Front. Struct. Civ. Eng. 2024, 18(1): 30–50
https://doi.org/10.1007/s11709-024-1041-y

RESEARCH ARTICLE

A comprehensive comparison of different regression techniques and nature-inspired optimization algorithms to predict carbonation depth of recycled aggregate concrete

Bin XI[a], Ning ZHANG[b], Enming LI[c*], Jiabin LI[d], Jian ZHOU[e], Pablo SEGARRA[c]

[a] *Department of Civil and Environmental Engineering, Politecnico Di Milano, Milano 20133, Italy*
[b] *Leibniz Institute of Ecological Urban and Regional Development (IOER), Dresden 01217, Germany*
[c] *Universidad Politécnica de Madrid–ETSI Minas y Energía, Madrid 28003, Spain*
[d] *Department of Civil Engineering, KU Leuven Campus, Brugge 8200, Belgium*
[e] *School of Resources and Safety Engineering, Central South University, Changsha 410083, China*
[*] *Corresponding author. E-mail: enming.li@alumnos.upm.es*

ABSTRACT The utilization of recycled aggregates (RA) for concrete production has the potential to offer substantial environmental and economic advantages. However, RA concrete is plagued with considerable durability concerns, particularly carbonation. To advance the application of RA concrete, the establishment of a reliable model for predicting the carbonation is needed. On the one hand, concrete carbonation is a long and slow process and thus consumes a lot of time and energy to monitor. On the other hand, carbonation is influenced by many factors and is hard to predict. Regarding this, this paper proposes the use of machine learning techniques to establish accurate prediction models for the carbonation depth (*CD*) of RA concrete. Three types of regression techniques and meta-heuristic algorithms were employed to provide more alternative predictive tools. It was found that the best prediction performance was obtained from extreme gradient boosting-multi-universe optimizer (XGB-MVO) with R^2 value of 0.9949 and 0.9398 for training and testing sets, respectively. XGB-MVO was used for evaluating physical laws of carbonation and it was found that the developed XGB-MVO model could provide reasonable predictions when new data were investigated. It also showed better generalization capabilities when compared with different models in the literature. Overall, this paper emphasizes the need for sustainable solutions in the construction industry to reduce its environmental impact and contribute to sustainable and low-carbon economies.

KEYWORDS recycled aggregate concrete, carbonation depth, nature-inspired optimization algorithms, extreme gradient boosting technique, parametric analysis

1 Introduction

Currently, countries around the world are actively working toward developing sustainable and low-carbon economies, recognizing the need to address climate change and reduce their environmental impact [1–5]. As a major contributor to energy consumption and carbon emissions, the construction industry has placed a significant emphasis on improving its energy efficiency and economic and environmental performance [6,7].

One of the key areas of focus has been the use of concrete, which is the most widely used material in the construction industry. However, the production of concrete requires large amounts of natural sand and gravel, leading to significant depletion of natural resources. It is estimated that over 400 million tons of natural aggregates (NA) are consumed each year, and this number is projected to increase to an astonishing 800

Article history: Received May 2, 2023; Accepted Jun 24, 2023

million tons in the next 20 years [8]. The excessive exploitation of NA can have a severe and lasting impact on the ecological environment. Another significant challenge facing the construction industry is the staggering amount of waste generated by urban renewal and expansion. According to 2014 statistics, the United States, the European Union, and China, which are three of the world's largest economies, generate more than 530 million, 850 million, and 1.13 billion tonnes of demolition waste respectively [9,10]. This not only puts a strain on the environment but also on the economies.

One solution to this problem is recycling waste concrete from construction and demolition waste to produce recycled aggregate (RA). The RA can then be used as a replacement for NAs in the production of RA concrete.

Studies have shown that the incorporation of RA can lead to a diminution of the mechanical properties of concrete [11,12]. In addition, researchers have found that the addition of RA can have an impact on the durability of concrete, particularly with respect to carbonation [13–16]. Concrete carbonation refers to the chemical reaction between carbon dioxide in the air and the calcium compounds within the concrete. This reaction decreases the pH value of the concrete, which can increase the corrosion of reinforcement and ultimately lead to the deterioration of reinforced concrete structures [5]. In addition, carbonation products (mainly $CaCO_3$) may cause microcracks within the concrete [13]. Xiao et al. [13] concluded that RA content (*RAC*) has a direct impact on the carbonation depth (*CD*) of RA concrete. They found that when the threshold value of the substitution rate of NA by RA is lower than 70% the *CD* of RA concrete increases with the increase of *RAC*. When the substitution rate exceeds 70%, further substitution leads to a reduction in the *CD*. Silva et al. [15] found that adding more *RAC* causes an increase in the *CD*. They also found that using all coarse grains of RA may lead to *CD*s that are up to twice those occurring in normal concrete. Additionally, they observed that the age and curing conditions of the concrete had little effect on the *CD*. Balayssac et al. [16] and Atiş [17] observed that the longer the curing time, the lower the degree of carbonation. Leemann and Loser [18] conducted a study and discovered that the high water absorption properties of RA had a significant impact on the compressive strength of RA concrete. However, a comparison between concrete made from dry aggregates and pre-saturated aggregates revealed minimal disparities in their carbonation resistance. Additionally, when the proportion of RA was increased from 25% to 50% by mass, there was no discernible enhancement in the concrete's carbonation factor [18]. A similar finding, that the appropriate addition of RA had little effect on the *CD*, can be found in Ref. [19]. The depth of carbonation of RA concrete was also related to the type of superplasticizer (*SP*) [20]. Zega and di Maio [21] examined the effect of exposing RA concrete samples containing 20% and 30% fine RA to air over periods of 310 and 620 d. Despite the extended exposures, there was little difference in the *CD*s.

From the literature review, it can be seen that the carbonation resistance of RA concrete is affected by various factors. And it can be indicated that the carbonation behaviors of RA concrete vary significantly among previous studies. Furthermore, most of the studies on RA concrete carbonation property were conducted through experimental trial-and-error methods, which are highly variable for different laboratories and different environments. The research is also time-consuming and expensive. Therefore, it is vital to establish accurate prediction models. Several mathematical theoretical models have been proposed to predict the carbonation resistance of RA concrete (see below). However, these models have limited ability to specifically discuss the effects of different factors on the carbonation results.

It has become increasingly common to use machine learning algorithms to predict various aspects of concrete performance, on the grounds that these techniques possess superior predictive capabilities compared to traditional mathematical models [22–26]. Nunez and Nehdi [22] proposed a gradient boosting regression tree (GBRT) model to predict the *CD* of RA concrete containing different mineral additives. Through a comprehensive analysis of the results, they found that the GBRT model achieved a high level of accuracy, as demonstrated by its low Root Mean Squared Error (*RMSE*) of 1.5139, Mean Absolute Error (*MAE*) of 0.948, and high coefficient of determination (R^2) of 0.9707. These results indicate that the GBRT model performed significantly better than traditional mathematical equations for predicting *CD*s. Furthermore, Liu et al. developed an Artificial Neural Network (ANN) model using 593 RA concrete *CD* data sets [23]. The ANN model was shown to have high accuracy and robustness in its predictions. However, it should be noted that ANN models are known to be "black boxes", meaning they lack interpretability in terms of understanding the relationships between the input and output variables. Despite the promising results of previous studies, the use of machine learning to predict *CD* of RA concrete is still relatively limited.

In this study, three representative regression techniques were employed, i.e., support vector regression (SVR), kernel extreme learning machine (KELM), and extreme gradient boosting (XGB). The role of SVR was mainly to find the best separation hyperplane in the feature space. KELM is mainly based on the theory of neural network and XGB is mainly inspired by the regression tree. Meanwhile, three types of nature-inspired meta-heuristic algorithms were adopted to optimize the hyper-

parameters in these regression techniques, i.e., genetic algorithm (GA), multi-universe optimizer (MVO) and sparrow search algorithm (SSA). GA is inspired by evolutionary theory, MVO is from physics and SSA simulates swarming behaviors. 682 *CD* data sets from published literature were collected and used to develop machine learning models for carbonation prediction for RA concrete. It should be noted that new input parameter combinations, i.e., RA water absorption, water-to-binder ratio (*WBR*), fine aggregate content (*FAC*), gravel content (*GC*), *RAC*, *SP*, carbon concentration (*CC*), and exposure time (*T*) were considered as potential influencing factors in the model. A total of nine prediction models are provided in this work, so as to provide a comprehensive comparison of different regression techniques and optimization algorithms as well as a better understanding of the relationships between different variables and the *CD*. The established prediction models were evaluated by eight mathematical indicators and Taylor Diagram. Finally, the most competitive prediction model was validated by parametric analysis from the perspective of physics.

2 Data description

As outlined in the introduction, a plethora of factors can impact the carbonation properties of recycled concrete. This study specifically examined eight key parameters that are commonly considered by researchers in this field. The authors compiled and extracted a database from by Moghaddas et al. [27]. This data set consists of 682 experimental samples gathered from 21 different papers and 11 different countries, as shown in Fig. 1. The study of carbonation performance of RA concrete has been conducted by a diverse range of countries worldwide, with a significant proportion of data coming from Portugal, China, and the UK. The data set includes variables such as RA water absorption (*RAWA*), *WBR*, *FAC*, *GC*, *RAC*, *SP*, *CC*, *T*, and *CD*. Table 1 and Fig. 2 present a statistical analysis and distribution of all variables. In the violin plot, the width of "violin" represents the density of data. It can be seen that the *T* ranges from 177 to 3650 d which indicates a broad database. Notably, *T* in the data reaches up to ten years, making it possible to consider the effect of long-term exposure on RA concrete carbonization, an aspect that has been previously understudied in models [28]. Potential relationships between different variables were also examined. Pearson correlation coefficient (*R*) was used to quantify the correlation between pairs of variables, with values ranging from −1 to 1 [29]. A value of 0 indicates no correlation, positive values indicate a positive correlation, and negative values indicate a negative correlation. The greater the absolute number, the stronger the correlation. As seen in Fig. 3, *WBR* and *RAC* show a higher positive correlation with *CD*, suggesting that these factors should be given particular attention. *T*, *CC*, *FAC*, and *RAWA* present a slight correlation with *CD*. *GC* and *SP* show a negative correlation. However, according to the review report from Silva et al. [15], *T*, *CC*, *FAC*, *RAWA*, and *SP* showed a significant effect on *CD* of *RA* concrete. It is important to note that *R* is

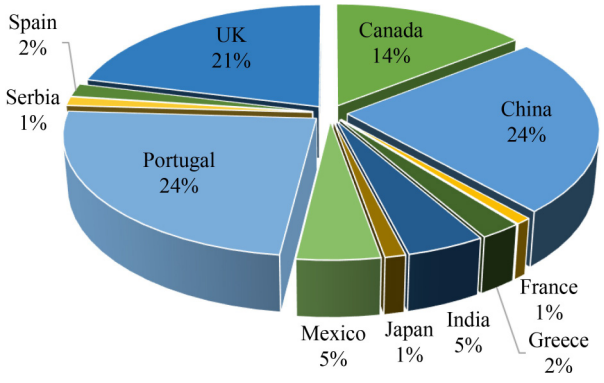


Fig. 1 Data set distribution from different countries.

Table 1 Data distribution and statistics of the whole data set

parameter	mean*	max*	min*	std*	25%*	50%*	75%*
RAWA (%)	5.8	16.6	0.2	2.4	4.7	5.3	6.3
WBR (kg/m^3)	0.5	1.0	0.3	0.1	0.4	0.5	0.5
FAC (kg/m^3)	645.6	998.0	357.7	171.0	550.0	625.0	787.0
GC (kg/m^3)	448.5	1311.0	0.0	436.0	0.0	454.9	846.5
RAC (kg/m^3)	586.7	1280.0	0.0	407.0	198.2	635.0	953.0
SP (kg/m^3)	0.9	7.3	0.0	1.8	0.0	0.0	0.7
CC (%)	5.3	50.0	0.1	6.5	3.0	3.5	5.0
T (d)	177	3650	7	539	28	56	91
CD (mm)	10.19	50.05	0.10	7.90	4.80	8.34	13.30

*Note: "mean", "max", "min", "std", "25%", "50%", and "75%" represent the average, maximum, minimum, standard deviation, percentile of 25%, percentile of 50% and percentile of 75% values of each variable, respectively.

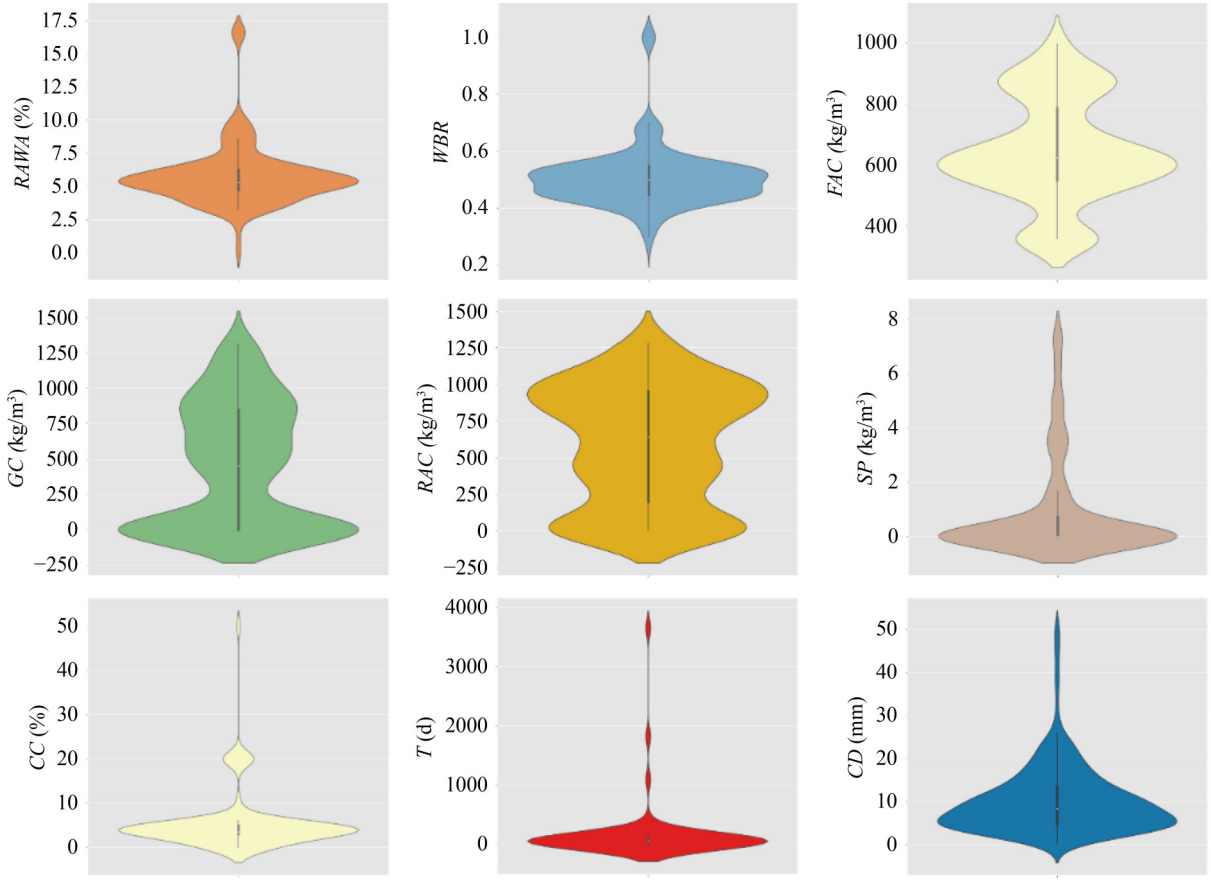


Fig. 2 Data distribution of influencing factors and *CD* by violin plot.

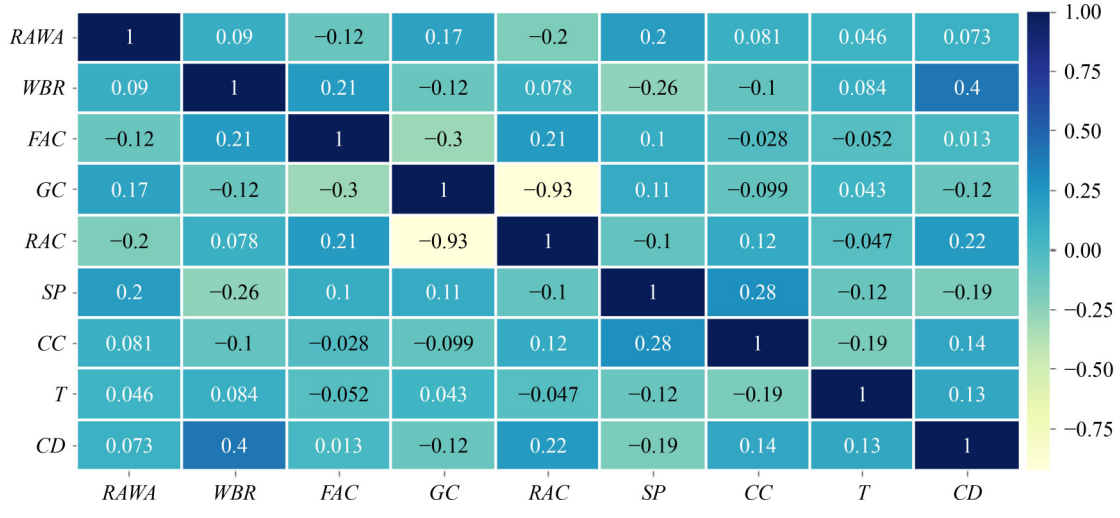


Fig. 3 Pearson correlation analysis.

restricted to measuring only linear relationships between two variables, is sensitive to outliers, assumes normality, does not account for causality, and is limited to bivariate analysis. Considering these limitations and the intricate relationship between these factors and *CD* of RA concrete, alternative correlation approaches were also employed, for comparison.

Mutual information (MI) correlation is a measure of the statistical dependence between two variables [30]. It quantifies the amount of information that is shared by two variables. From the analysis of MI shown in Fig. 4, it can be found that *WBR* has the lowest interactivity with *CD* and *FAC* has the highest MI with *CD*. Other factors have some MI with *CD*. These findings are different from the analysis of Pearson correlation between potential factors and *CD*. Therefore, the implementation of machine

learning techniques seems to be a desirable approach to integrating these factors and *CD*.

3 Methodology

In this study, three kinds of machine learning techniques were employed, i.e., SVR, KELM and XGB. The main role of SVR is to optimize the hyper-plane with the largest interval defined in the feature space. The KELM is modified based on classical extreme learning machine (ELM), while the main advantage of KELM over ELM is better generalization performance [31]. KELM achieves this by introducing a regularization term in the objective function (*OF*), which helps to prevent overfitting to the training data. XGB is a powerful tree-based ensemble machine learning algorithm that has gained popularity in recent years [32,33]. Compared to traditional tree-based models, XGB has higher generalization and prediction accuracy. Therefore, aforementioned representative machine learning techniques were used to implement regression tasks in this study. Meanwhile, three types of optimization algorithms were utilized to optimize the hyper-parameters in SVR, KELM and XGB. In some previous studies, only one kind of regression technique was optimized by optimization algorithms [32,34–37]. However, no obvious predictive performance difference could be observed between different optimization algorithms. In addition, most studies only focused on the same type optimization algorithm which hardly presented the superiority and inferiority of different regression techniques [38,39]. Regarding this, this study proposed to use three kind of different mechanism regression techniques combined with three different mechanism optimization algorithms to predict *CD* of RA concrete and their prediction abilities were evaluated by various evaluation indicators so as to conclude a competitive *CD* prediction model. Finally, the superior *CD* prediction model was used for reflecting the impact of influencing factors on *CD*.

3.1 Optimization algorithms

3.1.1 Genetic algorithm

The GA mimics the natural selection process and the genetic mechanism of biological evolution [40,41]. The algorithm begins by randomly generating an initial population of chromosomes, each of which comprises multiple genes that represent specific characteristics of data. These chromosomes are considered potential solutions to a given problem. The population is then updated through a series of iterations, utilizing three primary operators-selection, crossover, and mutation-to preserve the optimal chromosomes for the next generation. The crossover operator involves the exchange of segments of genetic code between two chosen individuals, leading to the creation of new individuals with a combination of their genes. The mutation operator, on the other hand, randomly alters the value of specific genes in an individual, maintaining genetic diversity within the population. The GA continually evaluates the generated solutions, comparing them to a fitness function, until a predetermined number of iterations or a predefined threshold is reached, resulting in the identification of the optimal solution. Figure 5 presents the flowchart of the GA.

3.1.2 Multi-universe optimizer algorithm

Mirjalili et al. [42] proposed the MVO algorithm, which is inspired by the interaction between universes through black holes, white holes, and wormholes. The algorithm is based on the multiverse theory, which posits that objects from parallel universes can be moved from one universe to another through white and black holes, and objects in each universe can travel through wormholes within it. In the MVO algorithm, each set of variables is considered a universe, and each variable within the universe is an object. The corresponding inflation rates, or fitness function values, are then calculated. The

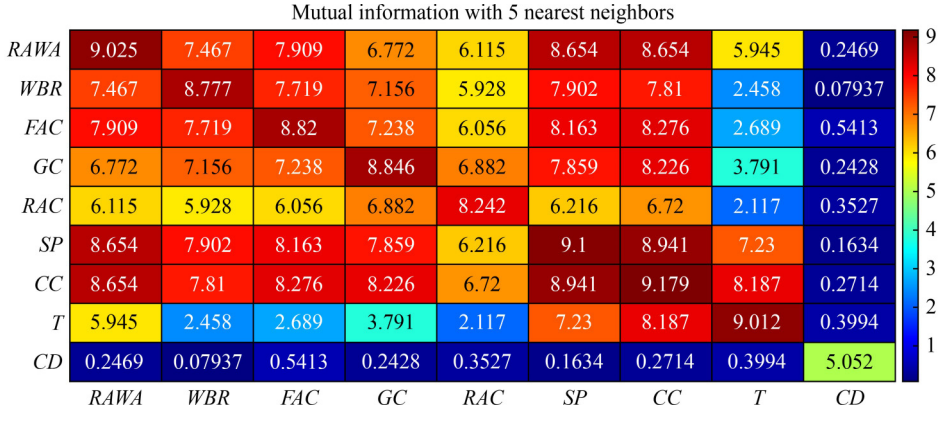


Fig. 4 MI with 5 nearest neighbors method.

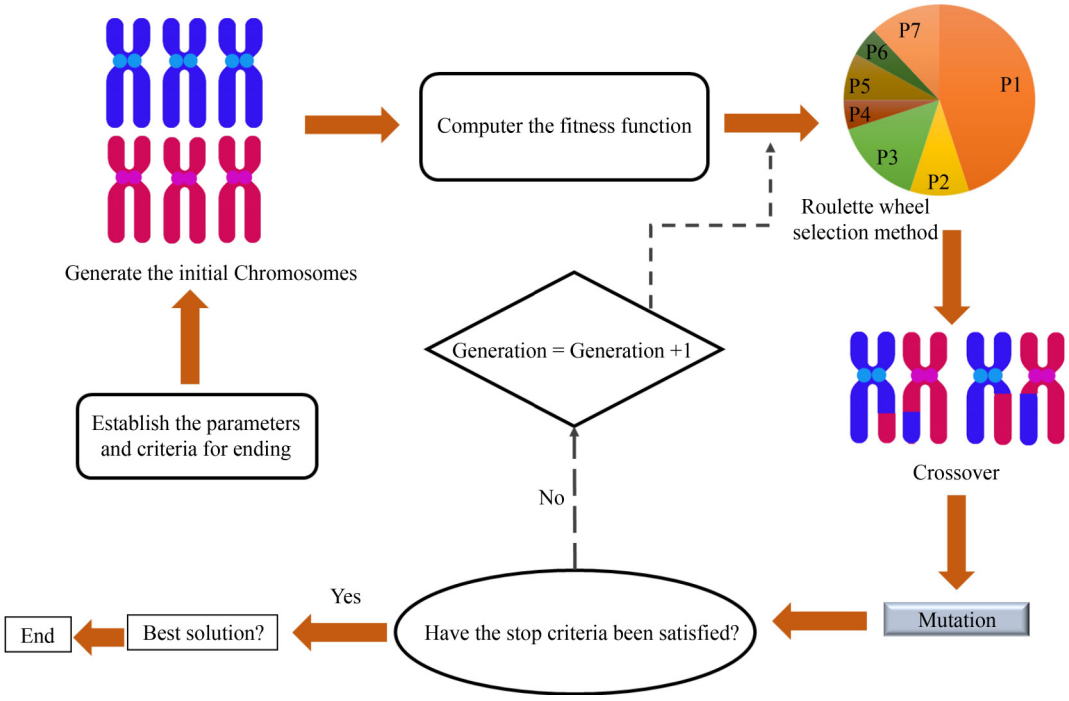


Fig. 5 Flowchart of the GA used for optimization.

universe's inflation rate is a defining characteristic, and its higher values reflect a greater capability for expansion. A white hole with a high inflation rate tends to send its objects to a black hole with a low inflation rate. This sudden change in universes raises the average population expansion rate and ensures the MVO's exploration capacity. By utilizing this approach, poor solutions and their mean objective values can be upgraded.

3.1.3 Sparrow search algorithm

In 2020, Xue and Shen [43] introduced the SSA as a novel optimization method that employs swarm intelligence. The algorithm is inspired by the feeding behavior of sparrows and their ability to avoid natural predators. The sparrow population is first divided into producers and scroungers during the foraging phase. The producer is responsible for locating food and directing the entire sparrow population; the scrounger uses an efficient optimization approach for designing machine learning models based on the knowledge provided by the producer to acquire food. Additionally, SSA incorporates a producer-scrounger model to detect early warning systems that represent sparrow anti-predation behavior.

3.2 ML methodology

3.2.1 Support vector regression

Cortes and Vapnik [44] advanced the field of machine learning by introducing SVMs for dichotomous techniques. The SVM model was predicated on the concept of a hyperplane that separates the data set into distinct classes. A variation of this model, SVR, was later developed specifically for regression tasks. It aims to identify an optimal hyperplane that minimizes the overall deviation of all data points from the hyperplane, and endeavors to fit the data as closely as possible while still preserving a desirable margin, as opposed to a hyperplane that simply splits two or more categories of data points.

3.2.2 Kernel extreme learning machine

KELM is a variation of the traditional ELM algorithm, with the key difference being that it utilizes kernel functions to transform the input data into a higher dimensional feature space, allowing for nonlinear decision boundaries [31]. The approach involves randomly initializing the weights of the hidden layer and utilizing the kernel trick to map the input data into a higher dimensional feature space, followed by determination of the weights of the output layer through the least squares method. KELM is computationally efficient as the training process only requires determination of the weights of the output layers, unlike traditional ELM. Additionally, it can handle nonlinear decision similar to SVMs through the use of the kernel trick. The output function of ELM is described in Eq. (1):

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i(x) = h(x)\beta, \tag{1}$$

where $\beta = [\beta_1, \beta_2, ..., \beta_L]$ represents the output weights

vector between the hidden layer of L nodes and the output node. $h(x) = [h_1(x), \ldots, h_L(x)]$ defines the hidden layer output matrix. $h(x)$ is a mapping feature because it maps data from the d-dimensional input space to the L-dimensional hidden layer feature space H.

3.2.3 Extreme gradient boosting

The XGB model is an implementation of gradient boosting to reduce the training loss function [33]. The XGB consists of a set of decision trees. The model expression can be represented as follows:

$$\tilde{y}_i = \sum_{k=1}^{K} f_k(x_i), \tag{2}$$

where $\tilde{y}_i$ represents the predicted value for input x_i, $f_k(x_i)$ is the kth tree.

Finally, the expression for OF is as follows:

$$OF = \sum_{i-1}^{n} L(\tilde{y}_i, y_i) + \sum_{k=1}^{K} \Omega(f_k), \tag{3}$$

where OF includes the regularization item Ω and the loss function L which denotes the gap between the predicted result $\tilde{y}_i$ and the actual result y_i.

The basic structure of XGB is shown in Fig. 6 and Ref. [33] provides a more comprehensive description and characterization of XGB.

4 Development of models

Figure 7 depicts the general framework of CD of RA concrete prediction model development. The precision of machine learning algorithms heavily depends on their intrinsic hyper-parameters.

Determining the appropriate combination of hyper-parameters for a given situation can be challenging, so optimization algorithms play a vital role in finding the best hyperparameters. Regarding this, GA, MVO, and SSA were used to identify the optimal hyper-parameters of KELM, SVM, and XGB models to construct accurate prediction models and determine the CD of RA concrete.

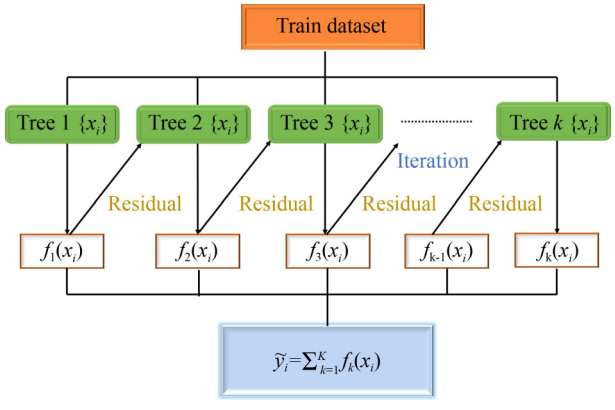


Fig. 6 The basic structure of the XGB algorithm.

To save computation time, the inputs (8 factors) and output (CD) are normalized to [0,1] by the equation below:

$$z = \frac{x - \mu}{\sigma}, \tag{4}$$

where z is transformed value; μ is the mean; σ is the standard deviation; x is the output value.

To create and evaluate models effectively, 80% of the data (550 data sets) was randomly selected as a training set and 20% (132 data sets) as a testing set. Figure 8 illustrates the distribution of data between the training and test sets. The results evince their comparability, attesting to the homogeneity of the training and test sets. 10-fold cross validation was then adopted in this study to increase the robustness of proposed prediction models where the training set was divided into ten equal sub-sets, with nine sub-sets being used as sub-training sets and the other being used as a validation set. Some initial model parameters were given to these sub-training sets and then a sub-model could be established and the prediction performance of a sub-model could be tested by a validation set. This process was repeated ten times and the average MSE of the ten validation sets was used as fitness value. By optimization from proposed meta-heuristic algorithms, the fitness value was updated until the iteration finished. Table 2 provides a detailed explanation of the hyper-parameters to be optimized and their search ranges for each model.

5 Evaluation indicators

To measure the overall performance of developed CD prediction models, a total of eight evaluation indicators were used in this study. These indicators include seven classical mathematical indicators, i.e., R^2, Variance Accounted For (VAF), MAE, Mean bias error (MBE), $RMSE$, Mean Absolute Percentage Error ($MAPE$) and R as well as a pre-defined indicator, $A10$. They can be calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y - y')^2}{\sum_{i=1}^{N} (y - \tilde{y})^2}, \tag{5}$$

$$VAF = \left[1 - \frac{var(y - y')}{var(y)}\right] \times 100\%, \tag{6}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y' - y|, \tag{7}$$

$$MBE = \frac{1}{N} \sum_{i=1}^{N} (y' - y), \tag{8}$$

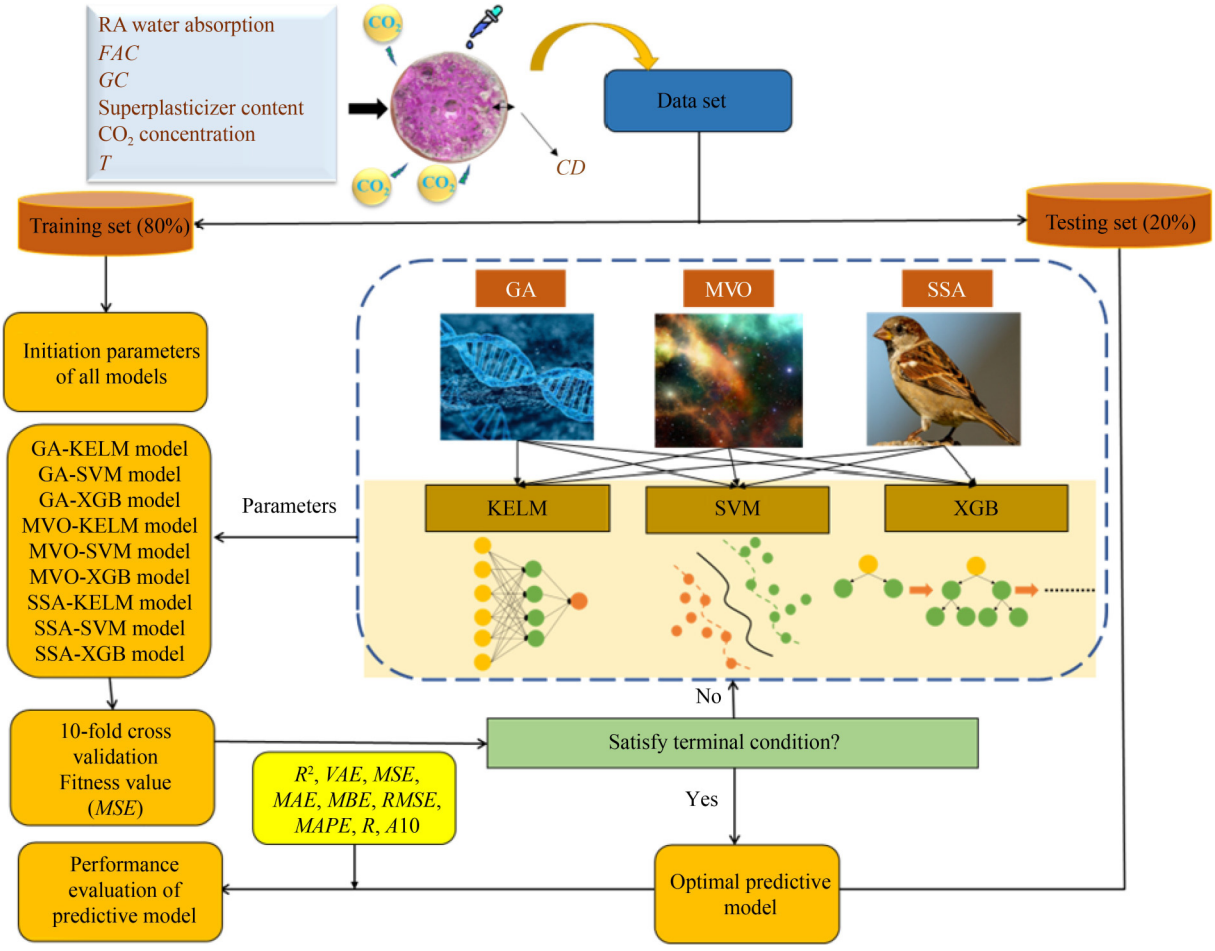


Fig. 7 The general framework of *CD* prediction model development and evaluation.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y-y')^2}, \tag{9}$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y'-y}{y}\right|, \tag{10}$$

$$R = \frac{N\sum yy' - \sum y\sum y'}{\sqrt{N\sum y^2 - \left(\sum y\right)^2}\sqrt{N\sum y'^2 - \left(\sum y'\right)^2}}, \tag{11}$$

$$A10 = size\left(0.8 < \left(\frac{y}{y'}\right) < 1.2\right), \tag{12}$$

where y represents the measured *CD*, y' and y represent the predicted and average value of measured *CD*, respectively. *NN* denotes the total number of samples, while i is the number in the present sample. The proposed comprehensive ranking system by Zorlu et al. [45] is used to give the overall performance of the model by considering all indicators. In this system, higher performing indicators receive higher ranking scores and the model's overall score is calculated by summing up each score. The model with the highest overall score is considered the most robust prediction model of *CD* of RA concrete.

6 Result and discussion

6.1 Development of hybrid carbonation depth of recycled aggregate concrete prediction models

For meta-heuristic algorithms, there are two parameters that influence the optimization performance, i.e., swarm size and iteration number where a swarm size and iteration that are too large increase the calculation cost while those that are too small induce under-fitting. After testing and comparison, the swarm size was equal to 30 and the iteration number was set to be 60 which was sufficient to complete the optimization process and these two significant values would be constant in each *CD* prediction model. The optimized parameters optimization algorithms are listed in Table 3. In the initialization stage, 30 combinations of hyper-parameters were generated and

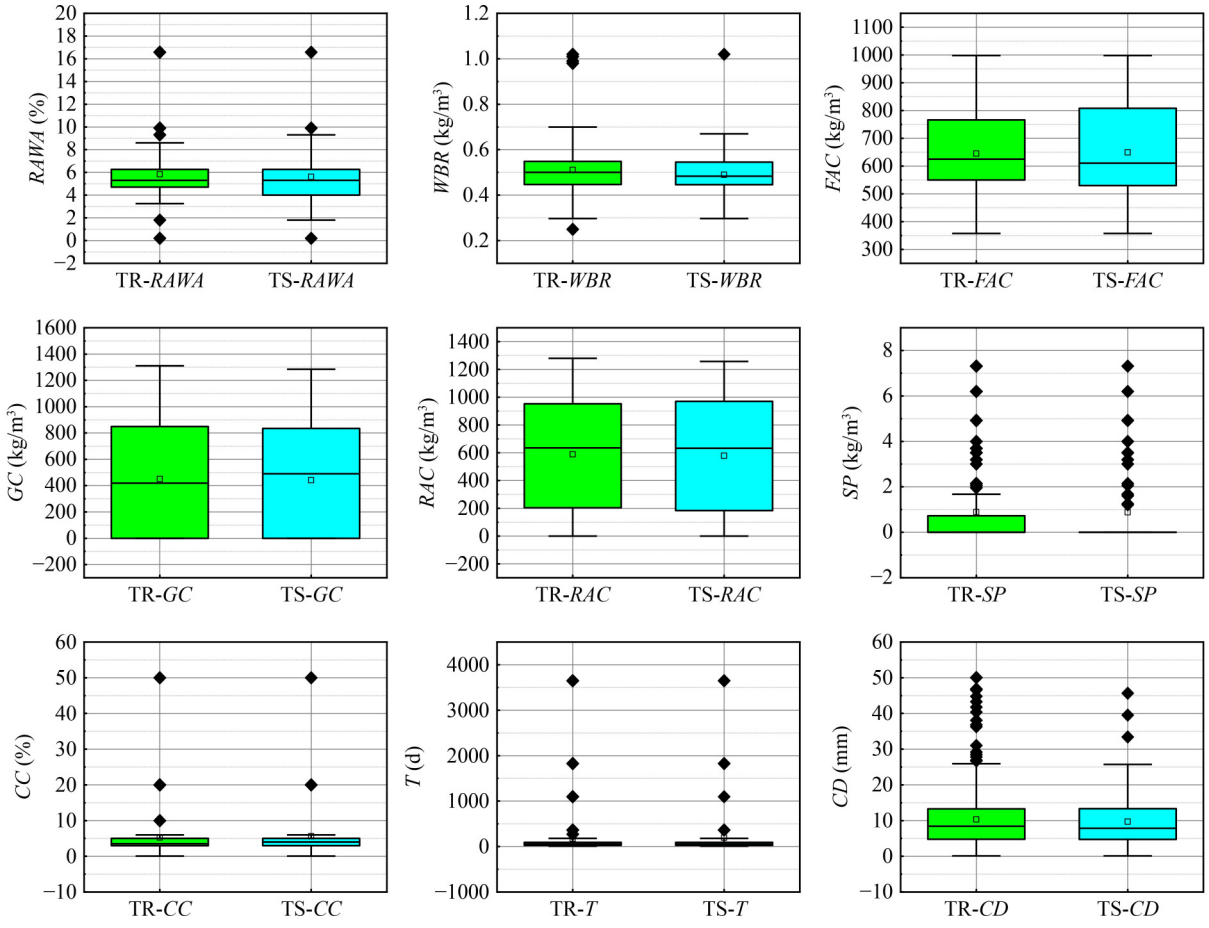


Fig. 8 Data distribution of inputs and *CD* for the training set and testing set.

Table 2 Hyper-parameters need to be optimized and corresponding searching ranges

regression technique	hyper-parameter	searching range	explanations
XGB	learning rate	[0,2]	A larger learning rate tends to induce a faster convergence speed. But it may not converge to the real 'best'. A smaller learning rate tend to mean to be more likely to find a more accurate optimal value, but the convergence speed will be slow.
	n_estimators	[400,800]	"n_estimators" represents the number of weak learners in the decision tree which is related to the complexity of the XGB model. Too small n_estimators is easy to be under-fitted while too large n_estimators will increase the complexity of XGB model.
KELM	regularization_coefficient	[1,500]	This involves a trade-off between training model error and training speed.
	kernel_width	[0.01,1]	"Kernel_width" denotes the influence of an individual training sample to training model and the complexity of the data distribution mapped to a new feature space.
SVR	penalty c	[0.01,500]	"Penalty C" signifies the tolerance for training error where too small C would cause under-fitting and too large C would decrease the generalization ability.
	gamma	[0.01,500]	"Gamma" represents the number of support vector where larger Gamma means fewer support vectors and it influences the training speed.

each combination of hyper-parameters generated a corresponding fitness value. By sorting, the best fitness values and corresponding hyper-parameters can be obtained. In the next step, the iteration progress was initially based on the best combination of hyper-parameters according to different optimization algorithms, until the iteration ends. The change of fitness values (MSE) by different optimization algorithms are shown in Fig. 9. It can be seen that XGB-based *CD* prediction models have lower initial and final fitness values compared with KELM-based and SVR-based *CD* prediction models. SSA-XGB has the lowest initial and final fitness values among all *CD* prediction models. For KELM-based prediction models, it can be found that SSA-KELM has the lowest fitness value when the optimization process ends. The same phenomenon can be observed from the SVR-based *CD* prediction models. Aforementioned facts suggest that the SSA optimization has stronger optimization abilities than GA and MVO.

Table 3 Parameters in optimization algorithms and optimized hyper-parameters in regression techniques

optimization algorithm	parameter	optimized hyper-parameter values
GA	pc=0.7; cross probability pm=0.3; mutation probability	XGB. learning rate: 0.3984 n_estimators: 482.9788
		KELM. regularization_coefficient: 361.1806 kernel_width: 0.1602
		SVR. penalty c: 46.7810 gamma: 9.8945
MVO	WEP_Max=1; maximum of wormhole existence probability WEP_Min=0.2; minimum of wormhole existence probability	XGB. learning rate: 0.5446 n_estimators: 744.5773
		KELM. regularization_coefficient: 321.1179 kernel_width: 0.1487
		SVR. penalty c: 190.2026 gamma: 3.8183
SSA	ST = 0.6; alert value PD = 0.7; proportion of producer sparrows JD = 0.3; proportion of scrounger sparrows SD = 0.2; proportion of alerter sparrows	XGB. learning rate: 0.3215 n_estimators:389.7850
		KELM. regularization_coefficient: 500.0000 kernel_width:0.1554
		SVR. penalty c: 119.8406 gamma: 3.4824

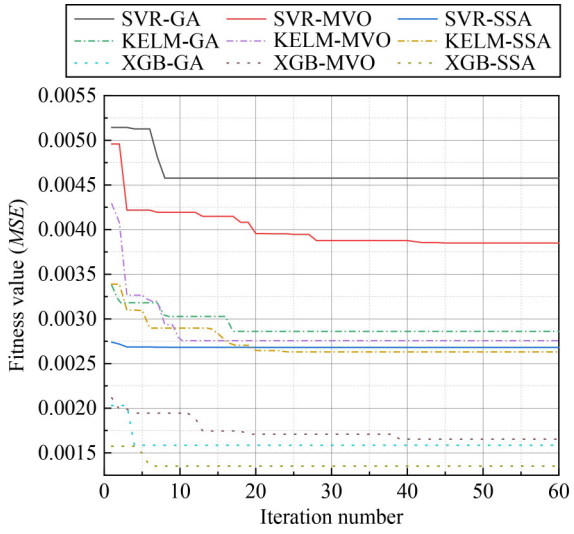


Fig. 9 The optimization process of different optimization algorithms.

6.2 Discussion and results

As mentioned above, the SSA produced smaller fitness value than GA and MVO for three regression techniques. However, it can hardly be concluded that SSA-based *CD* prediction models would bring the best overall prediction performance, because the model parameters obtained from 10-fold cross validation must still be checked by the original training set and testing set. Regarding this, the optimized model parameters were used to predict *CD* in the training set and testing set. The predicted results for training set and testing set, with prediction errors, are demonstrated in Figs. 10 and 11, respectively. In each figure, five fitting lines are depicted to illustrate the general distribution of measured *CD* and predicted *CD*. Meanwhile, the predicted error for each sample is also shown in Figs. 10 and 11. The performance of the training set is here discussed first in Fig. 10. For SVR-based *CD* prediction models, when the measured *CD* is lower than 30mm, most predicted errors are lower, between −5 and +5 mm, where one sample has more than 10mm error for MVO-SVR. For GA-SVR and SSA-SVR, the predicted errors for all measured samples, that are lower than 30 mm, are in the range of −8 to 5 mm. With the increase of *CD*, a few predicted errors are larger than 10mm but smaller than 15 mm for all SVR-based *CD* prediction models. For KELM-based approaches, it can be seen that when the measured *CD* is lower than 30 mm, all predicted errors are in the range of −10 to 5 mm. When the measured *CD* is larger than 30 mm, GA-KELM and MVO-KELM bring an error larger than 10 mm but for SSA-KELM approach, all errors are in the range of −10 and 5 mm. For XGB-based prediction scenarios, it can be seen that only one sample has error than 5 mm. Most samples presented minor prediction errors. Therefore, it can be concluded that XGB-based and SSA-based models seem to have stronger fitting abilities for the training data. In the next step, the performance from the testing set is discussed.

Unlike the performance of the training set, larger prediction errors are produced by the proposed nine prediction models as shown in Fig. 11. For SVR-based models, the prediction errors are lower than 10 mm when the measured *CD* is lower than 35 mm; however, when the measured *CD* is bigger than 35 mm, the prediction errors are much larger and even equal to 20 mm. For KELM-based models, a similar phenomenon occurs. In other words, the predictive ability for SVR-based and KELM-based models is limited when the measured *CD* is larger than 35 mm. For XGB-based approaches, better predictive performance can be obtained, as seen in Fig. 11. For GA-XGB, all prediction errors are in the range of −6 to 10 mm. And for MVO-XGB and SSA-XGB, there is one sample that has more than 10mm error. However, compared with GA-XGB, the range of errors from MVO-XGB and SSA-XGB are more centralized.

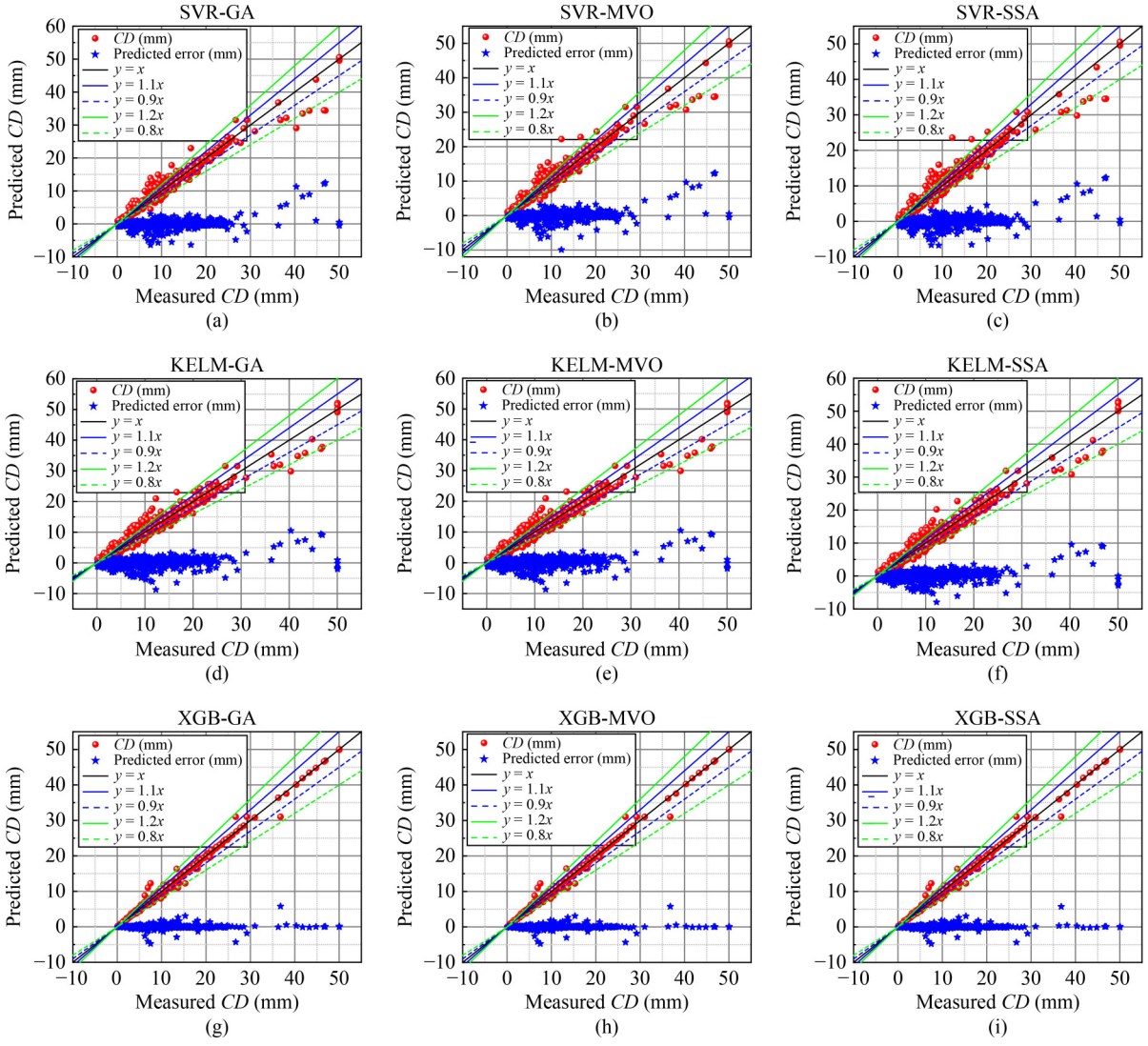


Fig. 10 Comparison between predictive *CD* and measured *CD* for the training data set for: (a) SVR-GA; (b) SVR-MVO; (c) SVR-SSA; (d) KELM-GA; (e) KELM-MVO; (f) KELM-SSA; (g) XGB-GA; (h) XGB-MVO; (i) XGB-SSA.

For quantifying the prediction performance, eight mathematical indictors as mentioned before are used to provide an overall prediction evaluation as shown in Tables 4 and 5. To intuitively compare these evaluation indicators, the ranking scores of each indicator and model are reflected by bar charts in Figs. 12 and 13.

For the ranking scores of the training set, it can be found that for SVR-based methods, GA-SVR has the highest-ranking score. For KELM-based models, SSA-KELM has the highest-ranking score as shown in Fig. 12. For XGB-based scenarios, XGB-MVO has the highest-ranking score. For the performance of testing set, GA-SVR is still the best among three SVR models. Similarly, SSA-KELM has better performance than the two other KELM-based models. For XGB-based approaches, GA-XGB brings the best prediction performance for the testing set. It is interesting that the optimization abilities for a certain optimization algorithm are not always the best for different regression techniques. However, XGB-based models present more robust *CD* prediction performance than the two other regression techniques. The final ranking scores of different models can be seen in Fig. 14. The ranking scores from high to low are: XGB-MVO, XGB-GA, XGB-SSA, SVR-GA, KELM-SSA, SVR-MVO(KELM-GA), KELM-MVO and SVR-SSA. The best overall prediction performance from XGB-MVO is R^2 of (0.9949 and 0.9398), R of (0.9949 and 0.9423), *RMSE* of (0.5764 and 1.7565), *VAF* of (99.4877 and 94.0487), *MAPE* of (0.0300 and 0.1863), *MAE* of (0.2241 and 1.0688), *MBE* of (0.0000 and −0.1915) and *A*10 of (0.9873 and 0.8258), for training and testing stages.

To further reflect the overall performance of proposed *CD* prediction models, a Taylor diagram is employed in this section, as shown in Fig. 15. The Taylor diagram employs three statistical indicators, namely *RMSE*, *R*, and standard deviation, to present a comprehensive evaluation. A smaller distance between the predicted

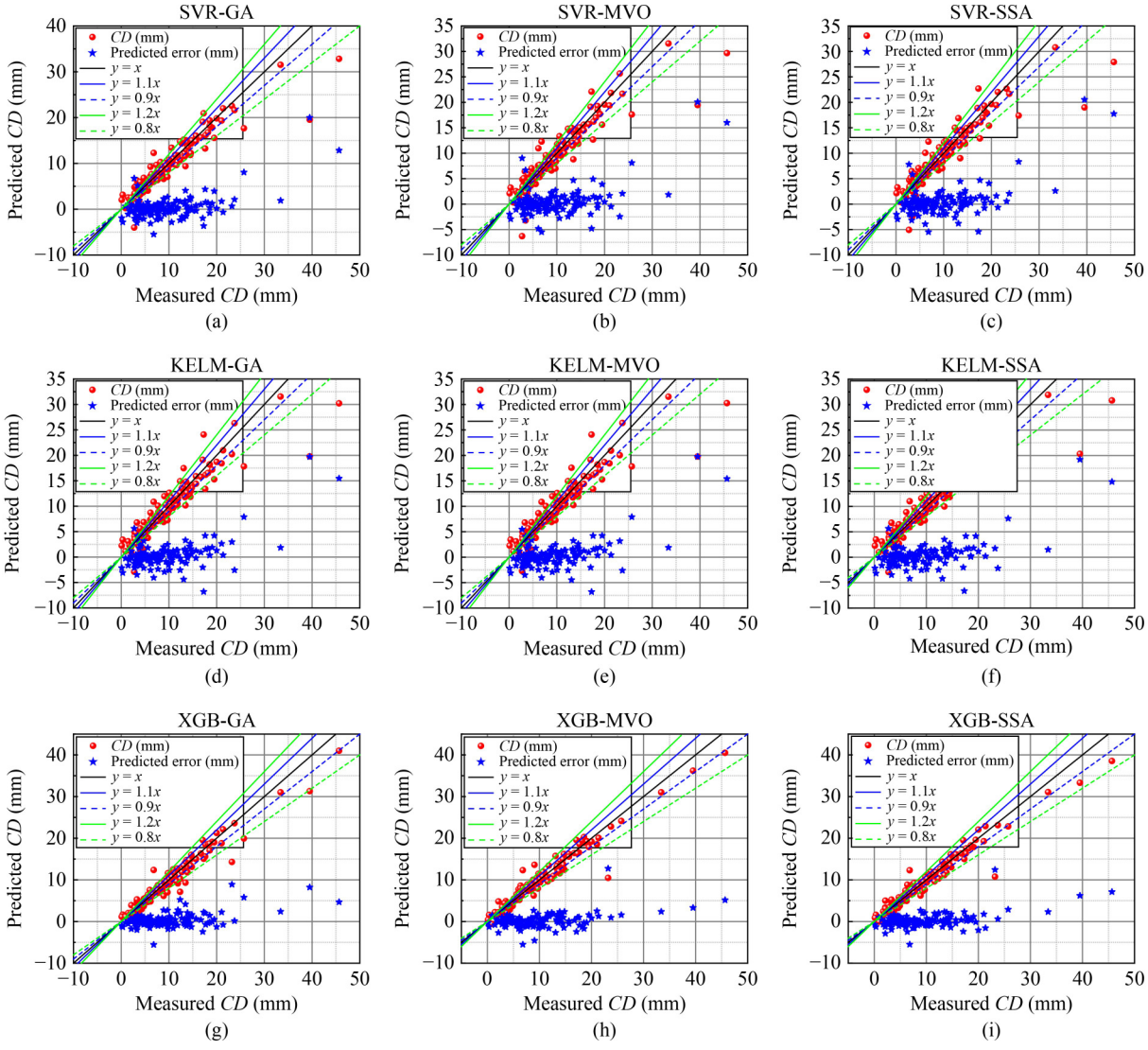


Fig. 11 Comparison between predictive *CD* and measured *CD* for the testing data set for: (a) SVR-GA; (b) SVR-MVO; (c) SVR-SSA; (d) KELM-GA; (e) KELM-MVO; (f) KELM-SSA; (g) XGB-GA; (h) XGB-MVO; (i) XGB-SSA.

point and the reference point (black dot) indicates better predictive performance. The analysis also shows that hybrid XGB models outperformed hybrid SVR and KELM models. Overall, the results suggest that SSA-XGB should be the recommended model for predicting *CD* due to its excellent performance on both the training and testing sets.

6.3 Model validation by parametric analysis

To verify the feasibility of the proposed *CD* of RA concrete predictive models, the most competitive prediction model is selected, i.e., XGB-MVO, to conduct parametric analysis. One influencing factor is selected to be variable and the other influencing factors is set to be constant. Moreover, the constant influencing factors were set at their average values, and the variable influencing factor changed from its lowest to highest value, with a defined interval where the former is determined by the original data distribution as shown in Table 1. The results presented in Fig. 16 provide important insights into the effect of various factors on the depth of carbonation in RA concrete by using XGB-MVO model. The following observations can be made from the data presented.

a) The *RAWA* of RA had a significant impact on the *CD* in RA concrete. When the *RAWA* exceeded 6%, the *CD* increased rapidly due to the increased porosity of the RA, which created more voids in the concrete matrix and reduced its compactness [46].

b) The *WBR* in concrete is another important factor affecting the *CD*. A higher *WBR* leads to a more porous concrete matrix, allowing carbon dioxide to penetrate more easily and increasing the *CD* [47].

c) The *CD* initially increased with increase of *FAC*, but after a certain point (around 600–700 kg/m^3), it started to decrease.

Table 4 Training set performance for different optimized *CD* prediction models

model	training set performance								
	R^2	*R*	*RMSE*	*VAF*	*MAPE*	*MAE*	*MBE*	*A*10	Score
SVR-GA	0.9613	0.9629	1.5840	96.1313	0.1422	0.8853	–0.0091	0.8545	
	(6)	(6)	(6)	(6)	(6)	(6)	(3)	(6)	45
SVR-MVO	0.9581	0.9590	1.6480	95.8127	0.1477	0.9326	0.0148	0.8473	
	(4)	(4)	(4)	(4)	(5)	(5)	(2)	(5)	33
SVR-SSA	0.9540	0.9551	1.7267	95.4033	0.1597	0.9747	0.0149	0.8382	
	(1)	(1)	(1)	(1)	(4)	(4)	(1)	(4)	17
KELM-GA	0.9556	0.9567	1.6965	95.5621	0.1696	1.0821	–0.0019	0.7873	
	(3)	(3)	(3)	(3)	(2)	(2)	(5)	(2)	23
KELM-MVO	0.9555	0.9566	1.6997	95.5455	0.1700	1.0834	–0.0021	0.7855	
	(2)	(2)	(2)	(2)	(1)	(1)	(4)	(1)	15
KELM-SSA	0.9601	0.9608	1.6091	96.0076	0.1617	1.0290	–0.0013	0.8091	
	(5)	(5)	(5)	(5)	(3)	(3)	(6)	(3)	35
XGB-GA	0.9948	0.9948	0.5791	99.4828	0.0319	0.2324	0.0000	0.9855	
	(8)	(8)	(8)	(8)	(8)	(8)	(9)	(8)	65
XGB-MVO	0.9949	0.9949	0.5764	99.4877	0.0300	0.2241	0.0000	0.9873	
	(9)	(9)	(9)	(9)	(9)	(9)	(9)	(9)	72
XGB-SSA	0.9948	0.9948	0.5797	99.4819	0.0351	0.2351	0.0000	0.9836	
	(8)	(8)	(7)	(7)	(7)	(7)	(9)	(7)	60

Table 5 Testing set performance for different optimized *CD* prediction models

model	testing set performance								
	R^2	*R*	*RMSE*	*VAF*	*MAPE*	*MAE*	*MBE*	*A*10	Score
SVR-GA	0.8517	0.8620	2.7558	85.5424	0.3869	1.5154	–0.4342	0.6970	
	(6)	(6)	(6)	(6)	(6)	(5)	(6)	(2)	43
SVR-MVO	0.8255	0.8333	2.9898	82.9741	0.4002	1.5874	–0.4661	0.7045	
	(2)	(2)	(2)	(2)	(3)	(1)	(2)	(5)	19
SVR-SSA	0.8181	0.8287	3.0521	82.2718	0.3953	1.5846	–0.4840	0.6894	
	(1)	(1)	(1)	(1)	(4)	(2)	(1)	(1)	12
KELM-GA	0.8391	0.8503	2.8707	84.3211	0.4006	1.5700	–0.4577	0.6970	
	(4)	(4)	(4)	(4)	(2)	(4)	(4)	(3)	29
KELM-MVO	0.8390	0.8503	2.8714	84.3138	0.4032	1.5710	–0.4582	0.6970	
	(3)	(3)	(3)	(3)	(1)	(3)	(3)	(4)	23
KELM-SSA	0.8498	0.8607	2.7742	85.3779	0.3932	1.4901	–0.4540	0.7197	
	(5)	(5)	(5)	(5)	(5)	(6)	(5)	(6)	42
XGB-GA	0.9401	0.9472	1.7516	94.2327	0.2186	1.0464	–0.3373	0.7955	
	(9)	(9)	(9)	(9)	(7)	(8)	(7)	(8)	66
XGB-MVO	0.9398	0.9423	1.7565	94.0487	0.1863	1.0688	–0.1915	0.8258	
	(8)	(7)	(8)	(8)	(9)	(7)	(9)	(9)	65
XGB-SSA	0.9373	0.9435	1.7916	93.8649	0.2152	1.0277	–0.2594	0.7727	
	(7)	(8)	(7)	(7)	(8)	(9)	(8)	(7)	61

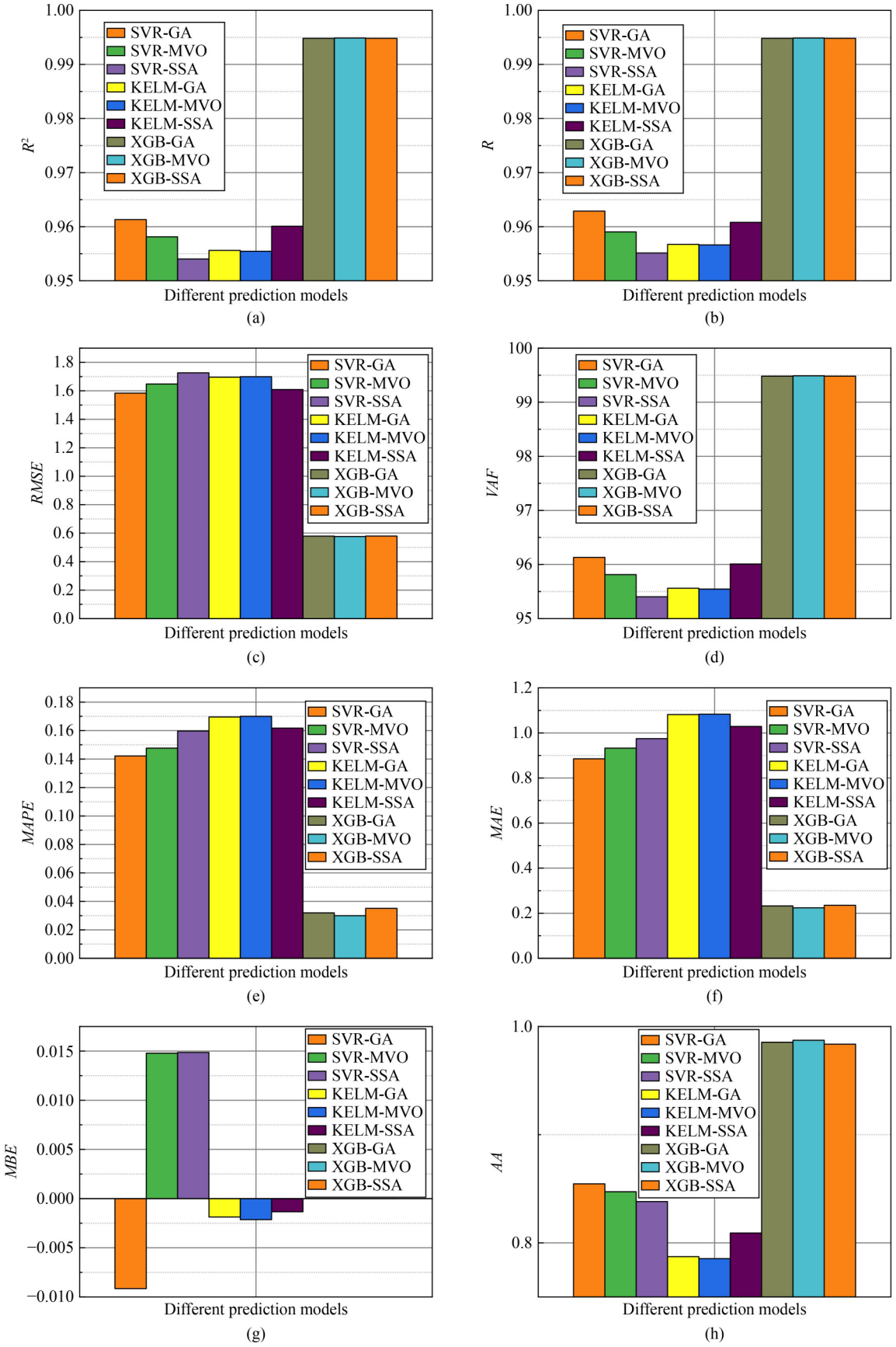


Fig. 12 Comparisons of different evaluation indicators for the training set for: (a) R^2; (b) R; (c) *RMSE*; (d) *VAF*; (e) *MAPE*; (f) *MAE*; (g) *MBE*; (h) *AA*.

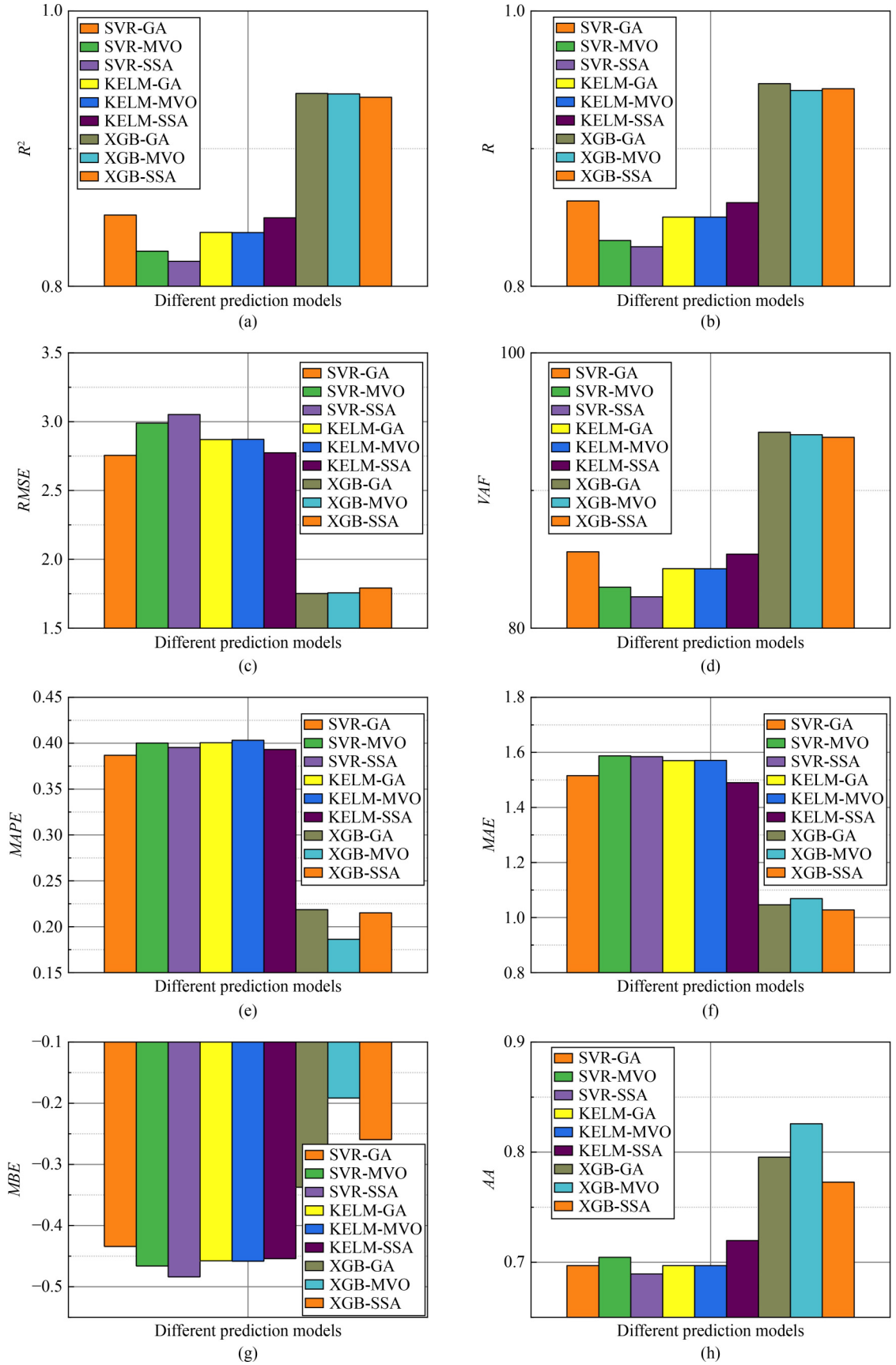


Fig. 13 Comparisons of different evaluation indicators for the testing set for: (a) R^2; (b) R; (c) *RMSE*; (d) *VAF*; (e) *MAPE*; (f) *MAE*; (g) *MBE*; (h) *AA*.

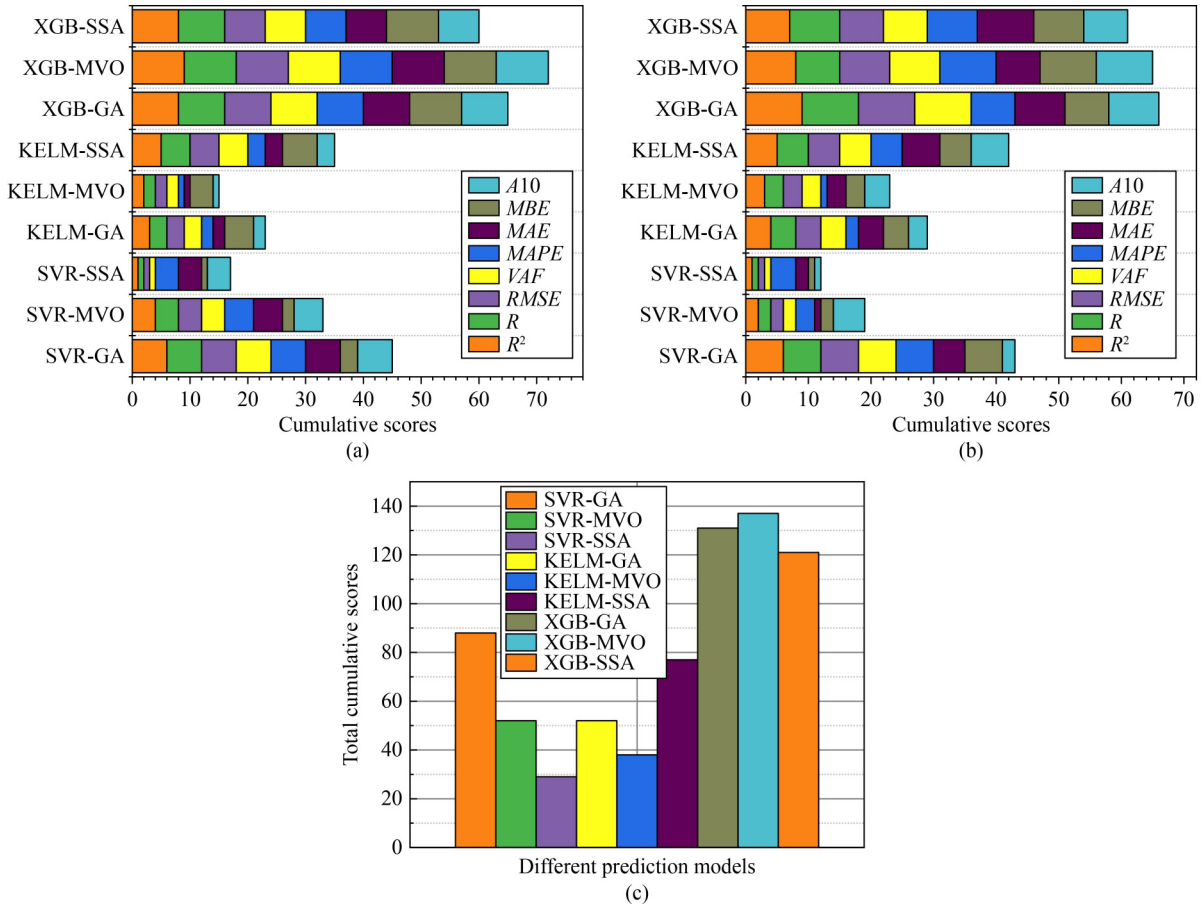
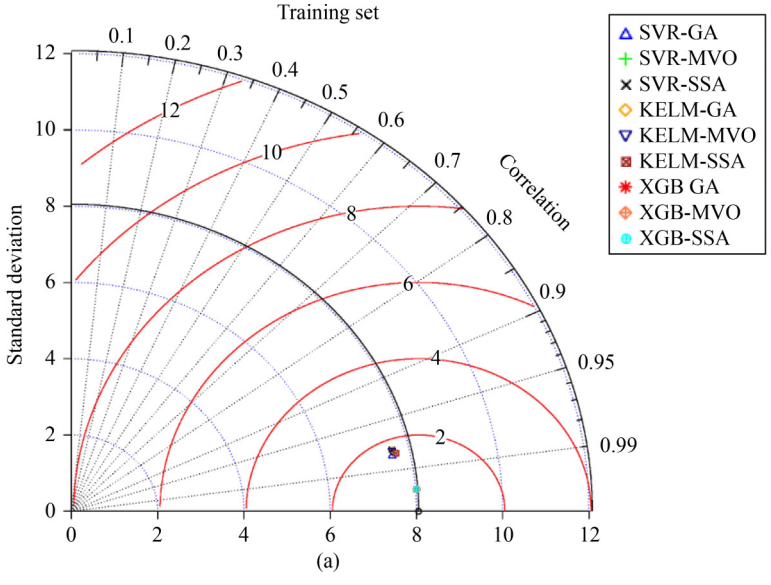


Fig. 14 Cumulative ranking scores of different *CD* prediction models: (a) training set; (b) testing set; (c) the rank of total cumulative scores.

d) Increasing the *GC* in RA concrete had a slight effect on the *CD* at lower content levels (0–1200 kg/m^3), but the increase of *CD* was not significant. However, when *GC* was greater than 1200 kg/m^3, the *CD* increased significantly.

e) The increase of *RAC* significantly increased the *CD* due to its higher porosity and roughness. The *CD* increased rapidly from 10 to 14 mm when *RAC* was between 0 and 1100 kg/m^3, and then further increased to 18 mm after *RAC* reached 1250 kg/m^3.



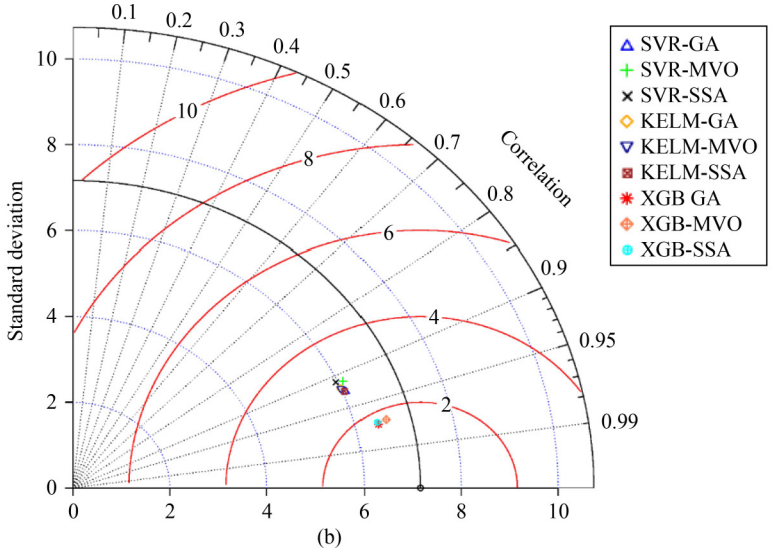


Fig. 15 Demonstration of different hybrid *CD* prediction models by Taylor diagram: (a) training set; (b) testing set.

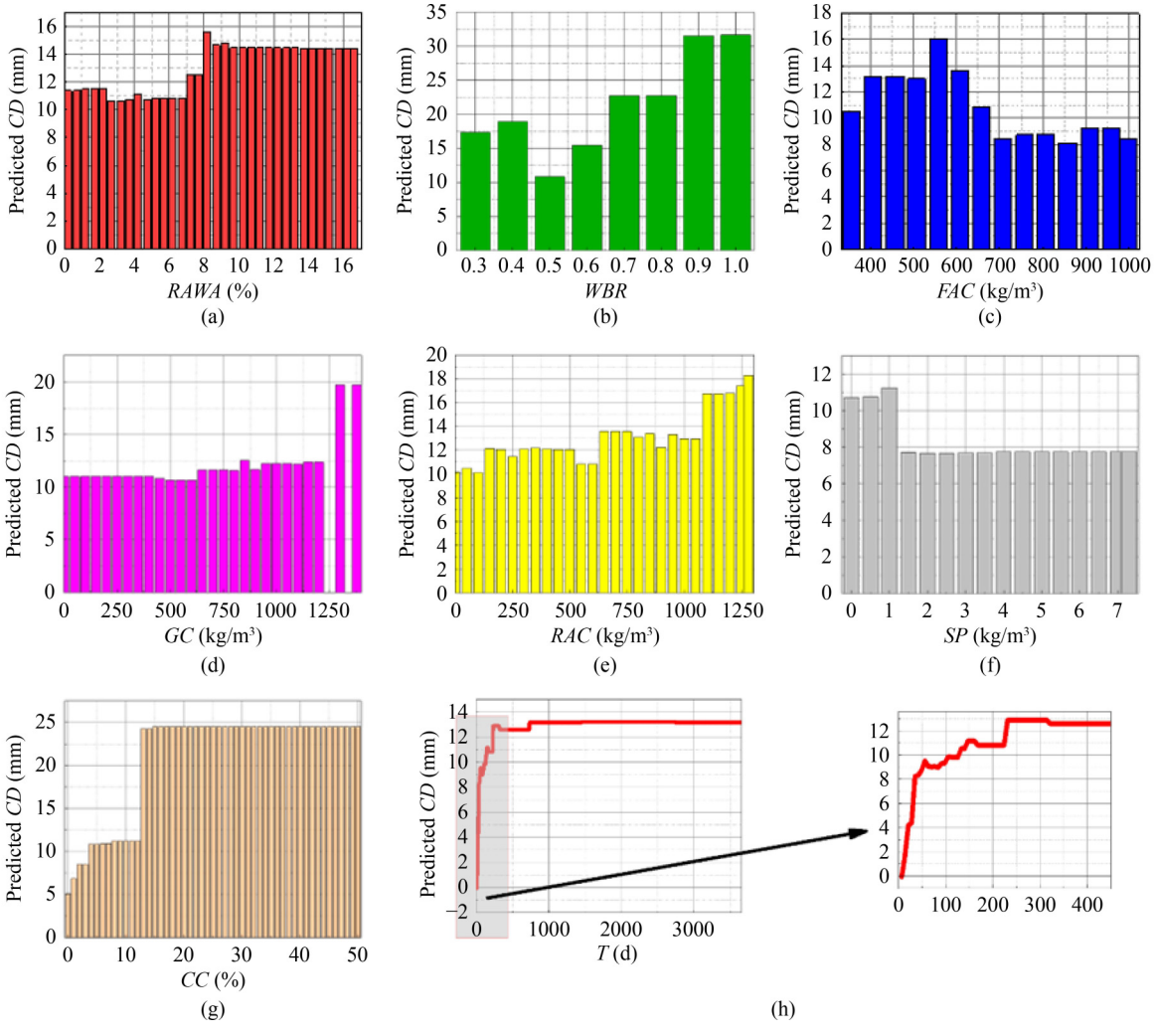


Fig. 16 Parametric analysis and evaluation based on the most competitive *CD* prediction model (XGB-MVO) of predicted *CD* vs: (a) *RAWA*; (b) *WBR*; (c) *FAC*; (d) *GC*; (e) *RAC*; (f) *SP*; (g) *CC*; (h) *T*.

f) The addition of *SP* to RA concrete significantly decreased the *CD*, especially when the increase was greater than 1 kg/m^3. This is because the higher content of *SP* improved the flowability and denseness of the RA concrete matrix, reducing its permeability [48].

g) The *CC* had a significant impact on the *CD* of RA concrete. The *CD* increased significantly with the increase in *CC* up to 12%, after which further increases had little effect on the *CD*. This can be attributed to the formation of a dense carbonation layer on the surface of the concrete, which increased its resistance to carbonation [49].

h) The *CD* of RA concrete increased rapidly in the early stages, but after one year of exposure, the change in *CD* was not significant. The increased *T* causes the carbonation layer and the layer decreases the adsorption rate on the concrete surface [50].

6.4 Comparison of the developed optimal model with literature models

In this section, the optimized model developed in this study is compared with other models from literature for predicting RA concrete *CD* [22,23,27,28]. Figure 17(a) shows a comparison of the *RMSE* and *MAE* values obtained by each model. It was observed that the RILEM 130-CSL model had the highest *RMSE* (11.95) and *MAE* (11.07), which probably can be attributed to the limited data size of 72, as shown in Fig. 17(b). The optimal model XGB-MVO predicted in this study had an *RMSE*

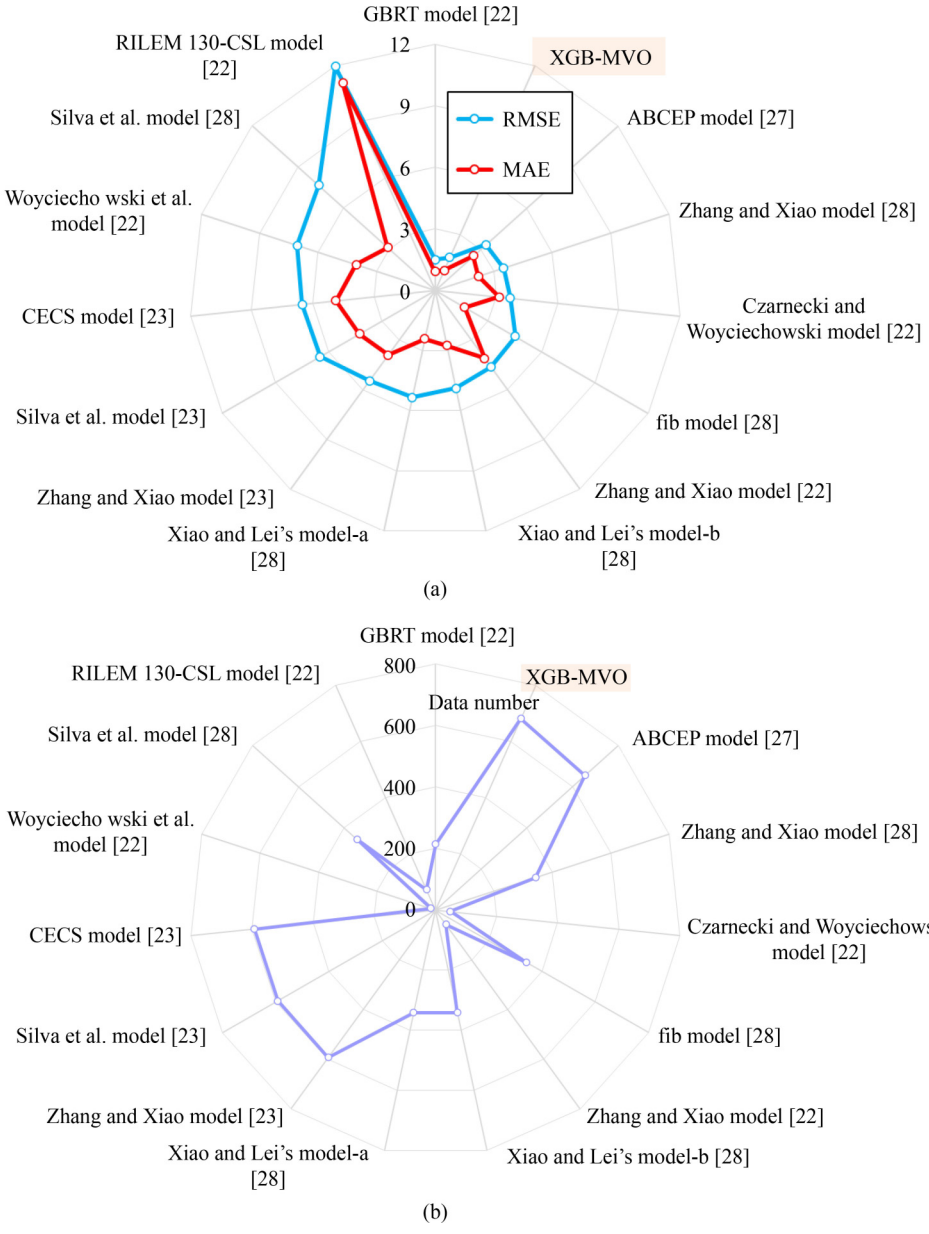


Fig. 17 Models' comparison: (a) *RMSE* and *MAE*; (b) data number in different models.

of 1.76 and *MAE* of 1.07, which was the second-best after the GBRT model, with an *RMSE* of 1.51 and *MAE* of 0.91. However, it is important to note that the GBRT model only used 217 data points (Fig. 17(b)), which may limit its generalizability. The prediction ability for *CD* of RA concrete by XGB-MVO has been validated by more data (682 data).

7 Conclusions

The monitoring of carbonation of concrete has always been a tough task because it is influenced by various factors. In this study, in order to further develop a model to accurately predict the *CD* of recycled concrete, a data set of the *CD* of RA concrete was obtained, based on the literature, and the effect of new combinations of influential factors on *CD* was simulated: *RAWA*, *WBR*, *FAC*, *GC*, *RAC*, *SP*, *CC*, and *T*. *CD* was the output. Meanwhile, three types of regression techniques and three types of meta-heuristic algorithms were combined and a total of nine *CD* prediction models were developed. It was found that the best competitive predictive performance was produced by XGB-MVO with R^2 of (0.9949 and 0.9398), *R* of (0.9949 and 0.9423), *RMSE* of (0.5764 and 1.7565), *VAF* of (99.4877 and 94.0487), *MAPE* of (0.0300 and 0.1863), *MAE* of (0.2241 and 1.0688), *MBE* of (0.0000 and −0.1915) and *A10* of (0.9873 and 0.8258) for training and testing set, respectively. The results demonstrate that the proposed predictive model effectively forecasts the *CD* of RA concrete. The predicted outcomes of the model align with previous research, attesting to its robustness. In addition, XGB-MVO also shows optimal predictive and generalization capabilities when compared with different models in the literature. The model's ability to estimate *CD* for unknown RA concrete samples highlights its potential in advancing the development and application of RA s in concrete, and furthering sustainable economic growth.

Acknowledgements Bin Xi and Enming Li want to acknowledge the funding supported by China Scholarship Council (Nos. 202008440524 and 202006370006). This research was partially supported by the Distinguished Youth Science Foundation of Hunan Province of China (No. 2022JJ10073), the Innovation Driven Project of Central South University (No. 2020CX040) and Shenzhen Science and Technology Plan (No. JCYJ20190808123013260).

Funding note Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Conflict of Interest The authors declare that they have no competing interests.

References

1. Li E, Xi B, Zhang N, Shi X, Zhou J, Segarra P, Wang H. Circular use of fine-grained tailings to underground mine wind walls. Circular Economy, 2023, 2(3): 100053
2. Khajuria A, Atienza V A, Chavanich S, Henning W, Islam I, Kral U, Liu M, Liu X, Murthy I K, Oyedotun T D T, Verma P, Xu G, Zeng X, Li J. Accelerating circular economy solutions to achieve the 2030 agenda for sustainable development goals. Circular Economy, 2022, 1(1): 100001
3. Xi B, Zhou Y, Yu K, Hu B, Huang X, Sui L, Xing F. Use of nano-SiO_2 to develop a high performance green lightweight engineered cementitious composites containing fly ash cenospheres. Journal of Cleaner Production, 2020, 262: 121274
4. Zeng X, Ogunseitan O A, Nakamura S, Suh S, Kral U, Li J, Geng Y. Reshaping global policies for circular economy. Circular Economy, 2022, 1(1): 100003
5. Zhang N, Xi B, Li J, Liu L, Song G. Utilization of CO_2 into recycled construction materials: A systematic literature review. Journal of Material Cycles and Waste Management, 2022, 24(6): 2108–2125
6. Zhang N, Zhang D, Zuo J, Miller T R, Duan H, Schiller G. Potential for CO_2 mitigation and economic benefits from accelerated carbonation of construction and demolition waste. Renewable & Sustainable Energy Reviews, 2022, 169(9): 112920
7. Zhang N, Duan H, Miller T R, Tam V W Y, Liu G, Zuo J. Mitigation of carbon dioxide by accelerated sequestration in concrete debris. Renewable & Sustainable Energy Reviews, 2020, 117: 109495
8. Visintin P, Xie T, Bennett B. A large-scale life-cycle assessment of recycled aggregate concrete: The influence of functional unit, emissions allocation and carbon dioxide uptake. Journal of Cleaner Production, 2020, 248: 119243
9. Villoria Sáez P, Osmani M. A diagnosis of construction and demolition waste generation and recovery practice in the European Union. Journal of Cleaner Production, 2019, 241: 118400
10. Lu W, Webster C, Peng Y, Chen X, Zhang X. Estimating and calibrating the amount of building-related construction and demolition waste in urban China. International Journal of Construction Management, 2017, 17(1): 13–24
11. Bai G, Zhu C, Liu C, Liu B. An evaluation of the recycled

aggregate characteristics and the recycled aggregate concrete mechanical properties. Construction & Building Materials, 2020, 240: 117978
12. Li X. Recycling and reuse of waste concrete in China. Resources, Conservation and Recycling, 2008, 53(1−2): 36−44
13. Xiao J, Lei B, Zhang C. On carbonation behavior of recycled aggregate concrete. Science China. Technological Sciences, 2012, 55(9): 2609−2616
14. Li L, Wu M. An overview of utilizing CO_2 for accelerated carbonation treatment in the concrete industry. Journal of CO_2 Utilization, 2022, 60(3): 102000
15. Silva R V, Neves R, de Brito J, Dhir R K. Carbonation behaviour of recycled aggregate concrete. Cement and Concrete Composites, 2015, 62: 22−32
16. Balayssac J P, Détriché C H, Grandet J. Effects of curing upon carbonation of concrete. Construction & Building Materials, 1995, 9(2): 91−95
17. Atiş C D. Accelerated carbonation and testing of concrete made with fly ash. Construction & Building Materials, 2003, 17(3): 147−152
18. Leemann A, Loser R. Carbonation resistance of recycled aggregate concrete. Construction & Building Materials, 2019, 204: 335−341
19. Lovato P S, Possan E, Molin D C C D, Masuero Â B, Ribeiro J L D. Modeling of mechanical properties and durability of recycled aggregate concretes. Construction & Building Materials, 2012, 26(1): 437−447
20. Matias D, de Brito J, Rosa A, Pedro D. Durability of concrete with recycled coarse aggregates: Influence of superplasticizers. Journal of Materials in Civil Engineering, 2014, 26(7): 06014011
21. Zega C J, di Maio Á A. Use of recycled fine aggregate in concretes with durable requirements. Waste Management, 2011, 31(11): 2336−2340
22. Nunez I, Nehdi M L. Machine learning prediction of carbonation depth in recycled aggregate concrete incorporating SCMs. Construction & Building Materials, 2021, 287: 123027
23. Liu K, Alam M S, Zhu J, Zheng J, Chi L. Prediction of carbonation depth for recycled aggregate concrete using ANN hybridized with swarm intelligence algorithms. Construction & Building Materials, 2021, 301: 124382
24. Liu B, Vu-Bac N, Rabczuk T. A stochastic multiscale method for the prediction of the thermal conductivity of Polymer nanocomposites through hybrid machine learning algorithms. Composite Structures, 2021, 273: 114269
25. Vu-Bac N, Rafiee R, Zhuang X, Lahmer T, Rabczuk T. Uncertainty quantification for multiscale modeling of polymer nanocomposites with correlated parameters. Composites. Part B, Engineering, 2015, 68: 446−464
26. Nariman N, Hamdia K, Ramadan A, Sadaghian H. Optimum design of flexural strength and stiffness for reinforced concrete beams using machine learning. Applied Sciences, 2021, 11(18): 8762
27. Moghaddas S A, Nekoei M, Mohammadi Golafshani E, Nehdi M, Arashpour M. Modeling carbonation depth of recycled aggregate concrete using novel automatic regression technique. Journal of Cleaner Production, 2022, 371(3): 133522
28. Zhang K, Xiao J. Prediction model of carbonation depth for recycled aggregate concrete. Cement and Concrete Composites, 2018, 88: 86−99
29. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. Noise Reduction in Speech Processing, 2009, 1–4
30. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 2004, 69(6): 066138
31. Shamshirband S, Mohammadi K, Chen H L, Narayana Samy G, Petković D, Ma C. Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: A case study for Iran. Journal of Atmospheric and Solar-Terrestrial Physics, 2015, 134: 109−117
32. Li E, Zhang N, Xi B, Zhou J, Gao X. Compressive strength prediction and optimization design of sustainable concrete based on squirrel search algorithm-extreme gradient boosting technique. Frontiers of Structural and Civil Engineering, 2023, 17(9): 1310−1325
33. Carmona P, Climent F, Momparler A. Predicting failure in the US banking sector: An extreme gradient boosting approach. International Review of Economics & Finance, 2019, 61: 304−323
34. Biswas R, Li E, Zhang N, Kumar S, Rai B, Zhou J. Development of hybrid models using metaheuristic optimization techniques to predict the carbonation depth of fly ash concrete. Construction & Building Materials, 2022, 346(7): 128483
35. Li E, Zhou J, Shi X, Jahed Armaghani D, Yu Z, Chen X, Huang P. Developing a hybrid model of salp swarm algorithm-based support vector machine to predict the strength of fiber-reinforced cemented paste backfill. Engineering with Computers, 2021, 37(4): 3519−3540
36. Vu-Bac N, Duong T X, Lahmer T, Zhuang X, Sauer R A, Park H S, Rabczuk T. A NURBS-based inverse analysis for reconstruction of nonlinear deformations of thin shell structures. Computer Methods in Applied Mechanics and Engineering, 2018, 331: 427−455
37. Vu-Bac N, Rabczuk T, Park H S, Fu X, Zhuang X. A NURBS-based inverse analysis of swelling induced morphing of thin stimuli-responsive polymer gels. Computer Methods in Applied Mechanics and Engineering, 2022, 397: 115049
38. Zhou J, Huang S, Qiu Y. Optimization of random forest through the use of MVO, GWO and MFO in evaluating the stability of underground entry-type excavations. Tunnelling and Underground Space Technology, 2022, 124(4): 104494
39. Zhou J, Shen X, Qiu Y, Shi X, Khandelwal M. Cross-correlation stacking-based microseismic source location using three metaheuristic optimization algorithms. Tunnelling and Underground Space Technology, 2022, 126(5): 104570
40. Ding S, Su C, Yu J. An optimizing BP neural network algorithm based on genetic algorithm. Artificial Intelligence Review, 2011, 36(2): 153−162
41. Hamdia K M, Zhuang X, Rabczuk T. An efficient optimization approach for designing machine learning models based on genetic algorithm. Neural Computing & Applications, 2021, 33(6): 1923−1933
42. Mirjalili S, Mirjalili S M, Hatamlou A. Multi-Verse Optimizer: A nature-inspired algorithm for global optimization. Neural Computing & Applications, 2016, 27(2): 495−513

43. Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm. Systems Science & Control Engineering, 2020, 8(1): 22−34
44. Cortes C, Vapnik V. Support-vector networks. Machine Learning, 1995, 20(3): 273−297
45. Zorlu K, Gokceoglu C, Ocakoglu F, Nefeslioglu H A, Acikalin S. Prediction of uniaxial compressive strength of sandstones using petrography-based models. Engineering Geology, 2008, 96(3–4): 141−158
46. Eckert M, Oliveira M. Mitigation of the negative effects of recycled aggregate water absorption in concrete technology. Construction & Building Materials, 2017, 133: 416−424
47. Ho D W S, Lewis R K. Carbonation of concrete and its prediction. Cement and Concrete Research, 1987, 17(3): 489−504
48. Papayianni I, Tsohos G, Oikonomou N, Mavria P. Influence of superplasticizer type and mix design parameters on the performance of them in concrete mixtures. Cement and Concrete Composites, 2005, 27(2): 217−222
49. Peter M A, Muntean A, Meier S A, Böhm M. Competition of several carbonation reactions in concrete: A parametric study. Cement and Concrete Research, 2008, 38(12): 1385−1393
50. Dias W P. Reduction of concrete sorptivity with age through carbonation. Cement and Concrete Research, 2000, 30(8): 1255−1261