

Machine-learning prediction of freezing temperature in Pb-contaminated soils remediated with biochar

Laila Safi

Received: 8 July 2026 / Accepted: 16 October 2026

© The Author(s) 2026

Abstract

Throughout the Pb-contaminated soil rehabilitation procedure, biochar may change the structure of the soil, which has a significant impact on soil freezing. This study examined the change in freezing temperature (ToF) during the biochar remediation of Pb-contaminated soils using machine learning techniques, such as Light Gradient Boosting (LGB), Extreme Gradient Boosting (XGB), Extra Tree Regression (ETR) and Random Forest Regression (RFR). The data used in this research is derived from previously conducted trials. The system's reliability is significantly impacted by hyperparameters, which must be selected using metaheuristic optimization approaches. To accomplish this, a variety of algorithms are used, including the Catch Fish Optimization (CFO). The information, which includes 360 ToF data points, came from studies that were submitted for publication. After studying each element, the study group came to the decision that 80% (288) of the dataset would be utilized for training and model construction, while the other 20% (72) of the dataset would be utilized for validating the model. Randomization, normalization, cross-validation, scaling, and feature significance evaluation were some of the pre-processing methods used to the dataset. The prediction efficacy of ETR(CFO) and RFR(CFO) in predicting ToF was assessed using metrics. With the data, ToF can be reliably predicted by both ETR(CFO) and RFR(CFO) over other baseline models. The assessment process assessed the reliability and plausibility of each model using evaluation criteria, logical analysis, ranking score, and model difference percentage. In achieving its primary target, the RFR(CFO) model performs better than the alternative model.

Keywords Soil · Biochar · Freezing Temperature · Pb-contaminated · Remediation · Estimation

Introduction

Literature review

Frozen soil refers to all types of rocks and soils with ice at or below 0 °C. As a result, it is very vulnerable to temperature fluctuations^{1,2}. The freezing temperature is the maximum temperature at which liquid water in the soil converts to ice. It is commonly employed as a physical indicator to determine whether the ground has frozen³. The freezing temperature profoundly influences the assessment of unfrozen water content, the calculation

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



Scientific Reports

<https://doi.org/10.1038/s41598-026-77024-4>

Published online: 24 October 2026

scientific reports

of water migration, and the analysis of frost heave and thaw settlement in frozen soil⁴. Soil freezing denotes the solidification of pore and bonded water within the soil matrix. Distilled water generally freezes at 0 °C; however, due to the effects of soil particle surface energy and a minor concentration of solute in water, the freezing point of soil frequently descends below 0 °C. The difference in freezing temperatures between soil and distilled water is known as freezing point depression⁵. At same ambient temperatures, the presence of heavy metals alters the soil structure. This might alter the dynamic viscosity of the pore fluid, changing the amount of unfrozen water in the soil^{6,7}. The freezing point of the soil drops when the concentration of heavy metals reaches a certain level^{8,9}. As ions in unfrozen water, heavy metal elements in soil mostly move to the lower soil profile, where they accumulate in deeper soil layers¹⁰. Investigating the freezing temperature of heavy metal-contaminated soil and its affecting factors is thus very important from a practical standpoint^{11–13}.

International scholars first investigated the formation process of freezing temperatures using thermodynamic theory^{14,15}. The results demonstrate that the main parameters affecting the freezing temperature are the free water content, void ratio, and soil surface properties^{16–18}. Several researches have conducted relevant experimental and theoretical studies on soil freezing temperatures. They considered strategies to observe subzero temperatures and the factors that influence them. Some scholars have examined the relationship between freezing temperature and the water content and soluble salts in Lanzhou loess¹⁹. The results demonstrate that the freezing point of the soil is closely correlated with its moisture and salinity levels. Additional research have examined the relationship between freezing temperatures and various soil types, salinity, salt concentration, and moisture content^{20,21}. The results demonstrate that elevated salt concentration reduces the freezing temperature of the soil, but increased water content elevates it. The quantity of soluble salts in the soil substantially influences its freezing point. Sodium sulfate crystallizes more efficiently at lower freezing points^{22–24}. Additionally, they performed tests to examine the freezing point of saline silty soil^{25,26}. The freezing temperature of saline silty soil is influenced by the free energy at the ice-water contact and pore radius. The initial freezing temperature of saline soil is greatly influenced by the properties of the salt solution. They found a relationship between the pore radius in saline soil and the free energy of the ice-water interaction.

ML is a strong field of artificial intelligence that lets computers learn from data and make accurate predictions without programming^{27–30}. It can be used in environmental engineering and other sciences with complicated, non-linear interactions^{31–34}. ML methods have proved successful in predicting heavy metal behavior and characteristics in environmental soils^{35–37}. These techniques outperform statistical methods in accuracy and efficiency^{38,39}. Thus, ML is essential for analyzing soil heavy metal pollution, mobility, and ecological hazards. Whether on-site or off-site, treating contaminated soil presents several obstacles, including type, longevity, source of contaminants, and high treatment costs. Organic and inorganic soil pollutants affected non-target soil organisms and humans via the food chain^{40–44}. Statistical methods and artificial neural network (ANN) simulation were used to assess heavy metal-contaminated soils. Self-organizing map (SOM), simple regression, and analysis of variances assessed soil quality and predicted ecological status. An ANN algorithm with contamination factor (CF) as target and heavy metal concentrations as inputs showed a strong $R=0.9974$. Statistical methods and ANN simulation can analyze and forecast soil contamination, allowing soil management decisions⁴⁵. Another study examined the use of graph neural network (GNN) technology to forecast heavy metals by extracting and learning information from large-scale networks. This research developed a rice heavy metal prediction model, *CoNet – GNN*, combining the co-occurrence network and GNN with 17 environmental parameters as input variables. Using a field research dataset, *CoNet – GNN* predicted with remarkable R^2 values of 0.872, 0.711, 0.683, 0.489, and 0.824. *CoNet – GNN* outperformed random forest, gradient boosting, and multilayer perceptron in all heavy metal prediction⁴⁶. Hyper-spectral technique was tested to quantify soil heavy metal concentration (Cd, Cr, and Pb) during phytoremediation. 108 surface soil samples (0–20 cm) were taken. Partial least squares regression (PLSR) and a mind evolutionary algorithm-optimized back propagation neural network developed inversion models. The *MEA – BP* neural network model enhanced estimate R^2 from 0.728, 0.737, and 0.675 to 0.873, 0.884, and 0.857⁴⁷. A database comprising contaminated and uncontaminated soil specimens with a broad range of soil water content, density, and pore-fluid

salinity is created. Two *ANN* models are constructed to (i) detect heavy metals in soil and (ii) identify their kind based on soil specimen complex permittivity. The first *ANN* model (identification) detects heavy metals 90% of the time. Second *ANN* model (classification) properly classifies heavy metals 95% of the time⁴⁸.

Main contribution

The earlier experimental study conducted by Han et al. offered significant understanding on the fluctuation of freezing temperature in the treatment of Pb-contaminated soils utilising biochar⁴⁹. Nonetheless, that research predominantly concentrated on empirically delineating the impact of singular variables and determining the established correlations among Pb contamination, biochar supplementation, and freezing dynamics. While experimental methods are crucial for comprehending soil freezing processes, they are typically constrained by the variety of conditions examined, the expense of experimentation, and the challenge of forecasting freezing temperature reactions to novel combinations of contamination levels, biochar amounts, and water conditions. Consequently, the current experimental investigations yield significant physical insights but fail to create a universal prediction model that can swiftly assess freezing temperatures for unexamined situations.

The current research does not assert originality from the exclusive use of Light Gradient Boosting (*LGB*), Extreme Gradient Boosting (*XGB*), Extra Tree Regression (*ETR*) and Random Forest Regression (*RFR*), as these methods are recognised machine learning techniques. The scientific advancement is in creating a prediction framework that utilises optimisation assistance and interpretability, merging the Catch Fish Optimisation Algorithm (*CFO*) with ensemble regression models to estimate the freezing temperature of Pb-contaminated biochar-amended soils. The suggested framework converts empirically derived correlations into a data-driven prediction instrument, facilitating the estimation of freezing behaviour without necessitating further laboratory tests for every conceivable combination of affecting factors.

The current study adds three factors to the previous experimental investigation. First, machine learning models quantify the nonlinear relationships between Pb content, biochar content, water content, and set temperature, allowing prediction beyond prior experimental situations. Second, SHAP study shows how Pb, water, and biochar content affect freezing temperature. This identifies multivariable dominant predicting elements to supplement experimental findings. Third, the proposed approach evaluates uncertainty using numerous statistical variables, including, to measure engineering prediction reliability.

To optimise the hyperparameter configuration of ensemble learning models in a difficult nonlinear soil-freezing scenario, *CFO* was used as an optimisation approach. Grid search and random search use specified parameter combinations and may need several model evaluations to explore high-dimensional parameter spaces. Bayesian optimisation approaches, including Optuna-based ones, efficiently explore parameters but depend on surrogate-model assumptions and acquisition tactics. *CFO* uses adaptive exploration and exploitation processes inspired by fish-searching behaviours to optimise hyperparameter combinations. However, *CFO* is used to examine its capacity to improve soil-freezing prediction performance, not to claim universal supremacy over all optimisation techniques. The optimised models were examined using numerous statistical indicators, cross-validation analysis, uncertainty assessment, and non-parametric statistical testing to confirm that the improvements were significant rather than random. We also used Wilcoxon signed-rank analysis to see if the suggested models and benchmark methods differed statistically. Thus, *CFO*'s improvement was assessed for predicted accuracy and statistical reliability.

Finally, the test samples were not used for model building or hyperparameter optimisation. The model performance was evaluated on a held-out testing subset within the experimental domain represented by the available dataset. The suggested model should be used as a predictive decision-support tool within the explored parameter range because the dataset was obtained from laboratory experiments and has few controlling factors. Future research should include more soil types, mineralogical features, thermal parameters, and field-scale observations to improve generalisation.

Methods

Data collection and pre-processing steps

To properly complete the data analysis and machine learning processes, both feature selection and dataset pre-processing are essential^{50,51}. The interdependence of these procedures is one of the causes of this. In addition to ensuring that the data is in an analysis-ready state, they also ensure that the most relevant qualities are employed when creating models for prediction⁵². This is crucial in addition to ensuring that the data is organized appropriately. Pre-processing the dataset aims to achieve many targets, such as gathering data for analysis, handling missing values, and maintaining consistency. Among the techniques utilized in this approach were data cleaning, randomization, normalization, cross-validation, feature significance assessment, subdivision, and the integration of several data sources. The study additionally utilized feature significance analysis to enhance interpretability, reduce overfitting, optimize model performance, and ultimately reduce computation time. In general, pre-processing and feature selection may significantly improve machine learning model performance by providing cleaner and more relevant data. Additionally, simplified data helps model training and speeds up calculation. This is because information that has been condensed is easier to understand. Better decision-making and more trustworthy information may be made possible by refined data and functional features.

The dataset utilised in this research was obtained from the experimental analysis conducted by Han et al.⁴⁹, which explored the freezing temperature reaction of Pb-contaminated soils treated with biochar. The initial tests assessed the combined effects of heavy metal levels, biochar incorporation, and moisture conditions on the freezing behaviour of soil. The researchers indicated that the concentration of heavy metals had a primary effect on freezing temperature, succeeded by water content and biochar content, and emphasised the interplay between biochar and heavy metals throughout the freezing process. Soil characteristics, treatment factors, and experimental conditions reported in Table 1⁴⁹.

The experimental data was derived from the published findings and organised into a structured dataset for machine-learning evaluation. The extraction process entailed converting the documented freezing temperature data into a digital format and linking each measurement to its relevant experimental parameters. The conclusive dataset comprises 360 observations with four input variables: lead content (L_C), biochar content (B_C), moisture content (W_C), and set temperature (ToS), with freezing temperature (ToF) designated as the predictive objective.

It is important to emphasise that the 360 observations denote experimental circumstances and measurements documented in the original paper, rather than 360 completely separate soil samples. Numerous observations may arise from interconnected treatment groups exhibiting the same or comparable soil composition and environmental circumstances. Consequently, the constructed models are to be regarded mostly as forecasting instruments within the experimental framework defined by the existing information, rather than as universal models suitable

Table 1 Soil characteristics, treatment factors, and experimental conditions⁴⁹.

Item	Information reported
Soil source:	Construction site of a university in Qilihe District, Lanzhou, China
Sampling depth:	40–70 cm below the surface
Soil preparation:	Air-dried, crushed, and passed through a 2 mm sieve
Contaminant:	Lead chloride, added using distilled water
Biochar:	Maize-stover biochar, produced at 500 °C under low-oxygen conditions
Water content:	10%, 15%, 20%, 25%, and 30%
Lead Usage:	0%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 1%, 1.5%, and 2%
Biochar content:	0%, 1%, 3%, 5%, and 7%
Freezing temperatures:	−5, −10, and −15 °C
Sample preparation:	Manual compaction to 90%; two layers of 40 mm each
Temperature recording:	Every 30 s
Freezing duration:	More than 12 h
Thawing:	12 h at room temperature

for all soil conditions. All 360 observations used in this study were digitized/extracted from Han et al.⁴⁹. 80% (288) of the collected data was used for training and model development after each dimension was evaluated, and the remaining 20% (72) was set aside for model validation.

Lead content is the amount of lead present in the sample. This might change how the sample freezes by changing the physical and chemical characteristics of the solution. The quantity of biochar utilized is known as the biochar content, and it is often stated as a percentage of the overall mass. This may affect the material's capacity to transport heat and hold onto water. The water content—the amount of water in the material—directly affects the freezing point by changing the amount of moisture available and the effects of phase transitions. Set temperature (ToS) represents the externally controlled ambient/freezing temperature imposed during the soil-freezing experiment. The environmental temperature was controlled at -5 , -10 , and -15 °C during the freezing tests. Thus, ToS should be interpreted as the prescribed ambient temperature of the freezing environment, rather than the initial temperature of the specimen, a cooling step, or the measured freezing temperature of the soil. It controls the temperature and cooling rate before to freezing.

The correlation among two grading criteria is measured using Kendall analysis, sometimes known as Kendall's tau. The effectiveness of a monotonic function in capturing the connection among the variables is evaluated. Instead of looking at values, Kendall's tau looks at the ordinal relationships that exist between two variables. In contrast, the linear connections among the variables are examined via Pearson's correlation. On the other hand, Pearson's correlation is used to examine linear correlations between variables. The following two are examples of potential values for the coefficient:

- $+1$ correlation: Positive numbers show that the two variables have a positive relationship. This implies that an increase in one parameter often leads to an increase in another.
- -1 correlation: A negative correlation, or relationship, among a decline in one measure and a gain in another is shown by negative numbers.

This is true when comparing the two variables. The Kendall analysis is shown in Fig. 1. Since specific variables could show low correlation values, a debate is necessary. As a result, the model-building procedure that makes use of these inputs has to be carried out as precisely as possible. According to the matrix, L_c/ToF has the sole negative correlation (-0.39), whereas ToS/ToF has the strongest positive correlation (0.8).

Table 2 displays a comprehensive statistical overview of the dataset, consisting of input attributes and the target parameter. Figure 2 also displays the depiction of a thorough relation comparison of the input and target parameters (Fig. 2a illustrates the distribution of L_c , while Figs. 2(b–d) show the corresponding distributions of B_c , W_c , and ToF , respectively). The forecast models' stability and generalizability were improved by the application

Fig. 1 Kendall-based correlation analysis highlighting inter-variable relationships

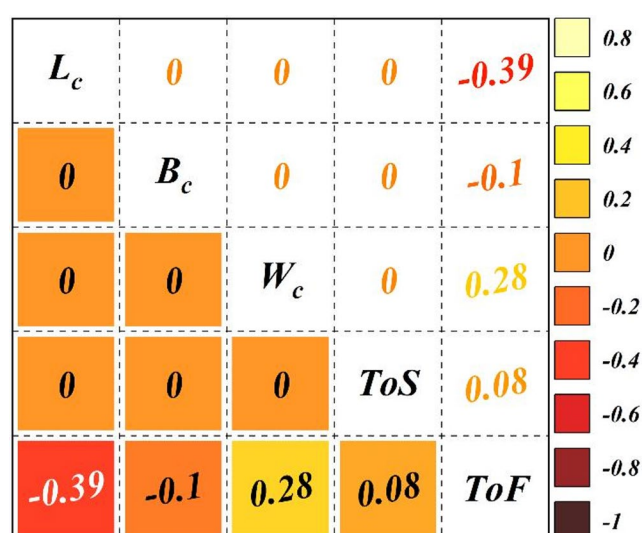
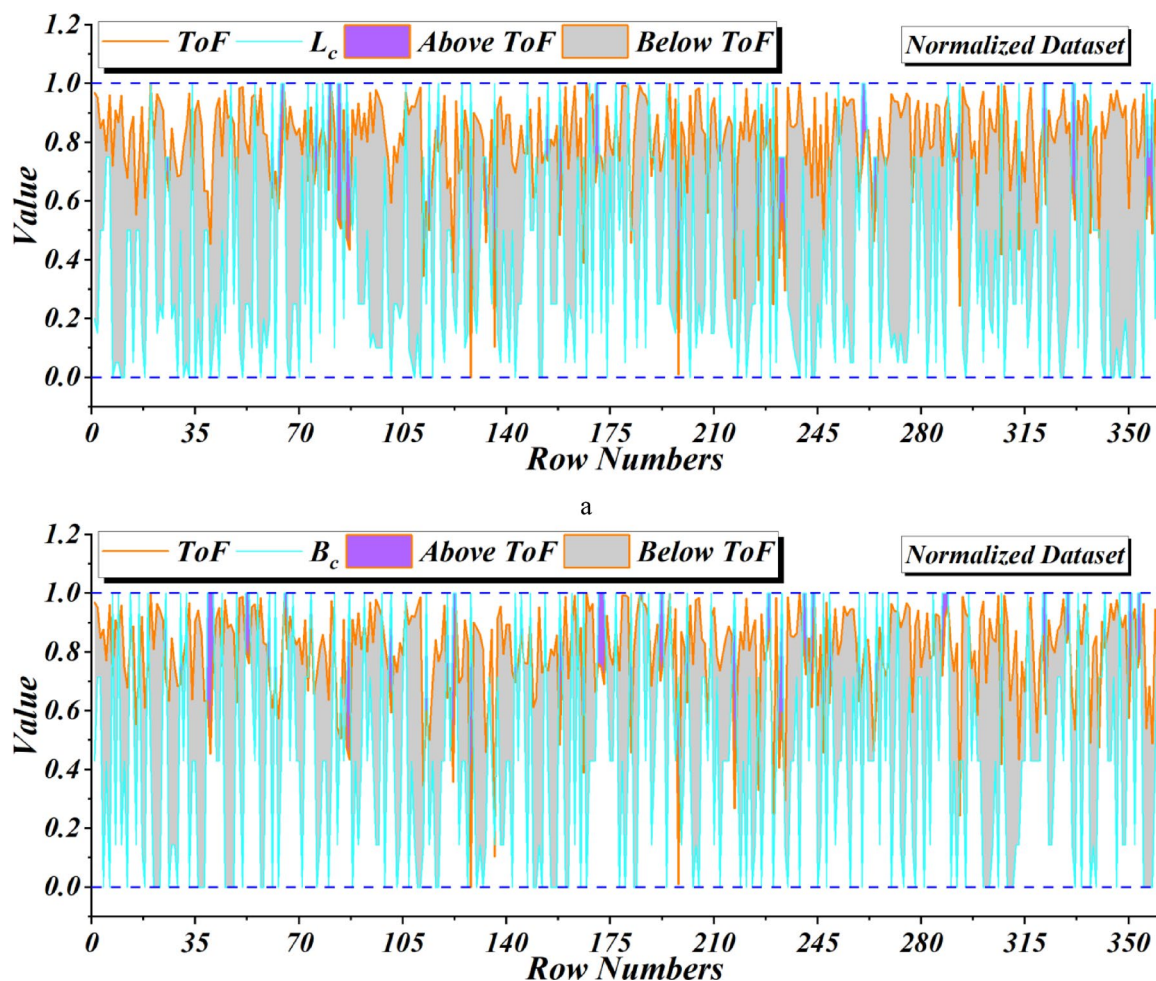


Table 2 Statistical properties of the dataset for input features and target variable.

Category	Statistics								
	Min.	Max.	Range (Max.-Min.)	St.D.	Var.	Skew.	Kurt.	Avg.	Med.
Input 1: Lead content, L_c, %									
Training portion	0	2	2	0.691	0.477	0.541	-1.144	0.799	0.5
Testing portion	0	2	2	0.691	0.477	0.628	-0.906	0.761	0.5
Input 2: Biochar content, B_c, %									
Training portion	0	7	7	2.683	7.201	0.169	-1.436	3.295	3
Testing portion	0	7	7	2.663	7.092	0.334	-1.346	3.069	3
Input 3: Water content, w_c, %									
Training portion	10	30	20	6.128	37.550	0.028	-1.003	19.809	20
Testing portion	10	30	20	6.047	36.569	-0.108	-0.927	20.764	20
Input 4: Set temperature, To_s, °C									
Training portion	-15	-5	10	2.517	6.334	0.020	0.982	-9.948	-10
Testing portion	-15	-5	10	2.421	5.859	-0.111	1.394	-10.208	-10
Target: Freezing temperature, To_F, °C									
Training portion	-4.12	0.011	4.131	0.699	0.489	-1.668	3.949	-0.842	-0.693
Testing portion	-3.12	-0.04	3.08	0.685	0.470	-1.136	0.769	-0.836	-0.659

**Fig. 2** Comprehensive relation comparison of input and target variable (a) distribution of L_c ; (b) distribution of B_c ; (c) distribution of w_c ; and (d) distribution of To_F

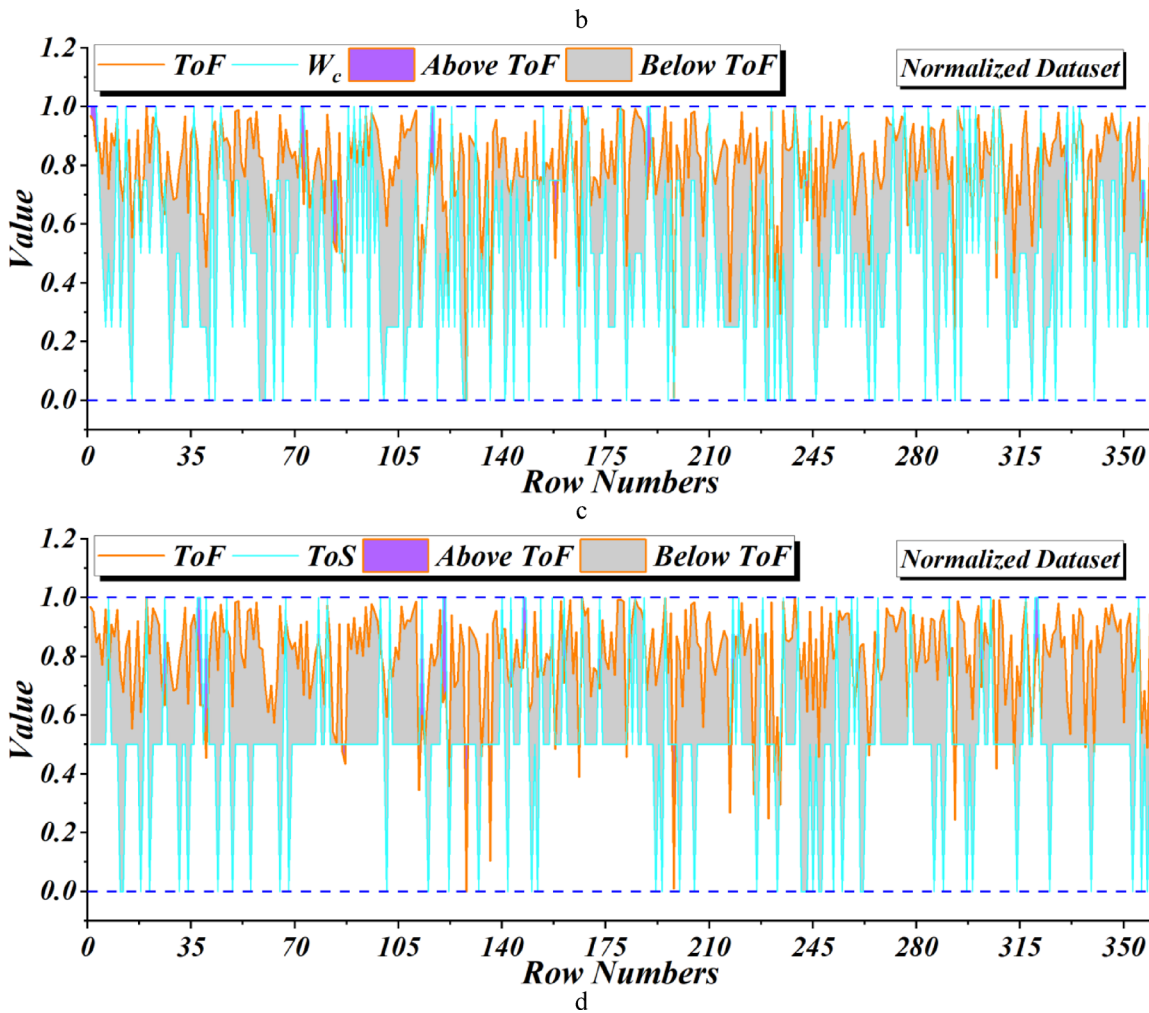


Fig. 2 (continued)

of randomization. To avoid overfitting and provide an unbiased assessment of the model, the dataset was divided into groups for training and testing at random. To enhance model performance, randomization was utilized throughout cross-validation and hyperparameter adjustment. Normalization ensures that all input features are normalized to a consistent scale, which helps to enhance models like as *ETR* and *RFR* that are sensitive to feature magnitude. It lessens the influence of characteristics with bigger magnitudes during model training and facilitates the precise determination of each feature's importance by normalizing each feature to a fixed range. Because the data was sorted and filtered, the methods were able to concentrate on finding pertinent patterns, which improved accuracy in prediction. Enhancing numerical stability by feature normalization accelerated training, especially when using *CFO* for hyperparameter optimization. Furthermore, the combination of normalized data and feature visualization enhanced comprehension and interpretation of model predictions, offering more accurate insights into the model's dynamics. The particular dataset was modified using Eq. (1) to improve dependability while remaining between 0 and 1.

$$g_j = \frac{g_i - g_{min}}{g_{max} - g_{min}} \quad (1)$$

G_{max} : the highest values of certain variable

G_{min} : the lowest values of certain variable

G_j : the feature's normalized value

G_i : the non-normalized value

Optimization and estimation algorithms

Catch fish optimization algorithm (CFO)

This study's suggested CFO draws inspiration from the traditional and straightforward fishing method. The fundamental idea of CFO is to use cooperative group efforts to catch as many fish as feasible. People may use gear and work together to fish in diverse bodies of water while telling each other about their catches. They may choose the most advantageous fish collecting locations and the finest fishing routes based on their knowledge. Below is a description of the precise implementation framework⁵³:

- The capacity for reasoning and using awareness varies between humans and animals. In order to find the best fishing locations, fishermen intentionally create favorable conditions in the water by using their knowledge and skills.
- The use of tools and teamwork are beneficial to humans. Fishermen work together, using nets to encircle and catch fish, significantly improving their catch rate.
- By forcing fish into a small circle, collaborative encirclement enables fishermen to optimize their harvest.

In summary, the CFO is implemented in two stages: the exploration phase, which comprises dispersed exploration for benefits and cooperative fishing, and the exploitation stage, which uses coordinated encirclement efforts to capture collective power. The metaheuristic algorithm explores the global scenario and ultimately determines the best course of action through continuous location updates via search agents. Each fisherman is a search agent in the following CFO model:

$$Fisher = \begin{bmatrix} Fisher_{1,1} & Fisher_{1,2} & \dots & Fisher_{1,d} \\ Fisher_{2,1} & Fisher_{2,2} & \dots & Fisher_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ Fisher_{N,1} & Fisher_{N,2} & \dots & Fisher_{N,d} \end{bmatrix}_{N \times d} \quad (2)$$

The resetting formulation of $Fisher$, a full matrix that contains the locations of N search agents in a dimensional space, is as follows:

$$Fisher_{i,j} = (ub_j - lb_j) \times r + lb_j \quad (3)$$

The j_{th} dimension's upper and lower bounds are represented by the variables ub_j and lb_j , the location of the i_{th} fisherman in the j_{th} dimension is represented by $Fisher_{i,j}$, and r is a random number between 0 and 1. Using the fitness value evaluation function $fobj$ and each fisherman's current position data, it calculates their fitness values. This results in the following fitness value matrix:

$$fit = fobjective(Fisher) = \begin{bmatrix} fit_1 \\ fit_2 \\ \vdots \\ Fit_N \end{bmatrix} \quad (4)$$

Fit_1 in the fitness matrix represents the fitness value of the first fisherman, fit_2 for the second fisherman, and so on. We use 0.5 to split the percentage of exploitation and exploration over the iterations. In the process, individuals look at the global during the first half of the phase (when $EFs/MaxEFs < 0.5$) and utilize it during the second half (when $EFs/MaxEFs > 0.5$).

CFO was chosen as a derivative-free global optimisation approach for calibrating the hyperparameters of the models. This selection is especially pertinent due to the nonlinear nature of the hyperparameter-performance correlation in ensemble tree models, which may exhibit uneven or non-smooth areas, rendering gradient-based

optimisation unsuitable. The *CFO* utilises two synergistic search strategies: an initial exploration phase characterised by widespread search and a later exploitation phase focused on collaboratively around viable solutions. This exploration–exploitation framework enables the algorithm to navigate a constrained hyperparameter space without necessitating analytical gradients or presuppositions about the nature of the goal function.

Extra tree regressor (*ETR*)

An enhancement to ensemble learning, the Extra Tree Regressor (*ETR*) builds on the Random Forest (*RF*) model and represents a substantial step forward. A collection of unpruned trees of regression, each constructed by utilizing a traditional top-down process, are used by the *ETR* technique. The *RF* model uses a two-step process for regression analysis called bagging and bootstrapping, which is different from the outlined technique. When developing individual trees, the *ETR* model uses a deterministic splitting technique. *ETR* randomly chooses a split point for every feature and then chooses the most effective split among these alternatives, in comparison with *RF*, which uses a selection method to find the optimum split from a random collection of features at each node. Equation (5) provides the following mathematical depiction of this process:

$$Split_{ETR} = \underset{f,s}{\operatorname{argmin}} [Error(f,s)] \quad (5)$$

In this case, the *ETR* algorithm's chosen split is shown by the variable $Split_{ETR}$. A feature is represented by the symbol f , and a randomly selected divided point for that feature is represented by the symbol s . The error reduction brought about by the split is calculated by the function $Error(f,s)$. The doublet of f and s that reduces this mistake is selected by the algorithm. It is usual practice to take an average of the votes cast by each tree in the ensemble during the bagging stage of the *RF* algorithm in order to arrive at the final prediction. Using a more diverse ensemble of unpruned trees, the *ETR* technique follows a similar process. In a nutshell, Eq. (6) is the mathematical expression of the complete *ETR* model's output:

$$Y_{ETR} = \frac{1}{N} \sum_{i=1}^N T_i(X) \quad (6)$$

The total amount of trees in the ensemble (represented by N), the i_{th} tree in the ensemble (represented by T_i), and the input feature vector (represented by X) all affect the expected output, which is represented by Y_{ETR} . By combining mathematical elements, the *ETR* algorithm provides a unique regression approach that strikes a compromise between accuracy and unpredictability.

Random forest regression (*RFR*)

Bagging and the random subspace approach are combined in the Random Forest (*RF*) algorithm, which is an ensemble learning technique. The basis classifiers in *RF* are decision trees (*DT*), which are very simple structures but might suffer from overfitting when input variables are complicated. Both the generalizability and classification performance of the *RF* algorithm are much better than those of conventional *DT* approaches. There are two ways in which the *RF* algorithm's unpredictability shows itself:

Initial way: The *RF* model has K *DT* as basis classifiers, which are derived from the out-ofbag (*OOB*) dataset, which is first randomly selected using the bootstrap approach to train a choice tree for each of the recently added data sets.

Second way: When *DT* are being built, a certain amount of features are chosen at random from all the initial features (N).

In a classification issue, each *DT* produces a classification outcome, and then the K classification outcomes are aggregated via voting to obtain the final *RF* classification outcome. The following is the *RF* model's ultimate decision:

$$H(X) = \operatorname{argmax} \sum_{i=1}^K I(h_i(X) = Y) \quad (7)$$

The vector that is input is X , and the right classification vector is Y . The Random Forest model, which is made up of DT , is symbolized by $H(X)$, and each DT is denoted by h_i . The amount of DT in the framework is denoted by K .

Implementation details of *CFO*-based hyperparameter optimization

To affirm the repeatability of the proposed optimization-assisted machine learning framework, this section fully outlines the implementation specifics of the *CFO*. This investigation uses *CFO* as a hyperparameter optimisation approach to ascertain the most effective configurations of the *RFR* and *ETR* models. The optimisation was executed entirely on the training dataset, with the independent validation dataset omitted from the hyperparameter selection procedure.

The *CFO*'s goal function was established according to the prediction error of the optimised model. The *RMSE* derived from cross-validation on the training dataset was reduced. A reduced fitness value signifies a superior amalgamation of model hyperparameters. In each iteration, the *CFO* produced potential solutions within established parameter limits, assessed their fitness values, and modified the population based on the algorithm's exploration and exploitation strategies. The optimised parameters for the *RFR* and *ETR* models were chosen according to their impact on model complexity, generalisation capacity, and predictive stability. The refined parameters encompassed the quantity of decision trees, the utmost depth of trees, the least number of samples necessary for node division, and the minimum samples needed for leaf nodes. The search limits were established based on widely accepted ranges for ensemble regression models, while guaranteeing enough exploration of the parameter space.

The optimisation process concluded when either the maximum iteration limit was attained or when no additional enhancement in the goal function was detected. To mitigate the impact of random variability, *CFO* optimisation was conducted separately on five occasions, with the most effective solution chosen as the ultimate hyperparameter configuration. The comprehensive optimisation configurations, search intervals, and ultimately chosen parameters are included in Table 3. All calculations were performed with Python 3.10.12. The machine learning models were constructed with Scikit-learn 1.3.2, whilst numerical computations and data manipulation were performed with NumPy 1.26.4 and Pandas 2.1.4, respectively. The optimisation procedure was executed

Table 3 *CFO* optimization settings and optimized hyperparameters of *RFR* and *ETR* models.

Category	Parameter	<i>RFR</i> (<i>CFO</i>)	<i>ETR</i> (<i>CFO</i>)
<i>CFO</i> settings	Population size	40	50
	Maximum iterations	200	200
	Number of independent runs	5	5
	Stopping criterion	RMSE	RMSE
	Objective function	Minimize 5-fold CV RMSE	Minimize 5-fold CV RMSE
Search space	Number of estimators	50—500	50—500
	Maximum depth	2—30	2—30
	Minimum samples split	2—20	2—20
	Minimum samples leaf	1—10	1—10
	Maximum features	0.3—1.0	0.3—1.0
Optimal parameters	Number of estimators	219	127
	Maximum depth	11	18
	Minimum samples split	2	4
	Minimum samples leaf	3	1
	Maximum features	0.9	0.85

with SciPy 1.11.4 together with conventional Python modules. (<https://github.com/LSafils/Pb-Contaminated-Soils-Remediated-with-Biochar>).

Algorithms' performance indicators

In order to assess the efficacy of the *ETR* and *RFR* models, a wide variety of quantitative variables were used, each of which offered a distinct point of view⁵⁴. These metrics are Coefficient of Determination (R^2), Variation Account Factor (*VAF*), Mean Forecast Error (*MFE*), Theil's U Statistic (*U*), Mean Squared Percentage Error (*MSPE*), Willmott's Index of Agreement (*WI*), Uncertainty with a 95% confidence level ($U_{95\%}$), and Scatter index (*ScI*).

$$R^2 = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^n (x_i - \bar{x})^2\right] \left[\sum_{i=1}^n (y_i - \bar{y})^2\right]}} \right)^2 \quad (8)$$

$$VAF = \left(1 - \frac{\text{var}(x_i - y_i)}{\text{var}(x_i)} \right) * 100 \quad (9)$$

$$MFE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i) \quad (10)$$

$$U = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2 + \frac{1}{n} \sum_{i=1}^n y_i^2}} \quad (11)$$

$$MSPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - y_i}{x_i} \times 100 \right)^2 \quad (12)$$

$$WI = 1 - \frac{\sum_{i=1}^n (y_i - x_i)^2}{\sum_{i=1}^n (|y_i - \bar{y}| + |x_i - \bar{x}|)^2} \quad (13)$$

$$U_{95\%} = 1.96 \sqrt{(SD^2 + RMSE^2)} \quad (14)$$

$$ScI = \frac{\sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n ((x_i - \bar{x}) - (y_i - \bar{y}))^2}}{\left(\frac{1}{n}\right) \sum_{i=1}^n y_i} \quad (15)$$

where,

x_i : the observed *ToF*

$\bar{x}$: the average of observed *ToF*

y_i : the predicted *ToF*

$\bar{y}$: the average of predicted *ToF*

n : the number of rows in the dataset.

SD: abbreviation form of the standard deviation.

Results, comparisons and discussions

This research gives the findings from the single and hybrid methods to showcase the specific ToF . Figure 3 shows the estimated ToF /observed ToF values throughout the training and testing stages of the suggested $ETR(CFO)$ and $RFR(CFO)$ techniques. Additionally, Fig. 3 displays the distribution of the residual ratio of ToF value, with the center representing the error ratio in the ToF concentration and the zero-error line. Figure 4 displays the $SHAP$ analysis used to evaluate each attribute's significance. The results of a five-fold cross-validation using U metric values are shown in Table 4. The prediction efficacy of models in predicting ToF was assessed using metrics including R^2 , VAF , MFE , U , $MSPE$, WI , $U_{95\%}$, and ScI (see Table 5), difference percentage (see Table 6), and rank score of metrics (see Table 7).

The results of the five-fold cross-validation considering the values of U are presented in Table 4. The average U values obtained by XGB , LGB , ETR , RFR , $ETR(CFO)$, and $RFR(CFO)$ were 0.0924, 0.0916, 0.0966, 0.0936, 0.0816, and 0.0733, respectively. The $RFR(CFO)$ model achieved the lowest average U , indicating the most favorable performance among the evaluated models based on this index. Furthermore, the standard deviations ranged from 0.00224 to 0.00291, suggesting relatively consistent results across the five folds. The optimized models, $ETR(CFO)$ and $RFR(CFO)$, also exhibited lower average U values than their corresponding unoptimized models, demonstrating the beneficial effect of the CFO optimization procedure.

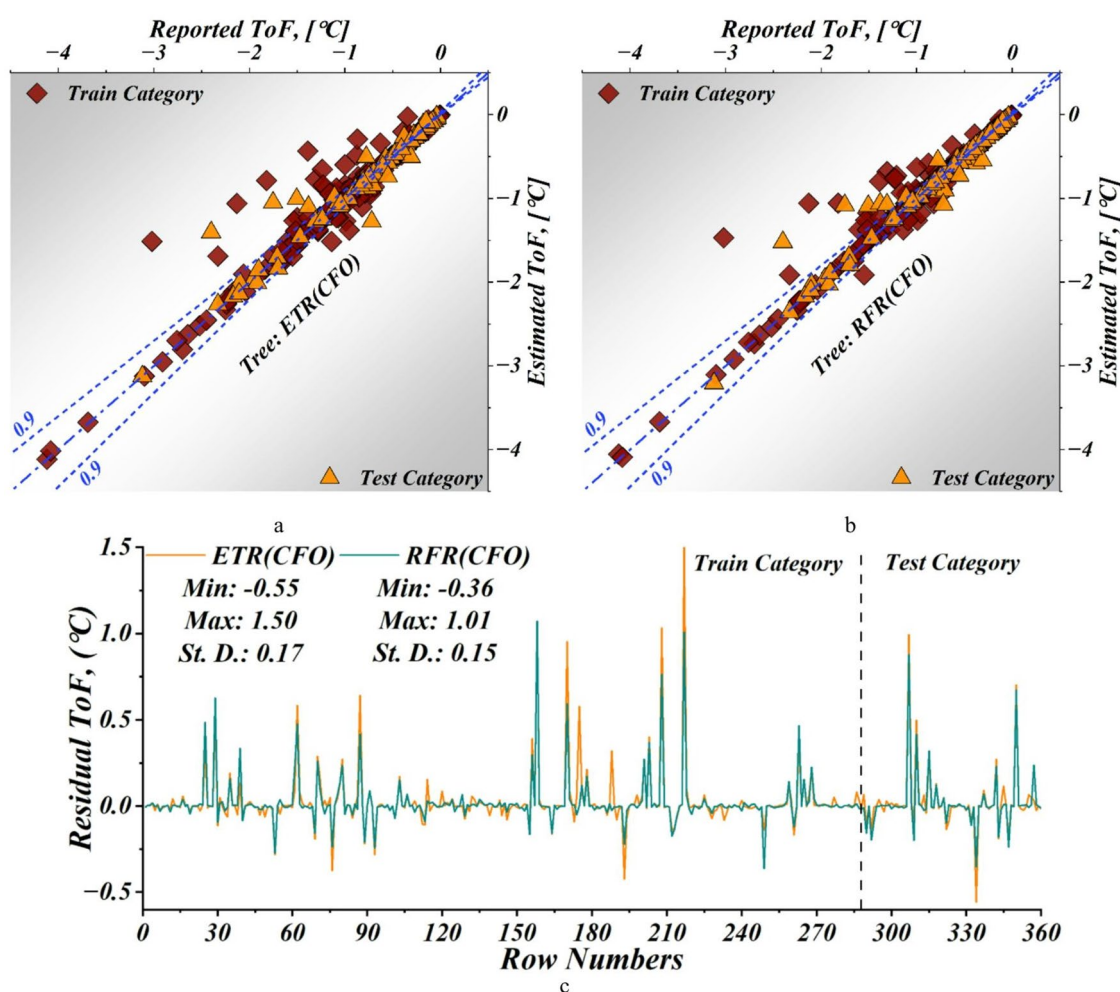


Fig. 3 Evaluation outcomes of developed models

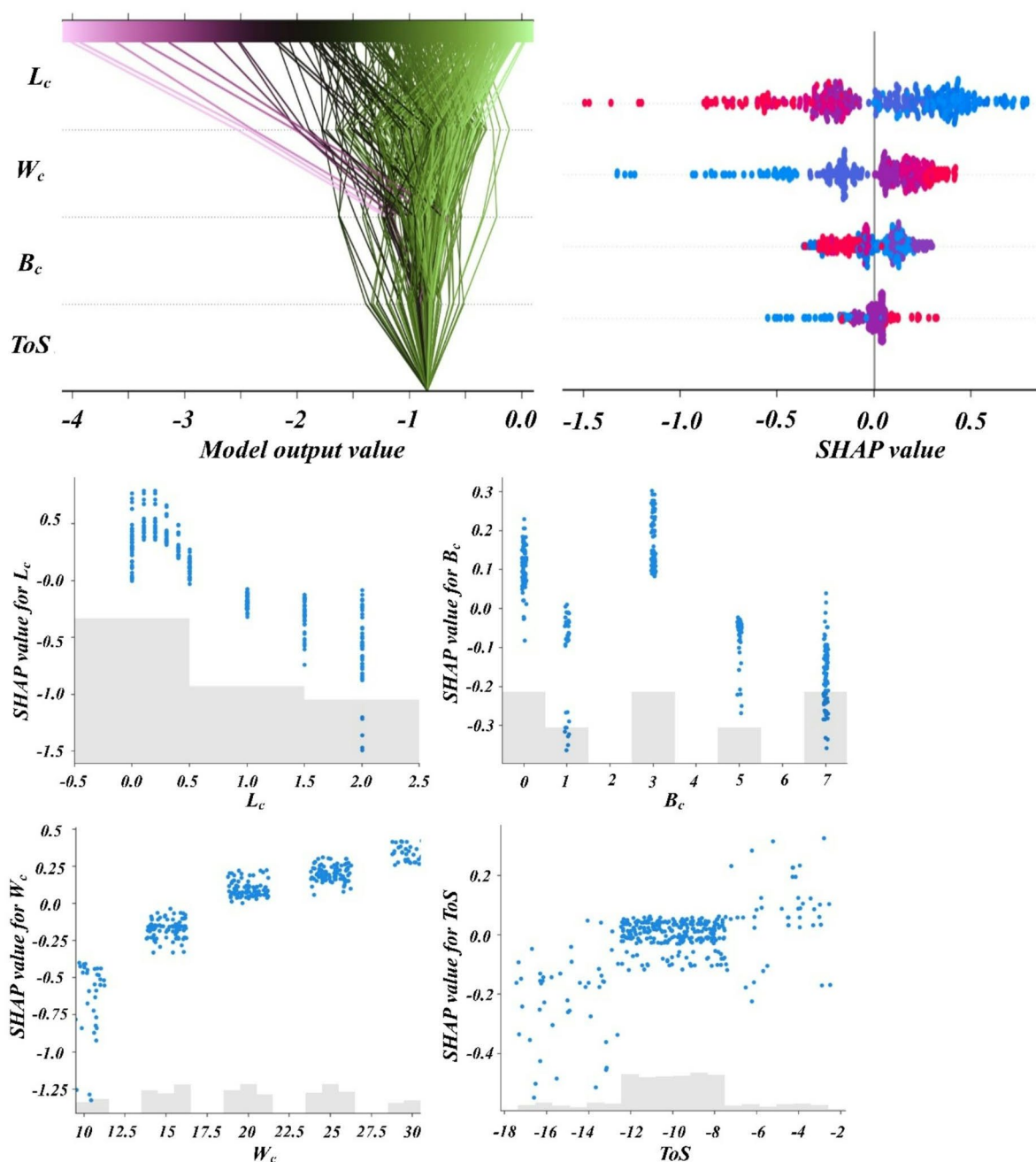


Fig. 4 SHAP analysis related to assess the importance of each attribute

Table 4 The results of 5-folds cross validation considering values of U .

Index	Framework					
	XGB	LGB	ETR	RFR	ETR(CFO)	RFR(CFO)
Fold 1	0.0953	0.0945	0.0996	0.0966	0.0791	0.0711
Fold 2	0.0902	0.0894	0.0942	0.0914	0.0836	0.0751
Fold 3	0.0882	0.0874	0.0921	0.0893	0.0855	0.0768
Fold 4	0.0951	0.0943	0.0994	0.0963	0.0793	0.0712
Fold 5	0.0931	0.0923	0.0973	0.0943	0.081	0.0728
Avg. $\pm$ St. D	0.0924 $\pm$ 0.00278	0.0916 $\pm$ 0.00276	0.0966 $\pm$ 0.00291	0.0936 $\pm$ 0.00282	0.0816 $\pm$ 0.0025	0.0733 $\pm$ 0.00224

Table 5 Statistical evaluation outcomes of developed and optimized forecasting models.

Category	Framework	R^2	VAF	MFE	U	MSPE	WI	U_95%	ScI
		High is better	High is better	Low is better	Low is better	Low is better	High is better	Low is better	Low is better
Train	<i>ETR(CFO)</i>	0.9385	93.8489	-0.0287	0.0816	224.667	0.9835	0.48418	0.2087
	<i>RFR(CFO)</i>	0.9501	95.004	-0.0241	0.0733	156.685	0.9866	0.43598	0.1878
	<i>ETR</i>	0.8728	87.279	-0.0313	0.0890	245.066	0.9147	0.5281	0.2276
	<i>RFR</i>	0.8836	88.354	-0.0310	0.0881	242.640	0.9175	0.5229	0.2254
	<i>XGB</i>	0.8929	89.272	-0.0303	0.0861	238.115	0.9305	0.5173	0.2217
	<i>LGB</i>	0.9022	90.22	-0.0299	0.0849	233.71	0.9454	0.5037	0.2171
	<i>SVR</i>	0.8673	86.73	-0.0311	0.0883	243.12	0.9089	0.5240	0.2258
	<i>LR</i>	0.8337	83.37	-0.0323	0.0919	252.91	0.8737	0.5450	0.2349
Test	<i>ETR(CFO)</i>	0.9301	92.9863	-0.0237	0.0862	372.402	0.9809	0.50691	0.2189
	<i>RFR(CFO)</i>	0.9425	94.212	-0.0174	0.0779	289.72	0.9844	0.45982	0.1983
	<i>ETR</i>	0.8650	86.477	-0.0286	0.1041	449.675	0.9122	0.6121	0.2643
	<i>RFR</i>	0.8765	87.617	-0.0273	0.0991	428.262	0.9155	0.5829	0.2517
	<i>XGB</i>	0.8841	88.53	-0.0272	0.0987	426.45	0.9291	0.5805	0.2507
	<i>LGB</i>	0.8941	89.39	-0.0270	0.0983	424.64	0.9429	0.5780	0.2496
	<i>SVR</i>	0.8595	85.93	-0.0308	0.1121	484.19	0.9064	0.6591	0.2846
	<i>LR</i>	0.8262	82.60	-0.0351	0.1278	552.11	0.8714	0.7515	0.3245

Table 6 Percent of differences between models with respect to *RFR(CFO)*.

Category	Framework	R^2	VAF	MFE	U	MSPE	WI	U_95%	ScI
Train	<i>ETR(CFO)</i>	1.236	1.231	16.028	10.172	30.259	0.315	9.955	10.014
	<i>ETR</i>	8.857	8.850	23.003	17.640	36.064	7.861	17.444	17.487
	<i>RFR</i>	7.526	7.527	22.258	16.799	35.425	7.531	16.623	16.681
	<i>XGB</i>	6.406	6.421	20.462	14.866	34.198	6.029	15.720	15.291
	<i>LGB</i>	5.309	5.303	19.398	13.663	32.958	4.358	13.445	13.496
	<i>SVR</i>	9.547	9.540	22.508	16.988	35.552	8.549	16.798	16.829
	<i>LR</i>	13.962	13.955	25.387	20.239	38.047	12.922	20.004	20.051
	<i>ETR(CFO)</i>	1.333	1.318	26.582	9.629	22.202	0.359	9.290	9.411
Test	<i>ETR</i>	8.960	8.944	39.161	25.168	35.571	7.917	24.878	24.972
	<i>RFR</i>	7.530	7.527	36.264	21.393	32.350	7.528	21.115	21.216
	<i>XGB</i>	6.606	6.418	36.029	21.074	32.062	5.952	20.789	20.901
	<i>LGB</i>	5.413	5.394	35.556	20.753	31.773	4.403	20.446	20.553
	<i>SVR</i>	9.657	9.638	43.506	30.508	40.164	8.608	30.235	30.323
	<i>LR</i>	14.076	14.058	50.427	39.045	47.525	12.970	38.813	38.891

With the data, T_{oF} can be reliably predicted by both *ETR(CFO)* and *RFR(CFO)*. R^2 values of 0.9501 in the training phase, showing a model difference percentage of 1.236%, and 0.94251 in the evaluation phase, indicating a model difference percentage of 1.3343%, show that the precision of the *RFR(CFO)* training and evaluation stages exceeded that of the *ETR(CFO)* technique. With training and evaluation values of 0.9385 and 0.9301, correspondingly, the *ETR(CFO)* showed worse R^2 outcomes. The *RFR(CFO)* model's strong functional dependability during the training and testing stages is shown by the *VAF* values of 95.004 and 94.2122. The *VAF* values of 93.8489 and 92.9863 were very comparable, according to the results of the *ETR(CFO)* approach. Based on expected performance assessment criteria, *RFR(CFO)* is a technique that outperforms *ETR(CFO)* in terms of precision and dependability. An extra assessment that focuses on measurement errors might improve the process's reliability. To achieve this, the previously specified criteria must be assessed. When contrasted to a larger value, a lower number in these metrics indicates a higher degree of error consistency. According to the statistics, the *RFR(CFO)* design achieved the lowest values while demonstrating remarkable reliability. During the training and evaluation stages, the *RFR(CFO)* produced lower *MSPE* metric outcomes, with indicated values of 156.685 and 289.72, accordingly. In comparison

Table 7 Rank of metrics of developed and optimized forecasting models.

Category	Framework	R^2	VAF	MFE	U	MSPE	WI	U _{95%}	Sci
Train	<i>ETR(CFO)</i>	7	7	7	7	7	7	7	7
	<i>RFR(CFO)</i>	8	8	8	8	8	8	8	8
	<i>ETR</i>	3	3	3	3	3	3	3	3
	<i>RFR</i>	4	4	4	4	4	4	4	4
	<i>XGB</i>	5	5	5	5	5	5	5	5
	<i>LGB</i>	6	6	6	6	6	6	6	6
	<i>SVR</i>	2	2	2	2	2	2	2	2
	<i>LR</i>	1	1	1	1	1	1	1	1
Test	<i>ETR(CFO)</i>	7	7	7	7	7	7	7	7
	<i>RFR(CFO)</i>	8	8	8	8	8	8	8	8
	<i>ETR</i>	3	3	3	3	3	3	3	3
	<i>RFR</i>	4	4	4	4	4	4	4	4
	<i>XGB</i>	5	5	5	5	5	5	5	5
	<i>LGB</i>	6	6	6	6	6	6	6	6
	<i>SVR</i>	2	2	2	2	2	2	2	2
	<i>LR</i>	1	1	1	1	1	1	1	1

to the other results, the training and evaluation stages of the *ETR(CFO)* showed lower dependability, with values of 224.667 and 372.402, accordingly.

The $U_{95\%}$ metric was used to quantify the overall uncertainty associated with the model predictions. It was calculated as Eq. 14. For the testing dataset, the $U_{95\%}$ values were 0.50691 °C for *ETR(CFO)* and 0.45982 °C for *RFR(CFO)*, indicating lower overall predictive uncertainty for *RFR(CFO)*. Accordingly, model predictions can be expressed together with a model-level uncertainty range as $ToF \pm U_{95\%}$. The lower $U_{95\%}$ value of *RFR(CFO)* corresponds to a narrower uncertainty range and supports its greater predictive reliability compared with *ETR(CFO)*. Nevertheless, these uncertainty ranges represent model-level uncertainty rather than observation-specific prediction intervals. The residual analysis also showed that some individual predictions may exhibit relatively large deviations from the observed values. Therefore, predictions associated with unusually large residuals should be interpreted cautiously and, where necessary, verified experimentally. In addition, the models should not be considered equally reliable when applied outside the ranges represented in the training dataset or under substantially different field conditions. Consequently, *RFR(CFO)* should be regarded as a decision-support tool, with experimental or field verification recommended when predictions are associated with large uncertainty or conditions beyond the model's training domain.

Table 5 presents a comparative evaluation of the proposed *RFR(CFO)* model against *ETR*, *RFR*, *XGB*, *LGB*, *SVR*, and *LR* using eight statistical indices. The comparison demonstrates that *RFR(CFO)* achieved the most favorable overall performance in both the training and testing phases. In the training phase, *RFR(CFO)* obtained the highest R^2 and VAF values of 0.9501 and 95.004%, respectively, while also achieving the highest WI(0.9866) and the lowest values of U (0.0733), $MSPE$ (156.685), $U_{95\%}$ (0.43598), and SCI (0.1878). Compared with the conventional *RFR* model, the proposed optimized model improved R^2 by approximately 7.53% and reduced $MSPE$ by 35.42%. Its advantages were also evident over *ETR*, *XGB*, *LGB*, *SVR*, and *LR*, with $MSPE$ reductions of approximately 36.06%, 34.20%, 32.96%, 35.55%, and 38.05%, respectively.

Similar results were obtained in the testing phase, confirming the predictive capability of *RFR(CFO)* on unseen data. The optimized model achieved $R^2 = 0.9425$, $VAF = 94.212\%$, and $WI = 0.9844$, together with the lowest U , $MSPE$, $U_{95\%}$, and SCI values. Compared with the conventional *RFR*, *RFR(CFO)* improved R^2 by approximately 7.53% and reduced $MSPE$ by 32.35%. The corresponding $MSPE$ reductions relative to *ETR*, *XGB*, *LGB*, *SVR*, and *LR* were approximately 35.57%, 32.06%, 31.77%, 40.16%, and 47.52%, respectively. Overall, the results indicate that the *CFO* optimization procedure enhanced the predictive accuracy of the *RFR* model and reduced forecasting errors

across the evaluated statistical indices. The consistency between the training and testing results further supports the effectiveness of *RFR(CFO)* for the forecasting task.

Figure 3c illustrates the residual distribution of the *ETR(CFO)* and *RFR(CFO)* models over all data, categorized into training and testing segments. Residuals represent the disparity between expected and actual values, with values approaching 0 indicating more accuracy. Both models exhibit random dispersion about zero without any discernible bias. Nonetheless, *RFR(CFO)* is the superior model due to its reduced variability (standard deviation = 0.15 compared to 0.17 for *ETR(CFO)*), a more constrained residual range (about −0.36 to 1.01 against −0.55 to 1.50 for *ETR(CFO)*), and fewer severe departures. *ETR(CFO)* exhibits more variability and elevated error spikes, particularly at certain sample points. *RFR(CFO)* exhibits more consistent and stable residual patterns across both training and testing datasets, indicating superior generalization and enhanced predictive reliability.

Shapley additive explanations, or *SHAP* analysis, give numerical values to each characteristic to determine its predictive relevance. This study helps stakeholders trust and understand the model's conclusions by showing how each characteristic affects projections. *SHAP* values help explain the results of any machine training model, even sophisticated ones like neural networks and boosting approaches. Because *SHAP* values answer questions thoroughly. Assigning priority to each parameter in a machine training model's projection may explain its results. Thus, model outcomes may be explained. The results from the *SHAP* research are shown visually in Fig. 4. Ultimately, *SHAP* analysis is a method that facilitates decision-making by offering self-evident and practical insights into the model's functioning and the aspects that contribute to its forecasts. The findings show that Lead Content is the most useful variable, followed by Water Content, and that the third most important parameter, Biochar Content, has a major impact on the target. In the end, Set Temperature was the criterion with the lowest values.

The *SHAP* findings establish a crucial link between the prediction model and the fundamental physics of soil freezing. The primary influence of liquid crystals is credible as the soil's freezing point is determined by both its thermal condition and the physicochemical properties of its pore water. Unlike pure water, soil pore water encompasses dissolved and mobile entities, and the existence of solutes diminishes the chemical potential of the liquid phase in comparison to pure water. Thus, the creation of ice necessitates a reduced temperature to achieve balance between the liquid and solid states. This occurrence, typically known as freezing-point depression, is closely linked to the concentration of solutes and the interaction between pore water and soil particle surfaces.

To ascertain the statistical significance of the disparities between the created models and benchmark methods, the Wilcoxon signed-rank test was performed utilizing the prediction errors derived from each model. This non-parametric statistical analysis was chosen as it does not necessitate the presumption of normally distributed errors and is appropriate for evaluating matched model predictions within the identical dataset. Figure 5 displays the paired p-values acquired throughout the training and testing phases. A p-value less 0.05 signifies that the disparity between two models is statistically significant, whereas a greater value implies that their predictive capabilities are statistically similar. The Wilcoxon analysis demonstrates the superiority of the suggested *CFO*-optimized models in comparison to conventional regression techniques. The lack of statistically significant disparities between *ETR(CFO)* and *RFR(CFO)* suggests that both models gain from the optimization process and offer similar prediction accuracy. Nonetheless, taking into account the previously documented assessment metrics, uncertainty analysis, and generalization performance, *RFR(CFO)* remains the favored model owing to its somewhat superior prediction accuracy and stability. The statistical evaluation therefore verifies that the noted advancements are not only a result of chance fluctuations but are linked to the augmented learning potential afforded by the optimization-assisted ensemble frameworks.

Figure 5 presents the pairwise Wilcoxon signed-rank test results at a significance level of 0.05. During the training phase[Fig. 5(a)], *ETR(CFO)* differed significantly from *LGB* ($p = 0.0391$), *SVR* ($p = 0.00416$), *LR* ($p = 0.00666$), and *XGB* ($p = 0.0468$), whereas no significant difference was observed against *RFR(CFO)*, *ETR*, or *RFR*. Similarly, *RFR(CFO)* showed significant differences from *LGB* ($p = 0.04889$), *LR* ($p = 0.0119$), *ETR* ($p = 0.01651$), and *XGB* ($p = 0.00949$). The statistical separation became more evident during independent testing[Fig. 5(b)]. *ETR(CFO)* exhibited significant differences from *RFR(CFO)*, *LGB*, *SVR*, *RFR*, and *XGB* ($p < 0.05$), while its comparisons with *LR* ($p = 0.0766$) and *ETR* ($p = 0.364$) were not significant. *RFR(CFO)* also differed significantly from *LGB*, *SVR*, *LR*, and *XGB*, but not

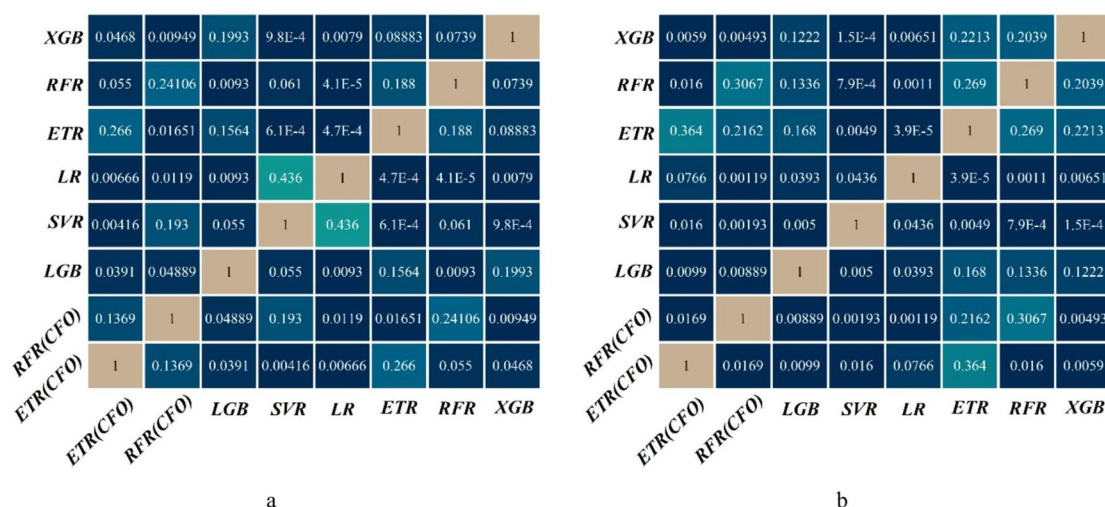


Fig. 5 Wilcoxon signed-rank test results for pairwise comparison of developed and benchmark models: **(a)** training phase and **(b)** testing phase.

from *ETR* or *RFR*. Overall, the testing results demonstrate that the optimized models generate statistically distinct prediction patterns from most benchmark approaches, particularly *LGB*, *SVR*, and *XGB*. Nevertheless, the Wilcoxon test indicates statistical differences rather than model superiority; therefore, these findings should be considered together with the performance metrics reported in Table 5.

Practical usages

The findings of this research hold significant practical ramifications for geotechnical and environmental engineering in cold regions. The formulated model offers a swift method for forecasting the freezing temperature (ToF) of Pb-affected soils amended with biochar, considering various combinations of lead concentration, biochar quantity, moisture level, and designated temperature. In real-world scenarios, engineers may juxtapose the anticipated ToF with the projected low ground temperature at a location to evaluate the probability of soil freezing and pinpoint potentially hazardous circumstances for roadways, foundations, and embankments. For instance, if the forecasted ToF approaches or exceeds the expected soil temperature, supplementary frost mitigation strategies, such as enhancing drainage, adjusting the biochar application rate, or implementing thermal safeguards, may be contemplated. The model may further serve to assess varying biochar quantities and moisture levels prior to doing comprehensive laboratory or field trials, facilitating the elimination of inappropriate circumstances and minimising both time and expenses associated with testing. Moreover, the sensitivity analysis offers supplementary assistance for decision-making by pinpointing significant factors, especially lead and water content, which may be emphasised for site characterisation, monitoring, and focused action. Consequently, the suggested model may function as an initial decision-support instrument for amalgamating frost-risk evaluation with remediation strategies in contaminated soils in cold regions.

Restrictions and possible solutions

The current study has a number of limitations that could affect how strong and generalizable the proposed models are. First, the dataset only has 360 data points from studies that have already been published or submitted. This may not fully capture the range of real-world soil conditions, such as differences in mineralogy, environmental factors, and testing procedures. To fix this problem, the dataset can be expanded with new experimental campaigns and multi-source data from different geographical areas. Furthermore, the utilization of literature-derived

data engenders discrepancies in experimental conditions, including fluctuations in cooling rates and sample preparation, which may introduce bias; standardization methodologies and the incorporation of metadata variables can assist in alleviating this concern. The research examines merely four input parameters (lead content, biochar content, water content, and set temperature), potentially neglecting other significant factors such as soil type and thermal properties. This indicates that subsequent studies should integrate a more extensive array of thermo-hydro-mechanical variables and utilize sophisticated feature selection methodologies. *RFR* and *ETR* are two models that are very good at making predictions, but they can still overfit and not generalize well. Nested cross-validation and external dataset validation can help with this. Also, using the *CFO* Algorithm to tune hyperparameters can lead to random changes and the possibility of converging to local optima. This can be fixed by comparing different optimization methods and using ensemble optimization methods. Even though *SHAP* (Shapley Additive Explanations) is used to make the models easier to understand, they are still mostly based on data and don't include physical freezing mechanisms. This shows that we need to use hybrid physics-informed machine learning methods. Furthermore, uncertainty quantification is confined to statistical metrics and fails to comprehensively represent predictive confidence across diverse conditions, which could be improved through probabilistic or Bayesian frameworks. Finally, because the models are based on normalized data from the lab, they may not work as well in the field. For this reason, it is suggested that future research focus on field-scale validation and the creation of useful decision-support tools.

Conclusions

This research presented a machine-learning system designed to forecast the freezing temperature (ToF) of Pb-contaminated soils amended with biochar by combining ensemble regression models with a metaheuristic optimisation approach. In contrast to traditional experimental methods necessitating comprehensive laboratory studies, the suggested framework offers a data-centric strategy for assessing soil freezing characteristics by utilising critical influencing factors such as lead concentration, biochar levels, moisture content, and regulated ambient/freezing temperature (ToS). The primary contribution of this study is the creation of enhanced models using Catch Fish Optimisation (*CFO*), along with an interpretability assessment utilising *SHAP* to ascertain the significance of input variables.

The findings indicated that the refined models effectively represented the nonlinear correlation among soil characteristics, freezing circumstances, and ToF . Among the established methodologies, *RFR(CFO)* demonstrated better predictive performance on the held-out testing subset relative to the optimised *ETR* model.

The examination indicated that lead concentration and moisture levels were the primary determinants influencing the anticipated freezing temperature, whilst the regulated ambient/freezing temperature offered supplementary insights into the thermal boundary conditions of the trials. These results validate the efficacy of refined ensemble learning techniques as powerful instruments for examining intricate freezing phenomena in polluted soils.

Notwithstanding the encouraging results, the established framework is constrained by the scale and variety of the existing literature-derived dataset and the lack of supplementary thermo-hydro-mechanical parameters that could affect soil freezing dynamics, including soil mineral composition, thermal conductivity, and pore attributes. Consequently, further research ought to integrate more extensive experimental datasets across various soil conditions and explore physics-informed or hybrid modelling approaches to enhance model transferability. Nonetheless, the suggested methodology offers a pragmatic decision-support instrument for the first evaluation of freezing dynamics in Pb-contaminated soils and can aid in enhancing remediation techniques and geotechnical frameworks in frigid regions.

Author contributions Laila Safi conceived and designed the study, collected and curated the data, developed the methodology, performed the formal analysis, implemented the computational models, validated the results, interpreted the findings,

prepared the figures and tables, wrote the original manuscript draft, critically revised and edited the manuscript, and approved the final version for publication.

Funding No funding.

Data availability The datasets used in this study are freely available from the corresponding author upon reasonable request for research purposes.

Code availability <https://github.com/LSafils/Pb-Contaminated-Soils>

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Fagerlund, G. Determination of pore-size distribution from freezing-point depression. *Mater. Constr.* **6**, 215–225 (1973).
2. Wu, X. & Zhao, Y. A novel heat pulse method in determining “effective” thermal properties in frozen soil. *Water Resour. Res.* **60**, e2024WR037537. <https://doi.org/10.1029/2024WR037537> (2024).
3. Ding, C. et al. Freeze–thaw controls on dissolved organic carbon export: A numerical investigation of hydrological mechanisms in a hillslope. *Water Resour. Res.* **62**, e2026WR043497 (2026).
4. Low, P. F., Anderson, D. M. & Hoekstra, P. Some thermodynamic relationships for soils at or below the freezing point: 1. Freezing point depression and heat capacity. *Water Resour. Res.* **4**, 379–394 (1968).
5. Zhan, Y. et al. Changes in meat consumption can improve groundwater quality. *Nat. Food* **6**, 703–714 (2025).
6. Dai, Z., Lu, X., Wang, L., Yang, S. & Wang, R. Shrinkage cracking characteristics and micro-mechanism of bentonite and glass-fiber-modified cement soil in dry environment. *Materials* **19**, 1671 (2026).
7. Yang, G. et al. New insights into linear-nonlinear tensile creep mechanisms of UHPC: Stress level-dependent creep-active site transitions. *Cem. Concr. Res.* **198**, 107988 (2025).
8. Lei, H., Bao, N., Ben-Dor, E., Peng, S. & Zhao, N. Soil bidirectional reflectance characteristics and applications in quantitative soil property retrieval: A systematic review. *CATENA* **273**, 110456 (2026).
9. Liu, J. et al. Pore-scale visualization and microscale barrier mechanisms of microplastics transport in bio-based hydrogel modified soils. *Water Res.* <https://doi.org/10.1016/j.watres.2025.125159> (2025).
10. Xu, T. et al. Warming increases CO₂ emissions in biochar-amended cropland soil. *Biochar* **8**, 106 (2026).
11. Han, Y. et al. Experiments on the initial freezing point of dispersive saline soil. *CATENA* **171**, 681–690 (2018).
12. Wang, Q., Qi, J., Wang, S., Xu, J. & Yang, Y. Effect of freeze-thaw on freezing point of a saline loess. *Cold Reg. Sci. Technol.* **170**, 102922 (2020).
13. Wang, S., Wang, Q., Xu, J. & Ding, J. Effect of freeze-thaw on freezing point and thermal conductivity of loess. *Arab. J. Geosci.* **13**, 206 (2020).
14. Song, Z. et al. Multiscale investigation into the hydrogel-stabilized river bank sand: Sight from NMR and CT technology. *Constr. Build. Mater.* **460**, 139900 (2025).
15. Elahi, A., Bidault, X. & Chaudhuri, S. Temperature-transferable coarse-grained model for poly (propylene oxide) to study thermo-responsive behavior of triblock copolymers. *J. Phys. Chem. B* **126**, 292–307 (2022).
16. Scherer, G. W. Crystallization in pores. *Cem. Concr. Res.* **29**, 1347–1358 (1999).
17. Setzer, M. J. Mechanical stability criterion, triple-phase condition, and pressure differences of matter condensed in a porous matrix. *J. Colloid Interface Sci.* **235**, 170–182 (2001).

18. Watanabe, K. & Mizoguchi, M. Amount of unfrozen water in frozen porous media saturated with solution. *Cold Reg. Sci. Technol.* **34**, 103–110 (2002).
19. Wan, X., Gong, F., Qu, M., Qiu, E. & Zhong, C. Experimental study of the salt transfer in a cold sodium sulfate soil. *KSCE J. Civ. Eng.* **23**, 1573–1585 (2019).
20. Bing, H. & Ma, W. Laboratory investigation of the freezing point of saline soil. *Cold Reg. Sci. Technol.* **67**, 79–88 (2011).
21. Cao, W., Liu, C., Chou, Y. & Peng, E. Experimental study on freezing temperature and its influencing factors of loess contaminated by heavy metals. *Case Stud. Therm. Eng.* **61**, 105034 (2024).
22. Han, Y. et al. Effect of freeze-thaw cycles on shear strength of saline soil. *Cold Reg. Sci. Technol.* **154**, 42–53 (2018).
23. Zhang, Y., Yang, Z., Liu, J. & Fang, J. Impact of cooling on shear strength of high salinity soils. *Cold Reg. Sci. Technol.* **141**, 122–130 (2017).
24. Zhang, M., Zhang, X., Lai, Y., Lu, J. & Wang, C. Variations of the temperatures and volumetric unfrozen water contents of fine-grained soils during a freezing–thawing process. *Acta Geotech.* **15**, 595–601 (2020).
25. Wan, X., Lai, Y. & Wang, C. Experimental study on the freezing temperatures of saline silty soils. *Permafrost. Periglacial Process.* **26**, 175–187 (2015).
26. Xu, X., Wang, Y., Bai, R., Zhang, H. & Hu, K. Effects of sodium sulfate content on mechanical behavior of frozen silty sand considering concentration of saline solution. *Results Phys.* **6**, 1000–1007 (2016).
27. Ghani, S., Thapa, I., Adhikari, D. K. & Waris, K. A. Soil categorization and liquefaction prediction using deep learning and ensemble learning algorithms. *Transp. Infrastruct. Geotechnol.* **12**, 22 (2025).
28. Mohammadi, M. Predicting Remaining Time in Object-Centric Process Executions Using Machine Learning Models. *J. Data Anal. Intell. Decis.* **2**, 143–155 (2026).
29. Rahmani Khorasani, V., Yamaghani, M. R. & Darvish Motevali, M. H. Preventing Customer Churn in E-commerce Based on Feature Selection Using Metaheuristic Methods and Deep Learning. *J. Eng. Manag. Soft Comput.* **12**, 62–82 (2026).
30. Esmacili Noroozi, M. R. & Zare Mehrjardi, F. Optimization of Steel Alloy Composition to Maximize Yield Strength Using a Machine Learning Model and the Cuckoo Optimization Algorithm (COA). *Eng. Manag. Soft Comput.* **12**, 131–143 (2026).
31. Guo, K. & Yang, G. Load-slip curves of shear connection in composite structures: prediction based on ANNs, Steel Compos. Struct. An. *Int. J.* **36**, 493–506 (2020).
32. P. Monfaredi, A. Memari, Mechanical and Structural Evaluation of Carbon Geocrete as an Alternative to Portland Cement Concrete, (2026).
33. Monfaredi, P., Emami, S. M. M. & Moghadam, A. S. Seismic behavior of hollow-core infilled steel frames; an experimental and numerical study. *J. Constr. Steel Res.* **192**, 107244 (2022).
34. M.S. Zadeh Naseri, I. Ershaghi, Permeability of Borehole Cements, in: SPE West. Reg. Meet., SPE, 2026: p. D031S009R004.
35. Thapa, I. & Ghani, S. Explainable AI-enhanced machine learning for CBR prediction in stabilized and unstabilized subgrade soils. *Appl. Soft Comput.* **178**, 113275 (2025).
36. Thapa, I. et al. Sustainable approach of strength measurement for soil's stabilized with geo-polymer with hybrid ensemble models. *Results Eng.* **27**, 106133 (2025).
37. Ghani, S., Thapa, I., Kumari, S., Correia, A. G. & Asteris, P. G. Revealing the nature of soil liquefaction using machine learning. *Earth Sci. Informatics* **18**, 198 (2025).
38. Zhuang, P.-Z. et al. CASM-U: a unified critical state model for unsaturated clays and sands. *Acta Geotech.* **20**, 211–230 (2025).
39. Zheng, Y., Baudet, B. A., Coop, M. R., Pereira, J.-M. & Delage, P. A study of pore properties in reconstituted Lucera clay at states of the critical state framework. *Géotechnique* **76**, 452–464 (2026).
40. Gautam, K. et al. A review on control and abatement of soil pollution by heavy metals: Emphasis on artificial intelligence in recovery of contaminated soil. *Environ. Res.* **225**, 115592 (2023).
41. Wang, M. et al. Pollution level mapping of heavy metal in soil for ground-airborne hyperspectral data with support vector machine and deep neural network: A case study of Southwestern Xiong'an, China. *Environ. Pollut.* **321**, 121132 (2023).
42. Luo, N. Methods for controlling heavy metals in environmental soils based on artificial neural networks. *Sci. Rep.* **14**, 2563 (2024).
43. Chu, Y., Liu, S., Cai, G. & Bian, H. Artificial neural network prediction models of heavy metal polluted soil resistivity. *Eur. J. Environ. Civ. Eng.* **25**, 1570–1590 (2021).
44. Shi, L. et al. Modeling phytoremediation of heavy metal contaminated soils through machine learning. *J. Hazard. Mater.* **441**, 129904 (2023).
45. El-Sharkawy, M. et al. Assessing and predicting soil quality in heavy metal-contaminated soils: Statistical and ANN-based techniques. *J. Soil Sci. Plant Nutr.* **23**, 6510–6526 (2023).
46. Li, P. et al. A field study to estimate heavy metal concentrations in a soil-rice system: Application of graph neural networks. *Sci. Total Environ.* **832**, 155099 (2022).

47. Wang, X. et al. A back propagation neural network model optimized by mind evolutionary algorithm for estimating Cd, Cr, and Pb concentrations in soils using Vis-NIR diffuse reflectance spectroscopy. *Appl. Sci.* **10**, 51 (2019).
48. Shang, J. Q., Ding, W., Rowe, R. K. & Josic, L. Detecting heavy metal contamination in soil using complex permittivity and artificial neural networks. *Can. Geotech. J.* **41**, 1054–1067 (2004).
49. Han, S., Cao, W., Chou, Y. & Peng, E. Experimental Study on the Change in Freezing Temperature During the Remediation of Pb-Contaminated Soils with Biochar. *Atmosphere (Basel)*. **15**, 1483 (2024).
50. Kumar, A., Sen, S., Sinha, S., Kumar, B. & Nidhi, C. Explainable LightGBM model for predicting compressive strength of silica fume modified high-volume fly ash concrete, *Asian. J. Civ. Eng.* **26**, 3951–3968 (2025).
51. Nidhi, C. & Singh, P. K. Explainable tree-based machine learning models for predicting gross calorific value of various waste from ultimate analysis, *Asian. J. Civ. Eng.* **27**, 3079–3098 (2026).
52. Samadi, N. et al. Machine learning assessment of crash severity in ADS and ADAS-L2 involved crashes with NHTSA data. *Safety* **12**, 2 (2025).
53. Jia, H., Wen, Q., Wang, Y. & Mirjalili, S. Catch fish optimization algorithm: a new human behavior algorithm for solving clustering problems. *Cluster Comput.* **27**, 13295–13332 (2024).
54. Kumar, A., Sen, S., Sinha, S., Kumar, B. & Nidhi, C. Prediction of compressive strength of concrete doped with waste plastic using machine learning-based advanced regularized regression models, *Asian. J. Civ. Eng.* **26**, 1723–1741 (2025).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Laila Safi¹

✉ Laila Safi
L.Safi.1998.L@gmail.com

¹ Department of Civil Engineering, Faculty of Engineering, Paktia University, Gardez, Afghanistan