Wang *et al. Biochar* (2026) 8:101
https://doi.org/10.1007/s42773-026-00611-1

Biochar

ORIGINAL RESEARCH Open Access

Achieving precise regulation of soil phosphorus availability by guiding the application of pristine biochars with machine learning techniques

Yuqian Wang[1,2], Junhui Yin[1], Xiao Yang[3], Bangxi Zhang[4], Qing Chen[5], Yutao Peng[1*] and Jia Liu[1*]

Abstract

Biochar plays a crucial role in regulating soil phosphorus (P) availability, yet its effectiveness is influenced by multiple factors, including biochar features and soil properties. Improper biochar application may reduce P availability towards plants or unintendedly increase the environmental risk of P leaching. The efficiency of biochar in regulating soil P availability can be predicted and quantified by analyzing the interactions between its physicochemical properties and soil conditions. This study employed machine learning models—Random Forest, Support Vector Regression, and Artificial Neural Networks—to predict biochar efficiency in soil P availability regulation (activation or passivation) using a dataset of 534 samples with 19 input features. Model optimization and evaluation revealed that the Random Forest model achieved the highest prediction accuracy ($R^2 = 0.9107$), outperforming the other two models. Mechanistic insights from feature importance analysis indicated that biochar pyrolysis temperature played a dominant role in influencing soil P availability. Moderate pyrolysis temperatures facilitated the formation of biochar with balanced porosity and surface reactivity, while biochar produced at higher temperatures favored for passivating soil availability. Furthermore, the biochar application rate, soil pH, and total soil P content are key factors influencing changes in soil available P following biochar amendment. Through a data-driven framework, this study demonstrated that pristine biochar could achieve or exceed the performance of modified biochar in P regulation, offering superior economic and environmental benefits. The findings integrated environmental science, soil chemistry, and data analytics, providing valuable guidance for precision agriculture and fostering sustainable agricultural practices by enhancing fertilizer efficiency and reducing environmental costs globally.

Highlights

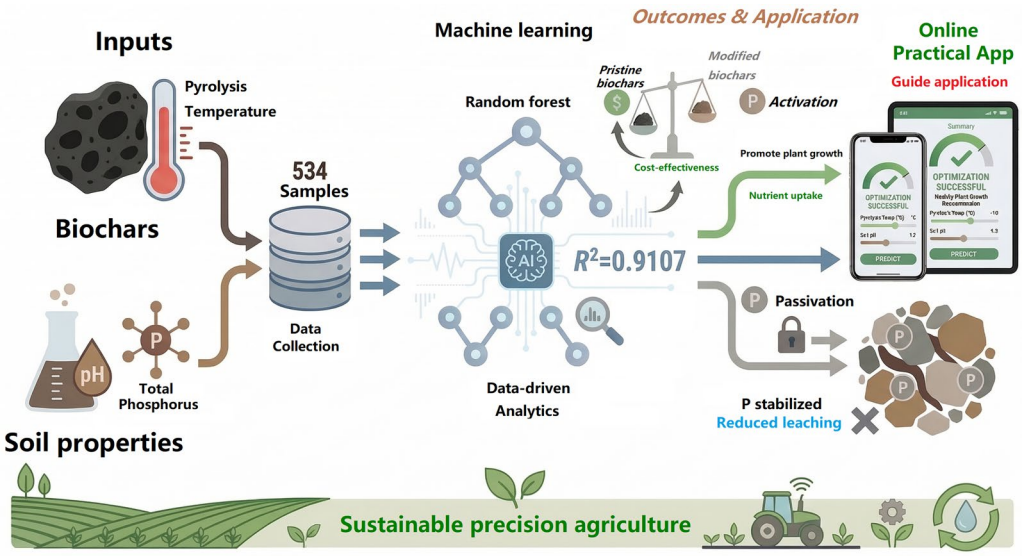
- Machine learning was integrated into frameworks to improve fertilizer efficiency and reducing environmental impact.
- The RF model exhibited best predictive performance ($R^2 = 0.9107$) compared to the SVR and ANN models.
- Pyrolysis temperature, biochar rate, soil pH, and total P significantly affect biochar's influence.

*Correspondence:
Yutao Peng
pengyt39@mail.sysu.edu.cn
Jia Liu
liujia67@mail.sysu.edu.cn
Full list of author information is available at the end of the article

Springer

Keywords Biochar, Soil phosphorus availability, Feature analysis, Machine learning, Performance prediction

Graphical Abstract



1 Introduction

Phosphorus (P) is a vital nutrient for plant growth. A deficiency of P can result in restricted plant growth and a reduction in agricultural yields (Etesami 2020). It is common practice among farmers to apply large quantities of P fertilizer to maintain the requisite P content of the crop. However, only 15% to 20% of the applied P is absorbed by the crop (An et al. 2023). The rest combines with the soil, forming a residual P pool or is fixed in the soil (Bindraban et al. 2020). Soil P exists in multiple chemical forms, among which only Ca-bound P and part of Al-bound P can be effectively utilized by crops, whereas Fe-bound and other complex forms are largely unavailable (Stávková and Maroušek 2021). Residual P in the soil is highly reactive and has the potential to be transported via surface runoff and leaching to groundwater, which can result in surface water pollution and lead to eutrophication (Ghodszad et al. 2021). Conversely, fixed P in the soil is less active and requires activation to enhance its efficacy in promoting the absorption of P by crops. Moreover, P fertilizers are primarily derived from natural phosphate rock, which takes millions of years to form through the decomposition of marine organisms (Alewell et al. 2020; De Boer et al. 2019). Some phosphate rock deposits may contain trace levels of radioactive contaminants, raising additional environmental and health concerns (Wang et al. 2023). Therefore, accurate prediction of soil P availability increase ratio (PAR, defined as the relative increase in plant-available P content after biochar amendment) has always been an urgent problem in soil P chemistry.

Biochar has been widely used to regulate soil P and interact with the P cycle. However, the effect of biochar application on the activity of soil P remains inconclusive. Studies have shown that the adsorption of phosphate by biochar renders it a P reservoir within the soil. Conversely, P-containing biochar can also be employed as a direct and efficacious P source, thus enhancing the soil P availability through direct input. (Zhang et al. 2016). Presently, research has utilized biochar in the production of slow-release P fertilizers to enhance P utilization efficiency in soils and mitigate P losses (Li et al. 2020a). The precise mechanism by which biochar alters the availability of P in soil is still under investigation. The efficiency of P fertilizer use is contingent upon a range of soil properties, including the ionic composition of the soil, its pH, and the quantity of organic matter present (Bornø et al. 2018). Biochar application can increase P availability by increasing soil pH and altering the activity or availability of cations, thereby reducing P adsorption or increasing P desorption in the soil (Xu et al. 2014). Other studies have also supported this conclusion. The use of biochar in acidic red soil has been evidenced to elevate the soil pH and

rapidly increase the soil available P content in the short term, and the time trend of this effect is related to the decomposition of biochar itself (Jin et al. 2019). Straw biochar has the characteristics of high negative charge and rich anionic functional groups, which can enhance the cation exchange capacity (CEC) of the soil. The anionic functional groups can compete with phosphate for adsorption sites in the soil, thereby inhibiting the adsorption of phosphate by the soil (Jiang et al. 2015). Furthermore, biochar can also release dissolved organic matter (DOM). The organic anions in the DOM components compete with the phosphate ions in the soil for exchange sites, effectively inhibiting the adsorption of P in acidic soils such as red soil and improving the efficiency of P fertilizer (Liu et al. 2018; Schneider And Haderlein 2016). Biochar may also affect the biological processes of soil microorganisms, converting fixed metal oxides or organic P into available P (Gao and DeLuca 2018). However, in consideration of the potential environmental risks, some studies have shown that biochar can also be employed to reduce the amount of available P in the soil. This is due to the high specific surface area and high concentration of basic cations in biochar produced by high-temperature pyrolysis, which increases the number of adsorption sites (Ghodszad et al. 2021). Additionally, some studies report that biochar application does not significantly affect soil P availability (Soinne et al. 2014; Bornø et al. 2018). This implies that the impact of biochar on soil P availability might be influenced by various factors. Biochar can affect the physical and chemical features of soil, and its impact on soil P availability depends not only on soil characteristics but also on the properties of the biochar itself. Consequently, to accurately ascertain the impact of biochar on soil P availability following its application, it is essential to quantify the relationship between biochar and the physical and chemical properties of the soil, as well as soil P availability. Moreover, for P-biochar technologies to be truly feasible, they must achieve both environmental sustainability and economic viability, particularly considering the production and application costs of biochar (Maroušek et al. 2023, 2024). Therefore, an integrated analytical approach is required to elucidate the interrelationships among multiple factors, aiming to maximize both ecological and economic benefits.

Previous studies have typically controlled for a single variable to investigate the effect of biochar on soil P availability, which is time-consuming, costly and difficult to fully account for all environmental factors (Gao et al. 2019). Traditional statistical methods such as meta-analysis can integrate the results of multiple studies and have shown that biochar could significantly increase plant-available P in biochar-amended soils (Glaser and Lehr 2019). However, due to the large number of parameters corresponding to biochar and the small sample size, narrow width, and high dimensionality of the data, meta-analysis is difficult to resolve the complex relationship between small samples and high-dimensional data (Wetterslev et al. 2017). Machine learning, as a prominent and emerging artificial intelligence method, is capable of handling high-dimensional and complex data, uncovering non-linear relationships that are difficult to solve with traditional methods, and providing rapid and effective prediction results (Li et al. 2022). Machine learning models such as Random Forest (RF), Support Vector Regression (SVR), and Artificial Neural Networks (ANN), which are commonly used on small datasets (Kokol et al. 2022), have been widely used to reveal the effects of biochar on soil nutrient cycling (Liu et al. 2019), soil microbial communities (Lei et al. 2023), and soil pollutants such as heavy metals (Sun et al. 2022). Palansooriya et al. (2022) utilized RF, SVR, and Neural Networks models to investigate the efficiency of biochar in immobilizing heavy metals in soil. Their findings highlighted that the RF model with the best tuned hyperparameters was the best algorithm for predicting heavy metal immobilization efficiency (test set $R^2 = 0.91$). Divband Hafshejani et al. (2024) employed RF and linear regression models to evaluate the influence of biochar on soil nitrate leaching. Their results showed that the RF model exhibited a superior fit ($R^2 = 0.9837$). However, research on applying machine learning to predict soil PAR with biochar remains largely unexplored. The predictive performance of different models in analyzing biochar and soil P availability data requires evaluation, while comprehensive datasets integrating biochar characteristics, soil properties, and available P dynamics are still lacking. Herein, we hypothesized that machine learning models could effectively predict the efficiency of biochar in regulating soil phosphorus availability by analyzing the interactions between biochar's physicochemical properties and soil conditions, thus optimizing biochar application for enhanced precision agriculture and sustainable soil management.

Therefore, to confirm our hypothesis, this study (i) collected a total of 534 data points on the efficiency of biochar in soil PAR from 32 research articles and examined the correlation between different characteristic variables and soil P availability; (ii) developed three machine learning models (RF, SVR, and ANN) for predicting the efficiency of biochar in soil PAR and compared their predictive performance; (iii) identified the main factors influencing soil PAR by biochar based on the optimized machine learning models; (iv) explored the interaction between input factors and soil PAR by biochar. The

model developed in this study can be used to predict the efficiency of biochar in regulating soil P availability based on biochar and soil characteristics. This offers guidance for applying biochar-enhanced P fertilizers in practical agricultural settings.

2 Materials and methods

2.1 Data collection

Data from relevant research articles published between 2014 and 2024 on the use of biochar to regulate soil P were collected. The article data used in this study were collected by searching for 'soil phosphorus and biochar' in Google Scholar, Web of Science, and China National Knowledge Infrastructure. Data were obtained directly from the article or supplementary material tables, or from graphs using Plot Digitizer 1.3. After an initial data screening, a total of 534 datasets containing soil available P content before and after biochar application were selected. The growth rate of soil P availability with biochar was calculated using Eq. 1 (Xu et al. 2024):

$$Ratio = \frac{C_{After} - C_{Before}}{C_{Before}} \tag{1}$$

where C_{Before} and C_{After} represent the soil available P concentrations measured before and after biochar application, respectively. Specifically, a ratio greater than 1 indicated that biochar enhanced P availability, whereas a ratio less than 1 suggested that biochar immobilized soil P, leading to passivation effects. To simulate the efficiency of biochar in soil PAR, a total of 22 factors were selected for model building and optimization: (i) Biochar characteristics, including pH value of biochar (pHBC), ash content (Ash, %), carbon content (CBC, %), electrical conductivity (EC, dS m^{-1}), nitrogen content (NBC, %), cation exchange capacity (CEC, $cmol_{(+)}$ kg^{-1}), phosphorus content (PBC, %), Olsen-P content (OPBC, mg kg^{-1}), specific surface area (SSA, m^2 g^{-1}); (ii) Soil properties, including soil pH (pHsoil), total phosphorus content (TP, g kg^{-1}), Olsen-P content (OPsoil, mg kg^{-1}), total nitrogen content (TN, %), total carbon content (TC, %), organic carbon content (OC, g kg^{-1}), soil texture, electrical conductivity (ECsoil, μS cm^{-1}); (iii) Experimental conditions, including biochar pyrolysis temperature (T, ℃), biochar pyrolysis time (Time, h), biochar application rate (Add, %), fertilizer type, incubation time. A summary of this data can be found in the supplementary materials.

2.2 Data preprocessing

Collinearity among input features can lead to unstable parameter estimates, unreliable models, and weakened predictive performance. By reducing collinearity between features, the model can be simplified, its interpretability enhanced, and the risk of overfitting reduced (Cheng et al. 2022). Therefore, a collinearity diagnosis based on the Variance Inflation Factor (VIF) was first performed on the 22 input features. Features with VIF > 20, specifically the pH and EC of biochar, and TN content of soil, were excluded. The formula for VIF is shown as Eq. 2 (Vu et al. 2015).

$$\mathrm{VIF}_j = \frac{1}{1 - R_j^2} \tag{2}$$

where R_j^2 represents the value obtained when feature X_j is treated as the dependent variable and regressed on all other features, reflecting the proportion of variance in X_j that can be explained by the other features.

Data preprocessing mainly includes feature coding, missing data handling, and data normalization. First, the categorical features of soil texture and fertilizer type were encoded and backfilled into the corresponding positions in the dataset. For missing data, the average value of the calculated feature was calculated and backfilled into the corresponding position in the dataset. This is a simple interpolation method and most studies use this type of interpolation method, replacing each missing value with the non-missing value of the quantitative or qualitative attribute (Emmanuel et al. 2021).

Data collected from different sources using different methods may contain outliers, and data normalization can mitigate the influence of outliers and feature dominance by achieving uniformity in numerical values (Singh and Singh 2022). The utilization of Eq. 3 facilitated the scaling of features within the range of 0–1:

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{3}$$

where x_i^* represents the normalized value of x_i, while x_{max} and x_{min} denote the maximum and minimum values of x_i, respectively.

2.3 Model development and evaluation

This study selected three machine learning methods: RF, SVR, and ANN. The prediction models for biochar's efficiency in soil PAR were trained and optimized using the Scikit-learn 1.2.1 and PyTorch 2.0.1 + cpu. Before model construction, the entire dataset was randomly divided into 80% training and 20% testing subsets. For hyperparameter optimization, RF required tuning the number of trees (n_estimators) and tree depth (max_depth). During the grid search process, five-fold cross-validation was used to evaluate the performance of each parameter combination to identify the optimal hyperparameters. After determining these two parameters, the maximum number of features used for node splitting (max_features) in each decision tree was further optimized. SVR

required adjustment of the kernel function, the epsilon (ε), and the regularization parameter (C). ANN required the learning rate, number of neurons and hidden layers to be adjusted.

After hyperparameters tuning, the three optimal models were selected and evaluated on the test set. The performance of the models was assessed using the coefficient of determination (R^2) and the root mean square error (RMSE), respectively. R^2 is a metric that measures the degree to which the model fits the data, with values ranging from 0 to 1; a value close to 1 indicates a stronger fit of the model to the data. RMSE is an indicator of the prediction error of the model, and its value determination is closely related to the specific dataset. In general, a smaller RMSE value indicates a smaller prediction error and better prediction performance of the model. The formulas for R^2 and RMSE are given by (Eqs. 4 and 5) respectively (Palansooriya et al. 2022):

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_t^i - y_p^i\right)^2}{N}} \tag{5}$$

where y_p^i and y_t^i represent the predicted and true values, y_m denotes the mean value, and N is the number of data samples in the dataset.

2.4 Influence factor analysis

To further explore the relationship between each input feature and the P growth rate, the Pearson correlation coefficient (PCC) was first used to assess the correlation between any two variables. The formula for PCC is as follows (Eq. 6):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{6}$$

where n is the sample size, x_i and y_i are the values of the two variables for the ith sample respectively, and $\bar{x}$ and $\bar{y}$ are the means of the two variables respectively.

Subsequently, two approaches were used to analyze the significance of input features in the dataset. Since the RF model itself has an algorithm for calculating variable importance measures, the feature importance analysis was first performed based on the feature importance ranking function in the RF model. SHapley Additive exPlanations (SHAP) is a game theory-based model interpretation method that calculates the marginal contribution of each feature across all possible feature combinations, providing consistent and locally accurate estimates of feature importance (Antwarg et al. 2021). To enhance the comprehensiveness and reliability of model interpretation, we further employed the SHAP method to assess the importance of each input feature in relation to the P growth rate. All computations and visualizations were conducted using IBM SPSS Statistics 27 and Python 3.7 with the following packages: Matplotlib 3.7.0, NumPy 1.23.5, Pandas 1.5.3, Seaborn 0.12.2, Scikit-learn 1.2.1, Keras 3.0.5, PyTorch 2.0.1+cpu, and SHAP 0.42.1. To ensure reproducibility, the random seed was fixed at 42 during dataset partitioning and model training. All quantitative variables were standardized to consistent units prior to analysis.

3 Results and discussion

3.1 Feature statistics and Pearson correlation matrix analysis

In recent years, biochar and biochar-based fertilizers have emerged as promising strategies for improving nutrient cycling and sustainability in agricultural systems (Minofar et al. 2025; Maroušek and Gavurová 2022). Improper biochar application may reduce P availability through nutrient immobilization and microbial community disruption, while exacerbating environmental risks via accelerated P leaching and mineralization-induced eutrophication. Moreover, the regulatory effect of biochar on soil P availability is influenced by multiple interacting factors and does not follow a simple linear relationship, exhibiting nonlinear and complex dynamic patterns that traditional statistical methods cannot fully capture. Therefore, we first employed Pearson correlation matrix analysis to examine the relationships among variables and subsequently applied machine learning models to capture the complex nonlinear patterns. Herein, in order to elucidate the potential interactions and dependencies among variables, the Pearson correlation matrix results are presented in Fig. 1. In addition, to provide a more comprehensive understanding of the original dataset, we selected several relatively important features based on the Pearson correlation matrix and subsequent feature importance analysis and visualized them using boxplots (Fig. 2). The box plot results implied that the pyrolysis temperature of biochar in the dataset of this study falls within the range of 200–700 ℃. The correlation coefficient between the efficiency of biochar in soil PAR and pyrolysis temperature is −0.18, revealing a weak negative correlation between these two variables, because high pyrolysis temperature produced biochar with high porosity and specific surface area—also consistent with the results shown in Fig. 1 (the Pearson correlation

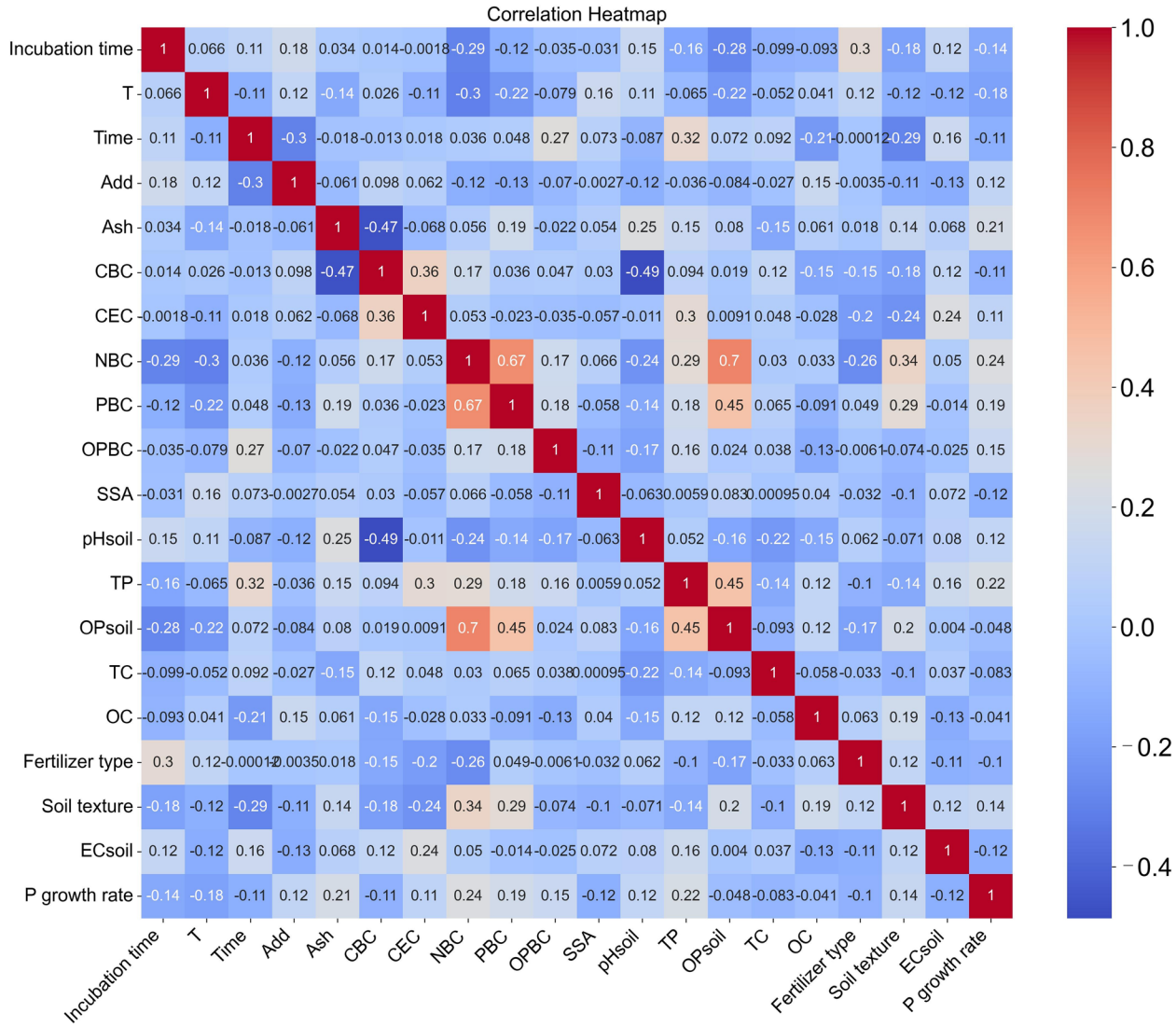


Fig. 1 The Pearson correlation matrix between two of the included features reveals the interrelationships among variables

coefficient between T and SSA is 0.16, indicating a positive correlation), which consequently increased its P adsorption capacity and caused a reduction in the effective P level within the soil (Ghodszad et al. 2021; Li et al. 2020a). This weak negative correlation suggests a potential trade-off, as biochar produced at relatively high pyrolysis temperatures may improve physical properties while simultaneously limiting P availability. Furthermore, the distinct pyrolysis temperatures resulted in disparate P forms and extractability in the biochar. The formation of stable P forms, such as pyrophosphate, is promoted by high temperatures. This resulted in a decrease in the biochar P releasing rate and P adsorption capacity of biochar (Li et al. 2019), which in turn led to a reduction in the efficiency of biochar in soil PAR. In the dataset of this study, the pH of the biochar is predominantly alkaline. As previous studies have shown, strongly alkaline biochar contains higher concentrations of alkaline cations, which can significantly increase soil pH. This process leads to the precipitation of soluble Al and Fe as insoluble Al and Fe hydroxides (Dume et al. 2017; Gao et al. 2019), which reduces the binding of P with Al and Fe, thereby decreasing P fixation in the soil and enhancing soil P availability. Additionally, the boxplots show that the pyrolysis duration of biochar in this study is mostly concentrated within 1–3 h. The Pearson correlation matrix revealed a weak negative correlation ($r = -0.18$) between the pyrolysis time of biochar and its efficacy in soil PAR. Previous studies have shown that biochar produced via fast pyrolysis typically exhibits higher CEC than that produced by

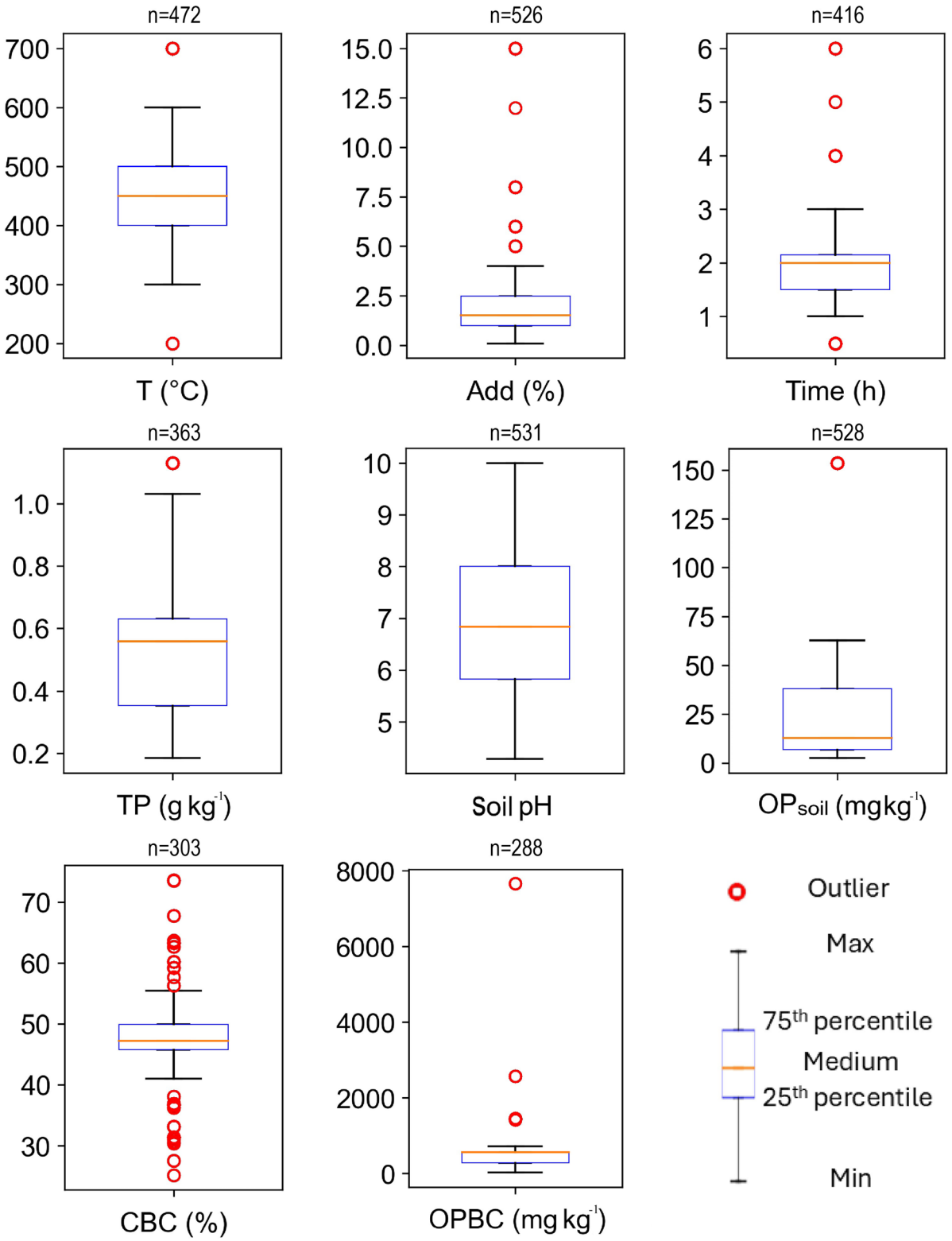


Fig. 2 Distribution of partial input features. In each boxplot, the five lines from top to bottom represent the maximum value, the third quartile, the median, the first quartile, and the minimum value of the data, respectively. Red circles denote outliers beyond this range, and n represents the sample size

slow pyrolysis, due to its faster heating rate and shorter reaction time (Lee et al. 2010), which enhanced the efficiency of biochar in soil PAR by reducing the fixation of P in the soil.

In this study, soil pH varies widely (4.29–10), covering weakly acidic red soil, paddy soil, and weakly alkaline fluvo-aquic soil. Soil pH plays a crucial role in influencing the effectiveness of biochar in enhancing soil P availability. Specifically, in acidic red soils, P tends to bind with Fe^{3+} and Al^{3+} to form insoluble compounds that limit its availability (Melese et al. 2015), biochar application can promote P release by adjusting pH and enhancing complexation capacity. In contrast, in alkaline and calcareous soils, P binds with $CaCO_3$ to form insoluble Ca–P compounds (Antoniadis et al. 2016); biochar can reduce P fixation through adsorption and complexation, thereby improving its availability. The total P content of the soil exhibited a range of 0.185–1.13 g kg^{-1}, while the Olsen-P content demonstrated a range of 2.48–153.52 mg kg^{-1}. The total P content in the soil was found to be positively correlated with the efficiency of biochar in soil PAR, with a correlation coefficient of 0.22. The above phenomenon may be attributed to the high content of total P in soil which provided more space for the biochar synergistic action of the physical, chemical and biological regulatory mechanisms. In alkaline soil, biochar application does not necessarily increase P availability; however, when biochar is derived from crop residues, P adsorption in alkaline soil slightly decreased with increased biochar application amount (Xu et al. 2014). When manure biochar was applied, the available P content in alkaline soil increased, along with the P concentration in plant tissues (Nahidan and Ghasemzadeh 2022; Gunes et al. 2014). Biochar derived from different raw materials exhibits distinct total P concentrations (Fig. S2). Animal manure retains a significant amount of P compounds during pyrolysis, as these compounds do not volatilize (Jin et al. 2016). Studies have shown that the total P content of manure-derived biochar is 53 to 105 times higher than that of lignocellulosic biochar, and manure-derived biochar

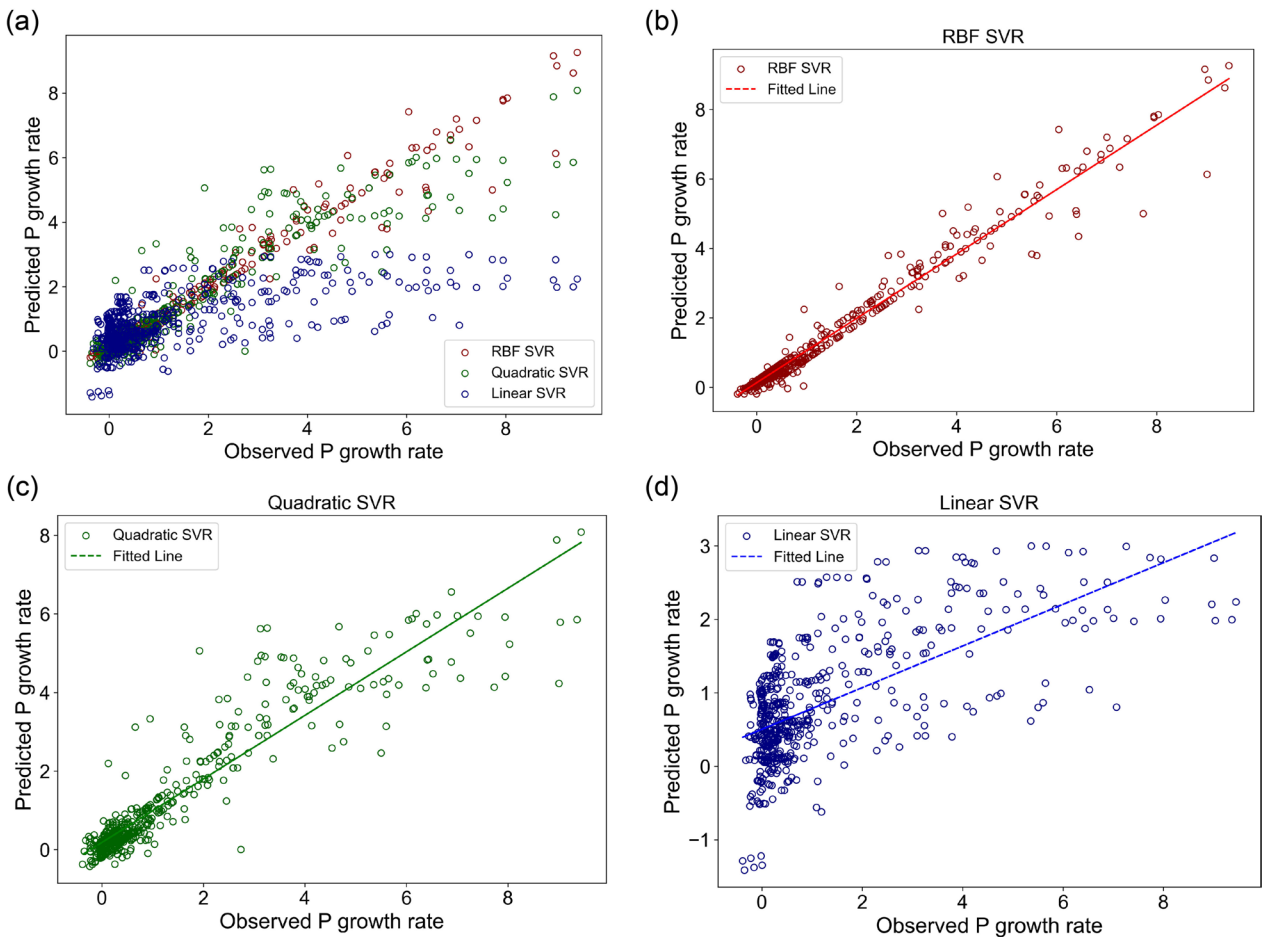


Fig. 3 In the optimization of hyperparameters for the SVR model, the predictive performance under three kernel functions evaluates model adaptability: **a** Overview, **b** RBF, **c** Quadratic, and **d** Linear

releases significantly more water-soluble P (Novak et al. 2018). To further investigate the fundamental relationships between biochar's soil PAR efficiency and its contributing factors, this study employed machine learning techniques, as described in the subsequent section.

Table 1 Optimal hyperparameters of machine learning models obtained through grid search to achieve the best predictive performance for each model

Target	Machine learning model	Hyperparameter	Optimal value
P growth rate	RF	n_estimators	200
		max_depth	15
		max_features	0.8
	SVR	Kernel	RBF
		C	50
		Epsilon	0.1
	ANN	Learning rate	0.01
		Number of neurons	32
		Number of hidden layers	2

Table 2 R^2 and RMSE prediction accuracy of RF, SVR and ANN algorithms on the efficiency of biochar in soil PAR

	RF	SVR	ANN
Train R^2	0.9721	0.9008	0.8947
Test R^2	0.9107	0.8146	0.8382
Train RMSE	0.3109	0.5865	0.5954
Test RMSE	0.5910	0.7254	0.8345

3.2 Optimization and comparison of machine learning model performance

Current dataset comprising 534 groups of data based on 19 input features was used to establish RF, SVR, and ANN models. Hyperparameters for each model were optimized to achieve the best performance. The initial step involved the selection of the kernel for the SVR model, and the value of C was set to 50, while that of ε was set to 0.1. The radial basis function (RBF), quadratic polynomial kernel, and linear kernel were selected to train and test the performance of the SVR model. The performance results of the SVR models with different kernels are shown in Fig. 3, revealing a significant difference in predictive performance among the kernels. The SVR model utilizing the RBF kernel achieved the best performance with an R^2 of 0.97. The performance of the SVR model using the quadratic kernel function was the second highest, with an R^2 of 0.86, while the SVR model employing a linear kernel exhibited the poorest performance, with an R^2 of only 0.34. Accordingly, the RBF kernel was selected for further hyperparameter optimization of the SVR model.

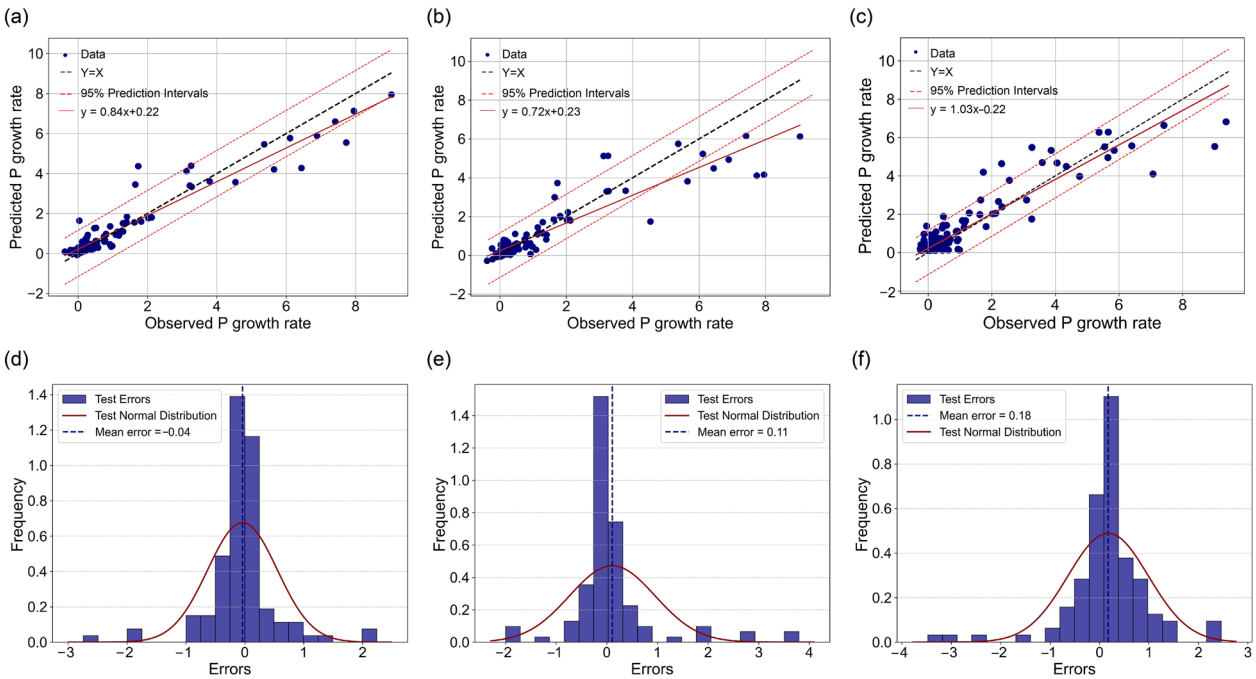


Fig. 4 Predictive performance of the **a**, **d** RF, **b**, **e** SVR, and **c**, **f** ANN model on the test set, for predicting the soil P availability changes in soil amended with biochar

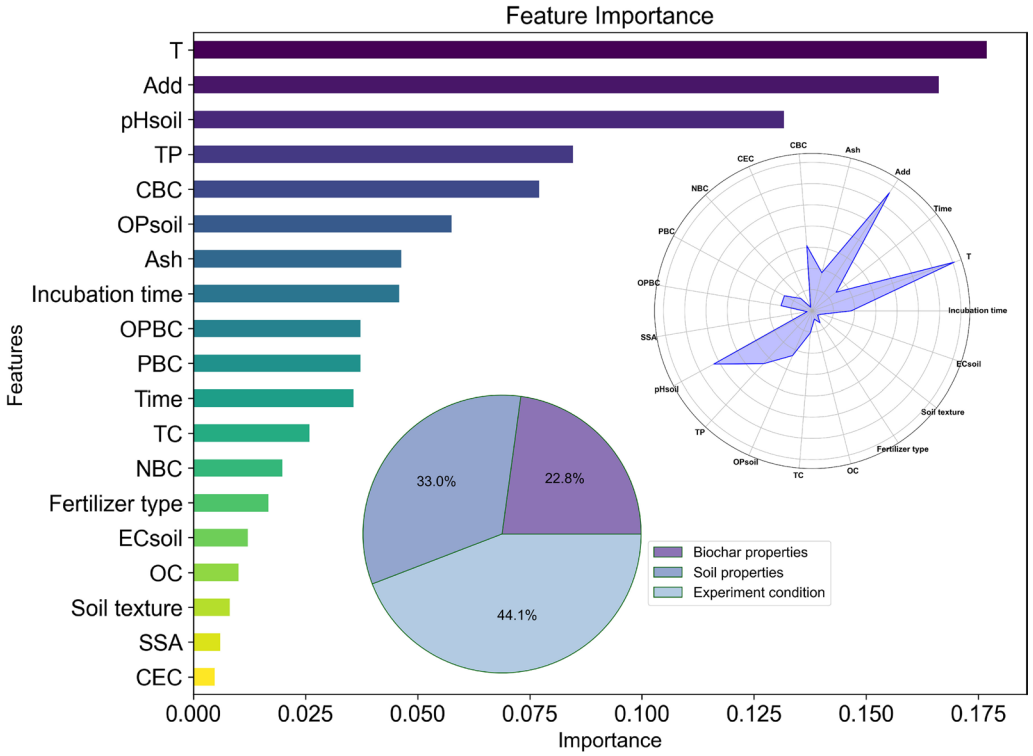


Fig. 5 Relative importance of input variables on P regulation efficiency using the RF model, to characterize the dominant factors influencing soil P availability

When constructing the ANN architecture, the number of hidden layers is typically considered a critical parameter. An excessive number of hidden layers relative to the complexity and size of the dataset can lead to reduced accuracy on the test set (Raut and Dani 2020). Considering the feature complexity and dataset scale in this study, the learning rate was set to 0.01, with 32 neurons per hidden layer. ANN models with one and two hidden layers were compared. The model with two hidden layers achieved a higher test set R^2 ($R^2 = 0.84$) than the one with a single hidden layer ($R^2 = 0.79$), indicating a better fit to the current dataset and feature complexity. The results suggest that this nonlinear system benefits from a two-hidden-layer architecture for improved stability and generalization (Thomas et al. 2017). Therefore, the ANN structure with two hidden layers was selected.

Subsequently, grid search was employed to select the optimal hyperparameter combinations for the three models, followed by five-fold cross-validation to minimize prediction errors. Table 1 illustrates the optimal hyperparameters identified for the machine learning models through grid search. The three models were updated with the optimal hyperparameters obtained, and the prediction performance of the three models on the

Table 3 Input feature importance based on RF model

Feature	Weight value
Incubation time	0.0459
T	0.1769
Time	0.0358
Add	0.1662
Ash	0.0463
CBC	0.0771
CEC	0.0047
NBC	0.0199
PBC	0.0372
OPBC	0.0372
SSA	0.0060
pHsoil	0.1317
TP	0.0847
OPsoil	0.0575
TC	0.0259
OC	0.0101
Fertilizer type	0.0167
Soil texture	0.0081
ECsoil	0.0121

test set is presented in Table 2 and Fig. 4. The prediction performance on the training set is illustrated in Fig. S3.

The RF model demonstrated the highest prediction accuracy for the test data, achieving an R^2 of 0.9107, which outperformed both the SVR and ANN models. Concurrently, the test set RMSE was 0.5910, representing the lowest value among the three models, and the error distribution exhibited tendency to align with a normal distribution (Fig. 4d). Error analysis revealed that the prediction results of all three models generally followed a normal distribution (Sun et al. 2022), but the SVR and ANN models exhibited larger errors compared to the RF model (Fig. 4d–f). Comparing the performance of the three algorithms in predicting the efficiency of biochar in soil PAR, the RF model was found to be the most optimal, followed by the ANN model, and lastly, the SVR model. This superior performance of the RF model could be attributed to its robustness against outliers and noise, its resistance to over-fitting, and its suitability for small datasets (Liu et al. 2012). This finding is consistent with those of previous studies, which have also demonstrated the superiority of RF in addressing lower values and small sample sizes (Zhu et al. 2019).

3.3 Analysis of input feature importance

The relative importance of each input characteristic to soil PAR efficiency was calculated using the RF model, as shown in Fig. 5. Thus, the key factors influencing the effectiveness of biochar in regulating soil PAR were identified, providing a theoretical basis for optimizing biochar production and application strategies. Experimental condition was the most influential factor, accounting for 44.1%, primarily due to the significant effects of biochar pyrolysis temperature and application rate on soil PAR efficiency. Soil properties ranked second, with a relative importance of 33.0%. Further analysis of feature importance revealed that biochar pyrolysis temperature was the most significant factor influencing biochar efficiency in soil PAR, accounting for 17.69% of the 19 features, followed by biochar application rate at 16.62% (Table 3). In addition, soil pH and total P content were the two most important soil-related features, with relative importance

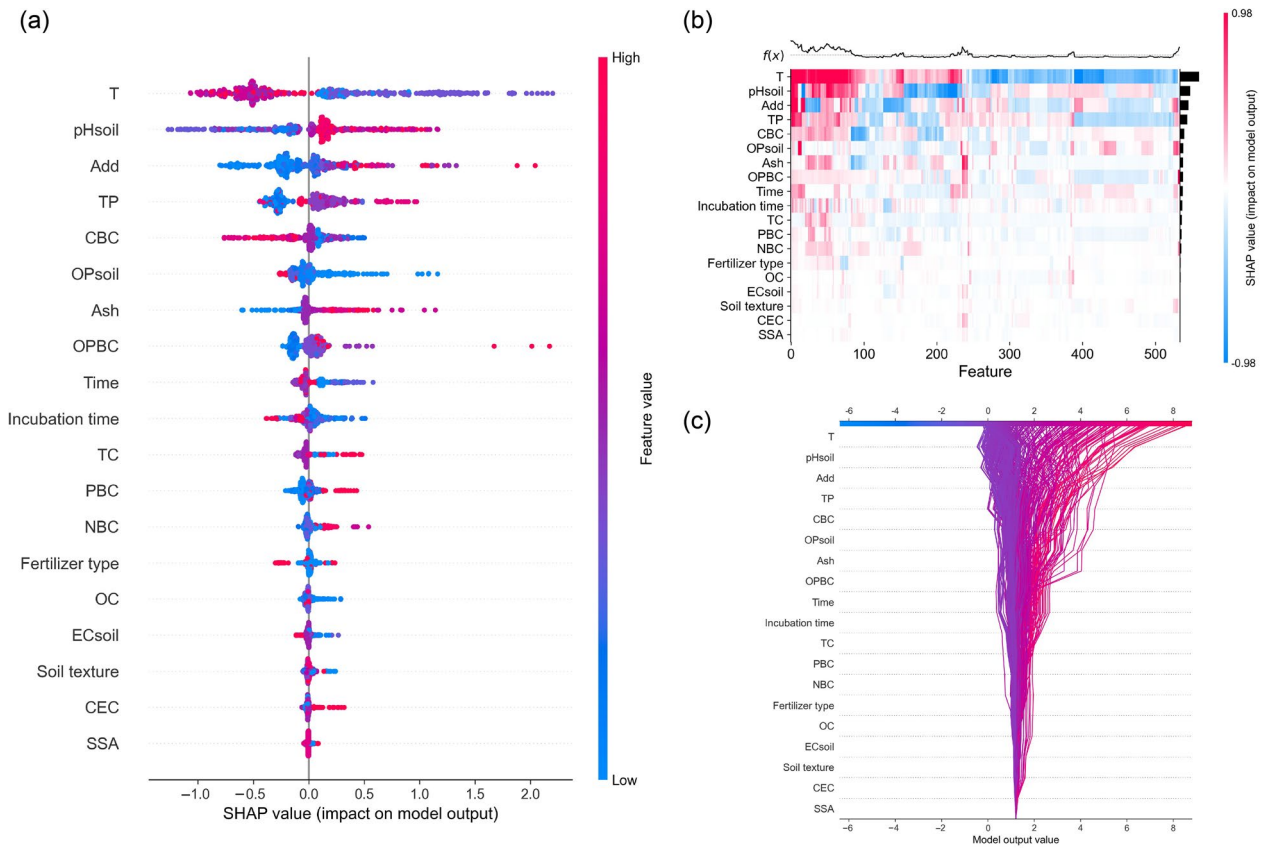


Fig. 6 SHAP analysis reveals the impact of features on model predictions. **a** The SHAP scatter plot illustrates the direction and magnitude of feature impacts on model outputs. **b** The heatmap displays the SHAP values of all features for each sample, with the curve at the top f(x) representing the trend of the average predicted values across all samples. **c** The SHAP decision plot demonstrates the influence of feature value changes on model outputs

of 13.17% and 9.47%, respectively. Soil pH directly or indirectly affects P solubility and speciation, while the total P reserve determines the potential availability of P. Soil P exists in various forms. Although most soils worldwide hold significant P reserves, the amount of inorganic phosphate available for plant use is extremely limited, often less than 1% (Zhu et al. 2018). Biochar can reduce the complexation of P with soil components such as Fe^{3+} and Al^{3+} (hydr)oxides, while also enhancing soil P mineralization (Yang et al. 2021). Furthermore, biochar-derived DOM could reduce P adsorption in soil through competition for adsorption sites and electrostatic repulsion, thereby enhancing soil P availability (Schneider and Haderlein 2016). Prior research suggested that biochar itself could serve as a P source (Dai et al. 2016). However, in our study, the total P and Olsen-P contents of biochar showed relatively low importance. This implied that the increase in soil available P following biochar application may be primarily attributed to the activation of endogenous soil P reserves, rather than direct P release from biochar.

The SHAP method results for the relative importance of input features in regulating P efficiency were comparable to the results obtained using RF for determining feature importance (Fig. 6). In the SHAP method, each point represented the SHAP value of a sample, with color indicating the magnitude of the feature value (Fig. 6a). The SHAP heatmap in Fig. 6b illustrates the distribution of SHAP values across samples (x-axis) for each input feature (y-axis), representing their impacts on the model's predictions. Compared to traditional feature importance methods, SHAP not only ranks input variables by importance but also reveals the direction and magnitude of their effects on model outputs, thereby providing complementary insights that significantly enhance the interpretability of the machine learning model in this study. Figure 6c depicts the decision-making process of the model for all samples. The results revealed that the pyrolysis temperature of biochar, soil pH, biochar application rate, and the total P content of the soil were the most influential features, exerting a considerable influence on the predictive outcomes. The pyrolysis temperature exerted a profound influence on the physical and chemical properties and stability of biochar, including the pH value, ash content, carbon stability, biochar porosity, and aromaticity (Zhang et al. 2015). The red dots (high temperature) for the pyrolysis temperature of biochar were primarily situated within the negative SHAP value area, whereas the blue dots (low temperature) were predominantly concentrated within the positive SHAP value area. This suggested that low temperature exerted a favorable influence on the model output. In contrast, biochar produced at high temperatures could diminish the labile P content of the soil, thereby curbing the influx of excessive P into water bodies through surface runoff. As for soil pH, SHAP analysis indicated that soils with neutral to slightly alkaline pH exhibited greater improvements in P availability following biochar application. In this study, the soil pH ranged from 4.29 to 10. In acidic soils, P tends to form insoluble complexes with Al^{3+} and Fe^{3+}, which reduces P availability (Penn And Camberato 2019). Moreover, the pH increase induced by biochar in such acidic conditions is limited, constraining its capacity to enhance P availability. Conversely, under neutral to slightly alkaline conditions, the activity of P-fixing metal ions is relatively low, and the alkalizing effect of biochar can further facilitate the transformation of P into more available forms, resulting in a more significant improvement in P availability. For total soil P content, red dots were concentrated in the positive SHAP value region, while blue dots clustered in the negative region. This suggested that higher total soil P positively influenced the model output, aligning with the findings on the relative importance of the RF model.

Moreover, the modeling results based on non-manure-derived biochar showed a similar pattern of feature importance in regulating soil PAR compared to the overall model (Fig. S6). This further supports the notion that the effectiveness of biochar in enhancing soil PAR is not solely attributed to its inherent P content, but is primarily driven by its ability to transform and mobilize existing soil P. From an economic and management perspective, this finding provides practical implications: in contrast to manure-derived biochar, which generally requires higher pyrolysis energy input, crop- and wood-derived biochars offer comparable regulatory efficiency with lower energy consumption and cost, making them more feasible and sustainable for large-scale agricultural deployment. It is also worth noting that unlike traditional RF-based feature importance, the SHAP approach not only delivers global importance rankings but also clarifies the exact direction and degree to which features influence prediction outcomes, offering more comprehensive and interpretable insights for decision-making and biochar management.

In addition to the pyrolysis temperature of biochar and the soil pH, biochar application rate is also considered one of the top three influencing factors. To reveal the nonlinear effects of interactions among key features on soil P availability, this study conducted a visual analysis of their interactions and impacts on the response variable. The results of the interaction between biochar pyrolysis temperature, pyrolysis time, application rate, and soil P content are comprehensively illustrated in Fig. 7. As shown in Fig. 7a, a synergistic relationship exists between pyrolysis temperature and

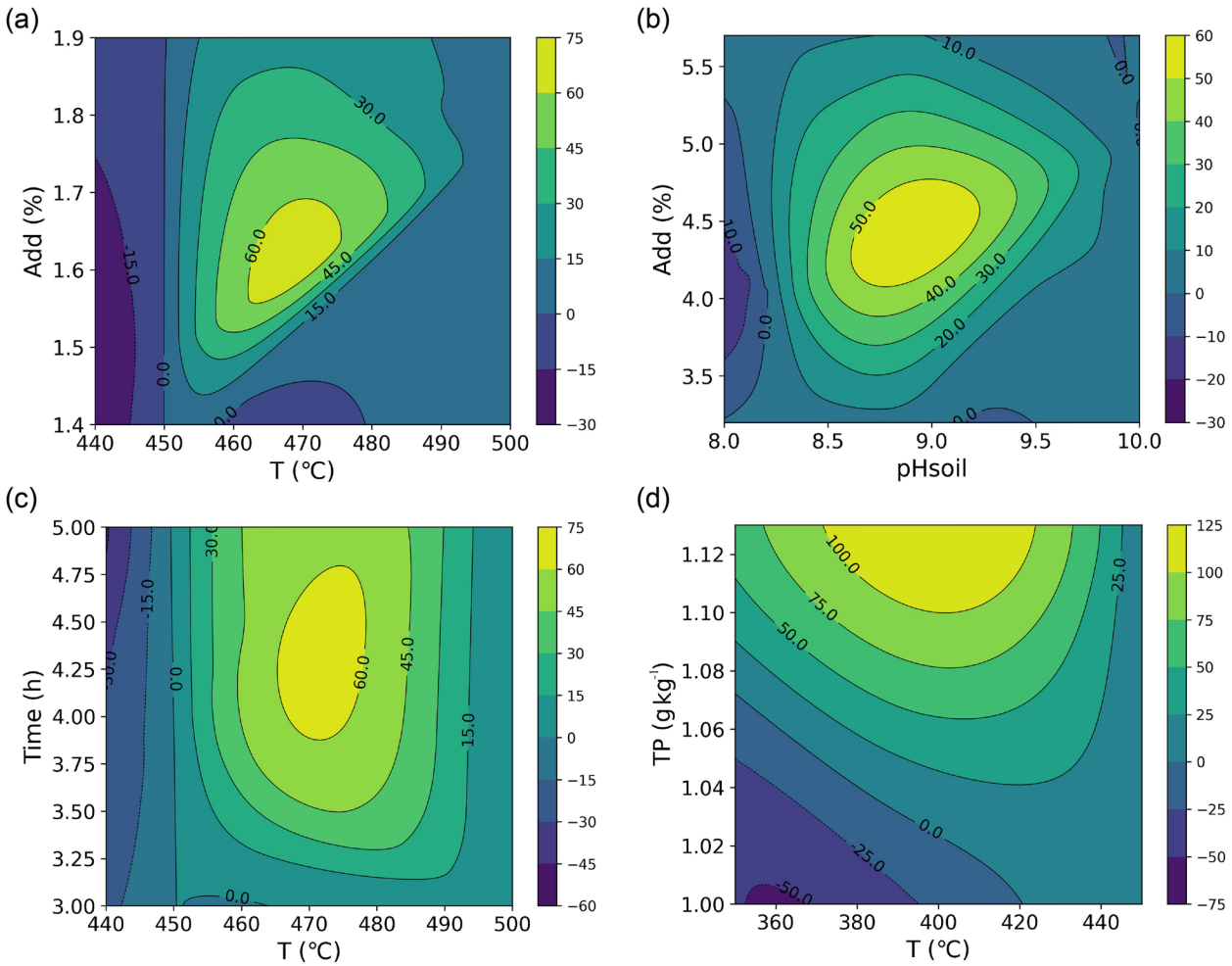


Fig. 7 Two-way interactive effects of key input variables on soil PAR efficiency. **a** Pyrolysis temperature vs. biochar application rate; **b** Soil pH vs. biochar application rate; **c** Pyrolysis temperature vs. pyrolysis duration; and **d** Pyrolysis temperature vs. soil total P content

biochar application rate, with optimal PAR efficiency observed at temperatures of 460–480 °C and application rates of 1.5–1.7%. This optimal range likely stems from biochar's balanced physicochemical properties at moderate temperatures, such as sufficient porosity for nutrient adsorption and retention, while avoiding excessive carbonization that might reduce reactive surface sites. From an economic and management perspective, maintaining biochar production within this temperature range also minimizes energy consumption and production costs, thereby improving the cost-efficiency of large-scale biochar utilization. When the pyrolysis temperature of biochar was in the range of 460–485 °C, a pyrolysis duration of 3.5–5 h was considered optimal (Fig. 7c). Under moderate pyrolysis temperatures, excessively prolonged pyrolysis time may lead to over-carbonization of the biochar structure (Liang et al. 2020), resulting in micropore collapse and reductions in SSA and porosity. Consequently, the adsorption capacity of biochar may decline, weakening its efficiency in P retention or release, which is consistent with the feature importance results presented earlier (Fig. 6a). Prolonged pyrolysis time may impair material performance and increase energy input, which could reduce economic feasibility in commercial production. Thus, optimizing both temperature and time is essential not only for functional efficiency but also for operational cost control. In addition, Fig. 7d highlights a distinct scenario where soil total P content influences the effectiveness of higher pyrolysis temperatures. Specifically, biochar pyrolyzed at 360–440 °C exhibited higher PAR efficiency when the soil total P level ranged from 1.04 to 1.13 g kg^{-1}, suggesting that the enhanced thermal stability and alkaline properties of high-temperature biochar may facilitate P mobilization in P-deficient soils by altering pH-dependent solubility

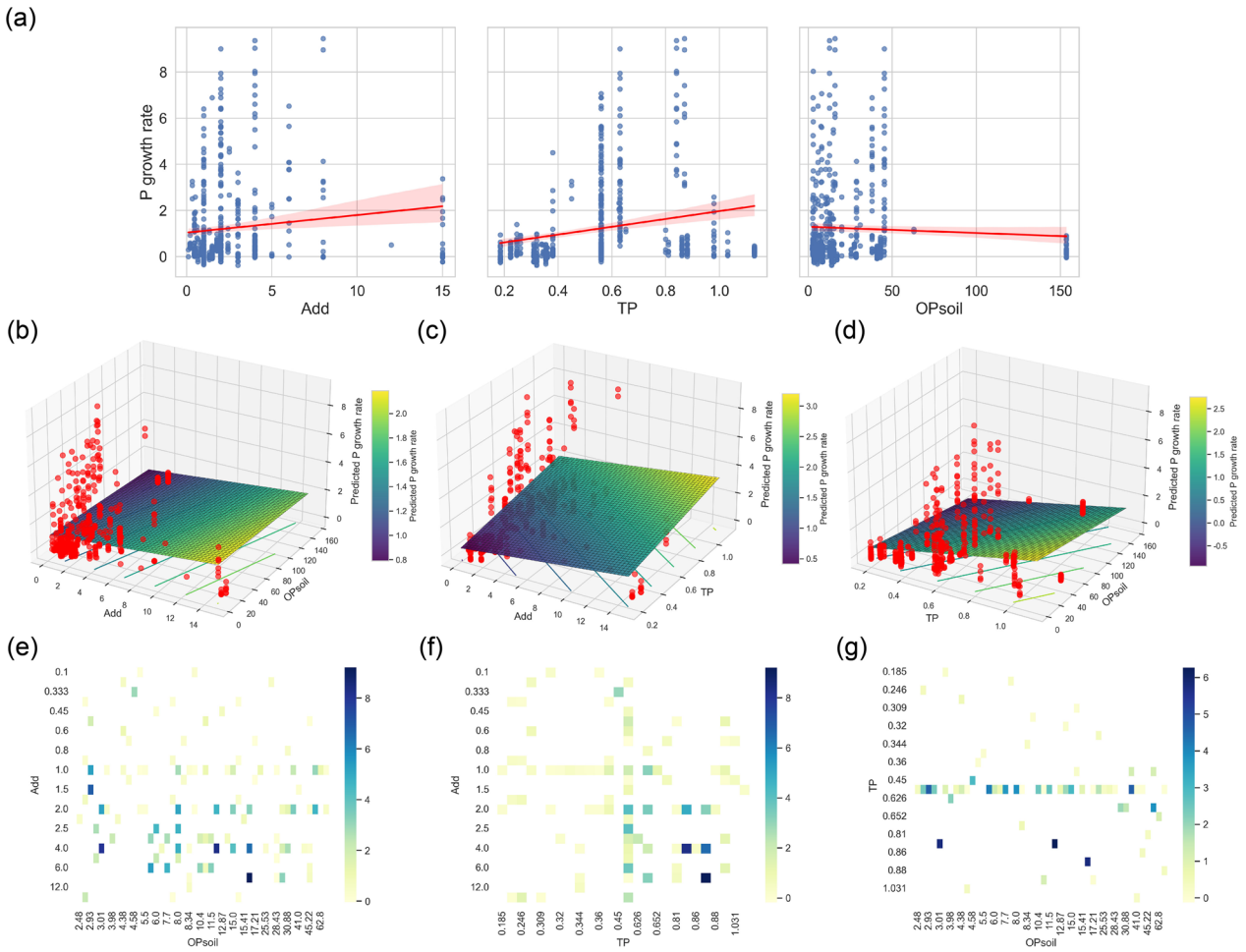


Fig. 8 Ordinary least squares linear regression was used to explore the interactive effects of Add, TP, and OPsoil on the output. **a** Linear regression fitting of P growth rate with Add., TP, and OPsoil; **b**–**d** 3D linear regression fitting plots showing the pairwise interactions among Add, TP, and OPsoil on P growth rate; **e**–**g** 2D heatmaps illustrating the effects of pairwise interactions among Add, TP, and OPsoil on the mean P growth rate

or chelating inhibitory metal ions. This temperature-dependent divergence underscored the importance of tailoring biochar production parameters to specific soil conditions: moderate-temperature biochar appears ideal for general applications in soils with adequate P, while high-temperature biochar may serve as a targeted amendment for P-deficient or metal-contaminated soils. In practice, increasing the biochar application rate and optimizing the pyrolysis temperature can enhance P release and utilization in low-P soils, whereas reducing the application rate in high-P soils may help avoid resource waste and potential environmental risks. These findings emphasize the nonlinearity of biochar-soil interactions, where the interplay of production variables and environmental factors dictates functional outcomes, necessitating a systems-based approach for optimizing biochar use in sustainable agriculture.

3.4 Machine learning for soil P availability prediction

In this research, machine learning algorithms (RF, SVR, and ANN) were developed and fine-tuned using a curated dataset comprising 534 data points sourced from 32 peer-reviewed studies, with computational durations ranging from seconds to minutes per iteration. These established ML models exhibited robust predictive performance in evaluating biochar's efficacy for soil PAR, achieving test R^2 values of 0.8146–0.9107, which underscored their reliability in capturing the complex relationships between biochar properties and P activation outcomes (Li et al. 2024). Given the dataset encompasses diverse biochar properties, soil environments, and experimental conditions, the machine learning framework demonstrates strong generalization capability. Compared with machine learning models, traditional direct prediction approaches, such as simple or generalized linear models (Fig. 8, using ordinary least squares regression to

Table 4 Machine learning models used in previous studies and the R^2 on the test set

Research object	Machine learning model	Test R^2	Reference
Soil phosphorus availability regulation by biochar	Random Forest	0.9107	This study
	Support Vector Regression	0.8146	
	Artificial Neural Network	0.8382	
Metal immobilization remediation by biochar	Random Forest	0.824	(Sun et al. 2022)
	Artificial Neural Network	0.841	
	Linear Model	0.507	
Relationships between biochar and anaerobic digestion	Gradient Boosting Decision Tree	0.84	(Zhang et al. 2023)
	Random Forest	0.80	
	eXtreme Gradient Boosting	0.80	
	Gradient Boosting Decision Tree	0.69	
	Random Forest	0.64	
	eXtreme Gradient Boosting	0.67	
Predict total yield and specific surface area of biochar	Random Forest	0.855	(Hai et al. 2023)
	Multiple Linear Regression	0.821	
	Decision Tree Regressor	0.559	
	K-nearest Neighbor	0.510	
	Random Forest	0.804	
	Decision Tree Regressor	0.618	
	K-nearest Neighbor	0.557	
	Multiple Linear Regression	0.292	
Heavy metal immobilization by biochar	Support Vector Regression	0.70	(Guo et al. 2023a, b)
	Linear Regression	0.69	
	Gradient boosting Decision Trees	0.82	
	Random Forest	0.85	
Predict biochar yield	Random Forest	0.59	(Ma et al. 2023)
	Gradient Boosting Decision Trees	0.84	
	eXtreme Gradient Boosting	0.79	
	Support Vector Regression	0.44	
	Radial Basis Function Neural Network	0.65	
	Levenberg–Marquardt Backpropagation Neural Network	0.88	
Biochar design for antibiotics adsorption	Random Forest	0.73	(Li et al. 2024)
	Gradient Boosting Regression	0.68	
	eXtreme Gradient Boosting	0.85	
	Random Forest	0.78	
	Gradient Boosting Regression	0.71	
	eXtreme Gradient Boosting	0.88	

examine the interactions of Add, TP, and OPsoil on the output), generally assume variable independence and fail to capture complex nonlinear relationships. When the number of interaction terms or sample size increases, explicitly incorporating multiple interactions substantially raises model complexity and may lead to multicollinearity or overfitting, limiting the ability to represent multifactorial interactions (Otse et al. 2025). In contrast, machine learning can simultaneously integrate 19 features across soil physicochemical properties, biochar characteristics, and environmental factors, capturing complex interactions and automatically identifying higher-order nonlinear effects. Consequently, machine learning demonstrates superior predictive performance and mechanistic insight, providing more precise guidance and hypotheses for subsequent mechanistic studies (Jorner et al. 2021). The RF model demonstrated the most optimal performance among the three models ($R^2 = 0.9107$, RMSE = 0.5910) (Fig. 4, Table 2). This is because the RF model employed an ensemble learning

approach, whereby each decision tree was trained and predicted independently, and the resulting average prediction was then produced. This approach makes the RF model more robust than those relying on a single training process (Chen et al. 2023).

Moreover, it is noteworthy that the model performance in this study was considered superior compared to research in other environmental fields (Table 4). This may be attributed to differences in dataset structure and target variables. Specifically, the present dataset integrated a broader range of biochar properties and soil physicochemical conditions, enhancing the model's ability to capture nonlinear interactions between biochar and soil P dynamics. Furthermore, the optimized hyperparameters and standardized quantitative variables in this study improved model generalization and minimized overfitting. The target variable, biochar-induced changes in soil P availability, was inherently influenced by multiple interacting mechanisms, which may produce stronger predictive relationships than those based on single-factor responses. However, the mainstream models used in this study also have limitations. For instance, the presence of a considerable number of missing values within the dataset could result in suboptimal model fitting, whereas a notable degree of data point dispersion could diminish the stability of the model (Wang et al. 2024a). To address this issue, Zhang et al. (2021) employed extreme gradient boosting (XGBoost) to analyze ground subsidence in anisotropic clay. XGBoost demonstrated remarkable predictive performance with sparse data (test set $R^2 = 0.9859$), owing to its capacity to handle missing values automatically and select the optimal split directions (Wang et al. 2019), thereby better managing imperfect datasets. Moreover, in scenarios demanding high prediction accuracy, Gradient Boosting Decision Trees (GBDT) consistently outperform RF, a trend substantiated by prior empirical studies (Wang et al. 2024b). However, both XGBoost and GBDT have limitations, including complex hyperparameter tuning and long computation times, which contributed to their lesser prevalence in comparison to mainstream models. Machine learning is data-driven, where data quality played a crucial role in model performance alongside dataset size. The quality of training data was often considered more important than the choice of algorithm (Idakwo et al. 2018). However, data quality could be influenced by variations in experimental conditions and methods used in published studies. Even with a large dataset, discrepancies in experimental design could lead to inconsistent results, increasing prediction uncertainty. Any variability, error, or biase in the training database will be reflected in the outcomes (Barragán-Montero et al. 2021). Therefore, ensuring consistency and reliability in data quality was critical for accurate prediction. Caution was required in selecting data from published sources, as well as identifying and correcting 'dirty data' (Guo et al. 2023a, b). Beyond data volume, the features included in our models do not fully capture the complexity of real-world processes governing soil P availability. For instance, soil redox conditions, which strongly influence P solubility and mobility, were not incorporated because key indicators such as Eh and soil moisture are not consistently reported in the existing literature. Likewise, dynamic processes affecting available P, including flooding, drought, and biochar aging, were not considered in our dataset. These factors can substantially modulate biochar effectiveness under field conditions, but their inclusion is challenging due to limited availability of high-quality data. Consequently, while the current models provide useful predictive insights, their applicability is inherently constrained by the underlying dataset, and caution is warranted when extrapolating results to more complex and dynamic field scenarios. Overall, future research on the efficiency of biochar application for soil PAR should focus on the following aspects: (i) increasing sample size of data through more field experiments and data collection to enhance model generalizability; (ii) improving data collection and preprocessing techniques to mitigate uncertainties in data quality, as well as the effects of missing data and outliers; (iii) exploring advanced models such as XGBoost and GBDT and optimizing existing algorithms to enhance model prediction accuracy and robustness; (iv) modern computational approaches, such as AI-driven digital twin frameworks, offer the potential to simulate complex systems and predict their techno-economic performance (Stefko et al. 2025; Chatterjee et al. 2025). Integrating these advanced methods with machine learning models could further provide insights into the cost-effectiveness, scalability, and sustainability of biochar applications under specific economic conditions.

In addition to model performance, analyzing the characteristics of soil PAR accuracy helped to understand the role of biochar in regulating P availability and was also of great significance for actual PAR. Our results showed a positive correlation between the total P content of the soil and the efficiency of biochar in soil PAR (Fig. 1, correlation coefficient = 0.22). The overall soil P content was a key factor, with the SHAP method showing that higher total P level in the soil improves biochar's effectiveness in enhancing soil PAR. In contrast, biochar's total P content and Olsen-P levels have a relatively minor impact (Figs. 6, 7). This suggested that the increase in soil available P following biochar application was primarily due to the conversion of native soil P to available P, rather than the release of P from biochar itself, as confirmed by previous studies (Peng et al. 2023). This finding provided guidance

for the practical application of biochar in conjunction with P fertilizers, particularly in soils with high total P but low P availability (Tesfaye et al. 2021). Additionally, biochar produced by high-temperature pyrolysis could reduce soil available P content (Fig. 7). This reduction is attributed to the high porosity and specific surface area of biochar, which enable the adsorption of excess available P in the soil—particularly in soils polluted by excessive phosphate fertilization—thereby mitigating environmental risks such as eutrophication. The porous structure and high specific surface area of biochar enhance the formation of soil macroaggregates and promote P accumulation within them (Cao et al. 2021). Optimizing biochar application rate and pyrolysis temperature according to soil total P levels can enhance P management efficiency. Biochar can also influence the distribution of P within soil aggregates by altering the community structure and activity of microorganisms, such as phosphate-solubilizing microbes (Lu et al. 2023a). Future studies may further explore this through soil microbial processes. Additionally, biochar with high ash content could increase soil available P content. However, due to some missing data on biochar ash content in the dataset used in this study, model performance and feature importance results may be affected to some extent. Adjusting the application rate and pyrolysis temperature of biochar based on the total soil P content could enhance the efficiency of biochar in soil PAR. It is also noteworthy that biochar application rates in incubation studies do not directly reflect recommended field rates, which should be determined by balancing cost and agronomic efficiency in practical use. Based on this, future research should focus on field trials to systematically assess the temporal dynamics of soil available P following biochar application, while incorporating crop type and climatic conditions as key influencing factors. Additionally, future research could focus on quantifying P activation or passivation through more detailed indicators, such as the transformation rates of P fractions, to further elucidate the complex biogeochemical mechanisms underlying soil P dynamics.

To facilitate the use of the prediction model developed in this study by researchers and practitioners, a graphical user interface (GUI) was built using Python 3.7 and Streamlit, incorporating the RF model proposed in this work (https://growthrateprediction-fmqs6nvuwtoyh7hs7evag2.streamlit.app, Fig. S8). By default, the GUI initializes each input variable with the mean value of the corresponding feature from the dataset. Categorical variables, such as soil texture and fertilizer type, can be selected via dropdown menus. It is worth noting that the more complete the input information, the higher the prediction accuracy. In addition to prediction functionality, the GUI integrates key adjustable parameters such as biochar pyrolysis time and temperature, providing practical guidance for optimizing soil P availability through biochar application. Meanwhile, it is important to note that our online model can achieve precise predictions under certain environmental conditions: soil CEC ranges from 15.2 to 128.7 cmol kg^{-1}, available soil phosphorus ranges from 26.8 to 2570 mg kg^{-1}, and maximum cultivation time of 390 days. As for the biochar characteristics: pyrolysis temperature ranges from 200 to 700 °C, pyrolysis time ranges from 0.5 to 6 h, application rate ranges from 0.1% to 15%, biochar pH ranges from 4.05 to 11.56, biochar ash content ranges from 1% to 62.5%, biochar carbon content ranges from 25.2% to 73.6%, biochar nitrogen content ranges from 0.0266% to 3.55%, biochar phosphorus content ranges from 0.012% to 2.6%, and biochar specific surface area ranges from 3 to 282 m^2 g^{-1}.

In relevant research or engineering contexts, users can tailor input variables based on specific soil and biochar characteristics to generate optimized application strategies, thereby enhancing the efficiency of biochar in regulating soil available P while significantly reducing time and resource costs.

3.5 Economic advantages of machine learning assisted raw biochar applications

The application of biochar to cropland soils demonstrates both environmental and economic sustainability in improving soil and water quality. Taking crop residue-derived biochar as an example, China produces approximately 600–800 Tg of crop straw annually (Chen et al. 2019), which is commonly managed through open burning or straw incorporation. However, straw burning releases large quantities of greenhouse gases (GHGs) and causes severe air pollution (Shi et al. 2023), while straw incorporation can also produce GHGs during anaerobic decomposition and introduce risks of heavy metal and pesticide residues, as well as pest infestations from residual eggs (Shan et al. 2021; Yu et al. 2023). In contrast, pyrolyzing crop residues into biochar represents an effective resource utilization strategy, with conversion yields typically ranging from 30% to 60% (Zhang et al. 2020; Cueva et al. 2022; Chandra and Bhattacharya 2019). For instance, wheat straw pyrolyzed at 300 ℃ produces a 46.96% biochar yield with a carbon content of 61.48% (Zhang et al. 2020). Each kilogram of straw burned emits approximately 1.239–1.670 kg CO_2-eq of GHGs (Shi et al. 2023), whereas long-term field and laboratory studies have shown that biochar application significantly enhances soil organic carbon sequestration rates by 31.8–47.8%, equivalent to 369.8–556.6 kg C fixed annually (Li and Tasnady 2023). Moreover, when applied at a rate of 2–4%, biochar can reduce annual P losses by 10–25% per hectare (Wang et al. 2022), directly

contributing to eutrophication mitigation and improved soil P use efficiency. According to World Bank data, global phosphate fertilizer prices have surged in recent years, reaching 172.82 USD t^{-1} in 2022 (Wendimu et al. 2023). In comparison, the production cost of biochar ranges from 0.24 to 0.61 USD kg^{-1} (Sakhiya et al. 2022). Assuming a neutral scenario with a biochar application rate of 3 t ha^{-1} and a P fertilizer input of 42.70 kg ha^{-1} (Li et al. 2020b), biochar could reduce P loss by approximately 17.5%, equivalent to saving 7.47 kg ha^{-1} yr^{-1} of phosphate fertilizer, translating into a cost saving of roughly 21.5 USD ha^{-1} yr^{-1}. The biochar investment cost is estimated at 1275 USD ha^{-1}, while the annual carbon sequestration benefit is approximately 2.26 t CO_2-eq ha^{-1}. Considering carbon credit values of 50–100 USD t^{-1} CO_2-eq, an additional revenue of 113–226 USD ha^{-1} yr^{-1} could be generated. Although the short-term fertilizer savings alone may not offset the initial biochar input cost, the combined long-term benefits in soil improvement, eutrophication mitigation, and carbon reduction indicate that biochar application holds substantial environmental and economic potential, offering a viable strategy for sustainable agricultural management.

In addition, traditional pristine biochar exhibits notable advantages in agricultural soil improvement and carbon sequestration, as it is derived from abundant feedstocks, features simple preparation processes, and maintains relatively low production costs, thereby demonstrating strong potential for large-scale application (Zhang et al. 2022). However, pristine biochar also presents several limitations. Biochar produced at low pyrolysis temperatures or from certain feedstocks tends to exhibit poor stability in soils and is susceptible to microbial degradation, which constrains its long-term carbon sequestration potential (Luo et al. 2023). Moreover, pristine biochar generally contains fewer surface functional groups and active sites, resulting in limited nutrient adsorption and contaminant immobilization capacities (Shaheen et al. 2025). Furthermore, substantial variations in physicochemical properties and field performance among biochars from different sources often lead to inconsistent application outcomes.

To enhance the performance of biochar, various modification strategies have been developed in recent years, including metal loading, phosphorylation, and functional composite approaches (Du et al. 2023). Such modification treatments can improve biochar's capacity for nutrient and heavy metal immobilization, increase its surface reactivity, and extend its service life (Gong et al. 2022; Khan et al. 2022). However, these benefits come at the cost of significantly higher production expenses and environmental impacts. For example, at a moderate pyrolysis temperature of 400 ℃, the production cost of Fe-modified biochar increases by approximately 45.93–70.05 USD t^{-1} compared with pristine biochar, Mg-modified biochar by about 46.59–69.12 USD t^{-1}, and La-modified biochar by 231.45–481.47 USD t^{-1} (Lu et al. 2023b). Moreover, functional composite biochars are even more expensive; for instance, starch-coated alkali-treated biochar-urea composites exhibit modification costs as high as 2051.1–2361.56 USD t^{-1} (Cheng et al. 2025). These modification processes often require additional chemical reagents and high energy input, which not only elevate production costs but also pose risks of wastewater generation and potential secondary pollution.

In contrast, the data used in this study were entirely derived from pristine biochar systems produced from crop residues, animal manure, and woody biomass, without any additional chemical modification processes. By integrating 534 datasets on soil P availability following pristine biochar application, a machine learning model was developed to analyze the relative importance of multiple input features, including biochar properties, soil characteristics, and experimental conditions. The optimized combination of key driving factors identified through machine learning enabled the pristine biochar to achieve, or even surpass, the overall performance of modified biochar in regulating P activation and immobilization. Meanwhile, pristine biochar can reduce modification-related costs by approximately 45–480 USD t^{-1} while maintaining lower environmental burdens and higher scalability. The application strategy of pristine biochar combined with machine learning not only demonstrates superior economic and sustainability advantages over modified biochar pathways but also exhibits greater potential in functionality and predictive precision.

4 Conclusion

This study confirmed that machine learning can effectively elucidate and predict the mechanisms governing biochar-induced regulation of soil P availability. The findings support our initial hypothesis that biochar application efficiency is primarily determined by the interplay between biochar physicochemical properties and soil conditions rather than by biochar P input alone. Furthermore, the success of the RF model highlights the potential of machine learning as a methodological framework for decoding complex soil-biochar-nutrient interactions, enabling the design of environmentally sustainable and economically viable biochar application strategies that promote precision and resource-efficient soil management.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s42773-026-00611-1.

Supplementary material 1

Acknowledgements
We are grateful to the editor and the anonymous reviewers for their valuable comments. We also would like to express our sincere gratitude to Prof. Dr. Chung-Yu Guan for his contribution and support to this work.

Author contributions
Yuqian Wang: Writing–review & editing, Writing–original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Junhui Yin: Writing–review & editing, Writing–original draft. Xiao Yang: Writing–review & editing, Writing–original draft. Bangxi Zhang: Writing–review & editing, Writing–original draft. Qing Chen: Writing–review & editing, Writing–original draft. Yutao Peng & Jia Liu: Writing–review & editing, Writing–original draft, Supervision, Project administration, Funding acquisition, Conceptualization. All authors read and approved the final manuscript.

Funding
We gratefully acknowledge the financial support of the National Natural Science Foundation of China (42207015), and the Research Funding of postdoctor who came to Shenzhen (szbo202323).

Data availability
The data will be made available on reasonable request, and the prediction data-code is provided in the supplementary information.

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]School of Agriculture and Biotechnology, Sun Yat-Sen University, Shenzhen 518107, Guangdong, China. [2]School of Pharmacy, Zunyi Medical University, Zunyi 563006, Guizhou, China. [3]Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China. [4]Soil and Fertilizer Research Institute, Guizhou Academy of Agricultural Sciences, Guiyang 550006, Guizhou, China. [5]Beijing Key Laboratory of Farmland Soil Pollution Prevention and Remediation, College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China.

Received: 23 August 2025 Revised: 18 November 2025 Accepted: 24 March 2026
Published online: 25 May 2026

References

Alewell C, Ringeval B, Ballabio C, Robinson DA, Panagos P, Borrelli P (2020) Global phosphorus shortage will be aggravated by soil erosion. Nat Commun 11(1):4546. https://doi.org/10.1038/s41467-020-18326-7
An R, Yu R-P, Xing Y, Zhang J-D, Bao X-G, Lambers H, Li L (2023) Enhanced phosphorus-fertilizer-use efficiency and sustainable phosphorus management with intercropping. Agron Sustain Dev 43(5):57
Antoniadis V, Koliniati R, Efstratiou E, Golia E, Petropoulos S (2016) Effect of soils with varying degree of weathering and pH values on phosphorus sorption. CATENA 139:214–219. https://doi.org/10.1016/j.catena.2016.01.008
Antwarg L, Miller RM, Shapira B, Rokach L (2021) Explaining anomalies detected by autoencoders using Shapley Additive Explanations. Expert Syst Appl 186:115736. https://doi.org/10.1016/j.eswa.2021.115736
Barragán-Montero AM, Thomas M, Defraene G, Michiels S, Haustermans K, Lee JA, Sterpin E (2021) Deep learning dose prediction for IMRT of esophageal cancer: the effect of data quality and quantity on model performance. Phys Med 83:52–63. https://doi.org/10.1016/j.ejmp.2021.02.026
Bindraban PS, Dimkpa CO, Pandey R (2020) Exploring phosphorus fertilizers and fertilization strategies for improved human and environmental health. Biol Fertil Soils 56(3):299–317
Bornø ML, Müller-Stöver DS, Liu F (2018) Contrasting effects of biochar on phosphorus dynamics and bioavailability in different soil types. Sci Total Environ 627:963–974. https://doi.org/10.1016/j.scitotenv.2018.01.283
Cao D, Lan Y, Sun Q, Yang X, Chen W, Meng J, Wang D, Li N (2021) Maize straw and its biochar affect phosphorus distribution in soil aggregates and are beneficial for improving phosphorus availability along the soil profile. Eur J Soil Sci 72(5):2165–2179
Chandra S, Bhattacharya J (2019) Influence of temperature and duration of pyrolysis on the property heterogeneity of rice straw biochar and optimization of pyrolysis conditions for its application in soils. J Clean Prod 215:1123–1139. https://doi.org/10.1016/j.jclepro.2019.01.079
Chatterjee S, Kliestik T, Rowland Z, Bugaj M (2025) Immersive collaborative business process and extended reality-driven industrial metaverse technologies for economic value co-creation in 3D digital twin factories. Oeconomia Copernicana. https://doi.org/10.24136/oc.3596
Chen J, Gong Y, Wang S, Guan B, Balkovic J, Kraxner F (2019) To burn or retain crop residues on croplands? An integrated analysis of crop residue management in China. Sci Total Environ 662:141–150. https://doi.org/10.1016/j.scitotenv.2019.01.150
Chen C, Wang Z, Ge Y, Liang R, Hou D, Tao J, Yan B, Zheng W, Velichkova R, Chen G (2023) Characteristics prediction of hydrothermal biochar using data enhanced interpretable machine learning. Bioresour Technol 377:128893. https://doi.org/10.1016/j.biortech.2023.128893
Cheng J, Sun J, Yao K, Xu M, Cao Y (2022) A variable selection method based on mutual information and variance inflation factor. Spectrochim Acta A Mol Biomol Spectrosc 268:120652. https://doi.org/10.1016/j.saa.2021.120652
Cheng Y, Dong C, Zhao X, Chen Q, Zhang Y, Li Y, Yu B (2025) Sustainable wheat production via starch-encapsulated alkali-modified biochar-urea: synergistic gains in yield, economics, and GHG mitigation. Chem Eng J 524:169045. https://doi.org/10.1016/j.cej.2025.169045
Cueva ZLL, Griffin GJ, Ward LP, Madapusi S, Shah KV, Parthasarathy R (2022) A study of chemical pre-treatment and pyrolysis operating conditions to enhance biochar production from rice straw. J Anal Appl Pyrolysis 163:105455. https://doi.org/10.1016/j.jaap.2022.105455
Dai L, Li H, Tan F, Zhu N, He M, Hu G (2016) Biochar: a potential route for recycling of phosphorus in agricultural residues. GCB Bioenergy 8(5):852–858
De Boer MA, Wolzak L, Slootweg JC (2019) Phosphorus: reserves, production, and applications. In: Ohtake H, Tsuneda S (eds) Phosphorus recovery and recycling. Springer, Singapore, pp 75–100
Divband Hafshejani L, Ali Naseri A, Hooshmand A, Soltani Mohammadi A, Abbasi F (2024) Prediction of nitrate leaching from soil amended with biosolids by machine learning algorithms. Ain Shams Eng J 15(7):102783. https://doi.org/10.1016/j.asej.2024.102783
Du L, Ahmad S, Liu L, Wang L, Tang J (2023) A review of antibiotics and antibiotic resistance genes (ARGs) adsorption by biochar and modified biochar in water. Sci Total Environ 858:159815. https://doi.org/10.1016/j.scitotenv.2022.159815
Dume B, Ayele D, Regassa A, Berecha G (2017) Improving available phosphorus in acidic soil using biochar. J Soil Sci Environ Manag 8(4):87–94
Emmanuel T, Maupong T, Mpoeleng D, Semong T, Mphago B, Tabona O (2021) A survey on missing data in machine learning. J Big Data 8:1–37
Etesami H (2020) Enhanced phosphorus fertilizer use efficiency with microorganisms. In: Meena RS (ed) Nutrient dynamics for sustainable crop production. Springer Singapore, Singapore, pp 215–245. https://doi.org/10.1007/978-981-13-8660-2_8
Gao S, DeLuca TH (2018) Wood biochar impacts soil phosphorus dynamics and microbial communities in organically-managed croplands. Soil Biol Biochem 126:144–150. https://doi.org/10.1016/j.soilbio.2018.09.002
Gao S, DeLuca TH, Cleveland CC (2019) Biochar additions alter phosphorus and nitrogen availability in agricultural ecosystems: A meta-analysis. Sci Total Environ 654:463–472. https://doi.org/10.1016/j.scitotenv.2018.11.124

Ghodszad L, Reyhanitabar A, Maghsoodi MR, Asgari Lajayer B, Chang SX (2021) Biochar affects the fate of phosphorus in soil and water: a critical review. Chemosphere 283:131176. https://doi.org/10.1016/j.chemosphere.2021.131176

Glaser B, Lehr V-I (2019) Biochar effects on phosphorus availability in agricultural soils: a meta-analysis. Sci Rep 9(1):9338

Gong H, Zhao L, Rui X, Hu J, Zhu N (2022) A review of pristine and modified biochar immobilizing typical heavy metals in soil: applications and challenges. J Hazard Mater 432:128668. https://doi.org/10.1016/j.jhazmat.2022.128668

Gunes A, Inal A, Taskin M, Sahin O, Kaya E, Atakol A (2014) Effect of phosphorus-enriched biochar and poultry manure on growth and mineral composition of lettuce (*Lactuca sativa* L. cv.) grown in alkaline soil. Soil Use Manag 30(2):182–188

Guo J, Sun M, Zhao X, Shi C, Su H, Guo Y, Pu X (2023a) General graph neural network-based model to accurately predict cocrystal density and insight from data quality and feature representation. J Chem Inf Model 63(4):1143–1156

Guo G, Lin L, Jin F, Mašek O, Huang Q (2023b) Application of heavy metal immobilization in soil by biochar using machine learning. Environ Res 231:116098. https://doi.org/10.1016/j.envres.2023.116098

Hai A, Bharath G, Patah MFA, Daud WMAW, K R, Show P, Banat F (2023) Machine learning models for the prediction of total yield and specific surface area of biochar derived from agricultural biomass by pyrolysis. Environ Technol Innov 30:103071. https://doi.org/10.1016/j.eti.2023.103071

Idakwo G, Luttrell J, Chen M, Hong H, Zhou Z, Gong P, Zhang C (2018) A review on machine learning methods for in silico toxicity prediction. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 36(4):169–191

Jiang J, Yuan M, Xu R, Bish DL (2015) Mobilization of phosphate in variable-charge soils amended with biochars derived from crop straws. Soil Tillage Res 146:139–147

Jin Y, Liang X, He M, Liu Y, Tian G, Shi J (2016) Manure biochar influence upon soil properties, phosphorus distribution and phosphatase activities: a microcosm incubation study. Chemosphere 142:128–135. https://doi.org/10.1016/j.chemosphere.2015.07.015

Jin Z, Chen C, Chen X, Jiang F, Hopkins I, Zhang X, Han Z, Billy G, Benavides J (2019) Soil acidity, available phosphorus content, and optimal biochar and nitrogen fertilizer application rates: a five-year field trial in upland red soil, China. Field Crops Res 232:77–87. https://doi.org/10.1016/j.fcr.2018.12.013

Jorner K, Brinck T, Norrby P-O, Buttar D (2021) Machine learning meets mechanistic modelling for accurate prediction of experimental activation energies. Chem Sci 12(3):1163–1175

Khan MN, Li D, Shah A, Huang J, Zhang L, Núñez-Delgado A, Han T, Du J, Ali S, Sial TA, Lan Z, Hayat S, Song Y, Bai Y, Zhang H (2022) The impact of pristine and modified rice straw biochar on the emission of greenhouse gases from a red acidic soil. Environ Res 208:112676. https://doi.org/10.1016/j.envres.2022.112676

Kokol P, Kokol M, Zagoranski S (2022) Machine learning on small size samples: a synthetic knowledge synthesis. Sci Prog (Lond) 105(1):00368504211029777

Lee JW, Kidder M, Evans BR, Paik S, Buchanan IA, Garten CT, Brown RC (2010) Characterization of biochars produced from cornstovers for soil amendment. Environ Sci Technol 44(20):7970–7974

Lei C, Lu T, Qian H, Liu Y (2023) Machine learning models reveal how biochar amendment affects soil microbial communities. Biochar 5(1):89

Li S, Tasnady D (2023) Biochar for soil carbon sequestration: current knowledge, mechanisms, and future perspectives. C 9(3):67

Li F, Liang X, Niyungeko C, Sun T, Liu F, Arai Y (2019) Chapter Two—Effects of biochar amendments on soil phosphorus transformation in agricultural soils. In: Sparks DL (ed) Advances in agronomy. Academic Press, Cambridge, pp 131–172

Li H, Li Y, Xu Y, Lu X (2020a) Biochar phosphorus fertilizer effects on soil phosphorus availability. Chemosphere 244:125471. https://doi.org/10.1016/j.chemosphere.2019.125471

Li H, Yang Z, Dai M, Diao X, Dai S, Fang T, Dong X (2020b) Input of Cd from agriculture phosphate fertilizer application in China during 2006–2016. Sci Total Environ 698:134149. https://doi.org/10.1016/j.scitotenv.2019.134149

Li Y, Rahardjo H, Satyanaga A, Rangarajan S, Lee DT-T (2022) Soil database development with the application of machine learning methods in soil properties prediction. Eng Geol 306:106769. https://doi.org/10.1016/j.enggeo.2022.106769

Li J, Pan L, Huang Y, Liu X, Ye Z, Wang Y (2024) Biochar design for antibiotics adsorption via a hybrid machine-learning-based optimization framework. Sep Purif Technol 348:127666. https://doi.org/10.1016/j.seppur.2024.127666

Liang Q, Liu Y, Chen M, Ma L, Yang B, Li L, Liu Q (2020) Optimized preparation of activated carbon from coconut shell and municipal sludge. Mater Chem Phys 241:122327. https://doi.org/10.1016/j.matchemphys.2019.122327

Liu Y, Wang Y, Zhang J (2012) 'New machine learning algorithm: Random forest' Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14–16, 2012. Proceedings 3. Springer, 246–252

Liu Y, Zhu Z, He X, Yang C, Du Y, Huang Y, Su P, Wang S, Zheng X, Xue Y (2018) Mechanisms of rice straw biochar effects on phosphorus sorption characteristics of acid upland red soils. Chemosphere 207:267–277. https://doi.org/10.1016/j.chemosphere.2018.05.086

Liu Q, Liu B, Zhang Y, Hu T, Lin Z, Liu G, Wang X, Ma J, Wang H, Jin H (2019) Biochar application as a tool to decrease soil nitrogen losses (NH_3 volatilization, N_2O emissions, and N leaching) from croplands: options and mitigation strength in a global perspective. Glob Change Biol 25(6):2077–2093

Lu J, Liu S, Chen W, Meng J (2023a) Study on the mechanism of biochar affecting the effectiveness of phosphate solubilizing bacteria. World J Microbiol Biotechnol 39(3):87

Lu X, Guo W, Wang B, Feng Y, He S, Xue L (2023b) Screening optimal preparation conditions of low-cost metal-modified biochar for phosphate adsorption and unraveling their influence on adsorption performance. J Clean Prod 425:138927. https://doi.org/10.1016/j.jclepro.2023.138927

Luo L, Wang J, Lv J, Liu Z, Sun T, Yang Y, Zhu Y-G (2023) Carbon sequestration strategies in soil using biochar: advances, challenges, and opportunities. Environ Sci Technol 57(31):11357–11372

Ma J, Zhang S, Liu X, Wang J (2023) Machine learning prediction of biochar yield based on biomass characteristics. Bioresour Technol 389:129820. https://doi.org/10.1016/j.biortech.2023.129820

Maroušek J, Gavurová B (2022) Recovering phosphorous from biogas fermentation residues indicates promising economic results. Chemosphere 291:133008. https://doi.org/10.1016/j.chemosphere.2021.133008

Maroušek J, Minofar B, Maroušková A, Strunecký O, Gavurová B (2023) Environmental and economic advantages of production and application of digestate biochar. Environ Technol Innov 30:103109. https://doi.org/10.1016/j.eti.2023.103109

Maroušek J, Gavurová B, Maroušková A (2024) Cost breakdown indicates that biochar production from microalgae in Central Europe requires innovative cultivation procedures. Energy Nexus 16:100335

Melese A, Gebrekidan H, Yli-Halla M, Yitaferu B (2015) Phosphorus status, inorganic phosphorus forms, and other physicochemical properties of acid soils of Farta district, Northwestern highlands of Ethiopia. Appl Environ Soil Sci 2015(1):748390

Minofar B, Milčić N, Maroušek J, Gavurová B, Maroušková A (2025) Understanding the molecular mechanisms of interactions between biochar and denitrifiers in N_2O emissions reduction: pathway to more economical and sustainable fertilizers. Soil Tillage Res 248:106405. https://doi.org/10.1016/j.still.2024.106405

Nahidan S, Ghasemzadeh M (2022) Biochemical phosphorus transformations in a calcareous soil as affected by earthworm, cow manure and its biochar additions. Appl Soil Ecol 170:104310. https://doi.org/10.1016/j.apsoil.2021.104310

Novak JM, Johnson MG, Spokas KA (2018) Concentration and release of phosphorus and potassium from lignocellulosic-and manure-based biochars for fertilizer reuse. Front Sustain Food Syst 2:54

Otse EJ, Obunadike GN, Abubakar A (2025) Linear regression approach to solving multicollinearity and overfitting in predictive analysis. J Sci Res Rev 2(1):108–117

Palansooriya KN, Li J, Dissanayake PD, Suvarna M, Li L, Yuan X, Sarkar B, Tsang DC, Rinklebe J, Wang X (2022) Prediction of soil heavy metal immobilization by biochar using machine learning. Environ Sci Technol 56(7):4187–4198

Peng Y, Chen Q, Guan C-Y, Yang X, Jiang X, Wei M, Tan J, Li X (2023) Metal oxide modified biochars for fertile soil management: effects on soil phosphorus

transformation, enzyme activity, microbe community, and plant growth. Environ Res 231:116258
Penn CJ, Camberato JJ (2019) A critical review on soil chemical processes that control how soil pH affects phosphorus availability to plants. Agriculture 9(6):120
Raut P, Dani A (2020) 'Correlation between number of hidden layers and accuracy of artificial neural network' Advanced Computing Technologies and Applications: Proceedings of 2nd International Conference on Advanced Computing Technologies and Applications—ICACTA 2020. Springer, 513–521
Sakhiya AK, Vijay VK, Kaushal P (2022) Efficacy of rice straw derived biochar for removal of Pb^{2+} and Zn^{2+} from aqueous: adsorption, thermodynamic and cost analysis. Bioresour Technol Rep 17:100920
Schneider F, Haderlein SB (2016) Potential effects of biochar on the availability of phosphorus — mechanistic insights. Geoderma 277:83–90. https://doi.org/10.1016/j.geoderma.2016.05.007
Shaheen SM, Ullah H, Wu Y, Mosa A, Fang Y, Shi Y, Liu J, Kumar M, Zhang H, Zhang B (2025) Remediation of emerging inorganic contaminants in soils and water using pristine and engineered biochar: a review. Biochar 7(1):34
Shan A, Pan J, Kang KJ, Pan M, Wang G, Wang M, He Z, Yang X (2021) Effects of straw return with N fertilizer reduction on crop yield, plant diseases and pests and potential heavy metal risk in a Chinese rice paddy: a field study of 2 consecutive wheat-rice cycles. Environ Pollut 288:117741. https://doi.org/10.1016/j.envpol.2021.117741
Shi W, Fang YR, Chang Y, Xie GH (2023) Toward sustainable utilization of crop straw: greenhouse gas emissions and their reduction potential from 1950 to 2021 in China. Resour Conserv Recycl 190:106824. https://doi.org/10.1016/j.resconrec.2022.106824
Singh D, Singh B (2022) Feature wise normalization: an effective way of normalizing data. Pattern Recogn 122:108307. https://doi.org/10.1016/j.patcog.2021.108307
Soinne H, Hovi J, Tammeorg P, Turtola E (2014) Effect of biochar on phosphorus sorption and clay soil aggregate stability. Geoderma 219–220:162–167. https://doi.org/10.1016/j.geoderma.2013.12.022
Stávková J, Maroušek J (2021) Novel sorbent shows promising financial results on P recovery from sludge water. Chemosphere 276:130097. https://doi.org/10.1016/j.chemosphere.2021.130097
Stefko R, Michalikova KF, Strakova J, Novak A (2025) Digital twin-based virtual factory and cyber-physical production systems, collaborative autonomous robotic and networked manufacturing technologies, and enterprise and business intelligence algorithms for industrial metaverse. Equilibrium (1689-765X) 20(1):389
Sun Y, Zhang Y, Lu L, Wu Y, Zhang Y, Kamran MA, Chen B (2022) The application of machine learning methods for prediction of metal immobilization remediation by biochar amendment in soil. Sci Total Environ 829:154668. https://doi.org/10.1016/j.scitotenv.2022.154668
Tesfaye F, Liu X, Zheng J, Cheng K, Bian R, Zhang X, Li L, Drosos M, Joseph S, Pan G (2021) Could biochar amendment be a tool to improve soil availability and plant uptake of phosphorus? A meta-analysis of published experiments. Environ Sci Pollut Res 28(26):34108–34120
Thomas A J, Petridis M, Walters SD, Gheytassi SM, Morgan R E (2017) 'Two hidden layers are usually better than one' Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings. Springer, 279–290
Vu DH, Muttaqi KM, Agalgaonkar AP (2015) A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. Appl Energy 140:385–394. https://doi.org/10.1016/j.apenergy.2014.12.011
Wang Y, Xiao B, Bi X, Li W, Zhang J, Ma X (2019) 'Prediction of sepsis from clinical data using long short-term memory and extreme gradient boosting' 2019 Computing in Cardiology (CinC). IEEE, 1–4
Wang Y, Zhang Y, Zhao H, Hu W, Zhang H, Zhou X, Luo G (2022) The effectiveness of reed-biochar in mitigating phosphorus losses and enhancing microbially-driven phosphorus dynamics in paddy soil. J Environ Manage 314:115087
Wang Z, Hill R, Williams G, Dwyer GS, Hu J, Schnug E, Bol R, Sun Y, Coleman DS, Liu X-M, Sandstrom MR, Vengosh A (2023) Lead isotopes and rare earth elements geochemistry of global phosphate rocks: insights into depositional conditions and environmental tracing. Chem Geol 639:121715. https://doi.org/10.1016/j.chemgeo.2023.121715
Wang M, Xie Y, Gao Y, Huang X, Chen W (2024a) Machine learning prediction of higher heating value of biochar based on biomass characteristics and pyrolysis conditions. Bioresource Technol 395:130364. https://doi.org/10.1016/j.biortech.2024.130364
Wang X, Gao Y, Hou J, Yang J, Smits K, He H (2024b) Machine learning facilitates connections between soil thermal conductivity, soil water content, and soil matric potential. J Hydrol 633:130950. https://doi.org/10.1016/j.jhydrol.2024.130950
Wendimu A, Yoseph T, Ayalew T (2023) Ditching phosphatic fertilizers for phosphate-solubilizing biofertilizers: a step towards sustainable agriculture and environmental health. Sustainability 15(2):1713
Wetterslev J, Jakobsen JC, Gluud C (2017) Trial sequential analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 17:1–18
Xu G, Sun J, Shao H, Chang SX (2014) Biochar had effects on phosphorus sorption and desorption in three soils with differing acidity. Ecol Eng 62:54–60. https://doi.org/10.1016/j.ecoleng.2013.10.027
Xu X, Li T, Cheng K, Yue Q, Pan G (2024) Geographical differences in the effect of biochar on crop yield and greenhouse gas emissions–a global simulation based on a machine learning model. Curr Res Environ Sustain 7:100239
Yang L, Wu Y, Wang Y, An W, Jin J, Sun K, Wang X (2021) Effects of biochar addition on the abundance, speciation, availability, and leaching loss of soil phosphorus. Sci Total Environ 758:143657. https://doi.org/10.1016/j.scitotenv.2020.143657
Yu F, Chen Y, Huang X, Shi J, Xu J, He Y (2023) Does straw returning affect the root rot disease of crops in soil? A systematic review and meta-analysis. J Environ Manage 336:117673. https://doi.org/10.1016/j.jenvman.2023.117673
Zhang J, Liu J, Liu R (2015) Effects of pyrolysis temperature and heating time on biochar obtained from the pyrolysis of straw and lignosulfonate. Bioresource Technol 176:288–291. https://doi.org/10.1016/j.biortech.2014.11.011
Zhang H, Chen C, Gray EM, Boyd SE, Yang H, Zhang D (2016) Roles of biochar in improving phosphorus availability in soils: a phosphate adsorbent and a source of available phosphorus. Geoderma 276:1–6. https://doi.org/10.1016/j.geoderma.2016.04.020
Zhang X, Zhang P, Yuan X, Li Y, Han L (2020) Effect of pyrolysis temperature and correlation analysis on the yield and physicochemical properties of crop residue biochar. Bioresource Technol 296:122318. https://doi.org/10.1016/j.biortech.2019.122318
Zhang R, Li Y, Goh ATC, Zhang W, Chen Z (2021) Analysis of ground surface settlement in anisotropic clays using extreme gradient boosting and random forest regression models. J Rock Mech Geotech Eng 13(6):1478–1484. https://doi.org/10.1016/j.jrmge.2021.08.001
Zhang W, Cho Y, Vithanage M, Shaheen SM, Rinklebe J, Alessi DS, Hou C-H, Hashimoto Y, Withana PA, Ok YS (2022) Arsenic removal from water and soils using pristine and modified biochars. Biochar 4(1):55
Zhang Y, Feng Y, Ren Z, Zuo R, Zhang T, Li Y, Wang Y, Liu Z, Sun Z, Han Y, Feng L, Aghbashlo M, Tabatabaei M, Pan J (2023) Tree-based machine learning model for visualizing complex relationships between biochar properties and anaerobic digestion. Bioresour Technol 374:128746. https://doi.org/10.1016/j.biortech.2023.128746
Zhu J, Li M, Whelan M (2018) Phosphorus activators contribute to legacy phosphorus availability in agricultural soils: a review. Sci Total Environ 612:522–537. https://doi.org/10.1016/j.scitotenv.2017.08.095
Zhu X, Wang X, Ok YS (2019) The application of machine learning methods for prediction of metal sorption onto biochars. J Hazard Mater 378:120727. https://doi.org/10.1016/j.jhazmat.2019.06.004