


Article

Machine Learning Application to Eco-Friendly Concrete Design for Decarbonisation

Abigail Lavercombe ¹, Xu Huang ² and Sakdirat Kaewunruen ^{1,2,*} 

¹ Department of Civil Engineering, School of Engineering, University of Birmingham, Birmingham B15 2TT, UK; AXL055@alumni.bham.ac.uk

² Laboratory for Track Engineering and Operations for Future Uncertainties (TOFU Lab), School of Engineering, University of Birmingham, Birmingham B15 2TT, UK; XXH689@student.bham.ac.uk

* Correspondence: s.kaewunruen@bham.ac.uk; Tel.: +44-121-414-2670

Abstract: Cement replacement materials can not only benefit the workability of the concrete but can also improve its compressive strength. Reducing the cement content of concrete can also lower CO₂ emissions to mitigate the impact of the construction industry on the environment and improve energy consumption. This paper aims to predict the compressive strength (CS) and embodied carbon (EC) of cement replacement concrete using machine learning (ML) algorithms, i.e., deep neural network (DNN), support vector regression (SVR), gradient boosting regression (GBR), random forest (RF), k-nearest neighbors (kNN), and decision tree regression (DTR). Not only is producing an optimal ML model helpful for predicting accurate results, but it also saves time, energy, and costs, compared to conducting experiments. Firstly, 367 pieces of experimental datasets from the open literature were collected, in which cement was replaced with any of the cementitious materials. Secondly, the datasets were imported into the ML models, whose parameters were tuned by the grid search algorithm (GSA). Then, the prediction performance, the coefficient of determination (R²), the prediction accuracy, and the root mean square error (RMSE) were employed to indicate the prediction ability of the ML models. The results demonstrate that the GBR models perform the best prediction of the CS and EC. The R² of the GBR models for predicting the CS and EC are 0.946 and 0.999, respectively. Thus, it can be concluded that the GBR models have promising abilities for design assistance in cement replacement concrete. Finally, a sensitivity analysis (SA) was conducted in this paper to analyse the effects of the inputs on the CS and EC of the cement replacement concrete. Pulverised fuel ash (PFA), blast-furnace slag (GGBS), Expanded perlite (EP), and Silica fume (SF) were noticed to affect the CS and EC of cement replacement concrete significantly.

Keywords: machine-learning-assisted design; eco-friendly concrete; cement replacement materials; embodied carbon; K-fold cross validation



Citation: Lavercombe, A.; Huang, X.; Kaewunruen, S. Machine Learning Application to Eco-Friendly Concrete Design for Decarbonisation. *Sustainability* **2021**, *13*, 13663. <https://doi.org/10.3390/su132413663>

Academic Editors: Michael Wistuba, Di Wang, Chiara Riccardi, Libo Yan, Zhanping You, Lily Poulidakos and Ana Jiménez del Barco Carrión

Received: 8 November 2021

Accepted: 8 December 2021

Published: 10 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Concrete is the most widely used artificial material, and it is the second most consumed resource in the world after water. More than four billion tonnes of cement are produced each year, which accounts for approximately 8% of global CO₂ emissions [1]. The most significant mechanical property of a concrete design is its compressive strength. The comprehensive strength is a reliable indicator of the overall performance of a concrete mix because it can be used to calculate other properties of concrete with high precision. Cement is a vital ingredient in concrete. Cement acts as the “glue”, which holds the fine and coarse aggregates together and gives the concrete strength. The primary raw materials of cement are limestone and clay, which are pulverised and blended with other elements, such as iron ore. These materials are then fed into a cylindrical kiln and are heated to approximately 1450 degrees Celsius. This process, known as “calcination”, generates more than 50% of the total CO₂ emissions of cementitious products [1].

Embodied carbon (EC) is the total CO₂ emitted when materials are produced. It can be estimated from the energy used to extract, process, and transport materials. The EC stated throughout this paper considers the carbon emissions from the manufacturing, transportation, and extraction processes of the supplementary cementitious materials (SMCs). When replacing the cement with other binding materials, the EC is reduced. In the majority of cases, the CS tends to decrease as the cement content drops. However, studies have demonstrated that some cementitious materials, such as ground granulated blast-furnace slag (GGBS) and metakaolin, produce a slightly increased CS when replacing cement in small amounts [2]. For this reason, it is not easy to calculate the CS and EC when reducing the cement content. This paper aims to assist in concrete design by producing high-quality ML models that can accurately predict the CS and EC of concrete with various cement replacement materials.

1.1. Cement Replacement Materials

Pulverised fuel ash (PFA) is a type of waste material that is collected from coal-burning power stations. It cannot be considered the best replacement material for cement as there are few coal-burning power stations left in the United Kingdom and, thus, it would have to be shipped in from elsewhere [3]. This not only increases the cost and time to construction, but also the total EC of the concrete because of the shipping emissions. GGBS is a byproduct from the production of iron. It is a cementitious material produced in blast furnaces, and its typical chemical composition is 40% calcium oxide, 35% silica, 13% alumina, and 8% magnesia [4]. It is a famous cement replacement material and has been researched in detail for many years. It is often combined with PFA at higher volumes of cement replacement [5]. Limestone fines (also known as limestone powder) can reduce the number of voids between aggregates and cement paste. This property makes it a useful material for reducing the CO₂ emissions of concrete, and it can be utilised as up to a 35% replacement within the British standard, BS EN 197-1 [6]. Silica fume (SF) is a pozzolanic material. It is a byproduct of the ferrosilicon industry that can enhance the mechanical properties of concrete. Employing SF increases the demand for water within the mix but it can produce dense and impermeable concrete [7]. Metakaolin is a relatively new cement replacement material. It was observed to increase and decrease the overall CS of concrete in various studies [8]. It should also be noted that, similarly to SF, metakaolin increases the demand for water within the concrete mix. Perlite is a naturally occurring material that is produced by the rapid cooling of volcanic lava. Expanded perlite (EP, also known as perlite powder) is perlite that has undergone a heating process that causes it to expand 4–20 times its original size [9]. EP is a widely available building material that has been reported to reduce CO₂ emissions significantly, and to improve the flowability of concrete slurry by means of replacing a part of the cement content [10,11]. A cement replacement of up to 40% of EP could produce 28-day compressive strength values comparable to traditional concrete [11]. Pumice is a pozzolanic material derived from volcanic sources, similar to perlite. It can be used in concrete as an aggregate, or as an SMC when ground into a powder. Studies have demonstrated that, when utilised at levels up to 25%, adding ground pumice as an SMC could improve the CS of concrete mixes at later ages [12].

1.2. Literature Review of ML-Assisted Prediction

Many articles were found from extensive research that provides details about using ML models to predict the mechanical properties of concrete containing cement replacement materials. For instance, four ML models were developed by Mohammed et al. [13] to predict the CS of mortar containing PFA. A total of 450 pieces of experimental CS data, on mortar with PFA ranging from 0% to 70%, were used to develop the ML models. A linear regression model (LR), a nonlinear regression model (NLR), an M5P-tree model (M5P), and an artificial neural network (ANN) were created. The ANN predicted the CS better than all of the other applied models, with a reliable correlation coefficient (R) of 0.934. Moreover, the CS of high-performance concrete with PFA and SF was investigated using

an ANN. The ANN used nonlinear modelling and employed a radial basis function (RBF). The ANN results showed a strong correlation between the experimental and predicted CSs, with the coefficient of correlation being 0.96 [14]. Furthermore, the CS of geopolymer concrete (GPC), with partial cement replacement with GGBS, SF, and natural zeolite, was experimentally studied. The replacement ranged from 0% to 30% in 5% increments. An ANN model to predict the CS of GPC containing these materials was also proposed, using the results collected during the study. To achieve the lowest absolute percent error, the ANN had two hidden layers, with six and five neurons, respectively. The ANN produced a training mean square error (MSE) of 3.5262 and a R of 0.985, validating the experimental results and the model itself [15]. Moreover, the CS of concrete with GGBS was modelled using the M5P model and an ANN. The ANN models hybridised with the M5P model produced architecture with approximately half the errors produced from the M5P model alone, in both the training and testing phases. The study also states that developing a predictive model is necessary for GGBS, as it is required in many design codes. This indicates that more cement replacement materials have the potentials to be included in codes as standards with time. This also evidences that the optimal ML models produced in this paper will become more useful over time [16]. Moreover, the CS of the mortar containing metakaolin was to be predicted employing the support vector machine, RF, decision tree, adaptive boosting (AdBoost), and kNN algorithms. The results show that AdaBoost and RF exhibited higher R^2 values than other types of ML models. The R^2 values of AdaBoost and RF were 0.9473 and 0.9439, respectively. Moreover, the CS of high-performance concrete was predicted using GBR, Gaussian process regression, and kernel ridge regression. The results demonstrate that the GBR model had the best prediction performance ($R = 0.965$) among the ML models [17].

However, only three types of cement replacement materials were considered in one article, at most. More cement replacement materials, such as SF, PFA, pumice, and GGBS, can be employed together as variables to enhance the practical applicability of an ML model for the better design instruction of cement replacement concrete. Additionally, there is no research related to the EC prediction of cement replacement concrete, which means that the EC of cement replacement concrete cannot be chosen in concrete design. Furthermore, the amount of cement should be considered an input to improve the prediction ability of ML models. Moreover, the proportions of coarse and fine aggregates need to be taken into account in the ML models, even though the EC of them are not considered in this paper.

This paper will build on the ideas addressed in previous literature and will fill in the gaps in the research by combining seven alternative cement replacement materials into six types of ML models. It is proposed, for the first time, that the EC of cement replacement concrete can reduce the negative influence of concrete on the environment. Firstly, 367 experimental datasets, containing the CSs and EC of cement replacement concrete, were collected from the open literature. Secondly, 12 variables of the datasets were considered as the inputs of the ML models. Then, SVR, RF, GBR, DTR, kNN, and DNN were utilised to predict the CS and EC of the cement replacement concrete, while the parameters of the ML models were tuned by the GSA. Finally, the R^2 and the RMSE values of the ML models were compared to identify the optimal ML model. This means that the concrete mixes containing at least one of the inputs can be designed to produce the expected CS while lowering the overall amount of EC produced by the concrete.

2. Materials and Methods

2.1. Data Collection

In this paper, 367 pieces of experimental datasets associated with the CS and EC of concrete containing cement replacement materials were collected from the open literature [2,6,7,18–31]. The datasets were collected from a large number of credible and published sources to ensure that the datasets were as accurate and robust as possible. Table 1 shows the EC values for each of the cementitious input materials of the ML models, according to British standards and the open literature. The EC of concrete mixes containing SMCs

is calculated according to Equation (1), in accordance with EN15978 [32]. The datasets are simplified into 12 input variables and 2 output variables, CS and EC, listed in Table 2. Regarding Table 2, the water binder ratio (WB), the amount of cement (C), the amount of superplasticiser (S), the amount of coarse aggregates (CA), and the amount of fine aggregates (FA) are essential concrete components. In addition, the variables associated with cement replacement materials include PFA, LP, SF, M, PP, and GP.

$$\text{Cumulative EC of a concrete mix} = EC_{process,m} + \left(\sum_{i=1}^n EC_i + EC_{transportation,i} \right) \quad (1)$$

where $EC_{process,m}$ indicates the EC produced by the manufacturing of SMCs ($m = C, PFA, SF, \text{etc.}$). EC_i and $EC_{transportation,i}$ represent the EC due to the manufacturing and transportation processes of the intermediate product, i , to produce SMCs.

Table 1. Embodied carbon of cement replacement materials.

Material	Embodied Carbon (kg CO ₂ /t)	Reference
Cement	860	[33]
Pulverised fuel ash	0.1	[34]
Ground granulated blast-furnace slag	79.6	[35]
Limestone powder	8	[36]
Silica fume	28	[37]
Metakaolin	330	[37]
Perlite powder	30	[38]
Ground pumice	30	[38]

Table 2. Inputs and outputs of the ML models.

Type of Parameters	Parameter	Symbol	Unit	Minimum	Maximum
Inputs	W/B ratio	WB	-	0.25	1
	Cement	C	%	0	100
	Pulverised fuel ash	PFA	%	0	70
	Ground granulated blast-furnace slag	GGBS	%	0	70
	Limestone powder	LP	%	0	43
	Silica fume	SF	%	0	25
	Metakaolin	M	%	0	50
	Perlite powder	PP	%	0	20
	Ground pumice	GP	%	0	25
	Coarse aggregate	CA	kg	0	1437.75
	Fine aggregate	FA	kg	0	1166
Outputs	Superplasticizer	S	-	0	1
	Compressive strength at 28 days	CS	MPa	3.3	106.5
	Embodied carbon	EC	kg CO ₂ /t	40.64	860.00

2.2. Prediction Procedure

Firstly, 367 datasets were normalised, employing Equation (2), to enhance the prediction ability of the ML models. The normalised data ranged from zero to one.

$$Y_n = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (2)$$

where Y_n is the experimental data after normalisation; y_{min} and y_{max} are the minimum and maximum experimental data, respectively; and y is the raw experimental data.

Secondly, the normalised datasets were input into the introduced ML models, i.e., GBR, DTR, RF, SVR, kNN, and DNN. Detailed information about the ML models can be found in the articles, [17,39–50]. The experimental datasets were randomly split at 8:2. A total of 20% of the datasets were randomly selected to test the generalisation ability of the ML models, while the rest of the datasets were utilised for training the ML models.

Then, the parameters of the ML models were tuned by the GSA to generate the best prediction ability. The GSA can search for the optimal parameter of multidimensional arrays from various directions. The theory of the GSA is to select the optimal parameter through an exhaustive analysis of a series of parameter combinations [51]. The GSA has been extensively employed to conduct ML parameter optimisation. For instance, a high prediction ability, 0.952 of the R^2 value, was obtained for predicting the CS [52]. The reasons for employing the GSA as the optimisation method are as follows [22,53]:

- Multiple parameters can be tuned simultaneously;
- It will not take a long time to conduct the GSA for fewer parameters;
- The global optimal solution can be obtained by employing the GSA.

Finally, the optimal ML models for predicting the CS and EC of concrete containing cement replacement materials were selected according to the prediction ability indicators, the R^2 and the RMSE, explained in Equations (3) and (4) [54,55]:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y'_i - y_i)^2}{n}} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y'_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

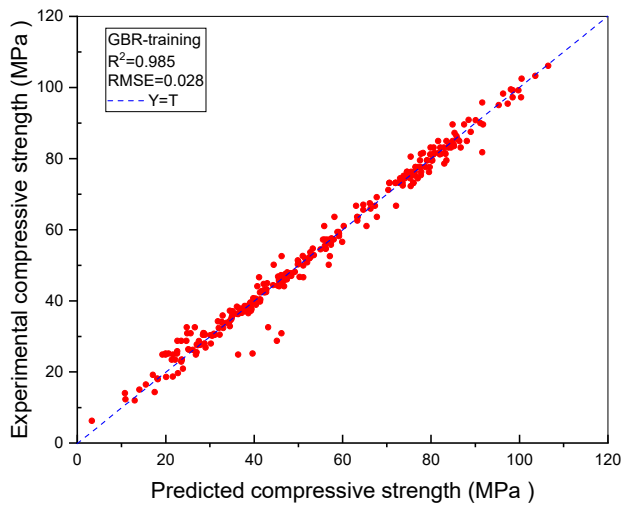
where n indicates the samples number; y'_i indicates the predicted value; and y_i represents the experimental value.

3. Results

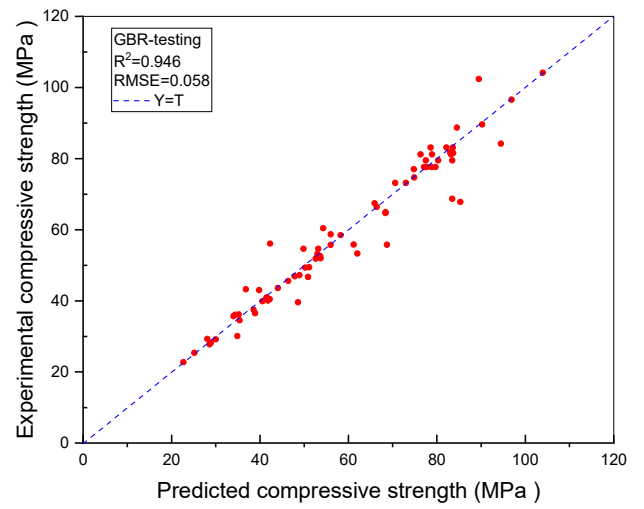
In this section, the prediction ability indicators of the aforementioned ML models are demonstrated. The relationship between the experimental and predicted properties (CS and EC) of cement replacement concrete for the training and testing datasets is demonstrated in Figures 1 and 2, respectively. As shown in Figures 1 and 2, the horizontal axis stands for the CS and EC generated by the ML models, respectively, while the vertical axis indicates the CS and EC of the experimental datasets collected from the open literature. Furthermore, the differences in the R^2 and the RMSE values of the ML models for predicting the CS and EC are shown in Table 3. Moreover, the tuned parameters of the ML models utilizing the GSA are exhibited in Table 4. The optimal ML model is defined when the model performs the highest R^2 and the lowest RMSE.

Table 3. The R^2 and RMSE values of the six types of ML models for predicting CS and EC.

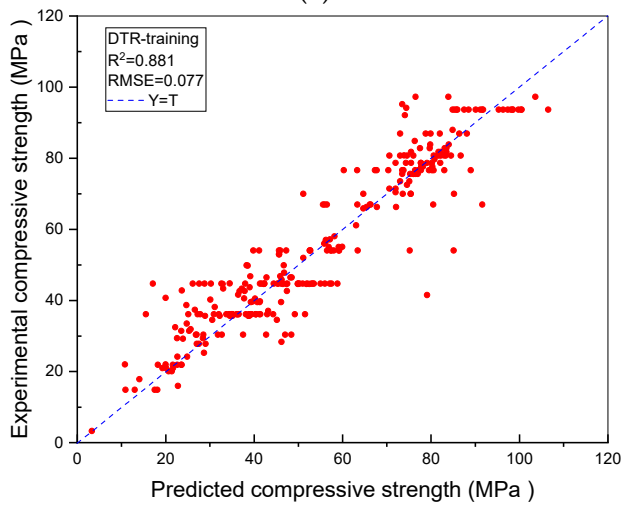
Algorithm	Dataset	CS Prediction Performance		EC Prediction Performance	
		R^2	RMSE	R^2	RMSE
GBR	Training	0.985	0.028	0.999	0.002
	Testing	0.946	0.058	0.999	0.012
DTR	Training	0.881	0.077	0.999	0.010
	Testing	0.876	0.093	0.998	0.015
DNN	Training	0.927	0.062	0.996	0.013
	Testing	0.892	0.077	0.995	0.015
SVR	Training	0.953	0.049	0.999	0.005
	Testing	0.924	0.057	0.985	0.026
RF	Training	0.977	0.030	0.998	0.010
	Testing	0.933	0.062	0.997	0.012
kNN	Training	0.924	0.063	0.969	0.038
	Testing	0.888	0.082	0.965	0.039



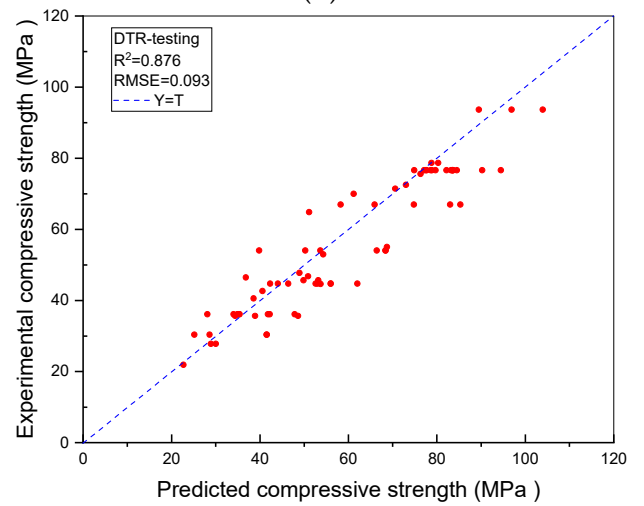
(a)



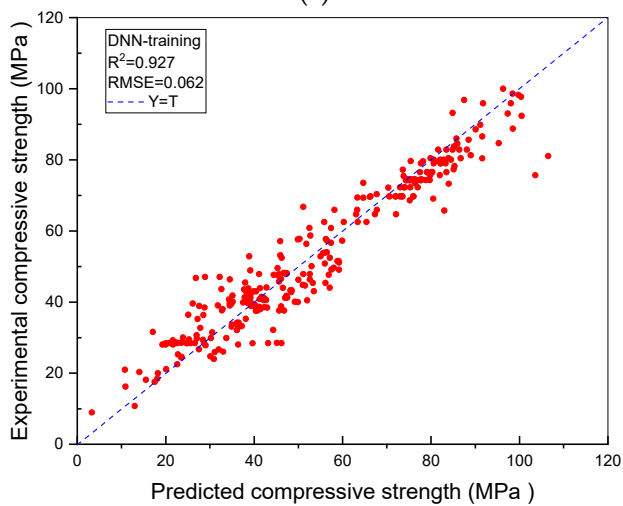
(b)



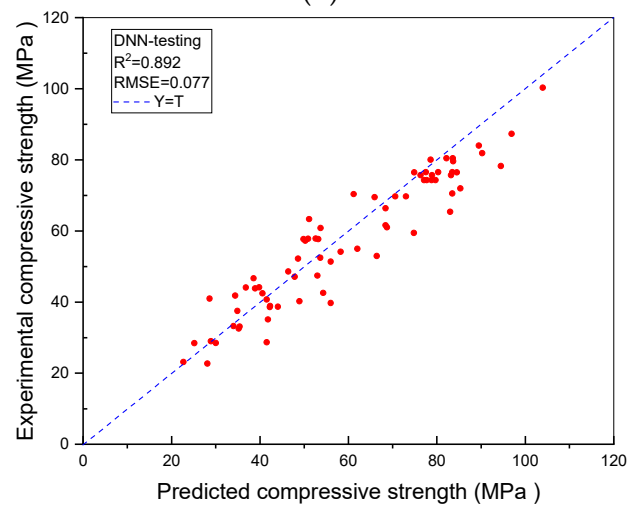
(c)



(d)



(e)



(f)

Figure 1. Cont.

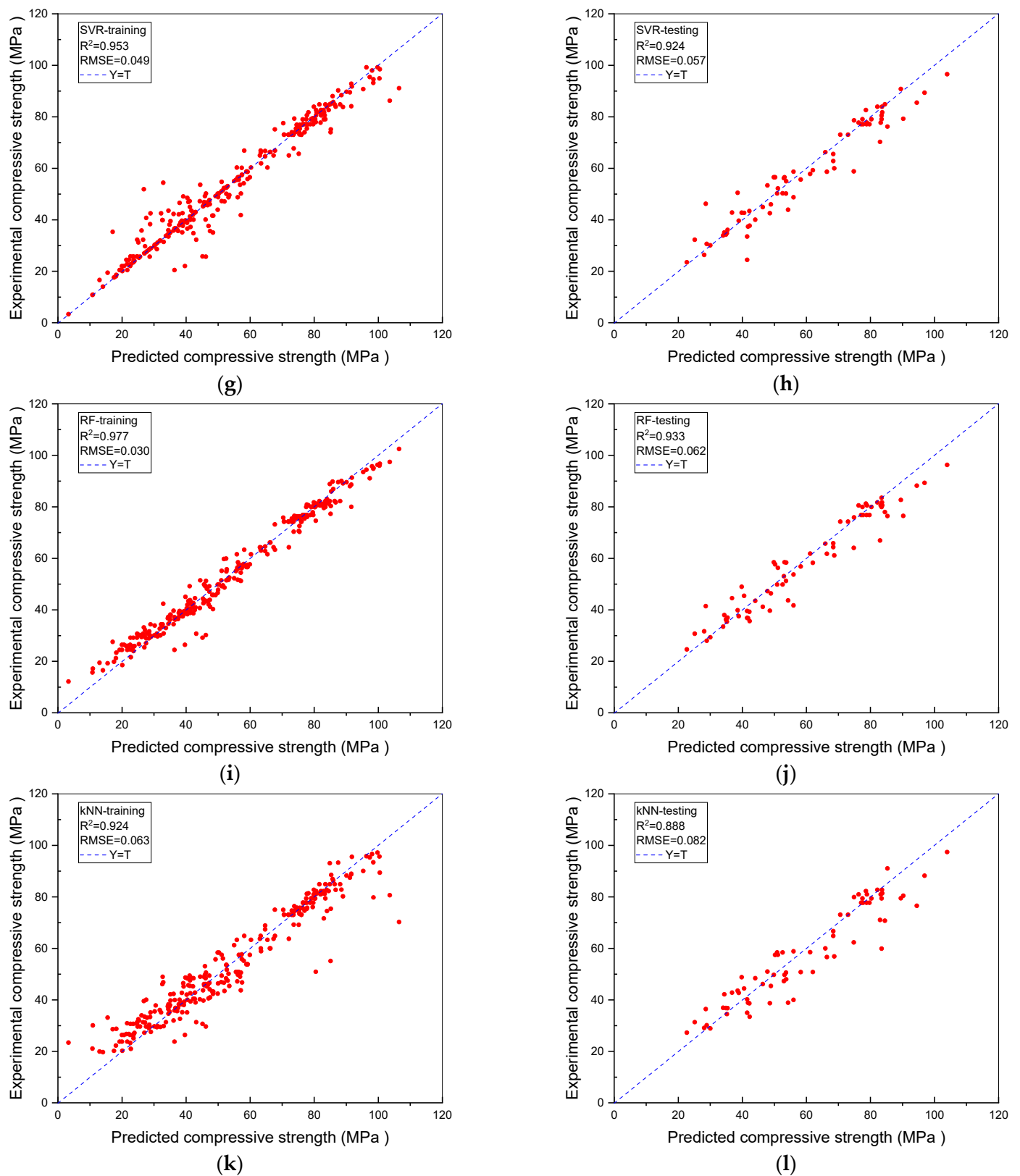


Figure 1. Predicted vs experimental CS values for the ML models: (a) GBR-training; (b) GBR-testing; (c) DTR-training; (d) DTR-testing; (e) DNN-training; (f) DNN-testing; (g) SVR-training; (h) SVR-testing; (i) RF-training; (j) RF-testing; (k) kNN-training; and (l) kNN-testing, with the corresponding R^2 and RMSE values.

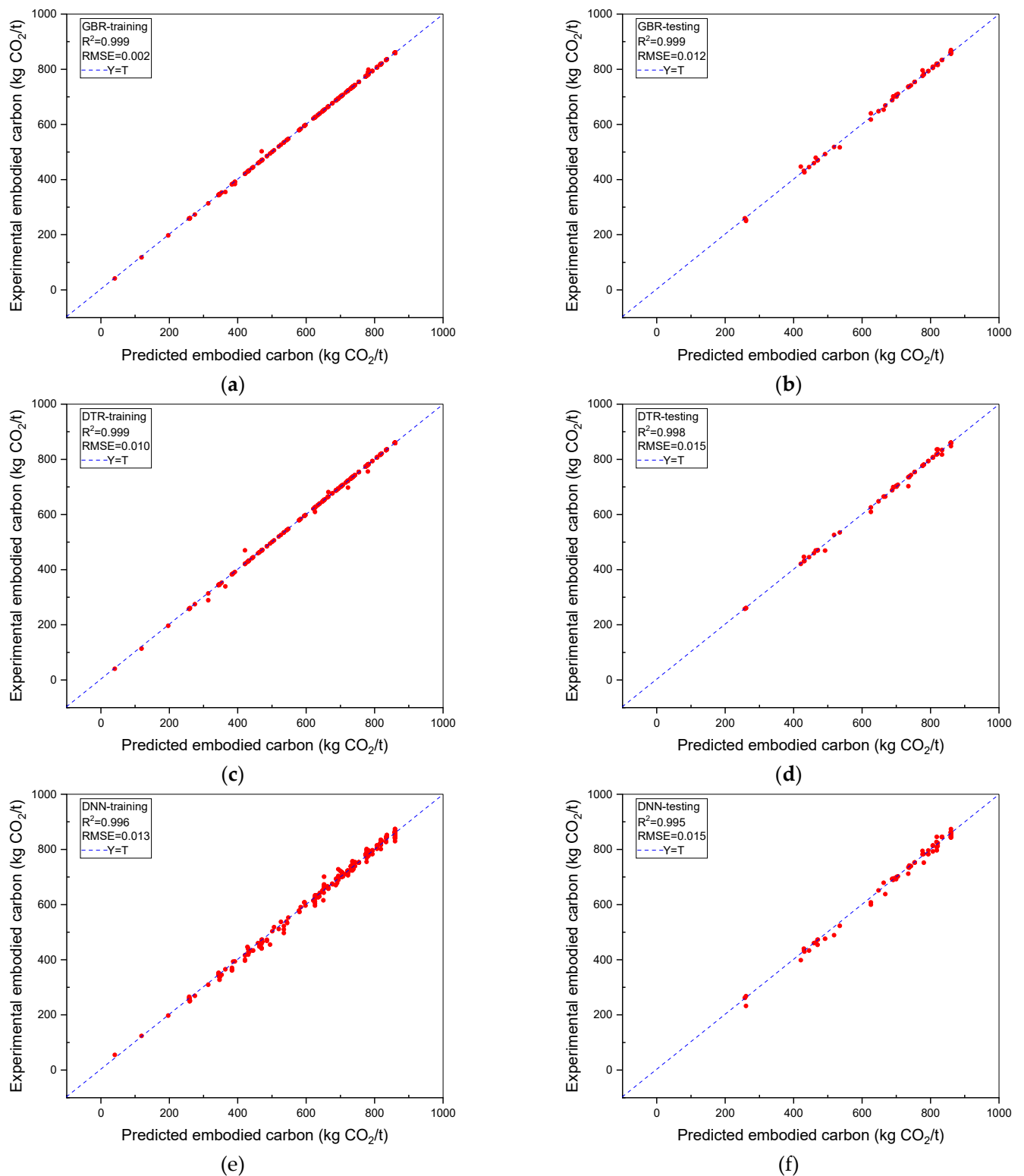


Figure 2. Cont.

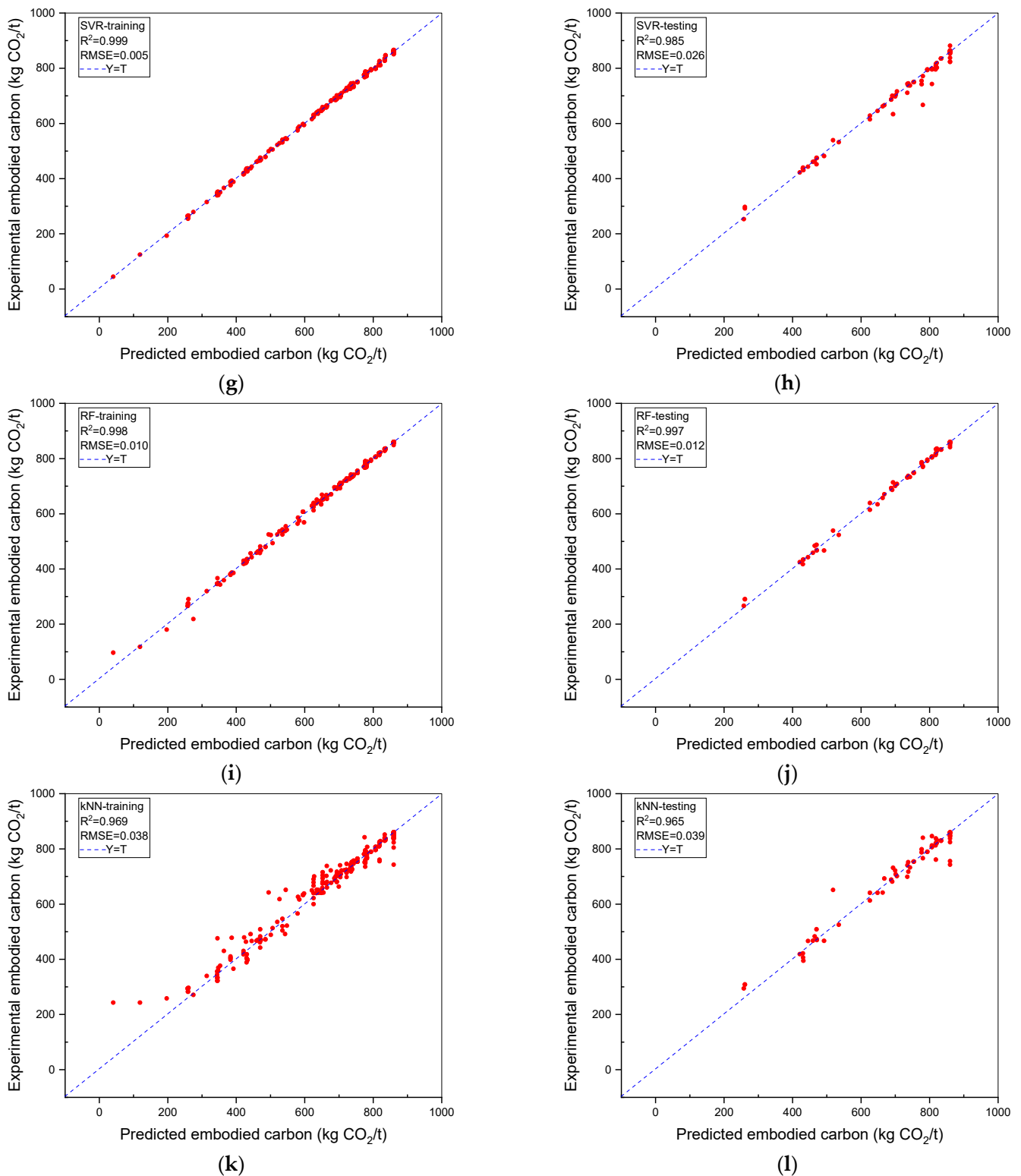


Figure 2. Predicted vs experimental EC values for the ML models: (a) GBR-training; (b) GBR-testing; (c) DTR-training; (d) DTR-testing; (e) DNN-training; (f) DNN-testing; (g) SVR-training; (h) SVR-testing; (i) RF-training; (j) RF-testing; (k) kNN-training; and (l) kNN-testing, with the corresponding R² and RMSE values.

Table 4. Tuned parameters of the ML models using the GSA.

Algorithms	Parameters	CS	EC
DNN	Hidden layers	3	3
	Hidden neurons	27–27–30	30–28–28
	Learning rate	0.1009	0.1000
	Activation function	Maxout	Maxout
GBR	Depth _{max}	11	11
	Split _{min}	0.001	0.001
	Learning rate	0.4	0.1
	Number of trees	21	100
DTR	Depth _{max}	5	21
	Split _{min}	1	1
	Leaf _{min}	1	1
	Gain _{min}	0.001	0.001
SVR	C _{penalty}	0.01	1.00
	Epsilon	0.001	0.001
	Gamma	6000.0004	10,000.0000
	Kernel type	Anova	Anova
RF	Depth _{max}	31	51
	Split _{min}	0.001	0.100
	Leaf _{min}	80	80
	Gain _{min}	1.0000	0.5005
	Number of trees	100	11
kNN	k	6	7

With regard to the ML-aided prediction of the CS, what can be clearly seen in Figure 1a,b is that the GBR model shows the highest R^2 (0.946) and the lowest RMSE (0.058) among the ML algorithms. Furthermore, the R^2 and RMSE values of the SVR and RF models are (0.924, 0.057) and (0.933, 0.062), respectively, which are similar to those of the GBR model (Figure 1g–j). As shown in Figure 1c–f,k,l, the R^2 and RMSE values of the DTR, DNN, and kNN are (0.876, 0.093), (0.892, 0.077), and (0.888, 0.082), respectively, representing a slightly low prediction performance and a poorer prediction accuracy than other ML models.

With regard to the prediction of the EC, the GBR model shows the best prediction performance ($R^2 = 0.999$) and prediction accuracy (RMSE = 0.012) among the ML models, as shown in Figure 2a,b. As demonstrated in Figure 2c–f,i,j, the R^2 and RMSE values of the DTR, DNN, and RF are (0.998, 0.015), (0.995, 0.015), and (0.997, 0.012), respectively. It can be interpreted that the DTR, RF, and DNN models have similar abilities for predicting the EC of cement replacement concrete. Moreover, the prediction ability of the SVR and kNN models is slightly lower than other ML models. The R^2 and RMSE values of the SVR and kNN models are (0.985, 0.026) and (0.965, 0.039), respectively (Figure 2g,h,k,l).

4. Discussion

The relationship between the predicted and experimental CSs and EC of cement replacement concrete employing the GBR models is illustrated in Figures 1 and 2. As demonstrated, the GBR models for predicting the CS and EC of cement replacement concrete perform better in terms of prediction ability than other types of ML models. In other words, the relationship between the outputs of the cement replacement concrete and the 12 variables can be precisely explained employing the GBR models. The better prediction ability of the GBR models can be attributed to the fact that the GBR model is a type of ensemble learning algorithm that has a remarkable generalisation capacity because of the employed boosting strategy. Weak learners can be generated by the boosting strategy. Higher weights will be distributed to the weak learners with promising prediction ability, while the weak learners with poor prediction ability will obtain lower weights. The robust

prediction ability of the GBR models will be generated by a strong learner made up of the weak learners. However, DNN, SVR, kNN, and DTR are individual ML algorithms, with relatively lower generalisation capacities than ensemble ML algorithms.

4.1. K-Fold Cross Validation

K-fold cross validation is employed in this paper to further investigate the prediction ability of the optimal ML models for predicting the CS and EC of cement replacement concrete. Moreover, the reliability of the optimal ML models are reported using K-fold cross validation [56]. Employing K-fold cross validation reduces the variance from the training and testing dataset selection. In this paper, 10-fold cross validation is utilised to demonstrate the prediction ability of the GBR models [57]. Firstly, the datasets are equally split into ten groups. Secondly, nine groups are used to train the GBR models, while the rest of the datasets are employed to conduct the validation of the GBR models. Then, the second step is repeated ten times. Finally, the prediction ability of the GBR models is generated by averaging the R^2 and RMSE values of the 10-fold cross validation [58].

Figure 3a–d demonstrate the R^2 and RMSE results of each fold for predicting the CS and EC in the 10-fold cross validation. It can be observed from Figure 3a that the R^2 values of the 10-fold cross validation for predicting the CS demonstrate slight fluctuations. For instance, the minimum R^2 is 0.939 at Fold 1, while the maximum R^2 is 0.951 at Fold 5. Moreover, the RMSE value exhibited in Figure 3b slightly decreases, from 0.246 to 0.223 between Folds 1 and 6. It then keeps constant at around 0.221, until Fold 10. As shown in Figure 3c, the R^2 values of the 10-fold cross validation for predicting the EC are maintained at approximately 0.997, while the RMSE values fluctuate between 0.012 and 0.014 from Fold 1 to Fold 10, demonstrated in Figure 3d. Furthermore, several statistical results, the average R^2 and RMSE values of the CS and EC predictions, are listed in Table 5. The average R^2 and RMSE values of the CS prediction are 0.9471 and 0.2270, respectively. Moreover, the standard deviations (SDs) of the R^2 and the RMSE for predicting the CS are 0.0037 and 0.0087, respectively, which means that the coefficient of variations (COVs) of the R^2 and RMSE for predicting the CS are only 0.4% and 3.8%, respectively. In addition, the average R^2 and RMSE values of the 10-fold cross validation for the EC prediction are 0.9967 and 0.0125, respectively, while the SDs are 0.0013 and 0.0007, indicating that the COVs of the R^2 and RMSE are 0.1% and 5.6%, respectively. On the basis of the average R^2 , the RMSE, and the COVs of the 10-fold cross validation for CS and EC prediction, it can be concluded that the prediction error of the GBR models is small; in other words, the excellent prediction ability of the GBR models is reliable.

Table 5. 10-fold cross-validation and statistical results.

Folds	CS Prediction Performance		EC Prediction Performance	
	R^2	RMSE	R^2	RMSE
Fold 1	0.939	0.246	0.995	0.014
Fold 2	0.942	0.240	0.995	0.013
Fold 3	0.946	0.231	0.997	0.012
Fold 4	0.949	0.225	0.997	0.012
Fold 5	0.951	0.223	0.997	0.012
Fold 6	0.950	0.223	0.997	0.012
Fold 7	0.949	0.222	0.997	0.012
Fold 8	0.949	0.221	0.995	0.013
Fold 9	0.946	0.220	0.999	0.012
Fold 10	0.950	0.219	0.998	0.013
Average	0.9471	0.2270	0.9967	0.0125
SD	0.0037	0.0087	0.0013	0.0007

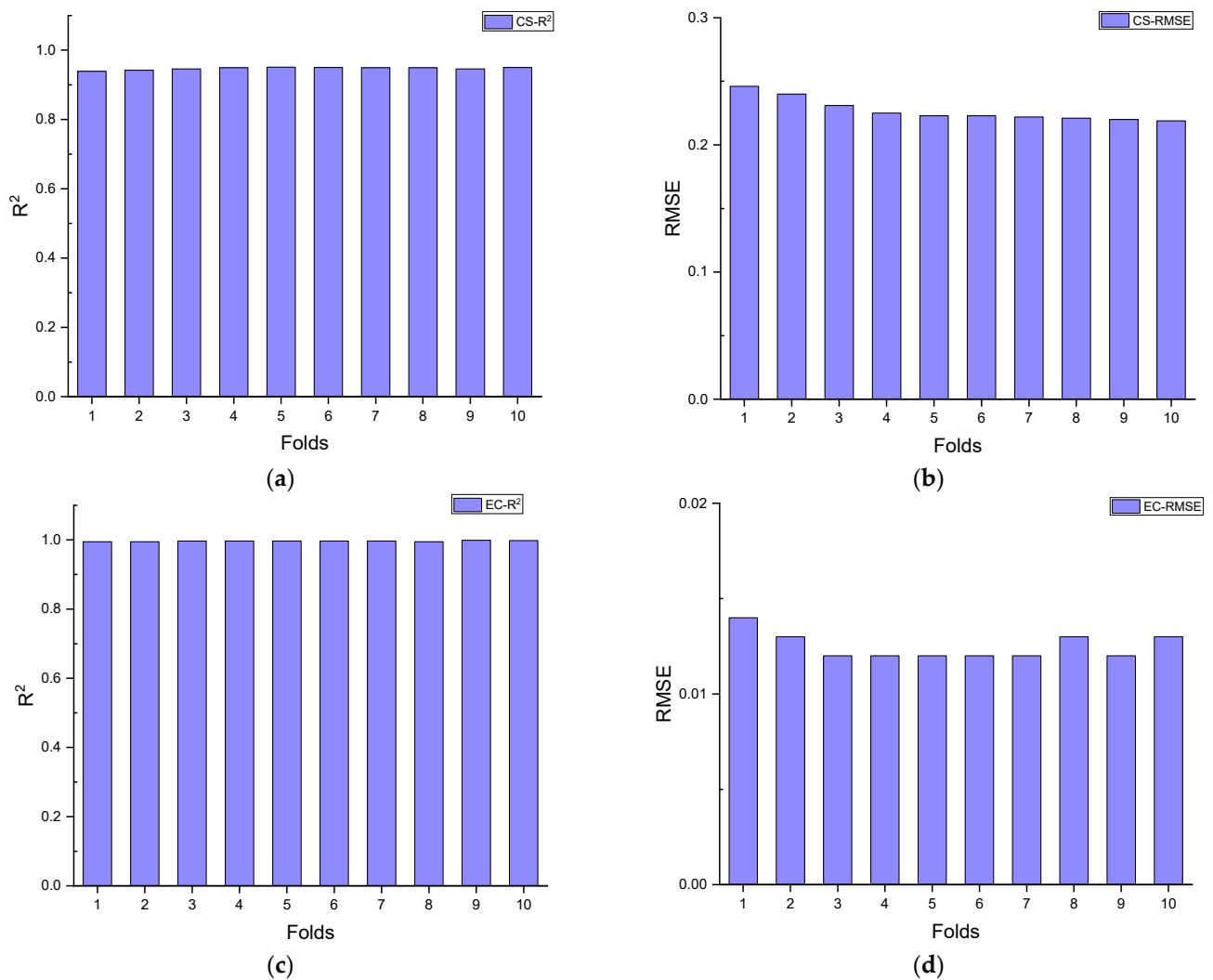


Figure 3. 10-fold cross validation results: (a) R² for predicting CS; (b) RMSE for predicting CS; (c) R² for predicting EC; (d) RMSE for predicting EC.

4.2. Sensitivity Analysis

Sensitivity analysis is a method used to evaluate how the changes in the outputs of ML models can be affected by their inputs [59]. In this paper, GBR ML models were selected to conduct the SA because they have the best performances for predicting the CS and EC of cement replacement concrete. In order to investigate the sensitivity of the chosen ML models, one type of cement replacement material is perturbed at a time, while the other five types of cement replacement materials are kept constant at their mean values. Then, the new datasets are introduced to the GBR models to predict the CS and EC of the cement replacement concrete [60]. After that, the corresponding sensitivity analysis parameter (SAP) of each input can be calculated using Equation (5):

$$SA_i = \frac{P_{max}(I_i) - P_{min}(I_i)}{\sum_i P_{max}(I_i) - P_{min}(I_i)} \times 100 \quad (5)$$

where $P_{max}(I_i)$ and $P_{min}(I_i)$ are the maximum and minimum predicted CS and EC values of the cement replacement concrete corresponding to the input, I_i . SA_i is the SAP of the input I_i .

Figure 4 represents the SAPs of the inputs, from which a pronounced influence, 19.96% and 30.44% of the PFA on the predicted CS and EC, respectively, can be observed. With regard to the SAPs of the inputs for predicting CS, similar SAPs of 15.25, 14.07, 12.37,

13.99, and 15.32% for GGBS, SF, LP, PP, and GP, respectively, are investigated. Furthermore, Figure 4 demonstrates the high SAPs of the inputs for predicting the EC: 24.05, 16.61, and 21.77% for GGBS, LP, and SF, respectively. Additionally, the SAPs for M, PP, and GP are lower than for GGBS, LP, and SF, which are 4.25, 1.58, and 1.30%, respectively. The SAP results of the inputs for predicting the CS and EC of cement replacement concrete indicate that PFA plays the most significant role in the CS and EC of cement replacement concrete. As such, GGBS, LP, and SF should be thoroughly investigated in predicting the CS and EC, employing the GBR models. According to these findings, PFA, GGBS, LP, and SF need to be carefully controlled in cement replacement concrete design because of their prominent effect on the CS and the EC.

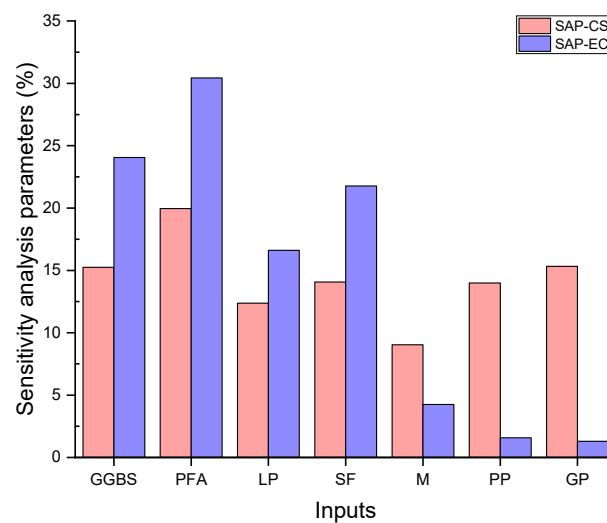


Figure 4. Sensitivity analysis parameter results of the inputs for the predicted CS and EC of cement replacement concrete.

5. Conclusions

This paper set out to predict the CS and EC of cement replacement concrete employing the ML models, with six kinds of algorithms, to aid in cement replacement concrete design. Thus, cement replacement concrete with the expected CS and the lowest EC can be designed by using the ML models produced in this paper. The ML models are employed to explain the relationship between 12 inputs and 2 outputs, the CS and the EC. Meanwhile, the GSA hyperparameter tuning method is utilised to optimise the parameters of the ML models. A 10-fold cross validation is employed to investigate the prediction ability of the optimal ML models. Finally, several key inputs are observed by applying SA.

With regard to the R^2 and RMSE values of the ML models, the prominent findings to emerge from this paper can be concluded as follows:

- This paper shows that the GBR ML models have the best ability to predict the CS and EC of concrete containing cement replacement materials, as indicated by the R^2 and the RMSE values of the CS prediction (0.946, 0.058), and of the EC prediction (0.999, 0.012). On the basis of the R^2 and the RMSE values, it can be stated that the GBR ML models have an excellent ability for predicting the CS and EC of cement replacement concrete using the 12 inputs;
- The average R^2 and RMSE values of the 10-fold cross validation for predicting CS are 0.9471 and 0.2270, respectively. Moreover, the average R^2 of the 10-fold cross validation for predicting the EC is 0.9967, while the RMSE value is 0.0125. The 10-fold cross-validation results indicate that the prediction error of the GBR models is very low. Hence, the promising prediction ability of the GBR models is robust;
- The R^2 and the RMSE values of the other five ML models (SVR, RF, DNN, kNN, and DTR) are compared with the GBR model. The results reveal that the GBR model, as

an ensemble ML algorithm, exhibits an outstanding superiority to other individual ML algorithms;

- The SAP results of the inputs note that PFA, GGBS, LP, and SF have stronger correlations to the CS and EC predictions of cement replacement concrete than other inputs. Thus, more attention should be paid to PFA, GGBS, LP, and SF in the ML-aided design of cement replacement concrete in order to reduce the EC.

Author Contributions: Conceptualization, A.L., X.H. and S.K.; data curation, A.L., X.H. and S.K.; formal analysis, A.L., X.H. and S.K.; funding acquisition, S.K.; investigation, A.L., X.H. and S.K.; methodology, A.L., X.H. and S.K.; project administration, S.K.; resources, S.K.; software, X.H. and S.K.; supervision, S.K.; validation, A.L. and X.H.; visualization, X.H.; writing—original draft, A.L. and X.H.; writing—review and editing, X.H. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by H2020-MSCA-RISE, grant number, 691135, and Shift2Rail, H2020-S2R Project No. 730849. The APC was funded by the MDPI's Invited Paper Initiative.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be made available upon reasonable request.

Acknowledgments: The authors are sincerely grateful to European Commission for the financial sponsorship of the H2020-RISE Project No. 691135, "RISEN: Rail Infrastructure Systems Engineering Network", which enables a global research network that tackles the grand challenge in railway infrastructure resilience and advanced sensing in extreme environments (www.risen2rail.eu, accessed on 30 October 2021) [61]. In addition, this project is partially supported by the European Commission's Shift2Rail, H2020-S2R Project No. 730849, "S-Code: Switch and Crossing Optimal Design and Evaluation".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AdBoost	Adaptive Boosting
ANN	Artificial Neural Network
CS	Compressive Strength
DNN	Deep Neural Network
DTR	Decision Tree Regression
EC	Embodied Carbon
EP	Expanded Perlite
GBR	Gradient Boosting Regression
GSA	Grid Search Algorithm
GGBS	Ground Granulated Blast-Furnace Slag
GPC	Geopolymer Concrete
kNN	k-Nearest Neighbours
LR	Linear Regression Model
ML	Machine Learning
M5P	M5P-tree model
MSE	Mean Square Error
NLR	Nonlinear Regression Model
OPC	Ordinary Portland Cement
PFA	Pulverised fuel ash
RF	Random Forest
R2	Coefficient of Determination
RMSE	Root Mean Square Error

RBF	Radial Basis Function
SVR	Support Vector Regression
SMC	Supplementary Cementitious Materials
SF	Silica Fume
SA	Sensitivity Analysis

References

- Suhendro, B. Toward Green Concrete for Better Sustainable Environment. *Procedia Eng.* **2014**, *95*, 305–320. [[CrossRef](#)]
- Gehlot, T.; Sankhla, S.S.; Parihar, S. Mathematical model to predict the chloride ion penetration and compressive strength of concrete incorporating GGBS. *Mater. Today* **2021**, in press. [[CrossRef](#)]
- Liu, M. Self-compacting concrete with different levels of pulverized fuel ash. *Constr. Build. Mater.* **2010**, *24*, 1245–1252. [[CrossRef](#)]
- Zhou, X.; Slater, J.R.; Wavell, S.E.; Oladiran, O. Effects of PFA and GGBS on Early-Ages Engineering Properties of Portland Cement Systems. *J. Adv. Concr. Technol.* **2012**, *10*, 74–85. [[CrossRef](#)]
- Lim, J.S.; Cheah, C.B.; Ramli, M.B. The setting behavior, mechanical properties and drying shrinkage of ternary blended concrete containing granite quarry dust and processed steel slag aggregate. *Constr. Build. Mater.* **2019**, *215*, 447–461. [[CrossRef](#)]
- Chen, J.; Kwan, A.; Jiang, Y. Adding limestone fines as cement paste replacement to reduce water permeability and sorptivity of concrete. *Constr. Build. Mater.* **2014**, *56*, 87–93. [[CrossRef](#)]
- Sarıdemir, M. Prediction of compressive strength of concretes containing metakaolin and silica fume by artificial neural networks. *Adv. Eng. Softw.* **2009**, *40*, 350–355. [[CrossRef](#)]
- Moradi, M.; Khaleghi, M.; Salimi, J.; Farhangi, V.; Ramezani-pour, A. Predicting the compressive strength of concrete containing metakaolin with different properties using ANN. *Measurement* **2021**, *183*, 109790. [[CrossRef](#)]
- Bageri, B.; Ahmed, A.; Al Jaberi, J.; Elkhatny, S.; Patil, S. Effect of perlite particles on the properties of oil-well class G cement. *J. Pet. Sci. Eng.* **2021**, *199*, 108344. [[CrossRef](#)]
- Kotwica, Ł.; Pichór, W.; Kapeluszna, E.; Różycka, A. Utilization of waste expanded perlite as new effective supplementary cementitious material. *J. Clean. Prod.* **2017**, *140*, 1344–1352. [[CrossRef](#)]
- Wang, X.; Wu, D.; Geng, Q.; Hou, D.; Wang, M.; Li, L.; Wang, P.; Chen, D.; Sun, Z. Characterization of sustainable ultra-high performance concrete (UHPC) including expanded perlite. *Constr. Build. Mater.* **2021**, *303*, 124245. [[CrossRef](#)]
- Mehrabi, P.; Shariati, M.; Kabirifar, K.; Jarrah, M.; Rasekh, H.; Trung, N.T.; Shariati, A.; Jahandari, S. Effect of pumice powder and nano-clay on the strength and permeability of fiber-reinforced pervious concrete incorporating recycled concrete aggregate. *Constr. Build. Mater.* **2021**, *287*, 122652. [[CrossRef](#)]
- Mohammed, A.; Rafiq, S.; Sihag, P.; Kurda, R.; Mahmood, W.; Ghafor, K.; Sarwar, W. ANN, M5P-tree and nonlinear regression approaches with statistical evaluations to predict the compressive strength of cement-based mortar modified with fly ash. *J. Mater. Res. Technol.* **2020**, *9*, 12416–12427. [[CrossRef](#)]
- Khan, M.I. Predicting properties of High Performance Concrete containing composite cementitious materials using Artificial Neural Networks. *Autom. Constr.* **2012**, *22*, 516–524. [[CrossRef](#)]
- Shahmansouri, A.A.; Yazdani, M.; Ghanbari, S.; Bengar, H.A.; Jafari, A.; Ghatte, H.F. Artificial neural network model to predict the compressive strength of eco-friendly geopolymer concrete incorporating silica fume and natural zeolite. *J. Clean. Prod.* **2021**, *279*, 123697. [[CrossRef](#)]
- Kandiri, A.; Mohammadi Golafshani, E.; Behnood, A. Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm. *Constr. Build. Mater.* **2020**, *248*, 118676. [[CrossRef](#)]
- Kaloop, M.R.; Kumar, D.; Samui, P.; Hu, J.W.; Kim, D. Compressive strength prediction of high-performance concrete using gradient tree boosting machine. *Constr. Build. Mater.* **2020**, *264*, 120198. [[CrossRef](#)]
- Selvarani, B.; Preethi, V. Investigational study on optimum content of GGBS and fibres in fibre non-breakable self compacting concrete. *Mater. Today* **2021**, in press. [[CrossRef](#)]
- Jaambavi, I.; Baby, J.E.; Akshaya, T.; Rajeswari, G. Mechanical properties of concrete by replacement of fly ash, Ggbfs, lime stone powder. *Mater. Today* **2021**, in press. [[CrossRef](#)]
- Gehlot, T.; Sankhla, S.S.; Parihar, S. Modelling compressive strength, flexural strength and chloride ion permeability of high strength concrete incorporating metakaolin and fly ash. *Mater. Today* **2021**, in press. [[CrossRef](#)]
- Elavarasan, S.; Priya, A.; Ajai, N.; Akash, S.; Annie, T.; Bhuvana, G. Experimental study on partial replacement of cement by metakaolin and GGBS. *Mater. Today* **2021**, *37*, 3527–3530. [[CrossRef](#)]
- Munjal, P.; Hau, K.K.; Arthur, C.C.H. Effect of GGBS and curing conditions on strength and microstructure properties of oil well cement slurry. *J. Build. Eng.* **2021**, *40*, 102331. [[CrossRef](#)]
- Suda, V.R.; Rao, P.S. Experimental investigation on optimum usage of Micro silica and GGBS for the strength characteristics of concrete. *Mater. Today* **2020**, *27*, 805–811. [[CrossRef](#)]
- Gupta, T.; Patel, K.; Siddique, S.; Sharma, R.K.; Chaudhary, S. Prediction of mechanical properties of rubberised concrete exposed to elevated temperature using ANN. *Measurement* **2019**, *147*, 106870. [[CrossRef](#)]
- Esfandiari, J.; Loghmani, P. Effect of perlite powder and silica fume on the compressive strength and microstructural characterization of self-compacting concrete with lime-cement binder. *Measurement* **2019**, *147*, 106846. [[CrossRef](#)]

26. Xie, J.; Huang, L.; Guo, Y.; Li, Z.; Fang, C.; Li, L.; Wang, J. Experimental study on the compressive and flexural behaviour of recycled aggregate concrete modified with silica fume and fibres. *Constr. Build. Mater.* **2018**, *178*, 612–623. [CrossRef]
27. Saridemir, M. Effect of silica fume and ground pumice on compressive strength and modulus of elasticity of high strength concrete. *Constr. Build. Mater.* **2013**, *49*, 484–489. [CrossRef]
28. Nili, M.; Afroughsabet, V. The long-term compressive strength and durability properties of silica fume fiber-reinforced concrete. *Mater. Sci. Eng. A* **2012**, *531*, 107–111. [CrossRef]
29. Wongkeo, W.; Thongsanitgarn, P.; Ngamjarurojana, A.; Chaipanich, A. Compressive strength and chloride resistance of self-compacting concrete containing high level fly ash and silica fume. *Mater. Des.* **2014**, *64*, 261–269. [CrossRef]
30. Sharaky, I.; Ghoneim, S.S.; Aziz, B.H.A.; Emara, M. Experimental and theoretical study on the compressive strength of the high strength concrete incorporating steel fiber and metakaolin. *Structures* **2021**, *31*, 57–67. [CrossRef]
31. Seelapureddy, J.; Bommisetty, J.; Rao, M.S. Effect of metakaolin and micro silica on strength characteristics of standard grades of self-compacting concrete. *Mater. Today* **2021**, *45*, 884–890. [CrossRef]
32. BSI. BS EN 15978: 2011. *Sustainability of Construction Works—Assessment of Environmental Performance of Buildings—Calculation Method*; BSI: London, UK, 2011.
33. Government Conversion Factors for Company Reporting of Greenhouse Gas Emissions. Available online: <https://www.gov.uk/government/collections/government-conversion-factors-for-company-reporting> (accessed on 30 October 2021).
34. Purnell, P.; Black, L. Embodied carbon dioxide in concrete: Variation with common mix design parameters. *Cem. Concr. Res.* **2012**, *42*, 874–877. [CrossRef]
35. Environmental Product Declaration of Ggbs. Available online: https://www.hanson.co.uk/en/system/files_force/assets/document/53/9b/environmental-product-declaration-hansonconcrete-c28-35-cemi.pdf?download=1 (accessed on 30 October 2021).
36. Bolte, G.; Zajac, M.; Skocek, J.; Ben Haha, M. Development of composite cements characterized by low environmental footprint. *J. Clean. Prod.* **2019**, *226*, 503–514. [CrossRef]
37. Homayoonmehr, R.; Ramezaniapour, A.A.; Mirdarsoltany, M. Influence of metakaolin on fresh properties, mechanical properties and corrosion resistance of concrete and its sustainability issues: A review. *J. Build. Eng.* **2021**, *44*, 103011. [CrossRef]
38. Hammond, G.; Jones, C.; Lowrie, E.F.; Tse, P. Embodied Carbon. The Inventory of Carbon and Energy (ICE). Version (2.0) 2011. Available online: <https://greenbuildingencyclopaedia.uk/wp-content/uploads/2014/07/Full-BSRIA-ICE-guide.pdf> (accessed on 30 October 2021).
39. Zhang, J.; Zhao, C.; Zhou, A.; Yang, C.; Zhao, L.; Li, Z. Aragonite formation induced by open cultures of microbial consortia to heal cracks in concrete: Insights into healing mechanisms and crystal polymorphs. *Constr. Build. Mater.* **2019**, *224*, 815–822. [CrossRef]
40. Chopra, P.; Sharma, R.K.; Kumar, M.; Chopra, T. Comparison of Machine Learning Techniques for the Prediction of Compressive Strength of Concrete. *Adv. Civ. Eng.* **2018**, *2018*, 5481705. [CrossRef]
41. Naderpour, H.; Mirrashid, M.; Parsa, P. Failure mode prediction of reinforced concrete columns using machine learning methods. *Eng. Struct.* **2021**, *248*, 113263. [CrossRef]
42. Chun, P.-J.; Ujike, I.; Mishima, K.; Kusumoto, M.; Okazaki, S. Random forest-based evaluation technique for internal damage in reinforced concrete featuring multiple nondestructive testing results. *Constr. Build. Mater.* **2020**, *253*, 119238. [CrossRef]
43. Zhang, J.; Ma, G.; Huang, Y.; Sun, J.; Aslani, F.; Nener, B. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. *Constr. Build. Mater.* **2019**, *210*, 713–719. [CrossRef]
44. Sun, J.; Zhang, J.; Gu, Y.; Huang, Y.; Sun, Y.; Ma, G. Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression. *Constr. Build. Mater.* **2019**, *207*, 440–449. [CrossRef]
45. Nguyen, T.D.; Tran, T.-H.; Nguyen, H.; Nhat-Duc, H. A success history-based adaptive differential evolution optimized support vector regression for estimating plastic viscosity of fresh concrete. *Eng. Comput.* **2021**, *37*, 1485–1498. [CrossRef]
46. Altay, O.; Ulas, M.; Alyamac, K.E. Prediction of the Fresh Performance of Steel Fiber Reinforced Self-Compacting Concrete Using Quadratic SVM and Weighted KNN Models. *IEEE Access* **2020**, *8*, 92647–92658. [CrossRef]
47. Kang, M.-C.; Yoo, D.-Y.; Gupta, R. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. *Constr. Build. Mater.* **2021**, *266*, 121117. [CrossRef]
48. Zhang, J.; Huang, Y.; Aslani, F.; Ma, G.; Nener, B. A hybrid intelligent system for designing optimal proportions of recycled aggregate concrete. *J. Clean. Prod.* **2020**, *273*, 122922. [CrossRef]
49. Tanyildizi, H.; Şengür, A.; Akbulut, Y.; Şahin, M. Deep learning model for estimating the mechanical properties of concrete containing silica fume exposed to high temperatures. *Front. Struct. Civ. Eng.* **2020**, *14*, 1316–1330. [CrossRef]
50. Ly, H.-B.; Nguyen, T.-A.; Mai, H.-V.T.; Tran, V.Q. Development of deep neural network model to predict the compressive strength of rubber concrete. *Constr. Build. Mater.* **2021**, *301*, 124081. [CrossRef]
51. Mezzatesta, S.; Torino, C.; De Meo, P.; Fiumara, G.; Vilasi, A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Comput. Methods Programs Biomed.* **2019**, *177*, 9–15. [CrossRef]
52. Feng, D.-C.; Liu, Z.-T.; Wang, X.-D.; Chen, Y.; Chang, J.-Q.; Wei, D.-F.; Jiang, Z.-M. Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Constr. Build. Mater.* **2020**, *230*, 117000. [CrossRef]
53. Xiao, T.; Ren, D.; Lei, S.; Zhang, J.; Liu, X. Based on grid-search and PSO parameter optimization for Support Vector Machine. In Proceedings of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 1529–1533.
54. Bilim, C.; Atis, C.; Tanyildizi, H.; Karahan, O. Predicting the compressive strength of ground granulated blast furnace slag concrete using artificial neural network. *Adv. Eng. Softw.* **2009**, *40*, 334–340. [CrossRef]

55. Atici, U. Prediction of the strength of mineral admixture concrete using multivariable regression analysis and an artificial neural network. *Expert Syst. Appl.* **2011**, *38*, 9609–9618. [[CrossRef](#)]
56. Ling, H.; Qian, C.X.; Kang, W.C.; Liang, C.Y.; Chen, H.C. Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment. *Constr. Build. Mater.* **2019**, *206*, 355–363. [[CrossRef](#)]
57. Xu, Y.; Ahmad, W.; Ahmad, A.; Ostrowski, K.A.; Dudek, M.; Aslam, F.; Joyklad, P. Computation of High-Performance Concrete Compressive Strength Using Standalone and Ensembled Machine Learning Techniques. *Materials* **2021**, *14*, 7034. [[CrossRef](#)] [[PubMed](#)]
58. Song, H.; Ahmad, A.; Farooq, F.; Ostrowski, K.A.; Maślak, M.; Czarnecki, S.; Aslam, F. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. *Constr. Build. Mater.* **2021**, *308*, 125021. [[CrossRef](#)]
59. Huynh, A.T.; Nguyen, Q.D.; Xuan, Q.L.; Magee, B.; Chung, T.; Tran, K.T.; Nguyen, K.T. A Machine Learning-Assisted Numerical Predictor for Compressive Strength of Geopolymer Concrete Based on Experimental Data and Sensitivity Analysis. *Appl. Sci.* **2020**, *10*, 7726. [[CrossRef](#)]
60. Tenza-Abril, A.; Villacampa, Y.; Solak, A.; Baeza-Brotons, F. Prediction and sensitivity analysis of compressive strength in segregated lightweight concrete based on artificial neural network using ultrasonic pulse velocity. *Constr. Build. Mater.* **2018**, *189*, 1173–1183. [[CrossRef](#)]
61. Kaewunruen, S.; Sussman, J.M.; Matsumoto, A. Grand Challenges in Transportation and Transit Systems. *Front. Built Environ.* **2016**, *2*, 2. [[CrossRef](#)]