024), as well as the deliberate introduction of structural defects including edge sites (Ouyang et al. 2019), curvature distortions (Zhang et al. 2020), and atomic vacancies (Zou et al. 2020), all of which promote electron transfer and catalytic reactivity. Consequently, the overall performance of biochar-based catalytic systems (BCSs) is governed by a complex interplay between intrinsic material properties, such as PFR concentration, pore structure, and doping characteristics (Zhang et al. 2021c, 2024), and operational conditions including pollutant concentration, oxidant dosage, and active site accessibility (Xu et al. 2020). This multivariate complexity results in substantial variability in reported degradation efficiencies and reaction kinetics across studies, underscoring the difficulty of rational catalyst design and highlighting the need for predictive, data-driven modeling approaches.

Recent computational advances have positioned machine learning (ML) as a powerful approach for uncovering complex relationships in large datasets and elucidating underlying mechanisms, offering advantages over traditional statistical methods (Zhong et al. 2021; Sun et al. 2025). Conventional algorithms like linear regression, decision trees, and stochastic gradient descent often showed limited accuracy and slower convergence when applied to small, heterogeneous datasets typical of published biochar studies (Latif et al. 2024). In contrast, integrated models such as Random Forests, Gradient Boosting, and Artificial Neural Networks offer improved performance due to their ability to model nonlinear relationships, manage high-dimensional data, and resist noise (Latif et al. 2024). However, these models require careful hyperparameter tuning to achieve optimal accuracy and generalizability. For instance, ML has been used to optimize biomass for the adsorption of antibiotics such as ciprofloxacin (Jiang et al. 2024), tetracycline, and sulfamethoxazole (Zhu et al. 2021), as well as to model adsorption kinetics (Jaffari et al. 2023; Liu et al. 2024) and reaction rate constants of emerging contaminants (Wang et al. 2024). Recent innovations like TabPFN, a transformer-based foundation model for tabular data, further enhance predictive capabilities (Hollmann et al. 2025). Pre-trained on over 130 million synthetic datasets derived from structural causal models, TabPFN employs in-context learning to approximate Bayesian inference and provides fast, accurate predictions in a single forward pass without extensive hyperparameter tuning. In environmental applications, integrated ML models have been widely adopted to study biochar systems. However, most existing studies focus on single pollutants or limited biochar characteristics to predict physicochemical properties and adsorption behavior. Despite the key role of PFRs in oxidant activation and ROS generation, they remain largely underrepresented in ML studies, likely due to challenges in quantification and limited data availability. This omission limits the ability of existing models to capture the full complexity of catalytic degradation. To address this, we integrate PFRs along with structural and operational descriptors into ML models, improving representation of nonlinear interactions in antibiotic degradation. This approach supports more efficient BCS design and enhances their real-world remediation performance.

This study presents a data-driven framework to evaluate and optimize the performance of BCSs for antibiotic removal. A comprehensive dataset was compiled from the published literature, incorporating catalytic properties, elemental composition, and reaction conditions. Data preprocessing ensured consistency across studies. Pearson correlation coefficients (PCC) were applied to assess the relationship among descriptors. Subsequently, multiple regression models were developed to predict the reaction rate constant (k) of antibiotics as the target variable. Feature importance analysis based on the trained ML models identified key predictors influencing degradation performance. Finally, a user-friendly graphical user interface (GUI) was developed to enable rapid prediction of reaction rates under varied experimental conditions, facilitating efficient catalyst screening and practical application.

2 Methodology

2.1 Literature review and data collection

A comprehensive dataset was constructed from 75 peer-reviewed studies retrieved from the Web of Science (Keywords: "biochar," "persistent free radicals," "antibiotic degradation"; see Table S1). The compiled data encompasses a wide range of antibiotics (e.g., sulfamethoxazole, ciprofloxacin, tetracycline; see full list in Table S2), ensuring the model captures the structural and compositional diversity of these pollutants. The target variable was the pseudo-first-order reaction rate constant (k), which was selected as the universal metric for evaluating and comparing degradation kinetics across different systems. These k values were either directly obtained from the literature or calculated from concentration–time data using Eq. 1:

$$\ln(C_t/C_0) = -k \times t \tag{1}$$

where C_0 and C_t are the antibiotic concentrations at the initial and reaction time (t), respectively. Experimental

data presented in tables or figures were digitized using WebPlotDigitizer.

To capture the multifaceted nature of the catalytic system, 19 variables were initially compiled from the literature and subsequently refined to 16 final input features based on data availability and redundancy screening. These features were grouped into three mechanistically meaningful categories guided by domain knowledge: (1) catalyst properties, including doping, graphitic structure (GS), specific surface area (SSA), average pore size (APS), total pore volume (TPV), Raman intensity ratio (I_D/I_G), pyrolysis temperature (PT), and persistent free radicals (PFRs); (2) elemental composition, including carbon (C), hydrogen (H), oxygen (O), nitrogen (N), and pH; and (3) reaction conditions, including oxidant concentration (C_{oxi}), biochar concentration (C_{bio}), pollutant concentration (C_{pol}). PFRs contents (spins g^{-1}) reported in the literature were primarily determined using Electron Paramagnetic Resonance (EPR) spectroscopy calibrated with 2,2-diphenyl-1-picrylhydrazyl (DPPH), which is the most widely adopted standard for quantifying and classifying environmentally persistent free radicals in biochar, soil, and ash matrices.

2.2 Data preprocessing and imputation

The detailed procedure of the ML exploration is illustrated in Fig. S1. The initial dataset comprised 389 records (data points) with 16 input features and the target variable *k* (Table S3). The chemical space of the curated dataset was defined by antibiotic class diversity to ensure the broad applicability of the developed models. Specifically, the dataset (Table S2) includes antibiotics from the tetracycline, fluoroquinolone, and sulfonamide families, which differ in structural frameworks and degradation behaviors. This diversity enables the model to capture general kinetic trends across multiple antibiotic classes rather than being restricted to a single compound type. To ensure the quality and reliability of the ML models, extensive data preprocessing was conducted.

First, features with > 50% missing data (e.g., "degradation efficiency", "g-factor", and "radical pathway") were excluded to avoid introducing excessive uncertainty. For the remaining features with missing values (Fig. 1a), a model-based interpolation method was employed (see Text S1), chosen for its ability to preserve underlying data structure over simpler mean/median imputation techniques because it predicts missing values based on the relationships between all other features, thereby preserving the underlying data structure and reducing bias (Hron et al. 2010).

To better reflect the non-linear and regime-dependent nature of radical formation during pyrolysis, we converted the continuous PFR spin density feature (spins g^{-1}) into four distinct categories: low, medium, high, and diminished. These groupings were determined through K-means clustering based on the relationship between pyrolysis temperature and PFR content, as illustrated in Fig. 1b and c and detailed in Text S2. We labeled samples without detectable signals as "No" to ensure they were treated as intentional control systems rather than missing data points. This four-tier classification allows for a more meaningful interpretation of PFR effects on reaction kinetics, as it accurately maps out the physical stages of radical evolution from initial growth to thermal decline without the loss of detail that occurs with fewer categories or the unnecessary complexity of a larger set.

Categorical features (Doping, GS, and the new PFRs category) were encoded. Doping and GS used label encoding, while the PFRs category used target encoding (replacing the label with the mean *k* of that category). Target encoding was selected for PFRs as it directly injects a supervisory signal related to the target variable, enhancing model learning for this critical feature.

A Pearson Correlation Coefficient (PCC) analysis was performed to assess inter-feature relationships and identify potential multicollinearity.

2.3 Feature and target transformation

To ensure model stability and performance, a two-step transformation was applied to all numerical features and the target variable. First, rather than removing numerical outliers, which could lead to a loss of realistic experimental variability, we applied the Yeo-Johnson method via a PowerTransformer. This approach was specifically chosen to normalize skewed distributions and minimize the influence of extreme values by approximating Gaussian behavior. Following this, the transformed features were scaled into the [0, 1] range using Min–Max scaling. This combined process ensures that all features contribute uniformly during training, preventing variables with larger magnitudes from disproportionately affecting the model's weight distribution.

For the target variable *k*, which exhibited significant right-skewness, logarithmic, square root, and Yeo-Johnson transformations were evaluated to stabilize variance. The Yeo-Johnson transformation (Eq. 2; λ parameter detailed in Text S3 and Table S4) was selected because it produced the closest to normal distribution (Fig. 1d–g) and resulted in subsequent model performance as validated by the RF model (Fig. 1h and i). The effectiveness of this normalization was confirmed via Q-Q plot and Shapiro–Wilk test (Fig. S2, Text S4).

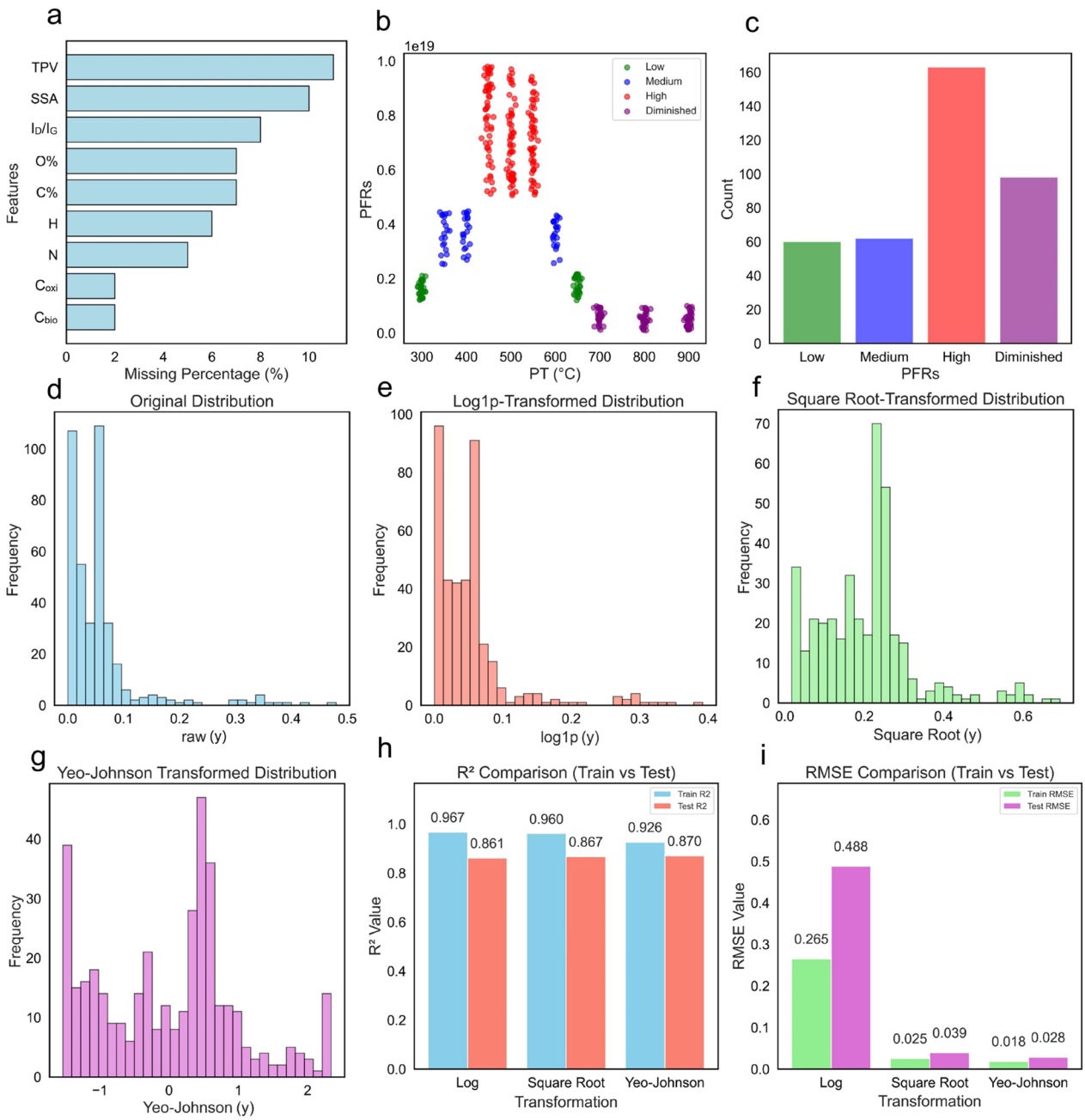


Fig. 1 Data preprocessing and transformation steps. **a** Missing value distribution. **b**, **c** PFRs categorization using *K*-means clustering. **d**–**g** Distributions of the target variable *k* after various transformations. **h**, **i** ML model performance across different transformations of *k*

$$y_i^{(\lambda)} = \begin{cases} \left((y_i+1)^{\lambda}-1\right)/\lambda & \text{if } \lambda \neq 0, y \geq 0 \\ \ln(y_i+1) & \text{if } \lambda = 0, y \geq 0 \\ -\left((-y_i+1)^{(2-\lambda)}-1\right)/(2-\lambda) & \text{if } \lambda \neq 2, y < 0 \\ -\ln(-y_i+1) & \text{if } \lambda = 2, y < 0 \end{cases} \quad (2)$$

where y_i is the original data and λ is the parameter of the Yeo-Johnson transformation.

2.4 Model development, training, and evaluation

An optimal data partitioning strategy was established through a sensitivity analysis, resulting in an 80/20 training–testing split with sevenfold cross-validation applied to the training subset. This configuration was selected to maximize model performance and generalization capability. The schematic of the subsequent model selection and evaluation workflow is presented

in Fig. S3. Using this optimized framework, a diverse suite of six ML algorithms was trained and evaluated: Random Forest (RF), eXtreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), Support Vector Regression (SVR), Multilayer Perceptron (MLP), and the transformer-based deep learning model (TabPFN). The inclusion of TabPFN is particularly noteworthy as a state-of-the-art method for tabular data, leveraging prior knowledge to achieve high performance with minimal hyperparameter tuning (Hollmann et al. 2025). Detailed model training procedures and hyperparameter spaces are provided in Text S5.

Model performance was comprehensively evaluated using multiple statistical metrics, including the Coefficient of Determination (R^2, Eq. 3), Root Mean Square Error (RMSE, Eq. 4), Mean Absolute Error (MAE, Eq. 5), Average Percentage Relative Error (APRE, Eq. 6), Mean Deviation (MD, Eq. 7), and Symmetric Mean Absolute Percentage Error (sMAPE, Eq. 8). All data preprocessing, model training, and evaluation procedures were carried out using Python (version 3.12).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_p^i - y_t^i\right)^2}{\sum_{i=1}^{N}\left(y_t^i - \overline{y_t}\right)^2} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_t^i - y_p^i\right)^2}{N}} \tag{4}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_t^i - y_p^i\right| \tag{5}$$

$$APRE(\%) = \frac{100}{N}\sum_{i=1}^{N}\frac{\left(y_p^i - y_t^i\right)}{y_t^i} \tag{6}$$

$$MD = \frac{1}{N}\sum_{i=1}^{N}\left(y_p^i - y_t^i\right) \tag{7}$$

$$sMAPE(\%) = \frac{100}{N}\sum_{i=1}^{N}\frac{\left|y_p^i - y_t^i\right|}{\left(\left|y_t^i\right| + \left|y_p^i\right|\right)/2} \tag{8}$$

Here, y_t^i and y_p^i represent the experimental and predicted values for the i-th observation, respectively; $\overline{y_t}$ denotes the mean of the experimental values, and N denotes the total number of observations. Higher R^2 values alongside minimized RMSE, MAE, APRE, MD, and sMAPE values signify superior model accuracy and predictive reliability (Chicco et al. 2021).

3 Results and discussion

3.1 Descriptive statistics and correlation analysis

Descriptive statistical analysis was performed on the raw dataset to gain preliminary insights into data characteristics before and after feature transformation and min–max scaling (Table S5). Fig. S4 further illustrates the distributions of all features across the raw, transformed, and scaled datasets, providing a comprehensive evaluation of the effects of preprocessing. The numerical features exhibited substantial variability. The specific surface area (SSA) ranged from 0.9 to 756 $m^2\ g^{-1}$, the average pore size (APS) from 0.0002 to 0.06 nm, and the total pore volume (TPV) from 0.002 to 2.23 $cm^3\ g^{-1}$. The pyrolysis temperature (PT) spanned 160–900 °C. While the Raman spectral intensity ratio (I_D/I_G) varied from 0.13 to 1.9. Persistent free radicals (PFRs) content ranged from 10^{16} to 10^{18} spins g^{-1}. The elemental compositions showed wide ranges: C (15–96%), H (0.19–13.5%), O (1.96–38.7%), and N (0.05–3%). Additional parameters included pH (3–12), oxidant concentration (C_{oxi}, 0–8 mg L^{-1}), biochar dosage (C_{bio}, 0.05–4 mg L^{-1}), and pollutant concentration (C_{pol}, 1–50 mg L^{-1}). The reaction rate constant (k) for antibiotic degradation varied between 0.0003 and 0.48. Categorical features included doping and graphitic structure (GS), each containing two categories.

The correlation heatmap (Fig. S5) was employed to examine the interrelationships among features. PCC revealed that total pore volume (TPV) and surface area (SSA) had the strongest positive correlation ($r=0.78$), highlighting the fundamental role of porosity in enhancing surface area, an essential attribute for catalytic performance (Muzyka et al. 2023). Substantial positive correlations were also observed between C_{pol} and C_{oxi} ($r=0.67$), and PFRs and pyrolysis temperature (PT) ($r=0.61$), indicating that thermal conditions may facilitate structural modifications conducive to PFR formation (Huang et al. 2019). The moderate correlation of TPV with H ($r=0.58$), with O with H ($r=0.52$), implies that hydrogen- and oxygen-containing functional groups contribute to pore development, with their co-degradation at elevated temperatures reflecting their thermally labile nature (Hawryluk-Sidoruk et al. 2024). Furthermore, the correlation between SSA and Doping ($r=0.51$) indicates that heteroatom doping may influence surface area properties (Tan et al. 2023). These results could be attributed to alterations in surface morphology and the reconstruction of internal pore architecture induced by the torrefaction process (Zhang et al. 2016). Conversely, a strong negative correlation between biochar concentration (C_{bio}) and nitrogen (N) ($r=-0.73$) suggests that higher biochar loading may enhance nitrogen volatilization or reduce nitrogen retention in the system. Similarly, the inverse relationship between pyrolysis temperature

(PT) and oxygen (O) content ($r = -0.45$) reinforces the expected loss of oxygenated functional groups at elevated thermal conditions (Yang et al. 2023). Collectively, these correlations underscore the coupled effects of material dosage and thermal processing on the structural and compositional transformation of biochar catalysts.

3.2 Development of predictive models

Various ML regression models, including TabPFN, RF, XGB, LGBM, SVR, and MLP, were developed and evaluated for their ability to predict the rate constant (k) of antibiotics. All models were trained on transformed and scaled datasets to ensure consistency and enhance predictive performance. The training process involved rigorous hyperparameter tuning to maximize accuracy and mitigate overfitting, as summarized in Table S6. Table 1 presents a comprehensive statistical summary of all the developed models during the training, testing, and overall evaluation phases. As shown, all the ML models exhibit a strong correlation between the predicted and experimentally observed values, indicating excellent predictive agreement. Among them, the TabPFN model demonstrates the highest accuracy, achieving an R^2 of 0.915 and 0.914 for the training and testing datasets, respectively. The model also exhibits low error metrics, with RMSE values of 0.020 (training) and 0.021 (testing), and MAE values of 0.008 and 0.009, further underscoring its robust performance. The APRE (%) values of −2.870 for training and 1.790 for testing further confirm the minimal systematic bias and strong generalization ability of the model. Figure 2 presents a comparative statistical evaluation of the algorithms using R^2, RMSE, MAE, and APRE %. The TabPFN model distinctly outperforms other algorithms, exhibiting the smallest error values and the highest correlation with experimental data. To further examine the consistency between the predicted and measured k values, a Taylor diagram was utilized (Fig. 3). In this diagram, models positioned closer to the reference point represent higher predictive accuracy, the TabPFN model is situated nearest to the observed data point for both stages, emphasizing its superior capability. In addition to statistical evaluation, graphical error analysis provides a visual assessment of performance, particularly when comparing multiple models. Several graphical techniques were employed to demonstrate the predictive reliability of the developed model. Figure 4 displays the cross-plots of predicted versus experimental k values for all ML models. The points are densely clustered around the X = Y line, reflecting high accuracy between predicted and actual values. Similarly, Fig. 5 illustrates the relative error distributions against the measured k values for the training and testing sets. The TabPFN data points align almost perfectly with the Y = 0 error line, indicating minimal relative deviation and robust predictive performance.

Comparatively, RF and XGB also demonstrated strong predictive performance relative to TabPFN, achieving R^2 scores of 0.94 (training) and 0.86 (testing) for RF, and 0.93 (training) and 0.84 (testing) for XGB, accompanied by relatively low RMSE values. In contrast, MLP and SVR yielded lower predictive accuracy ($R^2 < 0.81$ for training and < 0.79 testing) with higher RMSE (≥ 0.03), suggesting a reduced ability to generalize complex non-linear relationships within this relatively small and heterogeneous dataset (Jordan and Mitchell 2015; Padarian et al. 2019). While ensemble-based models, like RF, XGB, and LGBM, are generally considered robust under diverse data conditions due to their capacity to capture non-linear patterns while limiting overfitting (Zhu et al. 2019; Latif et al. 2024), they were not entirely immune to the challenges of

Table 1 Analytical review of the model outputs using statistical indicators

Model	TabPFN	RF	XGB	MLP	SVR	LGB
Train R^2	**0.915**	0.935	0.925	0.809	0.806	0.761
Test R^2	**0.914**	0.861	0.837	0.789	0.786	0.727
Train RMSE	**0.020**	0.018	0.011	0.030	0.031	0.035
Test RMSE	**0.021**	0.026	0.028	0.032	0.033	0.037
Train MAE	**0.008**	0.008	0.004	0.016	0.015	0.011
Test MAE	**0.009**	0.012	0.014	0.017	0.017	0.013
Train APRE (%)	**−2.870**	−4.410	−1.220	−8.390	−8.870	−6.290
Test APRE (%)	**1.790**	2.860	−4.740	1.970	1.040	7.550
Train MD	**−0.003**	−0.003	−0.001	−0.005	−0.005	−0.003
Test MD	**0.001**	0.001	−0.002	0.001	−0.002	0.004
Train sMAPE (%)	**10.314**	12.325	4.769	27.057	22.914	18.142
Test sMAPE (%)	**20.403**	20.381	22.375	28.542	29.637	24.647

Significant values are in bold

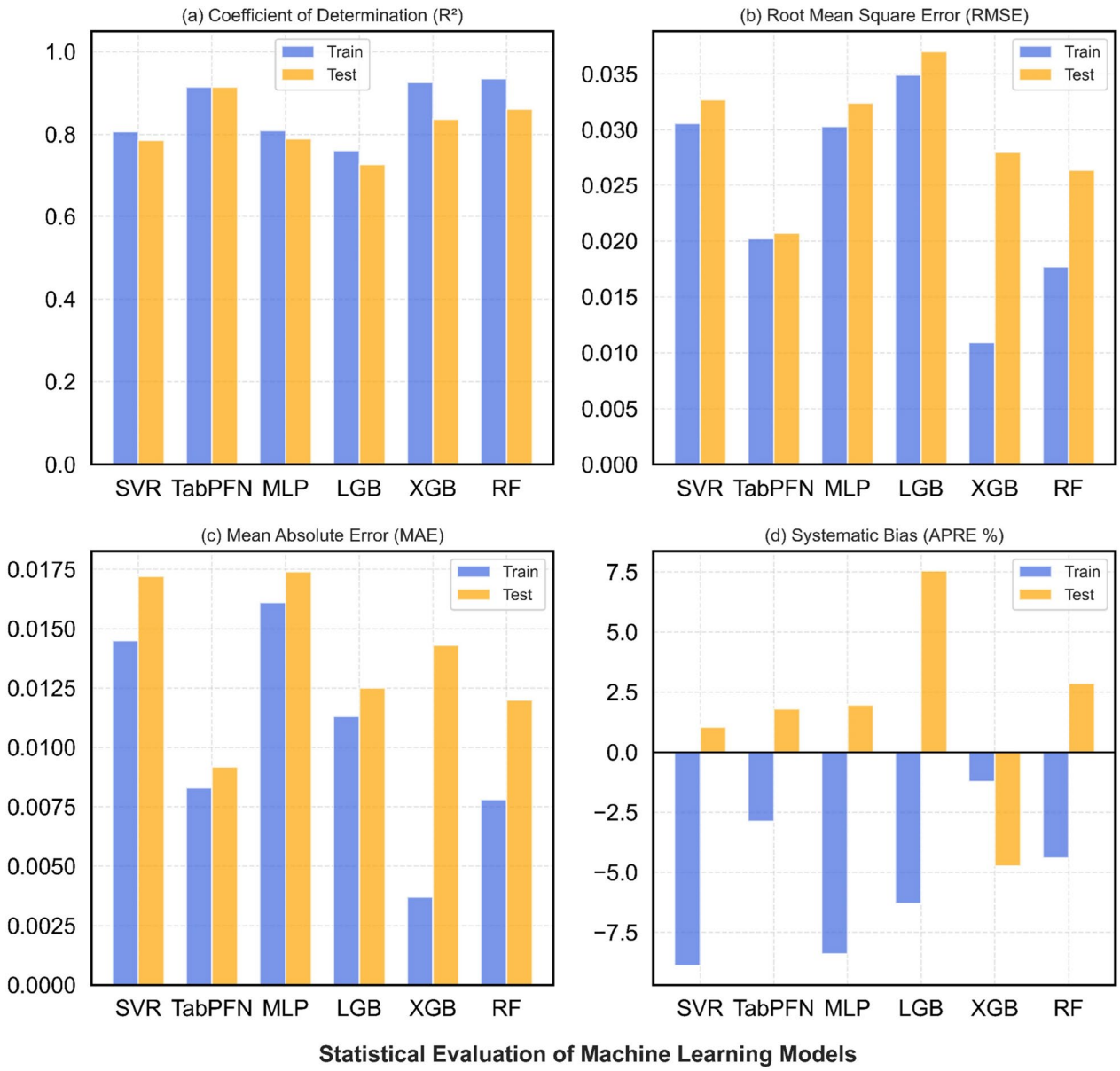


Fig. 2 Statistical evaluation of ML models

data sparsity and heterogeneity in this study. As indicated in Fig. 4 and Table 1, several of these models exhibited mild performance discrepancies between training and testing results. These gaps suggest a partial sensitivity to data variability despite the rigorous application of cross-validation and hyperparameter optimization. Notably, this observed overfitting was successfully mitigated by the TabPFN model, which demonstrated superior generalization and stability across both datasets, effectively bridging the performance gap seen in traditional ensemble methods.

TabPFN's superior performance likely stems from its foundation-model architecture, pre-trained on millions of synthetic datasets, allowing it to generalize effectively to small and heterogeneous tabular data without extensive fine-tuning (Hollmann et al. 2025). Its transformer-based framework efficiently models complex interdependencies, capturing non-linear relationships between catalyst properties, reaction parameters, and degradation reaction kinetics. These findings highlight TabPFN's strong potential for accurate and interpretable predictions in data-limited scientific applications,

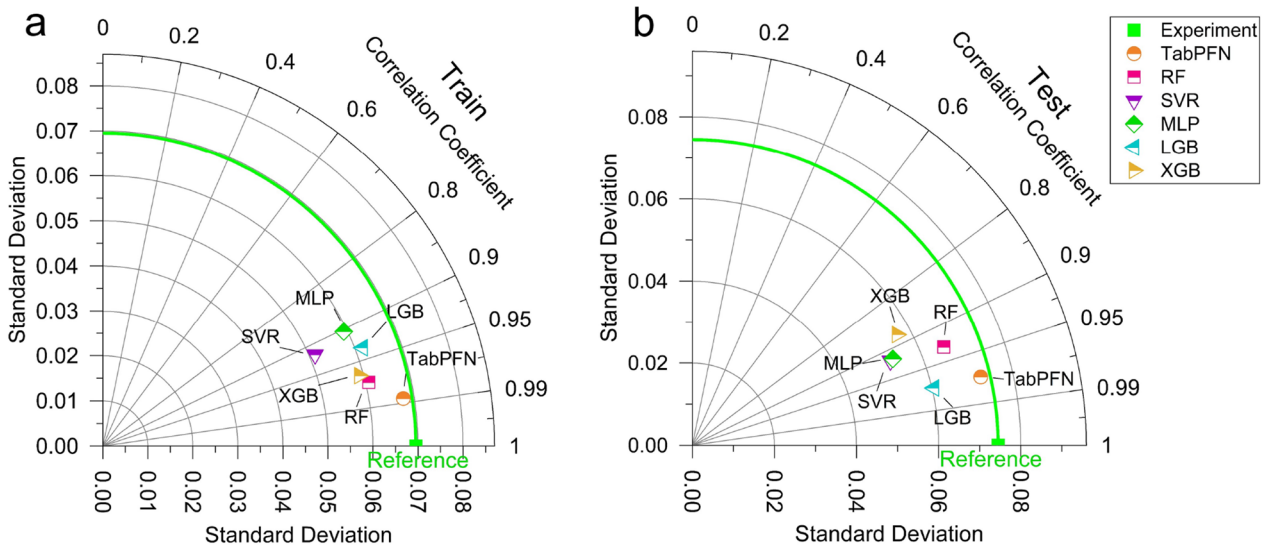


Fig. 3 Taylor diagram for **a** Training, and **b** Testing ML models

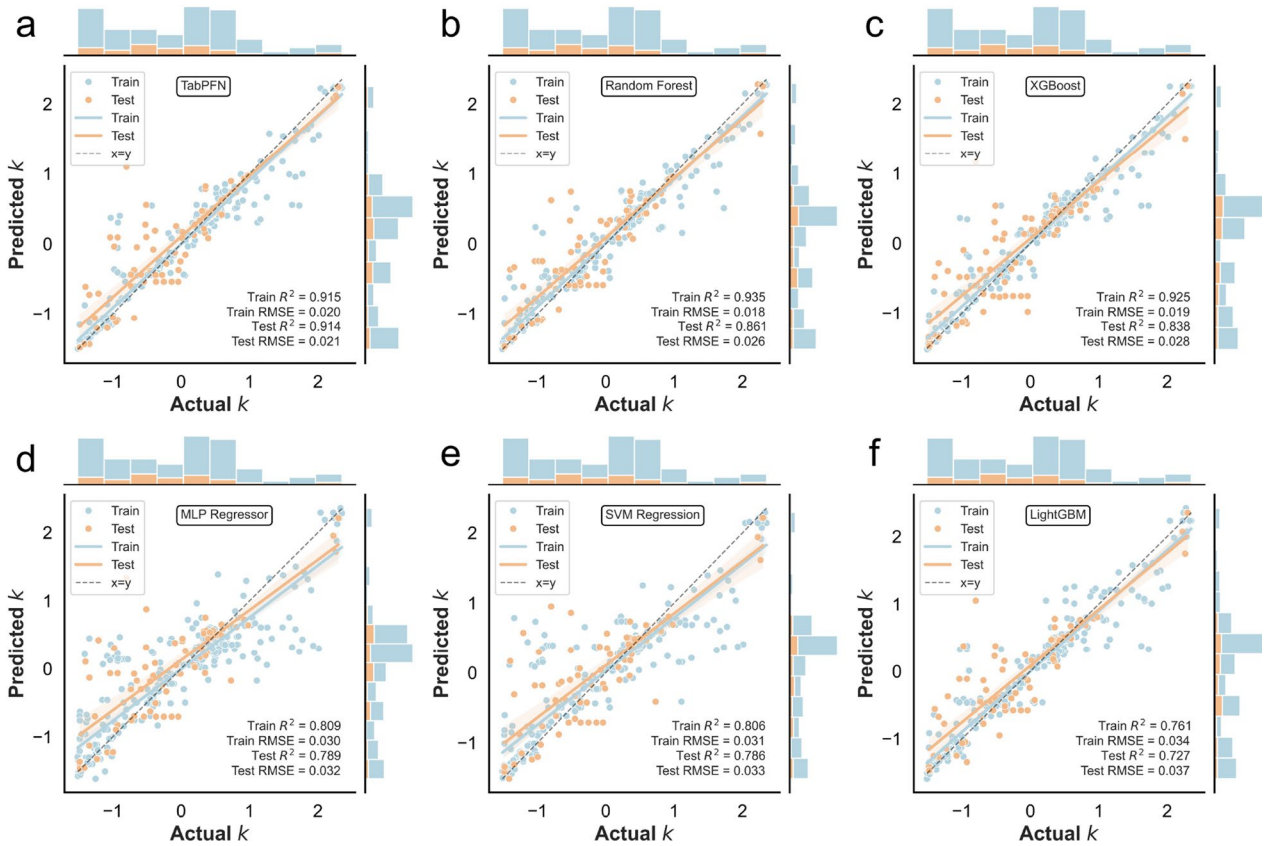


Fig. 4 Predicted vs. actual reaction rate constant (*k*) values for all developed models

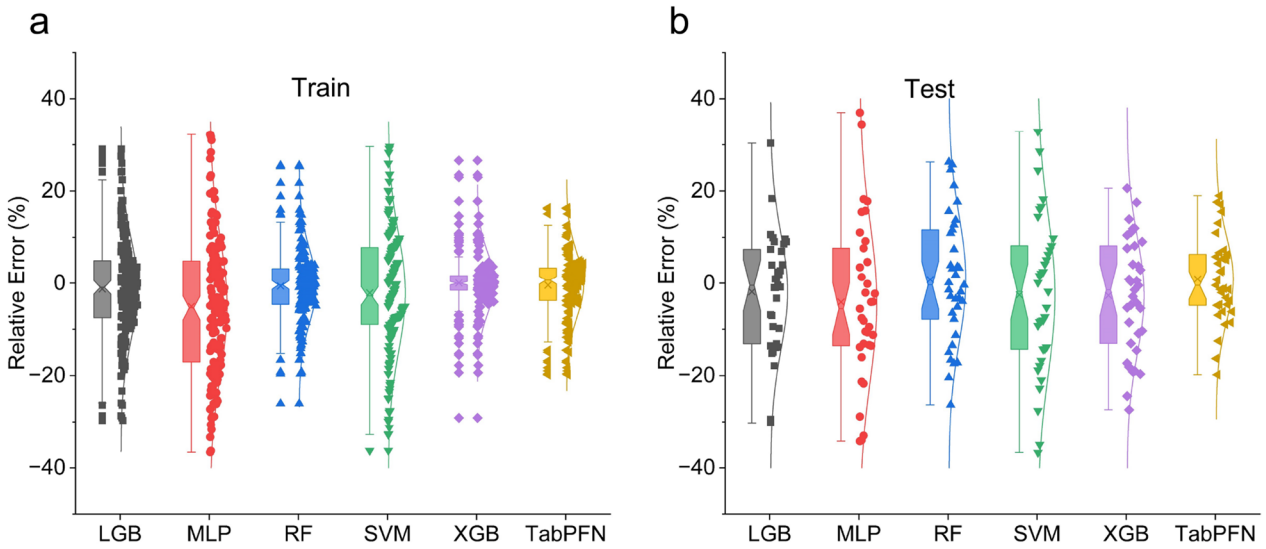


Fig. 5 Error distribution plot of the presented ML models to predict the reaction rate constant (*k*) values for **a** training and **b** testing

particularly for optimizing biochar-based catalytic systems in antibiotic degradation.

3.3 Interpretation and feature exploration

To elucidate the prediction mechanisms of the TabPFN model trained on the BCS dataset, we employed SHapley Additive exPlanations (SHAP), a model interpretability framework grounded in cooperative game theory. This approach quantified the individual contribution of each input variable to the target parameter *k*, providing both global and local interpretability. As illustrated in Fig. 6a, global feature importance was determined from the mean absolute SHAP values, revealing the most influential features to be PFRs, pore volume (TPV), oxidant concentration (C_{oxi}), pollutant concentration (C_{pol}), graphitic structure (GS), pore size (APS), biochar (C_{bio}), and the I_D/I_G ratio. These parameters collectively govern the catalytic activity and reaction dynamics of the system. Grouping the features into three functional categories revealed that catalyst properties contributed the largest share of the model's predictive capability (59.3%), followed by reaction conditions (25.9%) and elemental composition (14.8%). The bee swarm plot (Fig. 6b) visualizes the distribution and direction of feature effects across all samples. Features such as SSA, PT, pH, nitrogen, oxygen, carbon, hydrogen, and doping type exhibited relatively low mean SHAP values, suggesting a redundant influence on the target prediction within the context of this specific dataset. To maintain focus on the dominant drivers of reaction kinetics, only the highest-ranked features are discussed in detail in this section. The dependence patterns and marginal effects of the remaining, lower-importance variables are provided in the Supplementary Information (Figs. S6 and S7), which confirm their comparatively limited and less consistent influence on the predicted rate constant *k*. The SHAP-based heatmap (Fig. 6c) provides a holistic representation of feature contributions across samples, where color intensity reflects the magnitude and direction of each feature's impact on the predicted *k*. Features displaying consistently high-intensity vertical bands were identified as globally dominant, whereas those with dispersed or faint patterns indicated weaker or inconsistent contributions. Complementary to the SHAP analysis, permutation feature importance based on the mean decrease in R^2 (Fig. 6d) was conducted to further validate the feature ranking. The results showed that the largest decreases in R^2 occurred upon the removal of C_{bio} (0.1975), TPV (0.1616), C_{pol} (0.1411), and C_{oxi} (0.1340), emphasizing the dominant role of biochar dosages and reaction conditions in driving predictive performance. Moderate effects were observed for PFRs (0.0940), APS (0.0603), SSA (0.0573), and I_D/I_G (0.0389), while elemental composition features such as N (0.0246), O (0.0237), pH (0.0112), H (0.0097), and C (0.0048) exhibited comparatively minor influence. The lowest importance values were recorded for PT (0.0002) and Doping (−0.0022), indicating a negligible effect on model accuracy. Notably, C_{bio} exhibits a lower SHAP importance (ranked 7th) compared with its high permutation importance (ranked 1st), as shown in Fig. 6a and d, respectively. This discrepancy originates from the distinct principles used by the two methods to assess feature relevance. SHAP evaluates local marginal contributions to model predictions, which can lead to a reduced

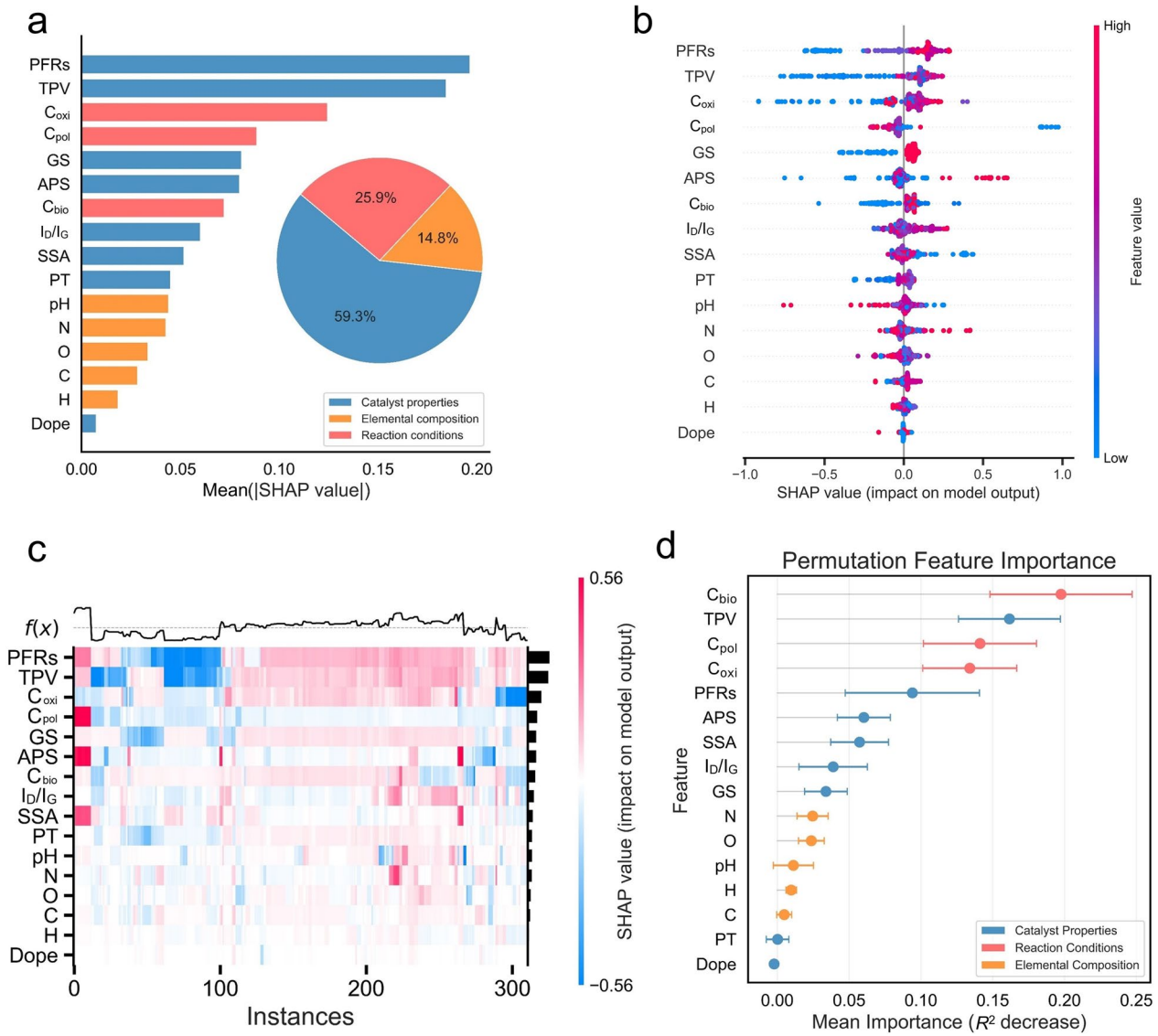


Fig. 6 Feature importance analysis using SHAP and permutation methods. **a** Global SHAP feature importance. **b** SHAP bee swarm plot showing individual feature effects. **c** SHAP heatmap illustrating feature contributions across samples. **d** Permutation importance highlighting key predictors influencing model performance

individual rank when a feature's predictive information is shared with correlated variables (Lundberg and Lee 2017). In contrast, permutation importance measures the global degradation in model performance when a feature is disrupted (Breiman 2001). The top ranking of C_{bio} in permutation importance indicates that, despite shared local influence, it is indispensable for maintaining the model's overall predictive stability.

To further investigate the relationships between features and the reaction rate, partial dependence plots (SPDPs) were generated. (Fig. 7) These plots illustrate the marginal effect of each feature on the predicted *k*, averaging out the influence of all others. These were accompanied by individual conditional expectation plots (IPDPs, Fig. 8), showing the variability of feature responses across samples.

Among all features, PFRs exerted the most pronounced influence on degradation, ranking first in importance (Fig. 6a). The corresponding SHAP dependence plot (Fig. 7a) demonstrates that higher PFR levels consistently enhance the predicted reaction rate constant *k*, indicating a strong positive contribution. This behavior is primarily attributed to the increased reactive oxygen species (ROS) generation, which directly promotes contaminant

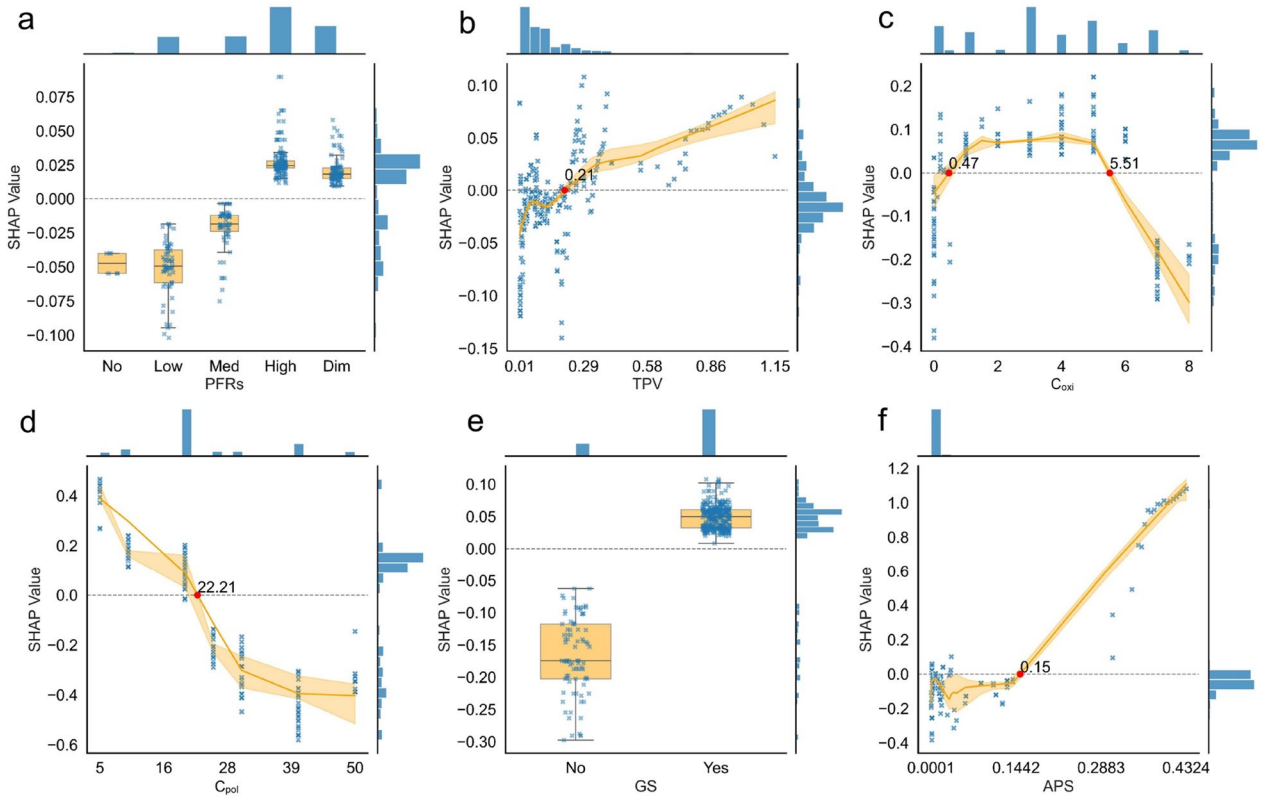


Fig. 7 SHAP partial dependence plots (SPDP) for representative variables. Each variable is plotted on a scatter plot and a box plot with whiskers of 1.5 times the interquartile ranges (**a**, **e**) or a regression line with an orange line of mean and the shade of SD (**b**, **c**, **d**, **f**). The distributions of SHAP and variable values are represented with a histogram on the right and top of each plot

degradation through radical-driven oxidation pathways. Previous studies have reported that PFR-rich biochar effectively activates oxidants and sustains continuous ROS production (Fang et al. 2014). Notably, at elevated pyrolysis temperatures (PT > 700 °C), PFRs abundance may decline due to radical recombination and loss of surface functional groups; however, degradation efficiency remains high. This indicates that catalytic performance is not solely governed by PFRs concentration. Instead, compensatory effects associated with high-temperature biochar formation collectively counteract the reduction in PFRs to sustain catalytic activity. These effects include enhanced aromatic condensation and electrical conductivity that facilitate electron transfer, the generation of defect-related active sites, and improved surface area and porosity that promote adsorption and mass transport (Ruan et al. 2019; Zhang et al. 2023). In contrast, low to moderate PFR levels exhibit negative contributions to *k*, which may result from insufficient radical density, limited redox cycling capability, or reduced ROS yield (Xie et al. 2024). Moreover, radicals formed at intermediate pyrolysis stages may possess greater structural stability but lower chemical reactivity, thereby limiting their catalytic effectiveness despite detectable PFR signals (Zhang et al. 2023). These trends are further corroborated by IPDP (Fig. 8a), which shows a monotonic increase in *k* with rising PFR content. Total pore volume (TPV) emerged as the second most impactful factor. Its influence was negligible or negative below a threshold of approximately 0.23 $cm^3\ g^{-1}$, becoming strongly positive above this value (Fig. 7b). This aligns with the role of porosity in enhancing adsorption capacity, oxidant transport, and catalytic activity (Li et al. 2020b). Figure 8b confirms a steady increase in *k* with increasing TPV, highlighting the importance of optimizing biochar's porosity through controlled pyrolysis to improve pollutant accessibility, oxidant diffusion, and electron transfer efficiency (Shan et al. 2024).

Oxidant (C_{oxi}) and pollutant (C_{pol}) concentrations ranked third and fourth in importance. Both exhibited non-linear relationships with *k*, emphasizing the need for precise control. For C_{oxi}, moderate concentrations (~0.5–5.5 mg L^{-1}) significantly enhanced degradation performance (Figs. 7c and 8c). This enhancement is attributed to the efficient generation of ROS, particularly sulfate ($SO_4^{\bullet-}$) and hydroxyl radicals ($^{\bullet}OH$), during

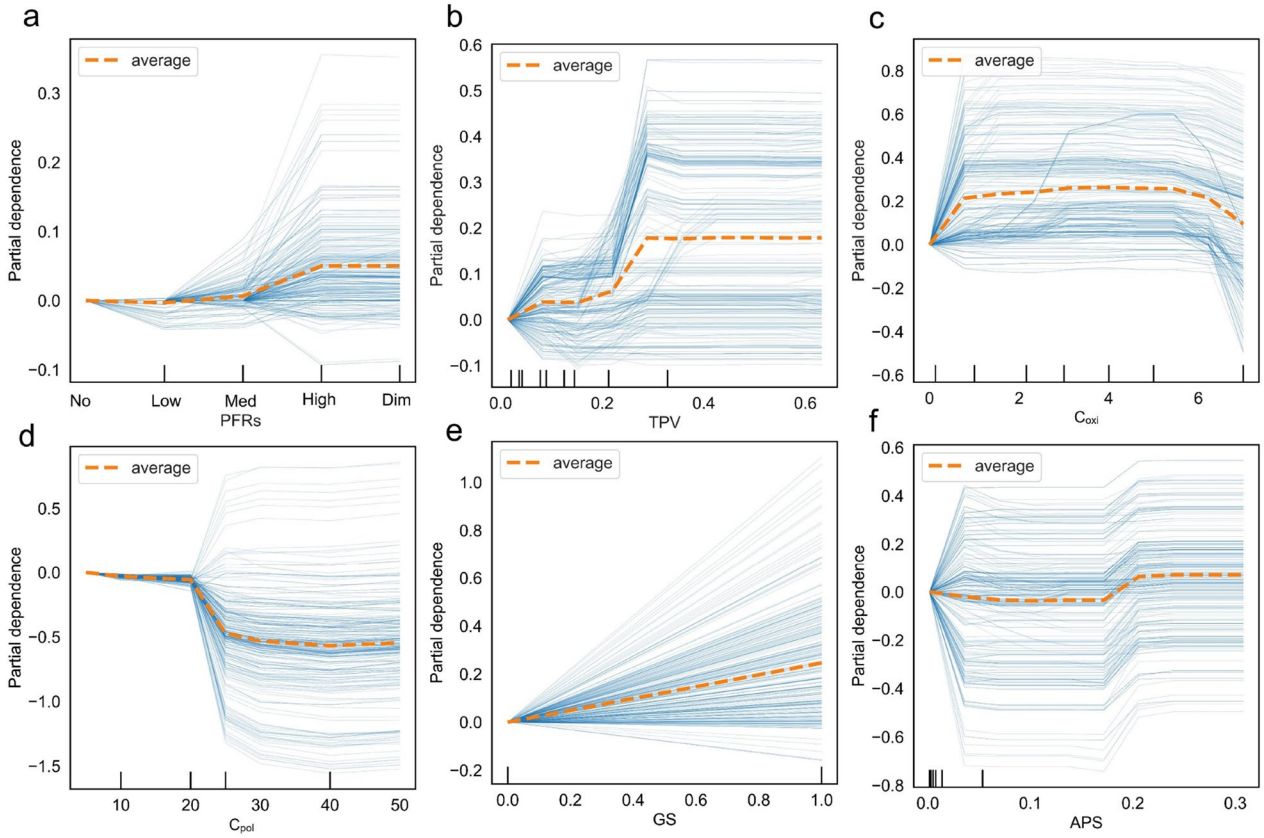


Fig. 8 Individual Conditional Expectation (IPDP) plots for the representative variables. Each line corresponds to an individual data point, highlighting potential varying effects within the dataset

biochar-mediated activation of oxidants such as persulfate or hydrogen peroxide. These radicals facilitate rapid and non-selective degradation of antibiotic molecules by abstracting hydrogen atoms and oxidizing functional groups, consistent with previous observations that optimal oxidant levels maximize ROS formation and catalytic efficiency (Li et al. 2020a). At higher oxidant concentrations, the degradation efficiency declined. This phenomenon is likely due to self-quenching or scavenging of radicals, wherein excess oxidant leads to non-productive radical–radical interactions that reduce the availability of active species for contaminant degradation (Klüpfel et al. 2014). Such behavior highlights the delicate balance required in oxidant dosing: sufficient to generate ROS, but not excessive to induce radical depletion. Pollutant concentration (C_{pol}) displayed an inverse relationship with *k*, revealing saturation effects at higher loads (Figs. 7d and 8d). At lower concentrations (< 22 mg L^{-1}), the degradation rate increased substantially, suggesting that active sites on the biochar surface remained readily available for oxidant interaction, promoting efficient ROS-mediated transformation of antibiotics. This observation is consistent with prior studies demonstrating that lower pollutant loads reduce competition for reactive sites and improve overall degradation efficiency (Zhu et al. 2020). As the pollutant concentration increased, similar to our results (~ 22–38.8 mg L^{-1}), the catalytic performance declined, due to surface site saturation and limited access of oxidants to target molecules, which reduces the generation and utilization of ROS (Li et al. 2021). Beyond ~ 38.8 mg L^{-1}, degradation plateaued, indicating that additional substrate molecules did not further increase ROS interactions. This observation is characteristic of equilibrium-limited kinetics, where surface saturation shifts the reaction from first-order to a zero-order-like kinetic regime.

Graphitic structure (GS) and average pore size (APS) were identified as the fifth and sixth most influential factors in predicting reaction kinetics (Fig. 6a). GS primarily facilitates electron transfer, which is essential for ROS generation and efficient catalysis (Tan et al. 2023). Our results indicate that the absence of GS resulted in negative contributions (median ≈ −0.2), whereas the presence of GS substantially increased the predicted *k*. This enhancement is attributed to the high electrical conductivity of GS, which promotes electron mobility across the

biochar surface, accelerating the activation of oxidants such as persulfate and hydrogen peroxide. Prior studies have similarly reported that biochar lacking GS exhibits reduced catalytic performance due to limited electron transfer and diminished ROS production (Zhang et al. 2015). APS contributes by influencing mass transport and accessibility of reactants to active sites. At low APS values (< 0.15 nm), the diffusion of pollutants and oxidants to active sites is restricted, limiting ROS-mediated degradation (Fig. 7f). As APS increases, larger pores enhance adsorption of antibiotics and facilitate oxidant diffusion, resulting in higher k values. This pattern is mirrored in the IPDP (Fig. 8f), confirming the progressive improvement in degradation efficiency with increasing APS. This mechanistic effect is supported by prior work demonstrating that increased pore size in biochar improves surface area, promotes pollutant loading, and enhances electron transfer, collectively boosting catalytic efficiency (Lu et al. 2017; Leng et al. 2021).

3.4 Development of GUI tool

A graphical user interface (GUI) web application was developed using Python (version 3.12) and the Flask web framework (version 3.0.3) to facilitate the practical application of the predictive model (Fig. S8). The GUI enables users to input parameters categorized into three empirical groups, including catalyst properties, elemental composition, and reaction conditions, to estimate the reaction rate constant (k) of antibiotics in a given system of biochar-based catalysts. Additionally, users may input an experimental k value for comparison with the model's prediction to assess accuracy. External datasets from training (Table S7) were employed for validation, where the model maintained prediction errors below 20% (Fig. S9), confirming its robust generalization capability. This tool provides a rapid and accessible means to evaluate new catalysts, reducing reliance on time-consuming quenching experiments and costly EPR analyses, thereby bridging predictive modeling with experimental catalyst design.

4 Conclusion

This study presents a data-driven predictive framework for the rate constant (k) of antibiotic degradation in biochar-based catalytic systems. Among all evaluated models, the transformer-based deep learning model TabPFN achieved the highest predictive accuracy ($R^2 = 0.92$ for training and 0.91 for testing; RMSE = 0.02–0.021), confirming its robustness in capturing complex non-linear relationships. Comprehensive evaluation using multiple statistical and graphical techniques validated the reliability and interpretability of the developed model. Among the sixteen examined features, PFRs, TPV, C_{oxi}, C_{pol}, GS, and APS were identified as the six most influential factors controlling k, with catalytic properties contributing 59.3% of the total importance, followed by reaction conditions (25.9%) and elemental composition (14.8%). To enhance accessibility and practical usability, a user-friendly web-based GUI was developed, allowing researchers to input experimental parameters and instantly predict k for new biochar catalysts.

Despite its strong predictive and interpretative performance, this study faced certain limitations primarily arising from the restricted size and heterogeneity of the compiled dataset, which was derived from diverse published studies. The limited data coverage restricts the model's ability to fully represent the variability of biochar properties, reaction environments, and their complex non-linear interactions. Consequently, some physicochemical parameters, such as C_{bio}, pH, and SSA exhibited lower feature importance (Fig. 6), which likely reflects insufficient data representation rather than a lack of mechanistic relevance. Nevertheless, partial dependence analyses (Figs. S6 and S7) revealed physically meaningful relationships: (i) increasing C_{bio} enhanced k through improved adsorption, electron transfer, and ROS generation (Hu et al. 2020); (ii) near-neutral pH (~ 7) maximized degradation efficiency, consistent with prior reports (Wang and Zhuan 2020); and (iii) higher SSA correlated with increased k, though this may be confounded by interdependence with PFRs and pyrolysis temperature. While PFRs were identified as the dominant drivers of catalytic oxidation, SSA and C_{bio} likely contributed via adsorptive removal mechanisms (Zhang et al. 2020).

Importantly, while the proposed model achieves strong predictive performance across multiple antibiotic families, it primarily captures generalized kinetic behaviors rather than detailed, class-specific degradation mechanisms. The incorporation of structurally distinct tetracyclines, fluoroquinolones, and sulfonamides improves robustness within these chemical domains. However, extrapolation to antibiotics with fundamentally different molecular structures should be approached with caution, as unrepresented reaction pathways or unique catalyst–substrate interactions may lead to increased predictive uncertainty. Future work should focus on expanding the chemical diversity of training datasets to enable the development of a more universally transferable, cross-category predictive framework.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s42773-026-00606-y.

Supplementary Material 1.

Acknowledgements
This study was supported by The National Natural Science Foundation of China (Grant Nos. 42307334 and 42277030), The "Qin Chuang Yuan" Innovation and Entrepreneurship Talent Program of Shaanxi Province (Grant No. QCYRCXM-2023-051), The Shaanxi Key Research and Development Program of China (Grant No. 2019ZDLNY01-02-01), The "One Hundred Talents" program of Shaanxi Province (Grant No. SXBR9171), The Shaanxi Science Fund for Distinguished Young Scholars (Grant No. 2019JC-18), The Open Fund of State Key Laboratory of Soil Erosion and Dryland farming on Loess Plateau (Grant No. A314021402-2021012) and The China Postdoctoral Science Foundation (Grant No. 2023M732875).

Author contributions
Junaid Latif: Data curation, formal analysis, writing – original draft, writing – review & editing. Na Chen: Funding acquisition, supervision, writing – review & editing. Jia Xie: Data curation, resources. Zheng Ni: Data curation, resources. Lang Zhu: Data curation, conceptualization, visualization. Azka Saleem: Data curation, software, visualization. Kai Li: Data curation, resources, methodology. Hanzhong Jia: funding acquisition, supervision, writing – review & editing.

Funding
The National Natural Science Foundation of China (Grant Nos. 42307334; 42277030); the "Qin Chuang Yuan" Innovation and Entrepreneurship Talent Program of Shaanxi Province (Grant No. QCYRCXM-2023-051); the Key Research and Development Projects of Shaanxi Province (Grant No. 2019ZDLNY01-02-01); the "One Hundred Talents" Program of Shaanxi Province (Grant No. SXBR9171); the Shaanxi Science Fund for Distinguished Young Scholars (Grant No. 2019JC-18); the Open Fund of State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau (Grant No. A314021402-2021012); and the China Postdoctoral Science Foundation (Grant No. 2023M732875).

Data availability
The datasets used and/or analyzed during the current study are provided in the Supporting Information. Additionally, the machine learning pipeline and GUI developed in this study are available as an open-source Python notebook on GitHub: https://github.com/junaid1990/antibiotic-k-prediction.

Declarations

Competing interests
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.

Author details
[1]State Key Laboratory of Soil and Water Conservation and Desertification Control, College of Natural Resources and Environment, Northwest A and F University, Yangling 712100, Shaanxi, People's Republic of China. [2]Key Laboratory of Low-Carbon Green Agriculture in Northwestern China, Ministry of Agriculture and Rural Affairs, Yangling 712100, Shaanxi, People's Republic of China. [3]College of Life Sciences, Northwest A and F University, Yangling 712100, Shaanxi, People's Republic of China.

Received: 4 November 2025 Revised: 12 February 2026 Accepted: 2 March 2026
Published online: 03 April 2026

References

Balaban NQ, Helaine S, Lewis K et al (2019) Definitions and guidelines for research on antibiotic persistence. Nat Rev Microbiol 17:441–448. https://doi.org/10.1038/s41579-019-0196-3
Breiman L (2001) Random Forests. Mach Learn 45:5–32
Brillas E (2022) A critical review on ibuprofen removal from synthetic waters, natural waters, and real wastewaters by advanced oxidation processes. Chemosphere 286:131849. https://doi.org/10.1016/j.chemosphere.2021.131849
Chicco D, Warrens MJ, Jurman G (2021) The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci 7:e623. https://doi.org/10.7717/peerj-cs.623
Fang G, Gao J, Liu C et al (2014) Key role of persistent free radicals in hydrogen peroxide activation by biochar: implications to organic contaminant degradation. Environ Sci Technol 48:1902–1910. https://doi.org/10.1021/es4048126
Hawryluk-Sidoruk M, Raczkiewicz M, Krasucka P et al (2024) Effect of biochar chemical modification (acid, base and hydrogen peroxide) on contaminants content depending on feedstock and pyrolysis conditions. Chem Eng J 481:148329. https://doi.org/10.1016/j.cej.2018.12.098
Hollmann N, Müller S, Purucker L et al (2025) Accurate predictions on small data with a tabular foundation model. Nature 637:319–326. https://doi.org/10.1038/s41586-024-08328-6
Hron K, Templ M, Filzmoser P (2010) Imputation of missing values for compositional data using classical and robust methods. Comput Stat Data Anal 54:3095–3107. https://doi.org/10.1016/j.csda.2009.11.023
Hu B, Ai Y, Jin J et al (2020) Efficient elimination of organic and inorganic pollutants by biochar and biochar-based materials. Biochar 2:47–64. https://doi.org/10.1007/s42773-020-00044-4
Huang D, Luo H, Zhang C et al (2019) Nonnegligible role of biomass types and its compositions on the formation of persistent free radicals in biochar: Insight into the influences on Fenton-like process. Chem Eng J 361:353–363. https://doi.org/10.1016/j.energy.2022.126128
Jaffari ZH, Jeong H, Shin J et al (2023) Machine-learning-based prediction and optimization of emerging contaminants' adsorption capacity on biochar materials. Chem Eng J 466:143073. https://doi.org/10.1016/j.cej.2023.143073
Jia H, Zhao S, Shi Y et al (2019) Formation of environmentally persistent free radicals during the transformation of anthracene in different soils: roles of soil characteristics and ambient conditions. J Hazard Mater 362:214–223. https://doi.org/10.1016/j.jhazmat.2018.08.056
Jiang S, Hou Y, Man Z et al (2024) Guiding experiment with machine learning: a case study of biochar adsorption of Ciprofloxacin. Sep Purif Technol 334:126023. https://doi.org/10.1016/j.seppur.2023.126023
Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. Science 349:255–260. https://doi.org/10.1126/science.aaa8415
Klüpfel L, Keiluweit M, Kleber M, Sander M (2014) Redox properties of plant biomass-derived black carbon (biochar). Environ Sci Technol 48:5601–5611. https://doi.org/10.1021/es500906d
Latif J, Chen N, Saleem A et al (2024) Machine learning for persistent free radicals in biochar: dual prediction of contents and types using regression and classification models. Carbon Res 3:39. https://doi.org/10.1007/s44246-024-00125-0
Leng L, Xiong Q, Yang L et al (2021) An overview on engineering the surface area and porosity of biochar. Sci Total Environ 763:144204. https://doi.org/10.1016/j.scitotenv.2020.144204
Li Y, Li J, Pan Y et al (2020a) Peroxymonosulfate activation on FeCo2S4 modified g-C3N4 (FeCo2S4-CN): mechanism of singlet oxygen evolution for nonradical efficient degradation of Sulfamethoxazole. Chem Eng J 384:123361. https://doi.org/10.1016/j.cej.2019.123361
Li Z, Sun Y, Yang Y et al (2020b) Biochar-supported nanoscale zero-valent iron as an efficient catalyst for organic degradation in groundwater. J Hazard Mater 383:121240. https://doi.org/10.1016/j.jhazmat.2019.121240
Li N, Li R, Duan X et al (2021) Correlation of active sites to generated reactive species and degradation routes of organics in peroxymonosulfate activation by Co-loaded carbon. Environ Sci Technol 55:16163–16174. https://doi.org/10.1021/acs.est.1c06244
Liang J, Xu X, Qamar Zaman W et al (2019) Different mechanisms between biochar and activated carbon for the persulfate catalytic degradation of sulfamethoxazole: roles of radicals in solution or solid phase. Chem Eng J 375:121908. https://doi.org/10.1016/j.cej.2019.121908
Liu G, Zhang Y, Yu H et al (2020) Acceleration of goethite-catalyzed Fenton-like oxidation of ofloxacin by biochar. J Hazard Mater 397:122783. https://doi.org/10.1016/j.jhazmat.2020.122783
Liu B, Xi F, Zhang H et al (2024) Coupling machine learning and theoretical models to compare key properties of biochar in adsorption kinetics rate and maximum adsorption capacity for emerging contaminants. Bioresour Technol 402:130776. https://doi.org/10.1016/j.biortech.2024.130776

Lu P, Huang Q, Chi Y, Yan J (2017) Preparation of high catalytic activity biochar from biomass waste for tar conversion. J Anal Appl Pyrolysis 127:47–56. https://doi.org/10.1016/j.jaap.2017.09.003
Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 30. https://arxiv.org/abs/1705.07874
Luo K, Yang Q, Pang Y et al (2019) Unveiling the mechanism of biochar-activated hydrogen peroxide on the degradation of ciprofloxacin. Chem Eng J 374:520–530. https://doi.org/10.1016/j.cej.2019.05.204
Luo X, Shen M, Liu J et al (2021) Resource utilization of piggery sludge to prepare recyclable magnetic biochar for highly efficient degradation of tetracycline through peroxymonosulfate activation. J Clean Prod 294:126372. https://doi.org/10.1016/j.jclepro.2021.126372
Muzyka R, Misztal E, Hrabak J et al (2023) Various biomass pyrolysis conditions influence the porosity and pore size distribution of biochar. Energy Oxf 263:126128. https://doi.org/10.1016/j.energy.2022.126128
Nathan C, Cars O (2014) Antibiotic resistance—problems, progress, and prospects. N Engl J Med 371:1761–1763. https://doi.org/10.1056/NEJMp1408040
Ohore OE, Addo FG, Zhang S et al (2019) Distribution and relationship between antimicrobial resistance genes and heavy metals in surface sediments of Taihu Lake, China. J Environ Sci China 77:323–335. https://doi.org/10.1016/j.jes.2018.09.004
Ouyang D, Chen Y, Yan J et al (2019) Activation mechanism of peroxymonosulfate by biochar for catalytic degradation of 1, 4-dioxane: important role of biochar defect structures. Chem Eng J 370:614–624. https://doi.org/10.1016/j.cej.2019.03.235
Padarian J, Minasny B, McBratney A (2019) Using deep learning to predict soil properties from regional spectral data. Geoderma Reg 16:e00198. https://doi.org/10.1016/j.geodrs.2018.e00198
Qin L, Yang L, Liu X et al (2021) Formation of environmentally persistent free radicals from thermochemical reactions of catechol. Sci Total Environ 772:145313. https://doi.org/10.1016/j.scitotenv.2021.145313
Ruan X, Sun Y, Du W et al (2019) Formation, characteristics, and applications of environmentally persistent free radicals in biochars: a review. Bioresour Technol 281:457–468. https://doi.org/10.1016/j.biortech.2019.02.105
Shan H, Wang W, Wang Z et al (2024) Unlocking the superior flexibility and enhanced catalytic oxidation performance of CoTiO3 nanofibrous membranes through zirconium doping. Sep Purif Technol 348:127714. https://doi.org/10.1016/j.seppur.2024.127714
Sun Z, Sun S, Wang S et al (2025) Prediction of uranium adsorption performance by machine learning for sustainable seawater extraction. Sci CHINA Chem. https://doi.org/10.1007/s11426-025-3233-y
Tan J, Chen X, Shang M et al (2023) N-doped biochar mediated peroxydisulfate activation for selective degradation of bisphenol A: The key role of potential difference-driven electron transfer mechanism. Chem Eng J 468:143476. https://doi.org/10.1016/j.cej.2023.148329
Wang J, Zhuan R (2020) Degradation of antibiotics by advanced oxidation processes: an overview. Sci Total Environ 701:135023. https://doi.org/10.1016/j.scitotenv.2019.135023
Wang R, Chen H, He Z et al (2024) Discovery of an end-to-end pattern for contaminant-oriented advanced oxidation processes catalyzed by biochar with explainable machine learning. Environ Sci Technol 58:16867–16876. https://doi.org/10.1021/acs.est.4c04714
Xie J, Latif J, Yang K et al (2024) A state-of-art review on the redox activity of persistent free radicals in biochar. Water Res 255:121516. https://doi.org/10.1016/j.watres.2024.121516
Xu L, Wu C, Liu P et al (2020) Peroxymonosulfate activation by nitrogen-doped biochar from sawdust for the efficient degradation of organic pollutants. Chem Eng J 387:124065. https://doi.org/10.1016/j.cej.2020.124065
Yang J, Zhang Z, Wang J et al (2023) Pyrolysis and hydrothermal carbonization of biowaste: a comparative review on the conversion pathways and potential applications of char product. Sustain Chem Pharm 33:101106. https://doi.org/10.1016/j.scp.2023.101106
Yi X, Lin C, Ong EJL et al (2019) Occurrence and distribution of trace levels of antibiotics in surface waters and soils driven by non-point source pollution and anthropogenic pressure. Chemosphere 216:213–223. https://doi.org/10.1016/j.chemosphere.2018.10.087
Yu J, Zhu Z, Zhang H et al (2020) Persistent free radicals on N-doped hydrochar for degradation of endocrine disrupting compounds. Chem Eng J 398:125538. https://doi.org/10.1016/j.cej.2020.125538
Yu T, Rajasekar A, Zhang S (2023) A decennial study of the trend of antibiotic studies in China. Environ Sci Pollut Res 30:121338–121353. https://doi.org/10.1007/s11356-023-30796-y
Zhang M, Bai X, Liu D et al (2015) Enhanced catalytic activity of potassium-doped graphitic carbon nitride induced by lower valence position. Appl Catal B Environ 164:77–81. https://doi.org/10.1016/j.apcatb.2014.09.020
Zhang S, Hu B, Zhang L, Xiong Y (2016) Effects of torrefaction on yield and quality of pyrolysis char and its application on preparation of activated carbon. J Anal Appl Pyrolysis 119:217–223. https://doi.org/10.1016/j.cej.2023.143476
Zhang Y, Yin M, Sun X, Zhao J (2020) Implication for adsorption and degradation of dyes by humic acid: light driven of environmentally persistent free radicals to activate reactive oxygen species. Bioresour Technol 307:123183. https://doi.org/10.1016/j.biortech.2020.123183
Zhang L, Cheng H, Pan D et al (2021a) One-pot pyrolysis of a typical invasive plant into nitrogen-doped biochars for efficient sorption of phthalate esters from aqueous solution. Chemosphere 280:130712. https://doi.org/10.1016/j.chemosphere.2021.130712
Zhang L, Zhang C, Lian K, Liu C (2021b) Effects of chronic exposure of antibiotics on microbial community structure and functions in hyporheic zone sediments. J Hazard Mater 416:126141. https://doi.org/10.1016/j.jhazmat.2021.126141
Zhang W, He Y, Li C et al (2021c) Persulfate activation using Co/AC particle electrodes and synergistic effects on humic acid degradation. Appl Catal B Environ 285:119848. https://doi.org/10.1016/j.apcatb.2020.119848
Zhang R, Zhang R, Zimmerman AR et al (2023) Applications, impacts, and management of biochar persistent free radicals: a review. Environ Pollut. https://doi.org/10.1016/j.envpol.2023.121543
Zhang P, Yang Y, Duan X, Wang S (2024) Oxidative polymerization versus degradation of organic pollutants in heterogeneous catalytic persulfate chemistry. Water Res 255:121485. https://doi.org/10.1016/j.watres.2024.121485
Zhong S, Zhang K, Bagheri M et al (2021) Machine learning: new ideas and tools in environmental science and engineering. Environ Sci Technol 55:12741–12754. https://doi.org/10.1021/acs.est.1c01339
Zhu X, Li Y, Wang X (2019) Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. Bioresour Technol 288:121527. https://doi.org/10.1016/j.biortech.2019.121527
Zhu G, Zhu J, Fu X et al (2020) Co nanoparticle-embedded N, O-codoped porous carbon nanospheres as an efficient peroxymonosulfate activator: singlet oxygen dominated catalytic degradation of organic pollutants. Phys Chem Chem Phys 22:15340–15353. https://doi.org/10.1039/D0CP00679C
Zhu X, Wan Z, Tsang DC et al (2021) Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption. Chem Eng J 406:126782. https://doi.org/10.1016/j.cej.2020.126782
Zou J, Yu J, Tang L et al (2020) Analysis of reaction pathways and catalytic sites on metal-free porous biochar for persulfate activation process. Chemosphere 261:127747. https://doi.org/10.1016/j.chemosphere.2020.127747