Mixture optimization for mechanical, environmental, and economic objectives in grouting slurry for repairing earthen sites

Kai Cui[1], Huian Shao[1*], Xiangpeng Yu[1], Qiangqiang Pei[2] and Shuai Zhang[3]

Abstract

Anchor and fissure grouting are used to repair earthen sites. However, the common method to obtain the compressive strength of grouting slurry would cause material, labor, and time losses. In addition the material properties, environmental and economic benefits have gained increasing attention. This study proposes a design framework for multi-objective proportioning optimization based on machine learning and metaheuristics. The results indicated that the eXtreme Gradient Boosting (XGBoost) model, whose hyper-parameters were optimized by a genetic algorithm, can accurately predicted the compressive strength of the slurries. The impact of the variables on development of compressive strength can explain the internal reaction mechanisms. The analytical framework based on meta-heuristic and technique for order of preference by similarity to an ideal solution (TOPSIS) provided Pareto-optimal solutions in design scenario of each sub-dataset. The framework proposed in this study can efficiently achieve mechanical, environmental, and economic design objectives of anchor grouting and fissure grouting slurries for repairing earthen sites.

Keywords Earthen sites, Grouting slurry, Oxide composition, Machine learning, Multi-objective optimization

Introduction

Earthen sites have valuable historical information and are distributed throughout China [1]. However, natural factors such as wind, water, temperature, and earthquakes have created erosive patterns that are widely distributed at these sites. The most common are large-scale earthen site cracking and small-scale erosion fissures, which severely decrease the structural safety of earthen sites [2, 3]. Anchor and fissure grouting are repair methods used for cracking and fissure erosion at earthen sites, respectively. Anchor grouting combines anchor rods, grouting slurries, and site soil to mobilize the entire earthen site to stabilize the damaged part. Fissure grouting injects grouting slurries that are compatible with the earthen site into the fissure to improve the overall mechanical strength. Both methods require suitable slurries characterized by a multiplicity of admixtures, which makes the prediction of their properties difficult [4–6]. Currently, three types of grouting slurries are commonly used to repair earthen sites in Northwest China: quicklime (Lime), calcined ginger nuts (CGN), and potassium silicate (PS). Their raw materials include traditional building materials such as Lime and CGN; natural materials such as loess (C), diatomite (Z), and quartz sand (S); supplementary cementitious materials such as fly ash (F) and bentonite (B); and binder materials such as PS, modified polyvinyl alcohol

*Correspondence:
Huian Shao
894126262@qq.com
[1] Western Center of Disaster Mitigation in Civil Engineering, College of Civil Engineering, Lanzhou University of Technology, No. 287 Langongping Road, Qilihe District, Lanzhou 730050, China
[2] National Research Center for Conservation of Ancient Wall Paintings and Earthen Sites, Dunhuang Academy, Dunhuang 736200, China
[3] College of Civil Engineering and Mechanics, Lanzhou University, Lanzhou 730000, Gansu, China

Cui *et al. Heritage Science*   (2024) 12:174

Page 2 of 21

solution (SH), and sticky rice paste (N) [7–21]. Among them, the preparation and utilization of some traditional building materials often involve large carbon emissions, whereas some newer materials have higher costs [22, 23]. Therefore, balancing the mechanical compatibility, environmental friendliness, and cost-controllable attributes of grouting slurries will become an important part of the research and promotion on repairing historically valuable earthen sites.

In recent years, data-driven machine learning (ML) algorithms have accurately predicted the mechanical strength of materials. They can particularly consider the non-linear relationship between strength and its influencing factors, which allows for improving the efficiency of experimental analyses [24–26]. A Çalışkan et al. developed regression methods based on extreme learning machines (ELM), support vector machines (SVM), and the grouping method for data handling (GMDH) to estimate the compressive strength of fly ash and nano-calcite-cemented mortars. These models not only predicted highly accurate compressive strengths and ultrasonic pulse velocities (UPV), but also minimized time, materials, labor, and cost [27]. Zhang et al. used a random forest (RF) optimized through a genetic algorithm (GA) to predict the compressive strength of alkali-activated materials. The optimization of the RF produced higher prediction accuracy than common algorithms, such as back propagation neural networks (BP), multiple linear regression (MLR), and SVM, and could also explain the effect of admixture compositions on mechanical strength [28]. Kang et al. used 11 ML models to predict the mechanical strength of steel fiber-reinforced concrete and found that boosting- and tree-based models provided high prediction accuracy, and the K-nearest neighbor, ridge regression, and lasso regressor performed the worst. In addition, the water-cement ratio and silica-ash content were the most influential factors in the prediction [29]. These studies showcase the utility of machine-learning-based research on the mechanical strength of concrete and cement mortar [30–32]. However, the earthen site restoration should adhere to the principle of "minimum intervention, maximum compatibility", using materials with physical, mechanical, and hydraulic properties to the site soil. This makes it impossible for the cement mortar which is commonly used in modern construction to be used in earthen site restoration engineering. Therefore, it is necessary to conduct the research on the prediction of the compressive strength of grouting slurries for repairing earthen sites. In addition, previous studies used the relative amount of each admixture as an input variable for their prediction models. They also mostly disregarded the effects of internal chemical

reactions and the compactness of each material particle on the mechanical strength of the admixture materials, which led to a focus on data-driven results and poor interpretability and clarity regarding the internal mechanisms involved. This approach may lead to theoretically feasible prediction results [33, 34]. Therefore, in the process of developing the dataset used for grouting slurries, it is necessary to both collect the data of each admixture dosage and consider the effects of the molar ratio of the main reacting oxides and the bulk density of the admixture materials on the compressive strength of the slurries.

It is necessary to develop targeted and intelligent design methods to comprehensively analyze the performance of different grouting slurries [35]. This study aims to establish a framework for designing the mixtures of grouting slurries. For this, a large dataset containing three sub-datasets was established, totaling 523 sets of data. Subsequently, after optimizing the hyper parameter by the GA, ten ML algorithms were used to predict the compressive strength of each sub-dataset, and the model with the highest accuracy was selected. Based on the relative importance and SHAP analysis, the effect of the input variables on the compressive strength was obtained. Finally, the Metaheuristic algorithm and TOPSIS were used to achieve multi-objective optimization of grouting slurries for repairing earthen sites, which must be analyzed after actual application. Compared with the existing research on optimizing admixture ratios using computational algorithms, this study provides three novelties. (1) An optimization model was developed for the ratios of admixture materials of anchoring grouting and fissure grouting slurries on earthen sites, which employs a dataset containing 523 sets of data. (2) This study considers the effect of the molar ratio of the main reacting oxides in grouting slurries and the bulk density of admixture materials on the mechanical strength of the slurries. These represent the effects of internal chemical reactions and the compactness of each material particle on the obtained compressive strength, respectively. (3) Detailed, objective functions for carbon emissions and costs during the production and transportation of each admixture were estimated; this, combined with the ML prediction model and multi-objective particle swarm optimization (MOPSO), permits the proposal of an intelligent framework for optimizing admixtures for repairing slurries with multiple objectives. This study provides insight into the prediction behavior of the compressive strength and mixture design optimization of anchor grouting and fissure grouting slurries for earthen sites, opening a new way for the application of slurries without the need for extensive experimental work.
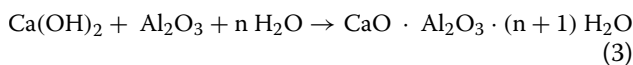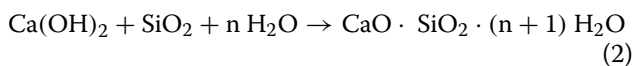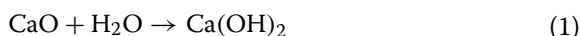
Cui *et al. Heritage Science* (2024) 12:174

Page 3 of 21

Data processing and analysis

Relevant data must be collected to establish analytical models for predicting the compressive strength of slurries. In response to the research results on slurries for anchor grouting and fissure grouting at earthen sites in Northwest China, this study collected the compressive strength data of Lime slurries (SH-(Lime+F+C) and N-(Lime+F+C)), CGN slurries (CGN-C, CGN-F, CGN-S, CGN-(F+S), N-(CGN-B-C)), and PS slurries (PS-C, PS-F, PS-(C+F), PS-Z). Table 1 lists the 523 data sets from different studies, including data sources, slurry names, numbers, admixture materials, ratios, and curing times. SSF in Table 1 represents the sodium fluorosilicate selected as the curing agent in the PS slurry. The liquid–solid ratio was calculated based on the sum of the masses of the solid materials in the slurries [36].

The total testing dataset was divided according to the different slurry types into sub-dataset_1, sub-dataset_2, and sub-dataset_3.

In order to more accurately reflect the influence of each admixture on the compressive strength of the slurry, this study will use the molar ratio of the main reacting oxides and the bulk density of each mixture material as input variables. Among them, the molar ratio of oxides can describe the impact of different effective components involved in chemical reactions within different data sources on mechanical strength, as well as reflect the composition and structure of reaction products, which is more helpful for sub-sequent mechanism analysis of prediction models [34]. The bulk density of admixtures can be used to describe the different physical properties of the same admixture from different sources. While combined with the amount of admixture added, it can more accurately describe the impact of this difference on compressive strength in the slurry.

The main chemical reactions generated during the mixing and curing of quicklime slurries are the following [37]:

$$CaO + H_2O \rightarrow Ca(OH)_2 \tag{1}$$

$$Ca(OH)_2 + SiO_2 + n\,H_2O \rightarrow CaO \cdot SiO_2 \cdot (n+1)\,H_2O \tag{2}$$

$$Ca(OH)_2 + Al_2O_3 + n\,H_2O \rightarrow CaO \cdot Al_2O_3 \cdot (n+1)\,H_2O \tag{3}$$

$$Ca(OH)_2 + CO_2 \rightarrow CaCO_3 + H_2O \tag{4}$$

The chemical reaction in Eq. (4) is difficult to quantify. Therefore, sub-dataset_1 chooses the CaO content (X1), $n(CaO)/n(H_2O)$ (X2), $n(SiO_2)/n(CaO)$ (X3), and $n(Al_2O_3)/n(CaO)$ (X4) to indicate the effect of the chemical reactions on the compressive strength. The molar ratio reflects the effect of the dosage on the final material

performance. Only the CaO in quicklime, the SiO_2 and Al_2O_3 in fly ash, and H_2O in the binders were considered in this sub-dataset. For the bulk density or dry density of each admixture material, its effect on the compressive strength must be analyzed in combination with the mass ratio of each admixture. Their respective expressions are ρ_b (Lime)$*\frac{m(Lime)}{m(Admixture)}$, ρ_b (FA)$*\frac{m(FA)}{m(Admixture)}$, and ρ_d (Soil)$*\frac{m(Soil)}{m(Admixture)}$, which are referred to as ρ_b (Lime) (X5), ρ_b (FA) (X6), and ρ_d (Soil) (X7) for concision. For the binders, in addition to X2, this study reflects the effect of their dosage on the compressive strength using the liquid–solid ratio (L/S) (X8). $m(Binder)/m(CaO)$ (X9) reflects the effect of polar hydroxyl groups in SH or polysaccharide substances in sticky rice paste to constrain and regulate the crystallization process of calcium-based materials. This indicates the mass ratio of the effective components in the binders to $m(CaO)$ [37, 38]. The effect of the content and arrangement of the active ingredients within the binders on the compressive strength is indirectly reflected by η (Binder) (X10), which indicates the viscosity of the binders [39]. The curing time (X11) reflects the effect of the curing conditions on the compressive strength. These 11 variables form the input variables of sub-dataset _1. The molar ratios of the oxides were converted from the mass-ratio data in the original literature. Table 2 presents the statistical results of the input and output variables for sub-dataset_1.

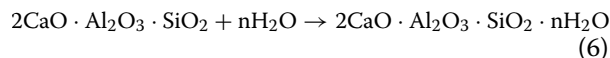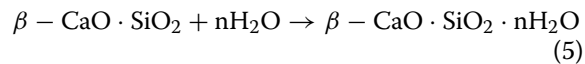The main chemical reactions of the slurries from sub-dataset_2 were added to Eq. (5) and Eq. (6) [40]:

$$\beta - CaO \cdot SiO_2 + nH_2O \rightarrow \beta - CaO \cdot SiO_2 \cdot nH_2O \tag{5}$$

$$2CaO \cdot Al_2O_3 \cdot SiO_2 + nH_2O \rightarrow 2CaO \cdot Al_2O_3 \cdot SiO_2 \cdot nH_2O \tag{6}$$

Table 3 presents the statistical results of the input and output variables for sub-dataset _2. In this table, SiO_2^* and $Al_2O_3^*$ indicate their content in CGN, whereas SiO_2 and Al_2O_3 correspond to fly ash. In addition, only the CaO in CGN and the H_2O in the water or binder were considered in this sub-dataset. In addition, in order to reduce the number of input variables in the model, ρ_b (FA) is used to describe the product value of the bulk density and mass ratio of fly ash and bentonite in the slurry. And ρ_d (Soil) is used to describe the product value of the density and mass ratio of site soil and quartz sand. The sub-dataset_2 contains limited data with bentonite and site soil. Bentonite has good lubricity and particle exchange properties similar to the fly ash. The site soil and quartz sand do not participate in chemical reactions within the slurry.

The solidification mechanism of the slurries in sub-dataset_3 is that of the silicate in PS gels with particles in fly ash and soil, whose combination forms a relatively

Cui *et al. Heritage Science*   (2024) 12:174

Page 4 of 21

Table 1 Database summary

Source	Slurry name	Number	Material	Ratio	Curing time
Current study	SH-(Lime + F + C)	90	Quick lime Fly ash Soil SH	0.10–0.30 0.10–0.30 0.40–0.80 0.50–0.58	7–180 days
Current study	N-(Lime + F + C)	60	Quicklime Fly ash Soil Sticky rice paste	0.10–0.30 0.10–0.30 0.40–0.80 0.58–0.70	7–90 days
Cui et al. [8]	SH-(Lime + F + C) N-(Lime + F + C)	6	Quicklime Fly ash Soil SH Sticky rice paste	0.10 0.10 0.80 0.48–0.52 0.40–0.50	28, 30 days
Cui et al. [9]	SH-(Lime + F + C)	30	Quicklime Fly ash Soil SH solution	0.30 0.20 0.50 0.5–0.54	7–90 days
Ren et al. [10]	CGN-C	7	Calcined ginger nuts Soil Water	0.50 0.50 0.61	3–120 days
Zhang et al. [11]	CGN-F	45	Calcined ginger nuts Fly ash Water	0.30–0.70 0.30–0.70 0.57–0.68	3–180 days
Ren et al. [12]	CGN-S	45	Calcined ginger nuts quartz sand Water	0.30–0.70 0.30–0.70 0.28–0.55	3–180 days
Wang et al. [13]	CGN-(F + S)	45	Calcined ginger nuts Fly ash quartz sand Water	0.30–0.70 0.15–0.35 0.15–0.35 0.40–0.62	3–180 days
Mao et al. [14]	N-(CGN-B-C)	20	Calcined ginger nuts Bentonite Soil Sticky rice paste	0.075–0.15 0.075–0.15 0.725–0.85 0.45–0.60	3–60 days
Yang et al. [15]	PS-F	10	Fly ash PS SSF	1 0.55–0.60 0.015	3–28 days
Li et al. [16]	PS-F	66	Fly ash PS SSF	1 0.60–1.20 0–0.04	3–90 days
Zhang et al. [17]	PS-(F + C)	4	Fly ash Soil PS	0.50 0.50 0.39	3–28 days
Li et al. [18]	PS-(F + C)	25	Fly ash Soil PS SSF	0.667 0.333 0.55–0.70 0–0.02	3–90 days
Li et al. [19]	PS-C	13	Soil PS SSF	1 0.7–1.035 0.01–0.04	3–45 days
Li et al. [20]	PS-C	25	Soil PS SSF	1 0.50–0.65 0–0.02	3–90 days
Li et al. [21]	PS-Z	32	Diatomite PS SSF	1 0.75–1.25 0.01–0.04	3–45 days

Cui *et al. Heritage Science* (2024) 12:174

Page 5 of 21

Table 2 Statistics information for sub-dataset_1

Variables	Notation	μ	Max	Min	σ
CaO content	X1	0.214	0.281	0.086	0.071
$n(H_2O)/n(CaO)$	X2	9.746	23.268	5.624	5.192
$n(SiO_2)/n(CaO)$	X3	0.414	0.610	0.190	0.147
$n(Al_2O_3)/n(CaO)$	X4	0.127	0.199	0.059	0.046
ρ_b (Lime) (g/cm^3)	X5	0.542	0.671	0.220	0.177
ρ_b (FA) (g/cm^3)	X6	0.132	0.246	0.077	0.060
ρ_d (Soil) (g/cm^3)	X7	1.050	1.440	0.708	0.247
L/S	X8	0.566	0.700	0.400	0.060
m(Binder)/m(CaO)	X9	0.057	0.169	0.027	0.037
η(Binder) (Pa·s)	X10	4.354	4.960	3.170	0.847
Curing time (day)	X11	61	180	7	58
Compressive strength (MPa)	Y	1.349	2.516	0.340	0.648

Table 3 Statistics information for sub-dataset_2

Variables	Notation	μ	Max	Min	σ
CaO content	X1	0.244	0.400	0.037	0.100
$n(H_2O)/n(CaO)$	X2	9.167	38.095	4.277	8.397
$n(H_2O)/n(SiO_2{}^*)$	X3	14.475	52.686	3.544	11.086
$n(H_2O)/n(Al_2O_3{}^*)$	X4	74.924	134.752	36.995	20.499
$n(SiO_2)/n(CaO)$	X5	0.562	2.051	0	0.569
$n(Al_2O_3)/n(CaO)$	X6	0.261	0.909	0	0.264
ρ_b (CGN) (g/cm^3)	X7	1.177	1.862	0.191	0.475
ρ_b (FA) (g/cm^3)	X8	0.183	0.574	0	0.175
ρ_d (Soil) (g/cm^3)	X9	0.560	1.573	0	0.481
L/S	X10	0.514	0.680	0.280	0.107
m(Binder)/m(CaO)	X11	0.036	0.588	0	0.127
η(Binder) (Pa·s)	X12	1.203	6.680	0.890	1.114
Curing time (day)	X13	65	180	3	60
Compressive strength (MPa)	Y	6.199	20.660	0.220	5.574

Table 4 Statistics information for sub-dataset_3

Variables	Notation	μ	Max	Min	σ
$n(SiO_2) / n(K_2O)$	X1	3.783	4.500	3.300	0.272
c(PS)	X2	0.146	0.180	0.030	0.036
η(PS) (Pa·s)	X3	2.015	2.650	1.300	0.395
SSF content	X4	0.018	0.040	0	0.008
L/S	X5	0.716	1.250	0.390	0.191
ρ_b (FA)	X6	0.452	0.860	0	0.389
ρ_d (Soil)	X7	0.501	1.450	0	0.543
Curing time (day)	X8	24	90	3	17
Compressive strength (MPa)	Y	3.989	23.100	0.110	4.537

stable structure [41]. Therefore, the effects of chemical reactions on the compressive strength can be disregarded in this sub-dataset. Table 4 presents the statistical results of the input and output variables for sub-dataset _3. In this table, c (PS) indicates the mass concentration of the PS. $n(SiO_2)/n(K_2O)$ indicates the modulus of PS, which refers to the molar mass ratio of SiO_2 to K_2O in potassium silicate solution. And this variable represents the content of effective bonding components in PS solution. The soil contains loess and diatomite. SSF is the firming agent used for the slurries.

The non-linear correlations between different variables in these tables were quantified through Pearson's correlation coefficient (ρ), using Eq. (7) [42].

$$\rho = \frac{\sum_{i=1}^{n}\left(P_1 - \overline{P_1}\right)\left(P_2 - \overline{P_2}\right)}{\sqrt{\sum_{i=1}^{n}\left(P_1 - \overline{P_1}\right)^2 \sum_{i=1}^{n}\left(P_1 - \overline{P_1}\right)^2}} \tag{7}$$

where P_1 and P_2 are the parameters considered for correlation, and $\overline{P_1}$ and $\overline{P_2}$ are their mean values, respectively. ρ ranges from −1 to 1, where 1, 0, and −1 respectively represent strong forward, uncorrelated, and strong inverse relational relationships between these parameters. Figures 1, 2, 3 show the distributions and correlations of the different input and output factors in each sub-dataset. In these figures, dark blue indicates a positive correlation, and yellow negative correlation.

In sub-dataset_1, the number of positive and negative correlations between the input and output variables is approximately half of each, with a strong positive correlation between X11 and Y and weak correlations between X3, X4, and Y. In sub-dataset_2, the input variables were mostly inversely correlated with Y, the values of which were generally small, whereas there was a strong positive correlation between X13 and Y. In sub-dataset_3, there were mostly positive correlations between the input variables and Y, especially for X5. However, X1 and X7 exhibited weak correlations with Y.

Development and performance of models

The ten ML algorithms for predicting compressive strength in this study were developed using MATLAB. These algorithms involve several types of learning including single, integrated, tree-based, and deep learning. The GA can optimize the hyper parameters of each model. Performance evaluation indices can be used to discern prediction models with high generalization and accuracy. Subsequently, a relative importance analysis and SHAP interpretation can identify the significant input variables for predicting the compressive strength and determine whether their effects are positive or negative [43, 44]. A schematic diagram of each machine-learning algorithm and GA is shown in Fig. 4 [27–29, 45–47].
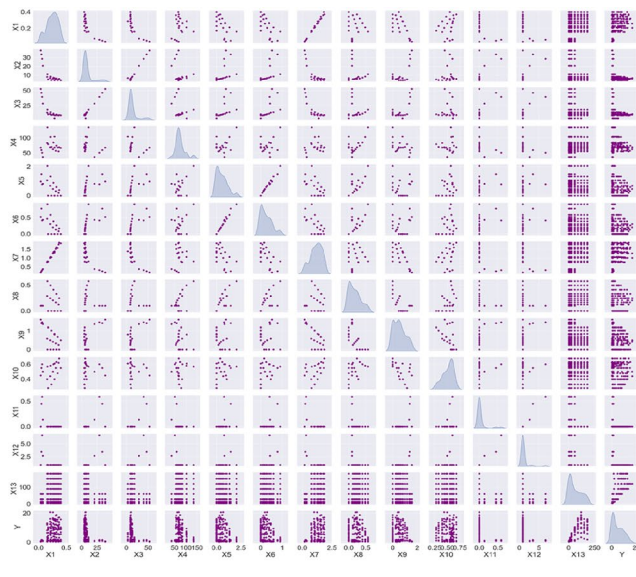
Cui *et al. Heritage Science*   (2024) 12:174
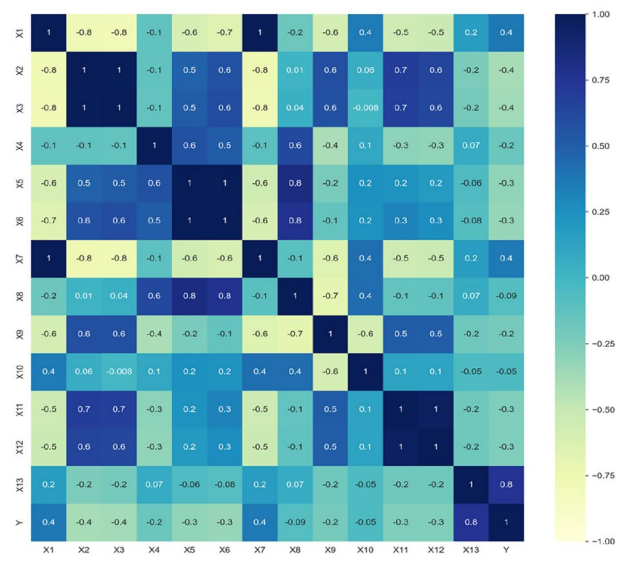
Page 6 of 21
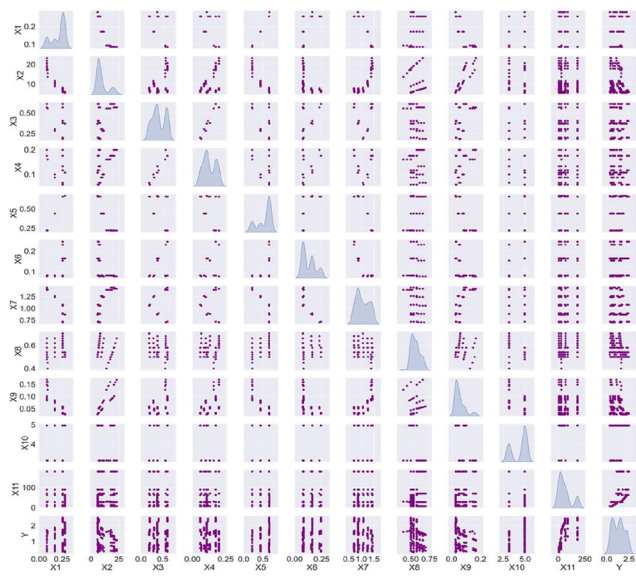
(1) Distribution of variables
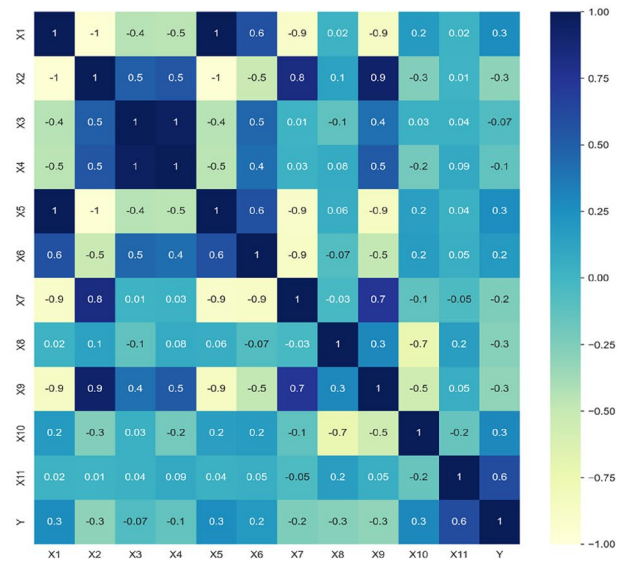
(2) Heat map of correlation matrix

Fig. 1 Distribution and correlation of variables for sub-dataset_1

(1) Distribution of variables
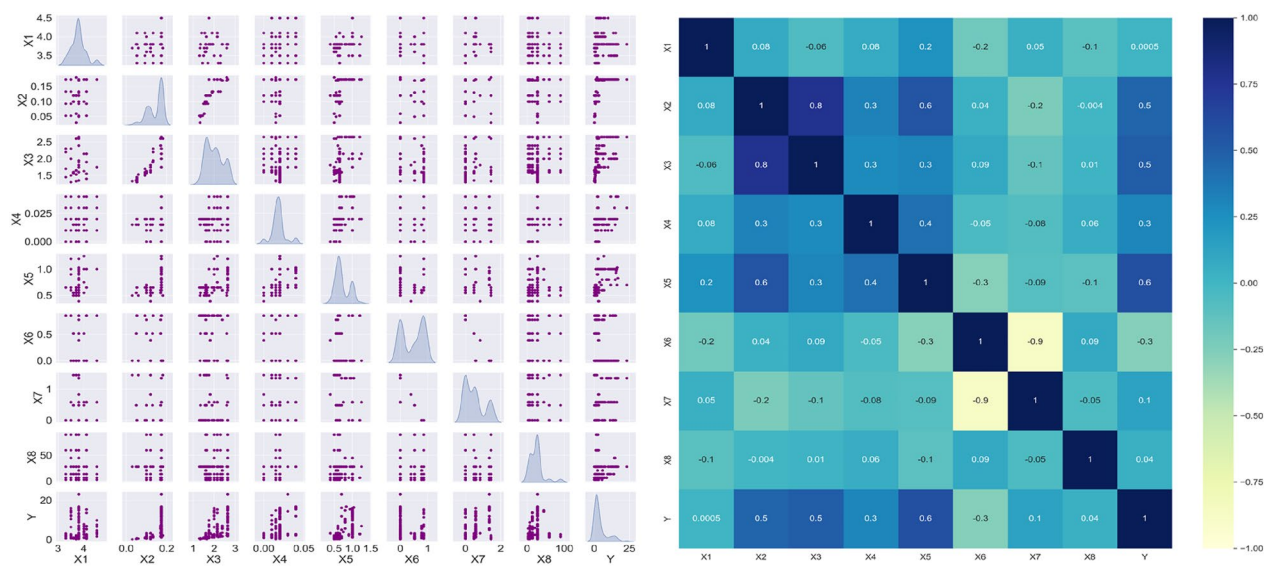
(2) Heat map of correlation matrix

Fig. 2 Distribution and correlation of variables for sub-dataset_2

Tuning of hyper-parameters

To guarantee high prediction accuracy with ML models, appropriate hyper parameter values must be selected. However, because the arithmetic principle of each algorithm varies, the best hyper parameter values must be obtained for different algorithms by using a controlled variable. In this study, GA was used to optimize the hyper parameters of each algorithm and provide their best combination. The optimization process requires the parameters of the GA to be set uniformly. For example, if the population size is set to 20, the maximum number of generations is set to 100, the crossover probability is set to 0.4, and the mutation probability is set to 0.05. Table 5 lists the best combinations of hyper parameter values

(1) Distribution of variables

(2) Heat map of correlation matrix

Fig. 3 Distribution and correlation of variables for sub-dataset_3

for the sub-datasets. For ANN, the hyper-parameters are sizes of hidden layer 1, 2 and 3 (NodeNum_1, Node-Num_2 and NodeNum_3); for BP, the hyper-parameters are weight parameters (W1, W2) and bias parameters (B1, B2); for SVR, the hyper-parameters are coefficient of the penalty term(c) and gamma value of gaussian kernel (g); for RF, the hyper-parameters are the total number of the trees (tree_num) and the minimum number of samples required to be at a leaf node (minleaf); for XGBoost, the hyper-parameters are maximum number of iterations (max_num_iters), maximum depth of the tree (params. max_depth) and learning rate (params.eta); for CNN, the hyper-parameters are result of neural network convolution on the input data (feature_map),reduction factor of learning rate (LearnRateDropFactor) and number of rounds to reduce learning rate (LearnRateDropPeriod); for LSTM, the hyper-parameters are size of hidden layer units (numHiddenUnits), LearnRateDropFactor, and LearnRateDropPeriod; for GRU, the hyper-parameters are numHidden Units, LearnRateDropFactor and Learn-RateDropPeriod; for RBF, the hyper-parameter is expansion speed (Spread); for ELM, the hyper-parameter is the number of hidden layer nodes (L).The hyper parameters of the BP are expressed in the matrix.
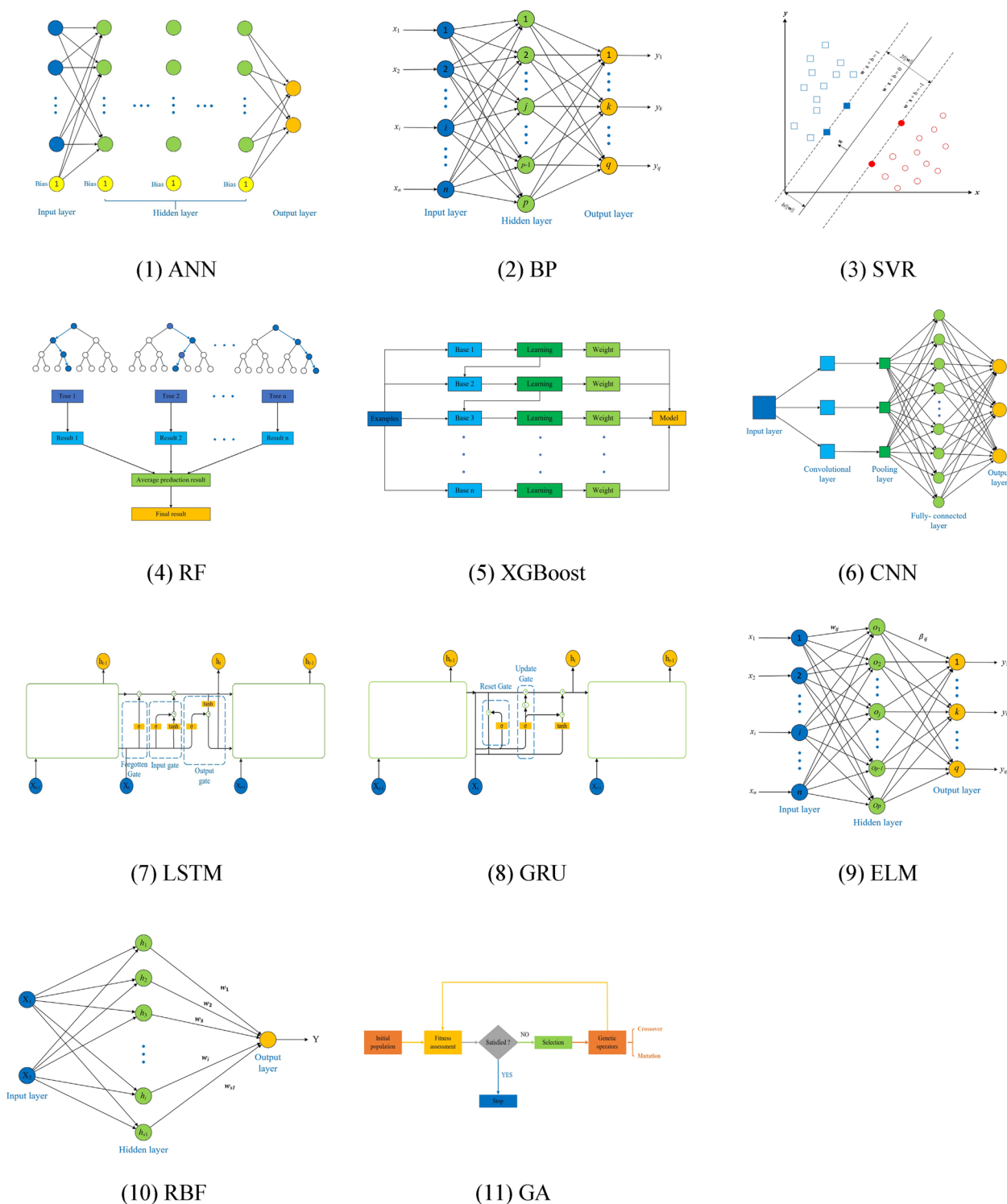
Assessment of models

The model prediction results are shown in Figs. 5, 6, 7. In Fig. 5, most data points in sub-dataset_1 are centralized within the 20% error region for all ten models, especially

for XGBoost; most points are almost on top of the diagonal. In Fig. 6, the data points in sub-dataset_2 are uniformly distributed on both sides of the centerline, except for some with small measurements. This shows the good predictive performance of these models. In Fig. 7, although the data points in sub-dataset _3 are more dispersed, the points in XGBoost continue to be generally close to the centerline, showing the great predictive accuracy of this model.

To evaluate the accuracy of the predictions for each model more intuitively, we used the metrics in Table 6 to assess these models. R^2 indicates the proximity between the predicted and measured values. The closer the value of R^2 is to 1, the higher the prediction accuracy of the model [34]. VAF indicates the degree of correlation between the predicted results of the model and the actual results. The closer the value of VAF is to 100, the higher the prediction accuracy of the model. RMSE and MAE indicate the deviation between predicted results of the model and the actual results. COV indicate the degree of dispersion between the predicted results of the model and the actual results. The closer the values of RMSE, MAE and COV are to 0, the higher the prediction accuracy of the model [42]. y_i, y_i^P, and $\bar{y}$ represent the measured value of number i, the predicted value of number i, and the average of all measured values, respectively.

Figures 8, 9, 10 show the radar plots of the values of the evaluation metrics for each sub-dataset, which contain the training and test sets. The models with high R^2

Cui *et al. Heritage Science*   (2024) 12:174

Page 8 of 21

(1) ANN

(2) BP

(3) SVR

(4) RF

(5) XGBoost

(6) CNN

(7) LSTM

(8) GRU

(9) ELM

(10) RBF

(11) GA

Fig. 4 Graphical representation of the ten ML models and GA

and VAF values and small RMSE, MAE, and COV values are considered high-performance models. The closer the point of each evaluation metric value is to the center, the higher the prediction accuracy of the representative model.

Cui *et al. Heritage Science*   (2024) 12:174

Page 9 of 21

Table 5 Hyper-parameters of the ML algorithm

Model	Hyper-parameter	Sub-dataset_1	Sub-dataset_2	Sub-dataset_3
ANN	NodeNum_1	17	30	89
	NodeNum_2	9	8	7
	NodeNum_3	9	8	7
BP	W1	Matrix 1*66	Matrix 1*90	Matrix 1*48
	W2	Matrix 1*6	Matrix 1*6	Matrix 1*6
	B1	Matrix 1*6	Matrix 1*6	Matrix 1*6
	B2	0.53	−0.12	−0.11
SVR	c	72	2	2.83
	g	0.09	0.25	0.71
RF	tree_num	49	667	205
	minleaf	1	2	1
XGBoost	max_num_iters	927	312	889
	params.max_depth	5	10	65
	params.eta	0.2	0.14	0.15
CNN	feature_map	97	126	111
	LearnRateDropFactor	0.1	0.1	0.1
	LearnRateDropPeriod	750	750	700
LSTM	numHiddenUnits	165	157	190
	LearnRateDropFactor	0.15	0.3	0.45
	LearnRateDropPeriod	750	650	750
GRU	numHiddenUnits	85	82	94
	LearnRateDropFactor	0.25	0.4	0.75
	LearnRateDropPeriod	700	750	550
RBF	Spread	13	57	33
ELM	L	34	42	43

XGBoost had the best evaluation metrics in each sub-dataset. In sub-dataset_1, for both the training and testing sets, the R^2 and VAF values are higher than 0.99 and close to 100, respectively. The RMSE and MAE values were both less than 0.05. The COV value in XGBoost was significantly lower than those in the other models. In the other two sub-datasets, the testing set of sub-dataset_3 contained the worst evaluation metrics, with R^2, RMSE, MAE, VAF, and COV values of 0.933, 1.105, 0.751, 93.35, and 31.40, respectively. However, the metrics were still significantly better than those of other models. In summary, XGBoost has the best prediction accuracy and generalizability, yielding more robust results with lower uncertainty. In contrast, some widely used models, such as BP, SVR, and RF, did not perform as expected.

Relative importance of the input variables

Based on the test results of the model performance evaluation, XGBoost with GA-optimized hyperparameters was selected to analyze the relative importance of the input variables. This study imported datasets and hyperparameters to obtain these importance scores for each input variable in the three sub-datasets. They were then used to assess the contribution of each input variable in this model. The importance scores are visually comparable in Fig. 11.

The most important input variables in sub-dataset_1 were $n(H_2O)/n(CaO)$, curing time, $m(Binder)/m(CaO)$, and L/S. Among these, $n(H_2O)/n(CaO)$ reflects the main chemical process of calcium hydroxide from quicklime and water in the slurries, which significantly impacts the compressive strength. The curing time reflects that the increase in the compressive strength of the grouting slurries is closely related to time. The $m(Binder)/m(CaO)$ and L/S ratios indicate that the effects of binders on the compressive strength are considerable.

Similarly, the most important input variables in sub-dataset_2 were the curing time, CaO content, $n(H_2O)/n(CaO)$, and $n(H_2O)/n(SiO_2^*)$. In sub-dataset_3, the values are L/S, c(PS), and the curing time. These results indicate that the calcium-based materials, binders, and curing time significantly impact the compressive strength. However, Fig. 11 shows that some input variables have a lower relative importance for the compressive strength in each sub-dataset. To simplify the analysis, this study neglected the variables with a small effect on the

Cui *et al. Heritage Science*   (2024) 12:174

Page 10 of 21
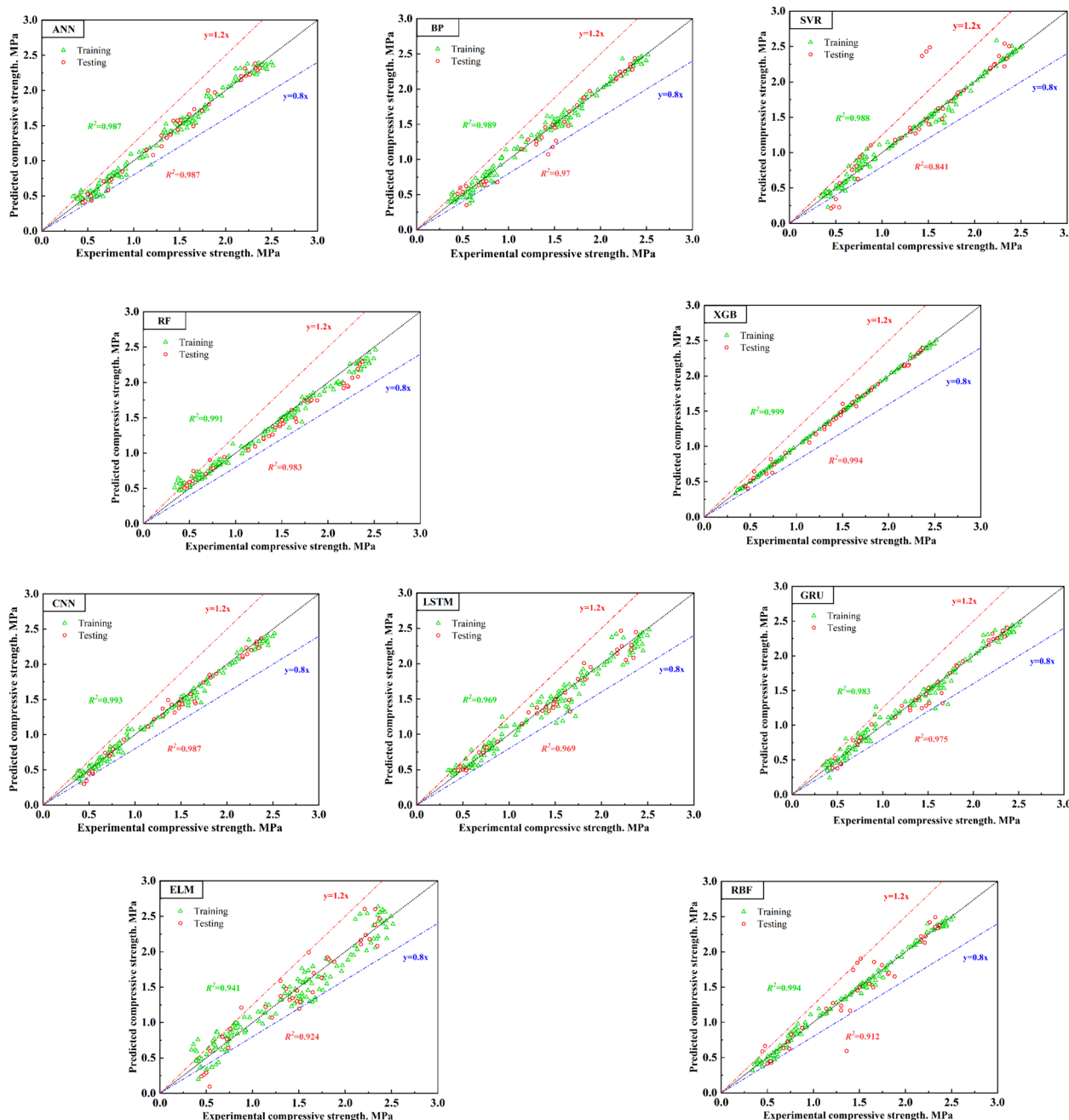
Fig. 5 Experimental and predictive values of the ten machine learning models for sub-dataset_1

compressive strength and employed the remaining variables for re-modeling. However, instead of directly selecting the input variables with a significant effect on the compressive strength; this process must retain at least one input variable for each admixture material to reflect its effect on the slurries. Figure 12 shows the final input variables with their relative importance in each

sub-dataset, which were obtained through re-modeling using GA-XGBoost.

SHAP analysis of the input variables

In the previous section, we selected the input variables for each sub-dataset. However, determining whether the effect of the variables on the compressive strength is positive or negative is difficult if only the relative importance

Cui *et al. Heritage Science*   (2024) 12:174

Page 11 of 21
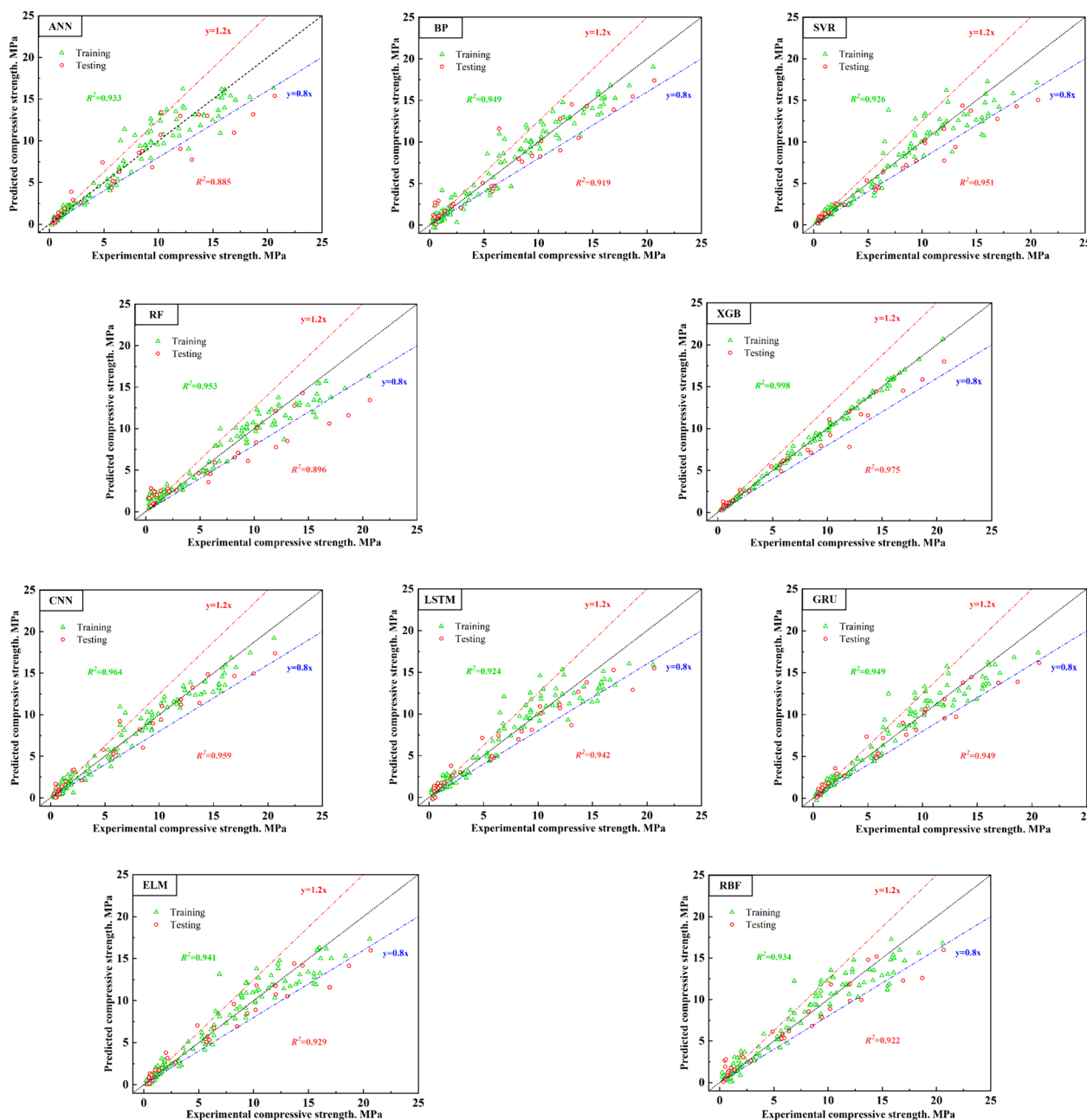
Fig. 6 Experimental and predictive values of the ten machine learning models for sub-dataset_2

analysis is considered. Therefore, the Shapley additive explanation was introduced [48].

The SHAP interpretation package in Python was used to obtain the analysis results. Figures 13, 14, 15 show the results of the global analysis and the SHAP value of each input variable in the three sub-datasets. The first plot of each figure presents the results of the global analysis. Each point corresponds to an input sample in the dataset and each row represents an input variable. Red points

indicate larger SHAP values, whereas blue indicates smaller ones. The other plots in Figs. 13, 14, 15 show the relationships between the variables and their SHAP values. The global analysis plots show that, except for a few variables in the top rows, the boundary of positive and negative contribution for other variables is not obvious, which indicates that the relationship between the input variables and the predictions is complex. The other plots with input variables and their SHAP values show the

Cui *et al. Heritage Science*   (2024) 12:174

Page 12 of 21
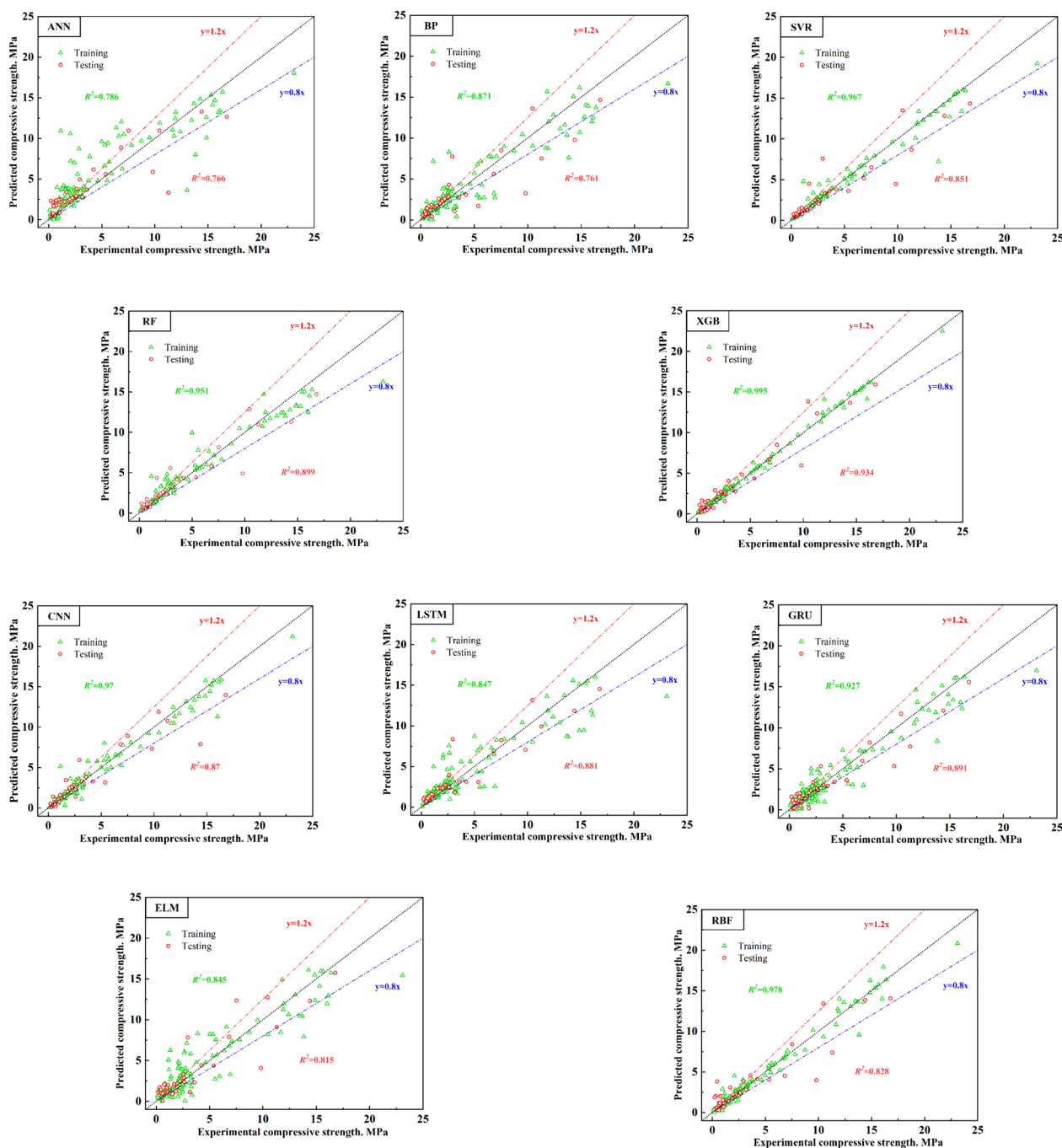
Fig. 7 Experimental and predictive values of the ten machine learning models for sub-dataset_3

overall trend of the effect of the input variable on the compressive strength.

In Fig. 13, the plot of the curing time shows that with increasing time, the SHAP value first increases and then gradually stabilizes. Figures 14, 15 show similar trends. However, the SHAP value in these two sub-datasets decreased in the latter stage when the curing time

increased. This indicates that the compressive strength of the slurries did not continuously increase over time. Therefore, determining the optimal curing time for each slurry was necessary. The variables L/S, $n(H_2O)/n(CaO)$ and $m(Binder)/m(CaO)$ have similar trends in the plots, showing a decreasing SHAP value when each variable increases. However, the downward trend in the latter
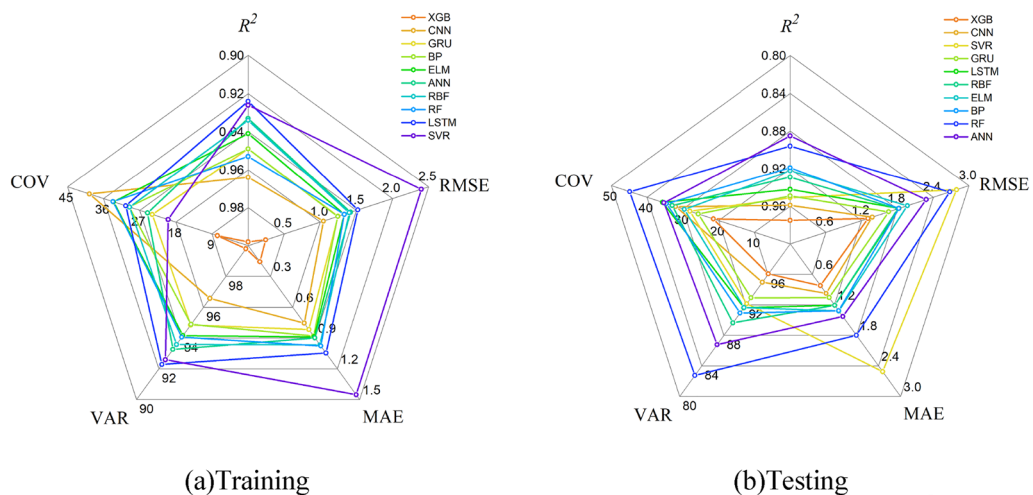
Cui *et al. Heritage Science* (2024) 12:174

Page 13 of 21

Table 6 Statistical parameters used to evaluate the performance of different machine learning algorithms

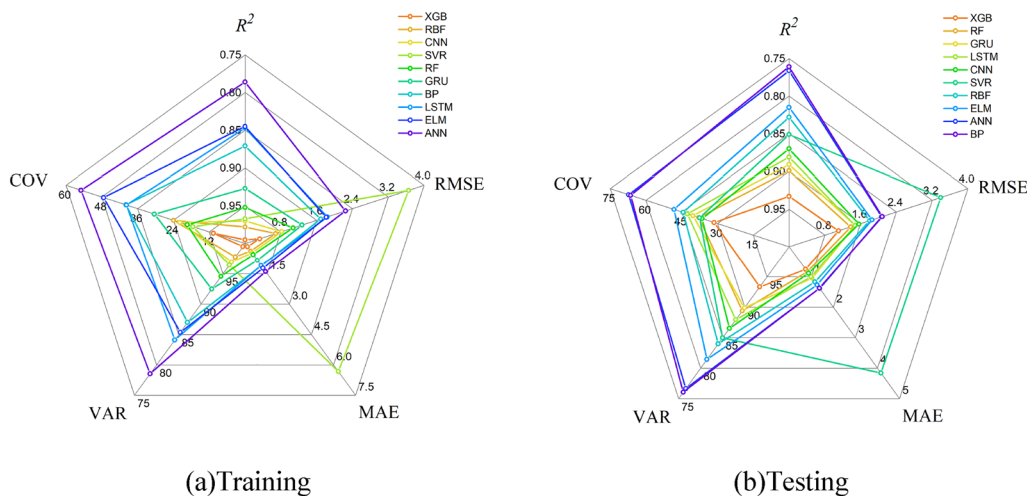| Metric	Ideal values	Formula
Coefficient of determination (R^2)	1	$\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2-\sum_{i=1}^{n}(y_i-y_i^P)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$
Root mean square error (RMSE)	0	$\sqrt{\frac{1}{n}\times\sum_{i=1}^{n}(y_i-y_i^P)^2}$
Mean absolute error (MAE)	0	$\frac{1}{n}\times\sum_{i=1}^{n}\left|y_i^P-y_i\right|$
Variance account for (VAF)	100	$\left(1-\frac{var(y_i-y_i^P)}{var(y_i)}\right)\times100\%$
Coefficient of variation (COV)	0	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{y_i^P}-\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{y_i^P}\right)^2}\times\frac{n}{\sum_{i=1}^{n}\frac{y_i}{y_i^P}}$

two variables was not linear. These results indicate that, as the binder dosage increased, a large amount of water was added to the slurries, which deteriorated their compressive strength. On the other hand, the effective component in the binders deteriorates the compressive strength of the slurries by affecting the structure of the generated calcium carbonate. In the plots of CaO content, n(H$_2$O)/n(CaO), and m(Binder)/m(CaO), the trends of the three variables indicate that as the dosage of quicklime increases, the compressive strength of the slurries increases. However, when the dosages of both quicklime and the binder increased, the compressive strength of the slurries decreased. In addition, the plot of ρ_d (Soil) shows that the SHAP value first increases and then decreases when the variable increases. This indicates that higher
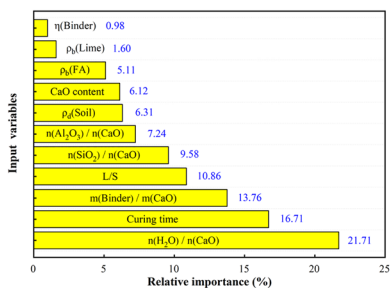

(a)Training   (b)Testing
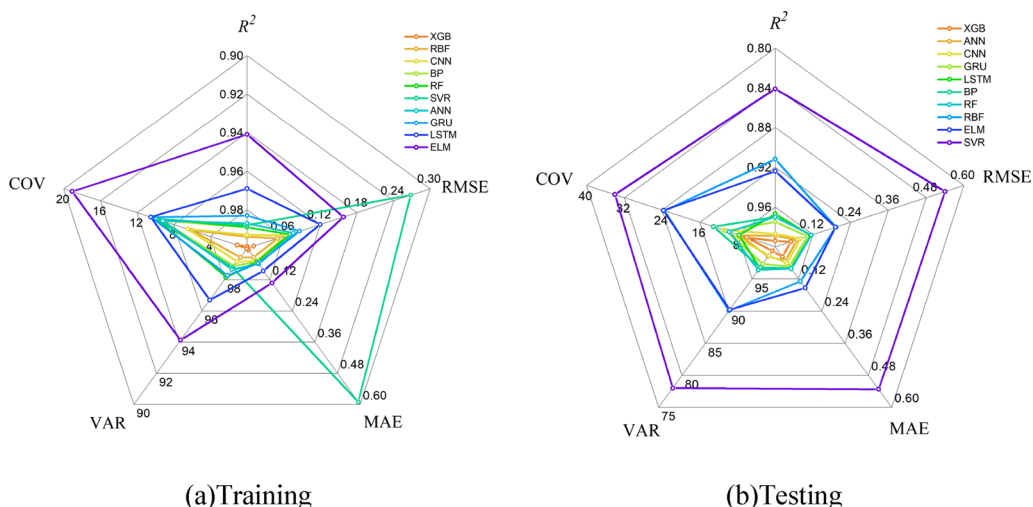
Fig. 8 Radar chart of the evaluation indexes of the ten ML models for sub-dataset_1

(a)Training   (b)Testing

Fig. 9 Radar chart of the evaluation indexes of the ten ML models for sub-dataset_2

Cui *et al. Heritage Science*   (2024) 12:174
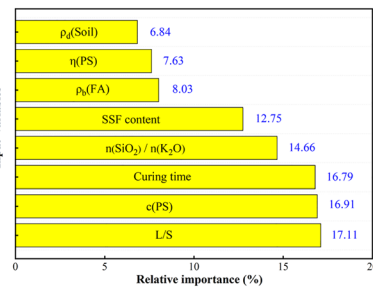
Page 14 of 21

(a)Training

(b)Testing

Fig. 10 Radar chart of the evaluation indexes of the ten ML models for sub-dataset_3

(1) sub-dataset_1
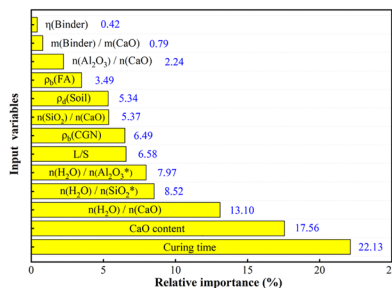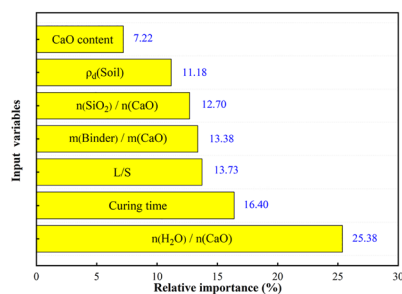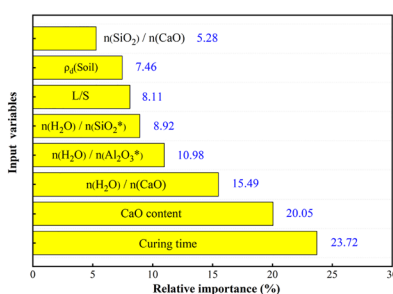
(2) sub-dataset_2

(3) sub-dataset_3

Fig. 11 Relative importance of input variables affecting the compressive strength

(1) Sub-dataset_1

(2) Sub-dataset_2
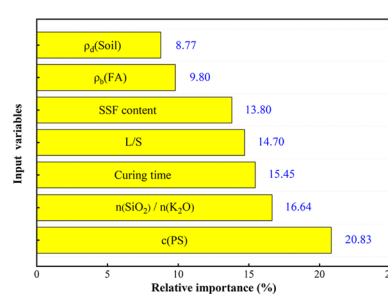
(3) Sub-dataset_3

Fig. 12 Relative importance of input variables affecting the compressive strength after optimization

loess dosages can increase the compressive strength of slurries within a certain range. Finally, the plot of $n(SiO_2)/n(CaO)$ shows that the SHAP value continuously increases as the variable increases, indicating that the

compressive strength of the slurries increases along with the dosage of fly ash.

Similarly, in the plots of Fig. 14, the variables $n(H_2O)/n(CaO)$, $n(H_2O)/n(Al_2O_3^*)$, and $n(H_2O)/n(SiO_2^*)$ show similar trends; the SHAP value first continues to

Cui *et al. Heritage Science*   (2024) 12:174

Page 15 of 21
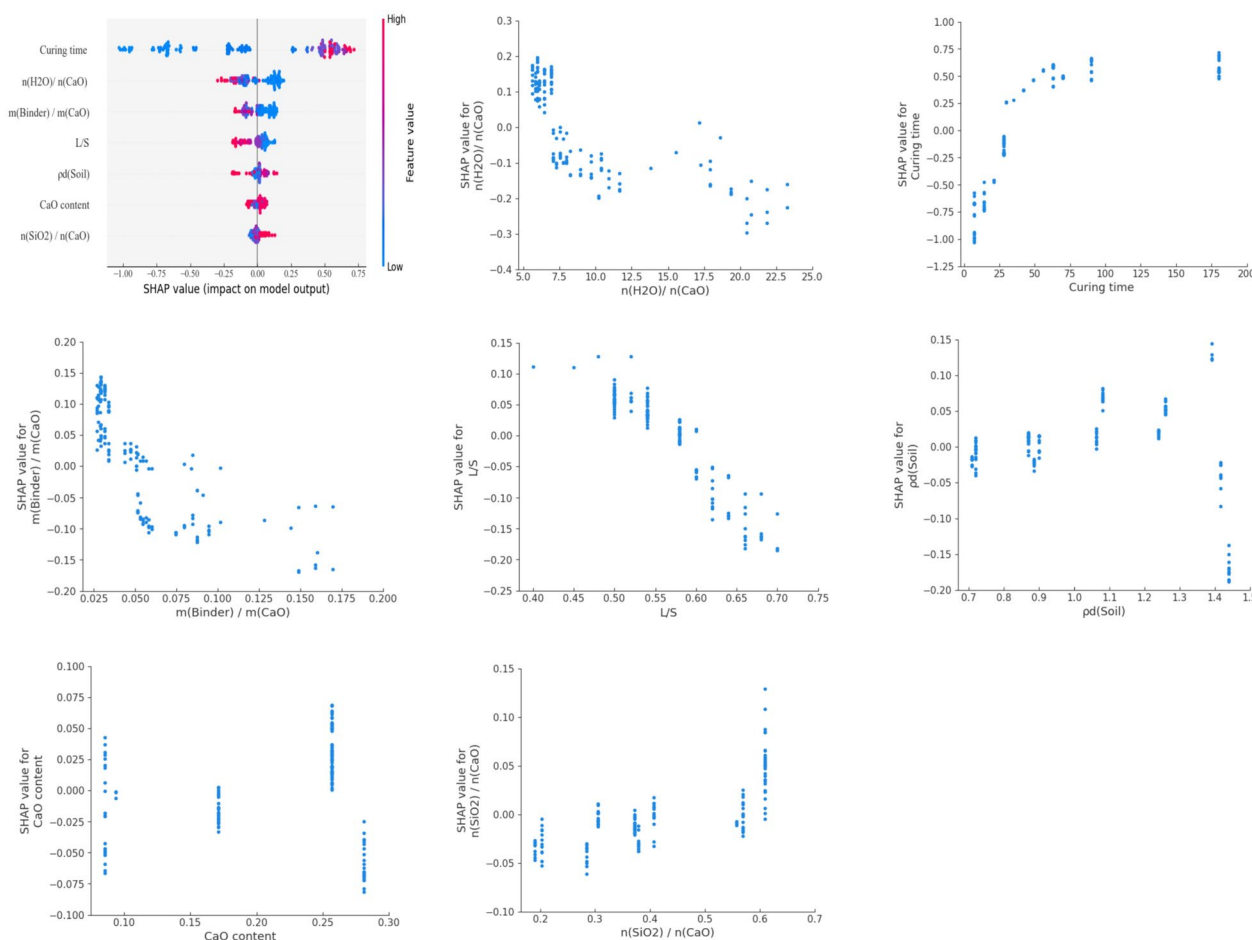
Fig. 13 SHAP values of the global analysis and seven input variables in sub-dataset_1

decrease and then gradually stabilizes. Combined with the trend of the CaO content, these results indicate that the compressive strength of the slurries did not continuously increase for higher CGN dosages. These results and the trend of L/S indicate that the dosage of water or binder can increase the compressive strength of the slurries within two specific ranges. The variables ρ_d (Soil) and $n(SiO_2)/n(CaO)$ have similar trends; the SHAP value first increases and then decreases when each variable increases. This indicates that an increase in the dosage of soil, quartz sand, fly ash, or bentonite can increase the compressive strength of slurries within a certain range.

In the plots of Fig. 15, the variables L/S, $n (SiO_2)/n (K_2O)$, and c (PS) have similar trends; the SHAP value first increases and then decreases when each variable increases, indicating that with an increase in the dosage, modulus, and concentration of PS, the compressive strength of slurries can increase within a certain range. The variables ρ_d (Soil) and ρ_b (FA) have similar trends; the SHAP value first decreases and then increases until

it decreases when each variable increases. This indicates that the fly ash and soil or diatomite have higher mechanical strength when utilized alone in PS slurries. Finally, the SSF content variable has a complex, changing trend, which indicates that values of SSF content within 0.01 and 0.03 in slurries lead to higher compressive strengths.

Multi-objective optimization
Objective functions
The optimization of the slurry ratio requires the analysis of three evaluation indicators: compressive strength, carbon emission coefficient, and material cost. In general, the objective function must be set to minimize carbon emissions and costs. However, the objective function of compressive strength must be discussed along with the utilization of slurries on earthen sites. Presently, research on fissure grouting slurries requires the compressive strength to be slightly higher than that of the site soil. Anchor grouting slurries have higher compressive strength requirements for them to be compatible with

Cui *et al. Heritage Science*   (2024) 12:174

Page 16 of 21
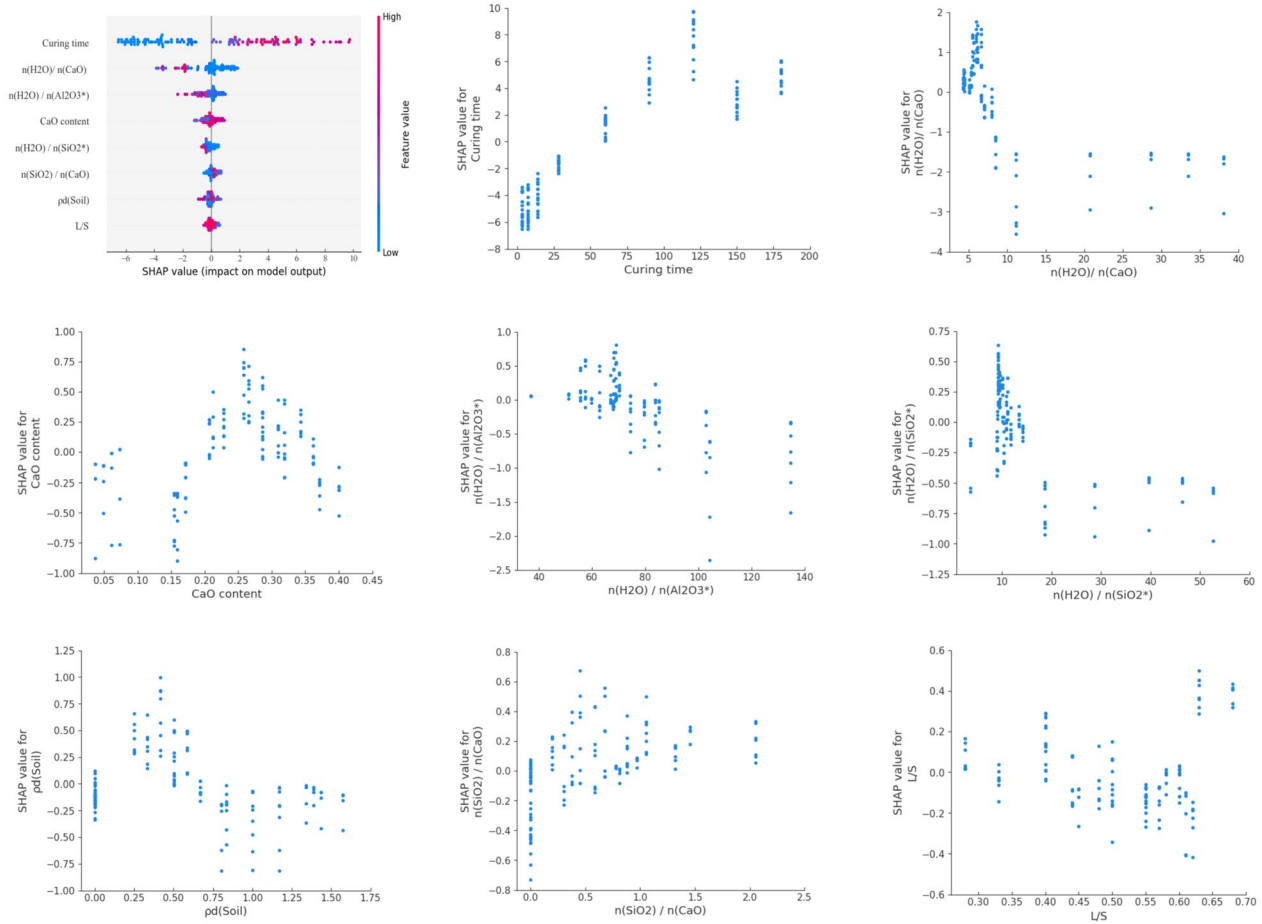
Fig. 14 SHAP values of the global analysis and eight input variables in sub-dataset_2

the anchor rods and site soil. However, no research has proven that the compressive strength of grouting slurries must be much greater than that of the site soil for optimal compatibility [5, 49]. In fissure grouting engineering, when the mechanical strength of the grouting slurry is less than that of the site soil, the part repaired by the slurries is more susceptible to erosion damage than the soil site, which causes secondary damage to earthen sites. Meanwhile, in anchor grouting engineering, if the mechanical strength of the grouting slurry is too small, the external forces acting on the earthen sites damage the slurry-soil interface before the slurry-rod interface is destroyed. This would lead to greater damage to the sites. In addition, the mechanical properties of earthen sites may vary significantly in different regions, geological conditions, and construction periods, leading to a more complex analysis of the mechanical compatibility between slurries and earthen sites [50, 51]. Therefore, this study favors slurries with higher compressive strengths, which are more widely utilized in the engineering of earthen site repairs. The optimization must maximize the

mechanical strength and minimize the carbon footprint and cost.

We modeled the objective function of the compressive strength of the slurry using the GA-XGBoost model. Because carbon emissions mainly originate from the production and transportation of admixtures in the slurries, the objective function of the carbon footprint (EC) can be expressed as follows [52]:
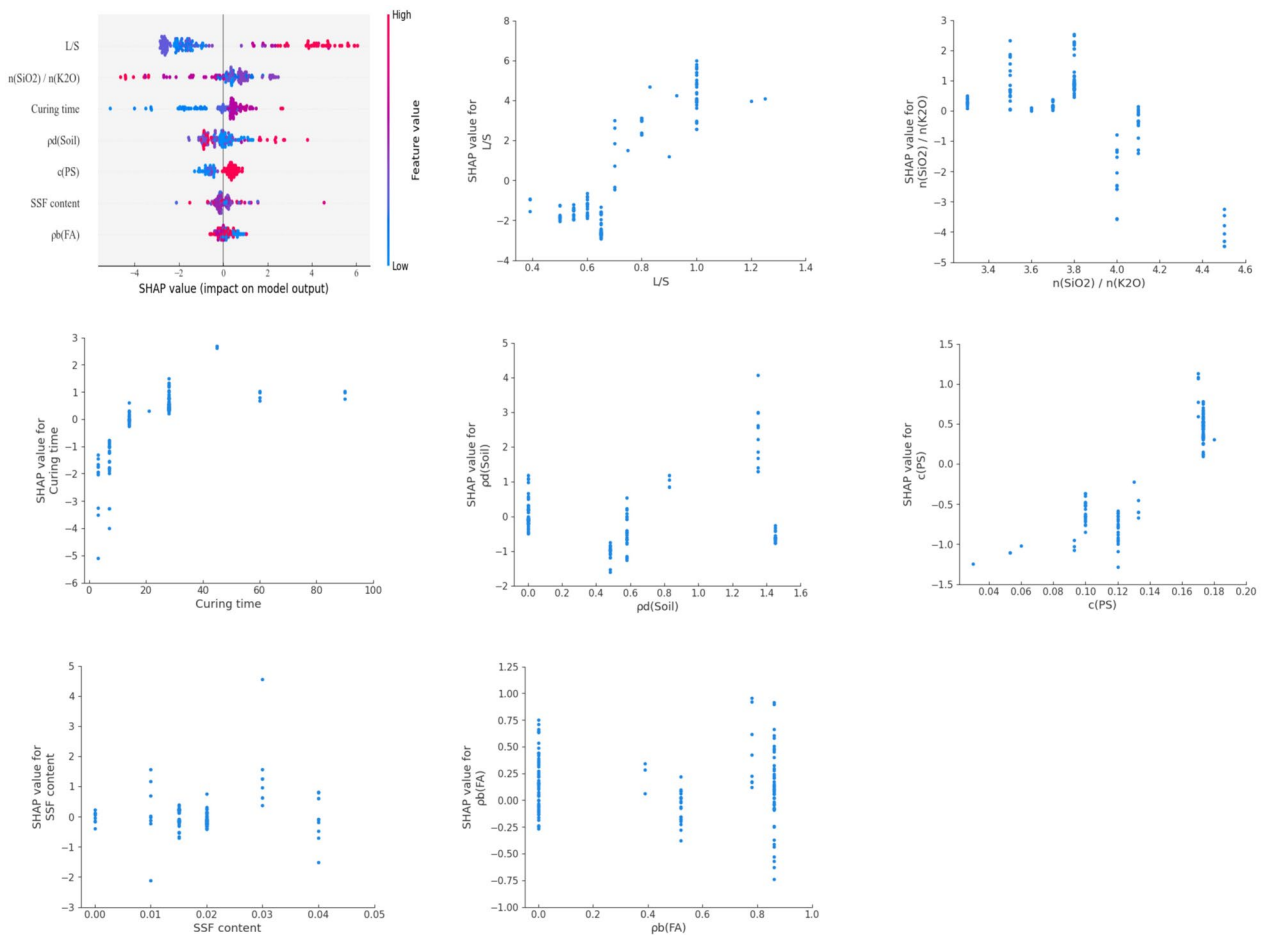
$$EC = EC_R + EC_T \tag{8}$$

$$EC_R = \sum_{i=1}^{n} Q_i EF_i \tag{9}$$

$$EC_T = \left(\sum_{i=1}^{n} Q_i D_i \right) \cdot EF_T \tag{10}$$

EC_R and EC_T represent the carbon footprints during production and transportation, respectively; Q_i represents the content of each admixture in the slurry;

Cui *et al. Heritage Science*   (2024) 12:174

Page 17 of 21

Fig. 15 SHAP values of the global analysis and seven input variables in sub-dataset_3

EF_i represents the carbon emission coefficient of each admixture; D_i represents the transportation distance of each admixture; and EF_T represents the carbon emission coefficient during transportation. The carbon emission coefficient of electricity consumption was 0.581 kgCO$_2$/(kW·h) based on the average value from the State Grid of China for 2022 [53]. However, several studies were conducted in Lanzhou, and many field tests did not specify their test sites. Therefore, to ease calculation and analysis, the transportation distance of each material was set according to the shortest distance from its origin to the Langongping Campus at the Lanzhou University of Technology. For the same material with different origins, the average distance was taken as the transportation distance. The carbon emission coefficient during the transportation was set to 0.18 CO$_2$/(ton ·km) [54].

Subsequently, the objective function for the production and transportation costs of the materials can be calculated as:

$$Price = \sum_{i=1}^{n} (Q_i P_i + P_{Ti}) \tag{11}$$

Q_i represents the content of each admixture in the slurry, P_i is the price of each admixture, and P_{Ti} represents the transportation cost of each admixture. Some materials require secondary processing after purchase, and the additional costs incurred during processing must be included. The transportation costs vary in different regions; therefore, the cost data do not vary linearly based on transportation mileage.

Table 7 lists the carbon emission factors and costs during the production and transportation of each admixture, where "\" denotes negligible parts. The prices of the raw materials were obtained based on purchasing experience and consultation with the material manufacturer. The transportation cost was obtained from consulting multiple transportation companies; this cost does not show linear changes based on the

Cui *et al. Heritage Science*   (2024) 12:174

Page 18 of 21

Table 7 Characteristic parameters for slurry mixture ingredients

Admixture materials	Price($/kg)		Carbon emission factor (kgCO$_2$/kg)	
	Production	Transport	Production	Transport
Water	< 0.001	\	0.168	\
Quicklime	0.111	0.048	1190	22
FA	0.083	0.028	8	3
Soil	0.145	0.090	\	128
SH	2.073(5%)	0.028	1550(5%)	3
Rice paste	0.094(3%)	\	383(3%)	\
CGN	4.147	0.111	930	200
Quartz sand	0.041	0.048	6.6	24
Bentonite	0.050	0.076	86	67
Diatomite	0.180	0.111	93	187
PS	0.885(26%)	0.041	272 (26%)	12
SSF	0.304 (40%)	0.041	70 (40%)	12

actual transportation mileage in different regions. The production price of the site soil was set based on labor costs from the experiences of on-site sampling, crushing, and screening of our research team. The adhesive concentration in the parentheses is the concentration at which the binders were purchased, which may need to be diluted during utilization. The carbon-emission coefficient of each material was obtained from the China Products Carbon Footprint Factors Database. Some materials, such as SH, CGN, PS, and SSF, are estimated by analyzing the possible carbon emissions generated through their production processes; therefore, the results may be inaccurate [21, 55–58].

Mixture optimization of the slurry

After determining the objective functions, the constraint conditions of each function must be set before conducting multi-objective optimization. The range constraints in this study varied within a range based on the number of slurry components in each sub-database. The proportional

constraints were selected based on the liquid–solid ratios of different slurries, as mentioned earlier.

In sub-dataset_1, binder (liquid) represents SH and sticky rice paste. Admixture (solid) represents quicklime, fly ash and soil. The ratio constraint is considered as follows:

$$0.10 \leq m(\text{Binder})/m(\text{Admixture}) \leq 0.70 \qquad (12)$$

In sub-dataset_2, binder (liquid) represents water and sticky rice paste. Admixture (solid) CGN, fly ash, quartz sand, bentonite and soil. The ratio constraint is considered as follows:

$$0.28 \leq m(\text{Binder})/m(\text{Admixture}) \leq 0.68 \qquad (13)$$

In sub-dataset_3, binder (liquid) represents PS. Admixture (solid) represents fly ash, soil and diatomite. The ratio constraint is considered as follows:

$$0.39 \leq m(\text{Binder})/m(\text{Admixture}) \leq 1.25 \qquad (14)$$

In addition, when analyzing the objective functions, the compressive strength must be considered in both the early and late stages to ensure experiment practicality.

In summary, three different multi-objective optimization scenarios were set for sub-dataset_1 sets for 28 days compressive strength, 63 days compressive strength, carbon emissions, and cost as objective functions, with 33 sets of data; sub-dataset_2 sets for 28 days compressive strength, 60 days compressive strength, carbon emissions, and cost as objective functions, with 20 sets of data; and sub-dataset_3, for only sets 28 days compressive strength (because of the lack of long curing-time data), carbon emissions, and cost as objective functions, with 85 sets of data. In these optimization scenarios, the compressive strength was calculated using the GA-XGBoost model, whereas the carbon emissions and cost were calculated using polynomials.

Optimization problems aim to maximize the compressive strength and minimize carbon emissions and cost. However, maximizing compressive strength may lead to

(1) sub-dataset_1   (2) sub-dataset_2   (3) sub-dataset_3
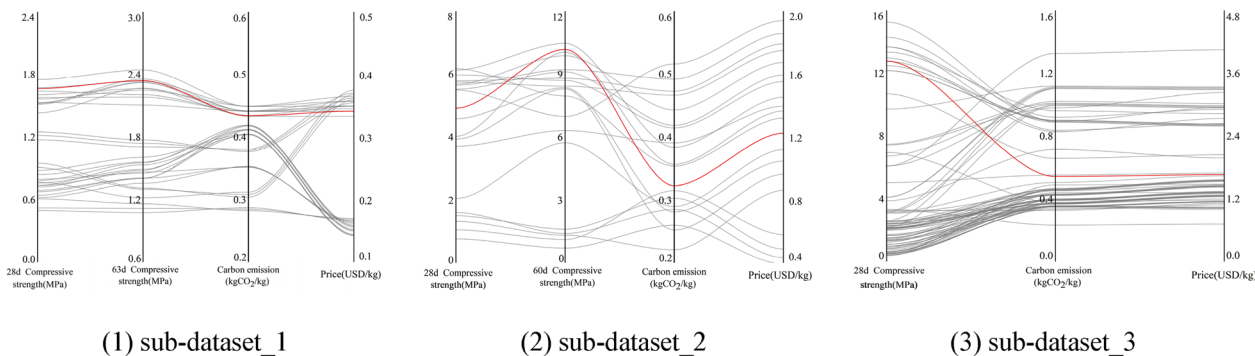
Fig. 16 Optimization results of the different slurry mixture designs

Cui *et al. Heritage Science* (2024) 12:174

Page 19 of 21

increasing the latter two; therefore, these objective functions have mutually constrained relationships, and optimal results cannot be achieved simultaneously [59].

The MOPSO results are represented using parallel coordinate plots, as shown in Fig. 16. The vertical axis represents each objective function, each curve represents a non-dominated solution in the Pareto solution set, and the red line represents the ratio with the highest relative closeness [60]. Table 8 presents the results of the optimal slurry ratio for each sub-dataset.

In sub-dataset_1, the optimal mass ratio of quicklime in the Lime slurry was 0.195, which means that the ratio of quicklime in the solid materials of the slurry was 0.3. This satisfies the requirement for the Lime slurry to produce a slight expansion through a high lime mass ratio, which can improve the mechanical compatibility between the slurry and earthen sites. Compared with the mass ratio with the highest compressive strength, the carbon emission coefficient of the optimal ratio decreased by 2.1% and the cost decreased by 1.6%. Meanwhile, the optimal ratio of the slurry resulted in a 12% increase in fly ash usage. Combining the SHAP analysis in Fig. 13 with the carbon emissions and cost values in Table 7, a high mass ratio of fly ash addition in the slurry can not only improve the mechanical strength of the slurry, but also increase the utilization rate of coal-fired by-products with low carbon. On the other hand, the optimal result of the mass ratio of site soil was decreased by 14%. This is because the collection of site soil has a high sampling cost and long-distance transportation.

In sub-dataset_2, the optimal mass ratios of CGN, fly ash, and quartz sand in the CGN slurry were 0.278, 0.208, and 0.208, respectively. Compared with the mass ratio with the highest compressive strength value, both the carbon emission coefficient and cost of the optimal mass ratio have decreased by 19%. Meanwhile, the optimal quality ratio of slurry the CGN decreased by 5.6%, while both the usage of fly ash and quartz sand increased

by 4.2%. As using CGN is an important factor affecting carbon emissions and costs in this slurry, controlling the usage of CGN is the most direct way to optimize carbon emissions and economic objectives. Based on the results of the SHAP analysis shown in Fig. 14, the optimal ratio meets the requirements for enhancing the mechanical strength of the slurry. In a study by Ren et al., the CGN slurry ratio with the best physical and mechanical properties is the mass ratio with the highest compressive strength value. But the optimal quality ratio in this article also meets the requirements of high mechanical strength, high density and low shrinkage rate for the anchoring slurry used in earthen sites.

In sub-dataset_3, the optimal slurry was PS-C. Compared with the mass ratio with the highest compressive strength value, the carbon emission coefficient of the slurry with the optimal mass ratio decreased by 39.3%, and the cost decreased by 43.6%. Meanwhile, the usage of PS decreased by 18.7%. Although the use of fly ash in the mass ratio with the highest compressive strength can reduce the carbon emissions and costs of the slurry, controlling the usage of PS is the optimal solution for optimizing the carbon emissions and economic objectives of the slurry. Then the usage of PS is an important factor affecting carbon emissions and costs in this type of slurry. Based on the SHAP analysis results in Fig. 15, the values of PS modulus, PS concentration, and SSF content of the optimal slurry meet meets the requirements for enhancing the mechanical strength of the slurry. In a study by Li et al., PS-C slurry not only meets the physical and mechanical properties of the repairing earthen site requirements, but also has good frost resistance, acid and alkaline resistance and water disintegration resistance.

The results above fully demonstrate that the proposed model framework can not only calculate solutions with better environmental and economic benefits, but also reduce the usage of the materials with high carbon emissions and high costs in the preferred solutions. Meanwhile, the optimized ratio results meet the actual engineering needs of repairing earthen site.

Conclusion

In this study, we propose an analytical framework for the intelligent mixed design of anchor and fissure grouting slurries used for repairing earthen sites in Northwest China. Based on this framework, the admixture ratio was optimized for three commonly used slurries, and different design schemes were proposed for the mechanical, environmental, and economic performance of the slurries. The following main conclusions were drawn:

(1) The relationship between the mechanical strength and its influencing factors in slurries is mostly non-

Table 8 Optimal solutions for the slurry mixture

sub-dataset_1		sub-dataset_2		sub-dataset_3	
Quicklime	0.195	CGN	0.278	c (PS)	0.173
FA	0.195	FA	0.208	$n(SiO_2)/n(K_2O)$	3.80
Soil	0.260	Quartz sand	0.208	m (PS)	0.402
SH	0.350	Water	0.306	Soil	0.575
28d CS (MPa)	1.663	28d CS (MPa)	4.917	SSF	0.023
63d CS (MPa)	2.340	63d CS (MPa)	10.185	28d CS (MPa)	12.780
Carbon emission (kgCO$_2$/kg)	0.434	Carbon emission (kgCO$_2$/kg)	0.323	Carbon emission (kgCO$_2$/kg)	0.541
Cost ($/kg)	0.341	Cost ($/kg)	1.224	Cost ($/kg)	1.660

Cui *et al. Heritage Science*   (2024) 12:174

Page 20 of 21

linear. This requires compressive strength prediction through ML algorithms. The GA optimized the hyper-parameters of each algorithm for sub-datasets, which allowed each model to be compared under relatively equal conditions, leading to better prediction accuracy with each optimized model. Among all the ML models, GA-XGBoost had significantly better evaluation metrics in all three sub-datasets, indicating extremely high accuracy and versatility in mechanical strength prediction.

(2) Based on the raw materials, mix composition, binder types, and curing conditions, the input variables affecting the mechanical properties were designed for modeling. A relative importance analysis showed that the calcium-based materials, binders, and curing time significantly impacted the compressive strength. However, SHAP analysis revealed that these impacts were sometimes negative.

(3) The data from the three sub-datasets used a design framework based on the combination of meta-heuristic optimization and TOPSIS decision to obtain multi-objective optimization results. This indicated good performance and met the slurry-optimization requirements in the original studies. By comparing the optimal solutions of the Pareto sets in each sub-dataset, the optimizations of compressive strength, carbon emissions, and cost are interdependent and cannot be achieved simultaneously.

Author contributions
Kai Cui: conceptualization, Writing–review & editing, Supervision, Funding acquisition, Project administration. Huian Shao: methodology, Software, Formal-analysis, Writing-Original Draft, Resources, Data Curation. Xiangpeng Yu: data Curation, Visualization. Qiangqiang Pei**:** investigation, Writing-Review & Editing. Shuai Zhang: software, Validation.

Availability of data and materials
If requested, the authors will provide the data or will cooperate fully in obtaining the data on which the manuscript is based for examination by the editors or their assignees.

Declarations

Ethics approval and consent to participate
Not applicable.

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.

References
1. Li ZX. Protection of ancient sites on the silk road. China Cultural Herit. 2004;03:118–22.
2. Peng EX, Li DD, Hu XY, Sheng Y, Chou YL, Cao W, Chen J, Li QF. Feasibility of ramming erosion area of earthen sites using solidified soil induced by an ancient curing agent, calcined ginger nuts. Constr Build Mater. 2023;370: 130442.
3. Li ZX, Wang XD, Sun ML, Chen WW, Guo QL, Zhang HY. Conservation of Jiaohe ancient earthen site in China. J Rock Mech Geotech Eng. 2011;3:270–81.
4. Wang N, Zhang JK, Chen WW, Liao RX, Li WH. Pulling resistance of two wooden bolts subjected to continuous loading in anchoring conservation of earthen sites. Int J Geomech. 2023;4(23):04023007.
5. Cui K, Yu XP, Huang JJ, Shao HA, Gu X. Early solidification process of the slurry with quicklime for fissure restoration of earthen sites. J Build Eng. 2022;53: 104530.
6. Cui K, Wang XH, Chen WW, Feng F, Cheng FQ. Researches on synergistic effect and mechanism of pile-slurry for reinforce-ment of fissure grouting in earthen sites. Adv Eng Sci. 2021;53:66–74.
7. Cui K, Yu XP, Pei QQ, Wang XH, Xu PF. Influence of grouting mode on the adhesion property of slurry-soil interface for fissure reinforcement of earthen sites. Rock Soil Mech. 2021;42:1501–11.
8. Cui K, Huang JJ, Chen WW, Wang DH, Han N. Research on selection of anchor slurry and performance of mixed quick lime in earthen sites. Rock Soil Mech. 2019;40(06):2183–91.
9. Cui K, Feng F, Chen WW, Wang XH, Cheng FQ. Study on the mechanical compatibility of fissure grouting slurry with quick lime and grouting technology optimization in earthen sites. Rock Soil Mech. 2019;40(12):4627–36.
10. Ren XR, Zhang JK, Wang N, Shan TT, Zhao LY. Age performance of grout mixed with calcined ginger nuts and earthen fabric in anchoring conservation of earthen sites. J Mater Sci Eng. 2017;35:62–6.
11. Zhang JK, Zhang LX, Zhao LY, Liu D, Guo QL, Pei QQ. Property changes of anchor grout calcined ginger nuts admixed with fly ash and quartz sand after accelerated ageing tests. J Cent South Univ. 2019;26:3114–25.
12. Ren XR. Study on the Age Performance of Modified Grout by Calcined Ginger Nuts in the Anchor Conservation of Earthen Sites. MA thesis. Lanzhou University. 2017.
13. Wang N, Zhang JK, Wang YC, Zhang H, Ma YJ, Zhao LY, Guo QL. Experimental study on mechanical properties of grout–soil interface in anchor system of rammed earthen sites. Int J Geomech. 2020;20:04020064.
14. Mao XF, Zhang YY, Hu FL, Li X, Lu JC. Experimental research on the mechanical properties of glutinous rice slurry modified calcined ginger nuts grouting material. Chin J Appl Mech. 2022;39(05):965–73.
15. Yang T, Li ZX, Chen WW. Physical and mechanical properties of PS-F grouting material. Dunhuang Res. 2005;04:40–50.
16. Li ZX, Yang T, Wang WF. Strength properties of PS-F grouting materials. Dunhuang Res. 2009;06:26–9.
17. Zhang JK, Chen WW, Li ZX, Guo QL, Sun ML. Laboratory tests on property and durability of PS—(C+ F) slurry used in anchoring conservation of earthen sites. Chin J Geotech Eng. 2015;37:1802–9.
18. Li ZX, Wang XD, Lin T. Experimentation of chemical consolidation on ancient earth structure sites of Jiaohe. Dunhuang Res. 1997;03:175–80.
19. Li ZX, Wang XD. New progress in research on the protection and reinforcement of ancient earthen architecture sites. Dunhuang Res. 1997;04:169–74.
20. Li ZX. Stability and strength issues of PS material reinforced soil and stone cultural relics. Dunhuang Res. 1996;03:96–111.

Cui *et al. Heritage Science*   (2024) 12:174

Page 21 of 21

21. Li ZX. Conservation of ancient sites on the silk road. Beijing: Science Press; 2003.
22. Fei Y, Han F, Cui ZJ. Reducing carbon emissions through industrial symbiosis: a case study of a large enterprise group in China. J Clean Prod. 2015;103:811–8.
23. Chen WC, Li L, Li ZX, Zhao LY, Shao MS, Afolagboye OL. Modification of traditional Chinese ginger nut and its mechanical behavior. Constr Build Mater. 2017;144:138–46.
24. Ahmed UH, Mohammed SA, Qaidi AMS, Faraj HR, Sor HN, Mohammed AA. Compressive strength of geopolymer concrete composites: a systematic comprehensive review, analysis and modeling. Eur J Environ Civ Eng. 2023;27(3):1383–428.
25. Emad W, Mohammed SA, Bras A, Asteris GP, Kurda R, Muhammed Z, Hassan AMT, Qaidi AMS, Sihag P. Metamodel techniques to estimate the compressive strength of UHPFRC using various mix proportions and a high range of curing temperatures. Constr Build Mater. 2022;349: 128737.
26. Ahmed UH, Mahmood JL, Muhammad AM, Faraj HR, Qaidi AMS, Sor HN, Mohammed SA, Mohammed AA. Geopolymer concrete as a cleaner construction material: an overview on materials and structural performances. Clean Mater. 2022;5: 100111.
27. Çalışkan A, Demirhan S, Tekin R. Comparison of different machine learning methods for estimating compressive strength of mortars. Constr Build Mater. 2022;335: 127490.
28. Zhang M, Zhang C, Zhang JF, Wang L, Wang F. Effect of composition and curing on alkali activated fly ash-slag binders: Machine learning prediction with a random forest-genetic algorithm hybrid model. Constr Build Mater. 2023;366: 129940.
29. Kang MC, Yoo DY, Gupta R. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Constr Build Mater. 2021;266: 121117.
30. Liang C, Qian C, Chen H, Kang W. Prediction of compressive strength of concrete in wet-dry environment by BP artificial neural networks. Adv Mater Sci Eng. 2018;2018:6204942.
31. Ma HX, Liu J, Zhang J, Huang J. Estimating the compressive strength of cement-based materials with mining waste using support vector machine, decision tree, and random forest models. Adv Civil Eng. 2021;2021:1–10.
32. Zhang JF, Ma GW, Huang YM, Sun JB, Aslani F, Nener B. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. Constr Build Mater. 2019;210:713–9.
33. Zhang L, Marani A, Nehdi ML. Chemistry-informed machine learning prediction of compressive strength for alkali-activated materials. Constr Build Mater. 2022;316: 126103.
34. Huo WW, Zhu ZD, Sun H, Ma BR, Yang L. Development of machine learning models for the prediction of the compressive strength of calcium-based geopolymers. J Clean Prod. 2022;380: 135159.
35. Liu KH, Zheng JK, Dong SX, Xie WL, Zhang XC. Mixture optimization of mechanical, economical, and environmental objectives for sustainable recycled aggregate concrete based on machine learning and metaheuristic algorithms. J Build Eng. 2023;63: 105570.
36. Yu XP, Kai Cui K, Hu X, Qiao Q, Zhang GJ. Properties and compatibility assessment of the slurry with quicklime for fissure restoration of earthen sites in Northwest China. Coatings. 2022;12(9):1255.
37. Liu J. The Study on the Performance of Fissure Grouting Material and Grouting Process to the Soil Ruins in the North-west Region. MA thesis. Lanzhou University. 2017.
38. Yang FW, Zhang BJ, Ma QL. Study of sticky rice–lime mortar technology for the restoration of historical masonry construction. Acc Chem Res. 2010;43(6):936–44.
39. Gołaszewski J. Influence of viscosity enhancing agent on rheology and compressive strength of super-plasticized mortars. J Civ Eng Manag. 2009;15(2):181–8.
40. Chen WW, Zhang KW, Zhang JK, Wang DY, Zhou J. Laboratory study on properties of calcined ginger nuts mixed with earthen sites'soil for grouting materials. J Sci Technol. 2018;49(06):1519–25.
41. Li ZX, Zhao LY, Sun ML. Deterioration of earthen sites and consolidation with PS material along silk road of China. Chin J Rock Mech Eng. 2009;28(05):1047–53.
42. Munir JM, Kazmi SMS, Wu YF, Lin XS, Ahmad RM. Development of a novel compressive strength design equation for natural and recycled aggregate concrete through advanced computational modeling. J Build Eng. 2022;55: 104690.
43. Zhang JF, Huang YM, Ma GW, Nener B. Mixture optimization for environmental, economical and mechanical objectives in silica fume concrete: a novel frame-work based on machine learning and a new meta-heuristic algorithm. Resour Conserv Recycl. 2021;167: 105395.
44. Ekanayake IU, Meddage DPP, Rathnayake U. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using shapley additive explanations (SHAP). Case Stud Constr Mater. 2022;16: e01059.
45. Khan K, Ahmad W, Amin NM, Aslam F, Ahmad A, Al-Faiad AM. Comparison of prediction models based on machine learning for the compressive strength estimation of recycled aggregate concrete. Materials. 2022;15(10):3430.
46. Deng FM, He YG, Zhou SX, Yu Y, Cheng HG, Wu X. Compressive strength prediction of recycled concrete based on deep learning. Constr Build Mater. 2018;175:562–9.
47. Chen N, Zhao SB, Gao ZW, Wang DW, Liu PF, Oeser M, Hou Y, Wang LB. Virtual mix design: prediction of compressive strength of concrete with industrial wastes using deep data augmentation. Constr Build Mater. 2022;323: 126580.
48. Wu YQ, Zhou YS. Hybrid machine learning model and shapley additive explanations for compressive strength of sustainable concrete. Constr Build Mater. 2022;330: 127298.
49. Cui K, Wang DH, Yu XP, Wu GP. Performance and mechanism of bolt and slurry bonds in anchor systems for earthen sites. Proceedings of the Institution of Civil Engineers-Geotechnical Engineering. 2020. 173(6): 485-499.
50. Cui K, Ge CR, Chen R, Yu XP, Wu GP. Temporal and spatial characteristics of rammed earth in arid area of Northwest China. Res Conserv Cave Temples Earthen Sites. 2022;1(02):51–62.
51. Du YM, Chen WW, Cui K, Zhang KW. Study on damage assessment of earthen sites of the Ming Great Wall in Qinghai Province based on Fuzzy-AHP and AHP-TOPSIS. Int J Archit Herit. 2019;14(6):1–14.
52. Neve S, Du J, Barhemat R, Meng WN, Bao Y, Sarkar D. Valorization of vetiver root biochar in eco-friendly reinforced concrete: mechanical, economic, and environmental performance. Materials. 2023;16(6):2522.
53. Ministry of Ecology and Environment the People's Republic of China 2022. Guidelines for Accounting and Reporting of Greenhouse Gas Emissions by Enterprises—Power Generation Facilities.000014672/2022-00506, Beijing.
54. Ministry of housing and urban-rural development of the People's Republic of China. Standard for building carbon emission calculation, GB/T 51366–2019. Beijing: China Architecture & Building Press; 2019.
55. Cai BF, Chen B, Chen SQ, Bai L, Bai YF, Cao LB. China Products Carbon Footprint Factors Database. China City Greenhouse Gas Working Group, China. 2022
56. Nagarkar R, Patel J. Polyvinyl alcohol: a comprehensive study. Acta Sci Pharm. 2019;3(4):34–44.
57. Zhang JK, Chen WW, Li ZX, Wang XD, Guo QL, Wang N. Study on workability and durability of calcined ginger nuts-based grouts used in anchoring conservation of earthen sites. J Cult Herit. 2015;16(6):831–7.
58. Ndiaye S, Toure OA, Sambe MF, Diop MGC, Prat L, Cassayre L. Exploration of a two-step aqueous process for the valorization of sodium fluorosilicate (Na_2SiF_6), an intermediate product of the fluorosilicic acid conversion. Waste Biomass Valorization. 2022;13:547–62.
59. Mahjoubi S, Barhemat R, Guo PW, Meng WN, Bao Y. Prediction and multi-objective optimization of mechanical, economical, and environmental properties for strain-hardening cementitious composites (SHCC) based on automated machine learning and metaheuristic algorithms. J Clean Prod. 2021;329: 129665.
60. Tian ZN, Zhang ZQ, Zhang KW, Tang XM, Huang SL. Statistical modeling and multi-objective optimization of road geopolymer grouting material via RSM and MOPSO. Constr Build Mater. 2021;271: 121534.

Publisher's Note