Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

Predicting rubberized concrete compressive strength using machine learning: A feature importance and partial dependence analysis

Mahdi Hasanipanah[1], Rini Asnida Abdullah[1], Mudassir Iqbal[2,3], Hai-Bang Ly[4*]
[1]Universiti Teknologi Malaysia, Johor Bahru, Malaysia.
[2]Jiao Tong University, Shanghai 200240, China.
[3]University of Engineering and Technology Peshawar, Pakistan.
[4]University of Transport Technology, Hanoi 100000, Vietnam.

Article info
Type of article:
Original research paper

DOI:
https://doi.org/10.58845/jstt.utt.2023.en.3.1.26-43

***Corresponding author:**
E-mail address:
banglh@utt.edu.vn

Received: 27/02/2023
Revised: 22/03/2023
Accepted: 24/03/2023

Abstract: Rubberized concrete is a material that is both ecologically friendly and sustainable, and it has been finding more and more usage in building applications recently. In this study, a machine learning model, namely LightGBM, is developed to predict rubberized concrete's compressive strength (CS) using 11 input parameters. The model's performance is measured using several different statistical criteria after being trained on a dataset containing 275 samples. In order to evaluate the impact that each input parameter has on the CS, feature importance and partial dependency plots (PDP) are used as analytical tools. According to the findings, the superplasticizer, chipped rubber, crumb rubber, coarse aggregate, fine aggregate, and water content all have a significant impact on the CS of rubberized concrete. On the other hand, the results indicate that the cement content, slag/fly ash content, and type of CS have a relatively minor effect. In addition to this, the PDP offers insights into the manner in which the input parameters affect the CS of rubberized concrete. Overall, the developed model and analytic techniques may be helpful in forecasting the CS of rubberized concrete and improving its mix design for various construction applications.

Keywords: Compressive strength; LightGBM; Rubberized concrete.

1. Introduction

Waste rubber tires are a mounting environmental concern that demands attention [1]. Made of synthetic rubber, a petroleum-based material that refuses to decompose, these tires pose a long-lasting threat to the environment [2]. Fortunately, the construction industry can play a vital role in reducing the impact by harnessing the power of recycled rubber in building materials like concrete and asphalt [3,4]. This not only lowers the number of raw materials required, but also minimizes the quantity of waste tires deposited into the surrounding environment. Rubberized concrete, a green alternative made by blending ground rubber tires into the concrete mixture, boasts remarkable properties, including enhanced shock absorption, sound-deadening, resistance to wear and abrasion, and lower thermal conductivity, leading to improved energy efficiency in infrastructure [5,6].

Despite its promise, the comprehensive understanding of rubberized concrete's properties and performance remains limited. Further research and experimentation are necessary to unlock its full

potential for use in construction. Compressive strength, denoted CS, is a measure of a material's capacity to withstand given stresses and persist without failure. This critical property, which determines a concrete's structural performance and resistance to cracking and deformation, must be evaluated to assess rubberized concrete's suitability for construction. The CS of rubberized concrete is influenced by several elements, including the kind of rubber particles used, their size, the quantity of rubber in the mixture, the length of time it is allowed to cure, and the other components that are included in the mixture.

Traditional approaches to determining the CS of rubberized concrete include performing compression tests on cylindrical or cubic specimens, and following standardized procedures for preparing and testing the specimens. In an effort to reduce the use of sand in concrete production, Pelisser et al. examine the feasibility of using used tire rubber as a substitute for sand aggregate [7], with a reduction of 14% (28 days) in CS compared to conventional concrete and a CS of 48 MPa for the mixture with the highest resistance, and improved mechanical properties due to alkaline activation and silica fume addition. Turatsinze et al. [8] showed that rubber aggregates in cement composites offered a more effective strategy for reducing fragility and avoiding shrinkage cracking, with a 20-30% replacement of natural sand aggregates. The advantages of fiber reinforcing and rubber integration may be realized concurrently due to the substantial reduction in the modulus of elasticity and substantial rise in the pseudo-strain corresponding to the peak load of the cement composites. Besides, Khaloo et al. [9] looked at the possibility of using tire rubber particles to substitute mineral aggregates in concrete, and found that although this did diminish the strength and modulus of elasticity, it also considerably decreased the brittle behavior, fracture width, and crack propagation velocity. Self-Compacting Concrete (SCC) with rubber aggregates as a partial replacement for natural aggregates was studied by Turatsinze and Garros [10], who found that the composite's modulus of elasticity was reduced while its strain capacity was increased, highlighting the composite's potential application in situations requiring high resistance to cracking due to imposed deformation. Rubcrete blends have fascinating features that may be beneficial in both structural applications and environmental remediation, as researched by Sgobba et al. [11], thanks to an acceptable CS. Overall, rubberized concrete made by combining recycled rubber and concrete aggregates was found to have lower strength than conventional concrete, rendering it unsuitable, particularly for structural applications. Rubberized concrete has been found to be more suitable for paving applications, as the required strength in such applications falls in a lower range [12,13]. When rubber particles are used as a partial substitute for traditional aggregates, the CS of the resulting rubberized concrete deteriorates in inverse proportion to the rubber percentage [14–17]. It has been discovered that the CS of rubberized concrete is drastically altered by incorporating rubber particles in concrete. As rubber content rises, CS of the rubberized concrete lowers. The reason is that rubber particles are less dense than typical aggregates [18], resulting in a less compacted and weaker concrete structure. The extent of the reduction in CS varies depending on the type and content of the rubber particles used.

Although the CS of rubberized concrete is often studied via experimental techniques, these approaches have drawbacks, such as high cost, lengthy testing, and variable findings influenced by several variables, such as the type of rubber used, mixing procedure, and curing time. Therefore, a reliable and accurate method using the constituent materials' properties and mixture proportions to predict rubberized concrete's CS is necessary. This could help reduce the time and cost involved in testing and provide an effective way of assessing rubberized concrete's CS for practical applications.

Recently, there has been a growing interest

in using machine learning (ML) techniques in civil engineering. These approaches have been successfully applied in a wide range of applications, including structural engineering [19], geotechnical engineering [20–24], material sciences [25,26]. Due to its capacity to evaluate massive quantities of complex data and generate accurate predictions that lead to better design and decision-making, ML has shown to be a beneficial tool for civil engineering applications. Light gradient boosting algorithm (LightGBM), a recently developed robust ML model, can be highly advantageous when solving complex problems. With its ability to handle large datasets and optimize complex engineering processes, it is well-suited to prediction tasks. However, it is important to note that LightGBM has not been used specifically to predict the CS of rubberized concrete.

Therefore, the primary purpose of this research is to create an ML model to reliably forecast the CS of rubberized concrete, which can potentially contribute to the development of sustainable and environmentally friendly construction materials. The use of ML techniques can effectively address the limitations of experimental approaches and provide a fast and cost-effective means of predicting concrete strength. A comprehensive dataset is compiled from a review of previous studies and experimental results to achieve this goal. The dataset, covering 275 samples, is then pre-processed to remove outliers and ensure data quality. This study represents a contribution to the field of concrete materials science by adding a number of new samples to the dataset, compared with [26], and testing the feasibility of using LightGBM for predicting the CS of rubberized concrete. This algorithm is evaluated and finely tuned to select the best-performing model. The model is trained on the pre-processed data and validated using a range of metrics to ensure its accuracy and reliability. The developed model provides a reliable method for estimating rubberized concrete's CS, thus enabling the development of new construction materials that are more sustainable and cost-effective.

2. Database collection

The compiled database consists of a total of 275 samples, which are sourced from 19 papers in the relevant literature, including Guneyisi et al. [18], Geosglu and Guneyisi [27], Batayneh et al. [2], Zheng et al. [28], Ganjian et al. [29], Ozbay et al. [30], Paine et al. [31], Ghedan and Hamza [32], Bala et al. [33], Fiore et al. [34], Geosglu et al. [35], Kumar et al. [36], Mohammed and Azmi [37], Almaleeh et al. [38], Bharathi et al. [39], Abusharar [40], Ishwariya [41], Liu et al. [42], and Asutkar et al. [43]. The input space comprises 11 inputs, and one output is represented by the CS (FC) feature. Table 1 provides an easy-to-refer-to summary of the database's information. The specimen type variable is an important additional variable considered in this study, used to indicate the size of the concrete samples being tested for CS. This variable ranges from 0 to 1, where a value of 0 represents cubic concrete samples with dimensions of 100^3 mm, and 1 represents cubic concrete samples with dimensions of 150^3 mm.

For each input variable, a histogram showing the range of values is shown, along with its cumulative distribution, is illustrated in Fig. 1. The histogram facilitates the observation of the concentration of values and any outliers present in the data. The cumulative distribution, on the other hand, allows for the analysis of the distribution's shape and skewness, which is crucial in statistical analysis.

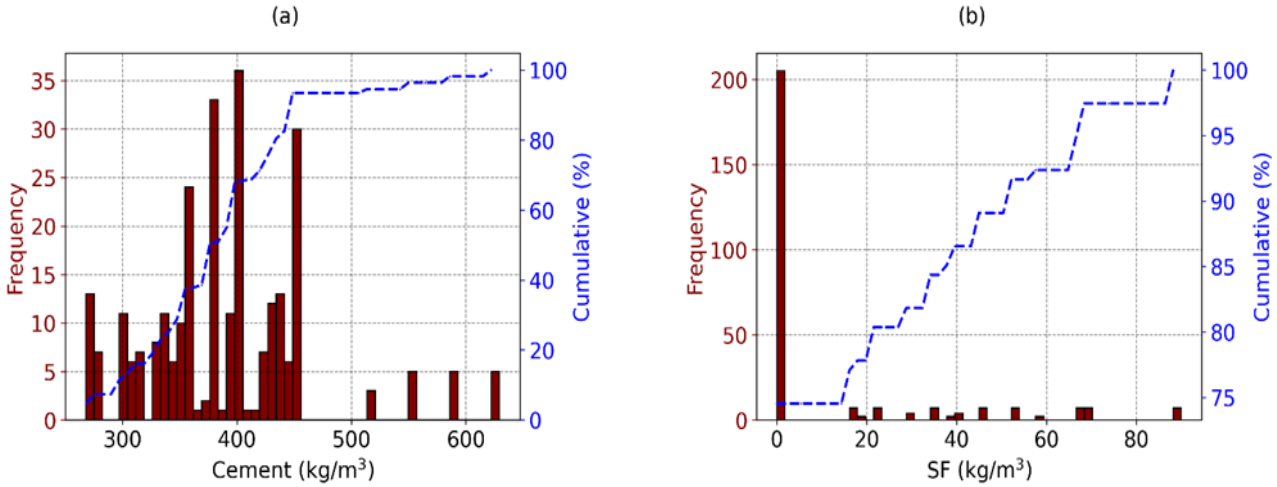
The correlation analysis of all the features present in the data space is depicted in Fig.2. The visual representation provides insights into the relationship between the inputs and the output and helps identify the significant factors affecting the output. Correlation analysis is a commonly used tool for exploratory data analysis and feature selection in data science and ML. It is observed that the highest correlation concerning CS is exhibited by variable SP, with a correlation

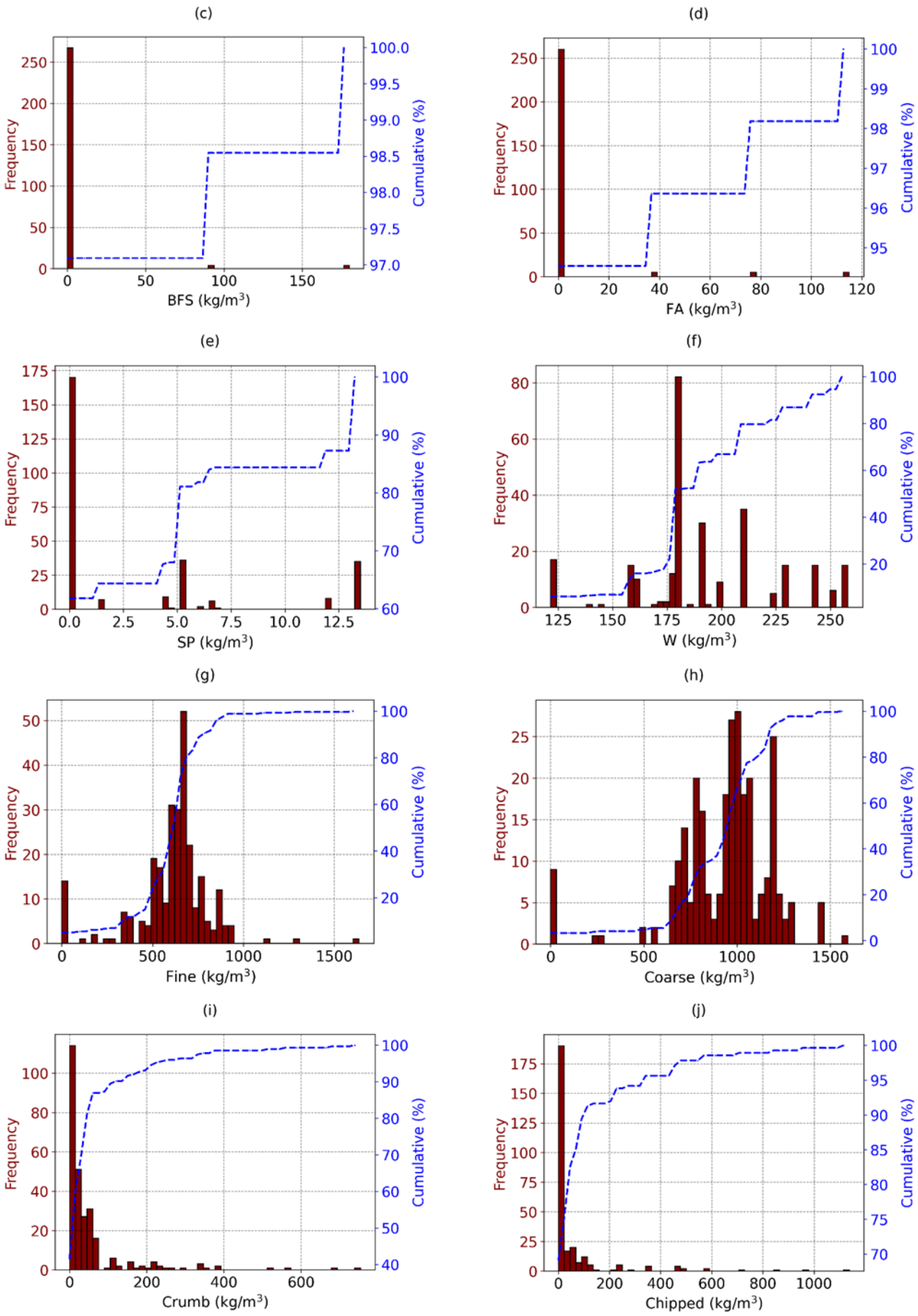
coefficient (R) of 0.56, followed by SF, with a coefficient of 0.39, and fine aggregate, with a coefficient of 0.38. The correlation between input variables is demonstrated to be highest between crumb rubber and chipped rubber, with a correlation of R=0.73, followed by the correlation between SF and SP, which has a correlation of R=0.63. According to the definition provided by Tabachnick et al. [44], pairs of parameters exhibiting strong correlations are characterized by an absolute value of R greater than 0.75. The correlation analysis performed in this study indicates that the input space exhibits suitable characteristics for creating ML model.

Table 1. Statistical characteristics of the input and output parameters in the database.

Parameter	Unit	Mean	Std	Min	Q25	Q50	Q75	Max
Inputs								
Cement (Cement)	kg/m^3	391.0	72.3	268.0	347.4	383.0	430.0	629.3
Silica Fume (SF)	kg/m^3	12.1	23.5	0.0	0.0	0.0	17.5	90.0
Blast Furnace Slag (BFS)	kg/m^3	3.9	24.0	0.0	0.0	0.0	0.0	180.0
Fly Ash (FA)	kg/m^3	4.2	18.9	0.0	0.0	0.0	0.0	115.0
Superplasticizer (SP)	kg/m^3	3.2	4.8	0.0	0.0	0.0	5.3	13.5
Water (W)	kg/m^3	192.4	32.9	121.5	180.0	180.0	210.0	258.0
Fine aggregate (Fine)	kg/m^3	609.0	209.0	0.0	544.1	645.5	692.5	1635.5
Coarse aggregate (Coarse)	kg/m^3	931.9	260.4	0.0	792.0	975.0	1069.2	1594.0
Crumb rubber (Crumb)	kg/m^3	53.2	99.8	0.0	2.7	24.3	54.5	754.0
Chipped rubber (Chipped)	kg/m^3	55.6	146.6	0.0	0.0	0.0	30.5	1138.0
Specimen type (Type)	-	0.7	0.5	0.0	0.0	1.0	1.0	1.0
Output								
Compressive strength (FC)	MPa	30.2	19.6	0.6	15.9	28.0	42.8	85.7

Std=Standard deviation;




(c)
Frequency
Cumulative (%)
BFS (kg/m³)
(d)
FA (kg/m³)
(e)
SP (kg/m³)
(f)
W (kg/m³)
(g)
Fine (kg/m³)
(h)
Coarse (kg/m³)
(i)
Crumb (kg/m³)
(j)
Chipped (kg/m³)

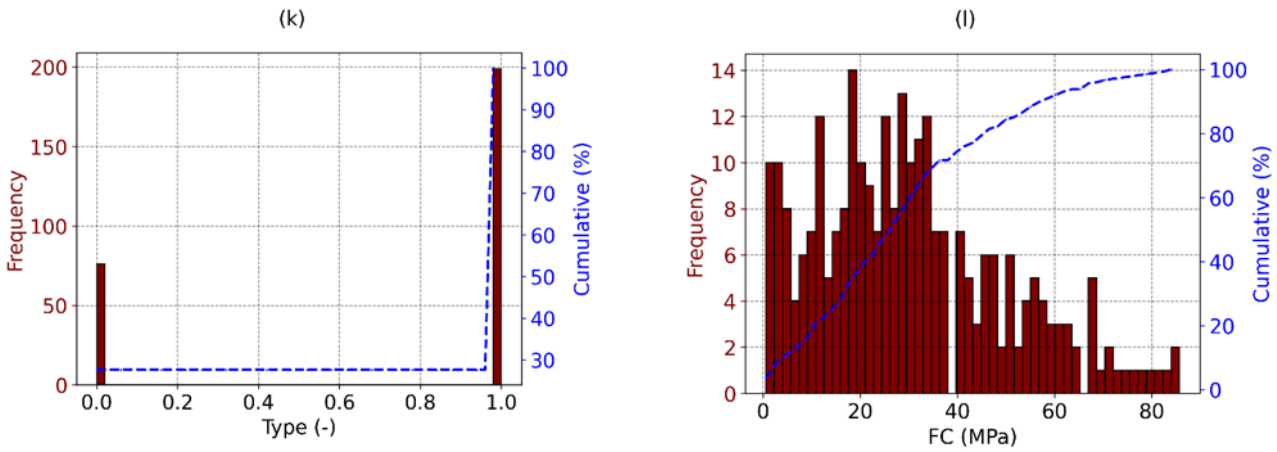


Fig. 1. Histograms of all the variables used in the data space

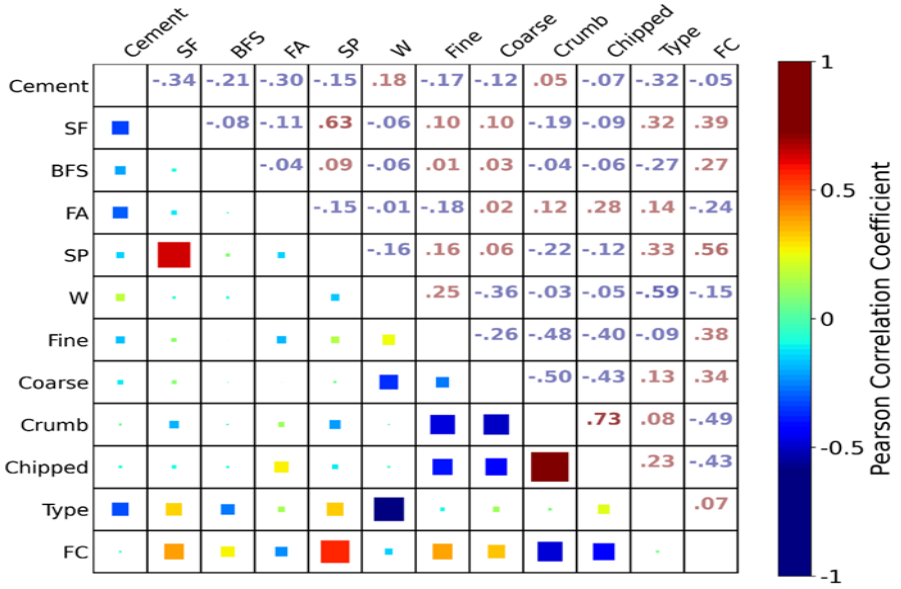


Fig. 2. Correlation of each input parameter and output parameter of the database

The division of the dataset into two sub-datasets has been carried out randomly, with the first part, which comprises 70% of the data, designated as the training part, and the second part, which consists of the remaining 30% of data, designated as the testing part. The selection of this split ratio, with the larger portion being assigned to the training part, has been made following suggestions found in the relevant literature [45], to ensure efficiency during both the training and testing phases.

3. Method

3.1. Light gradient boosting

LightGBM is an effective gradient-boosting (GB) framework that utilizes gradient-based learning algorithms to address the challenges of large-scale datasets [46]. It may be used for regression and classification issues. The "light" in LightGBM refers to its efficient implementation, which reduces the computational time required for training and inference compared to traditional GB algorithms.

The innovative method used by LightGBM is known as gradient-based one-side sampling (GOSS), which focuses on building models with more significant data instances and reducing the impact of noisy and irrelevant data. LightGBM uses a histogram-based algorithm to discretize continuous features, which helps to reduce the computational cost and increase the training efficiency. LightGBM also introduces a leaf-wise tree growth strategy that grows the tree from the

bottom to the top, selecting the best split based on the maximum reduction of the loss function. Unlike traditional depth-wise tree growth strategies that grow the tree in a level-by-level manner, leaf-wise tree growth leads to a more compact tree structure and can handle sparse and imbalanced data more effectively. Additionally, LightGBM employs a cache-conscious algorithm to reduce memory usage, and it supports parallel and GPU-based learning, making it more scalable and efficient than traditional GB methods. In summary, LightGBM's unique combination of one-sided sampling, leaf-wise tree growth, and cache-conscious learning makes it a highly efficient gradient-boosting framework for large-scale datasets.

3.2. Performance indices of models

The RMSE, MAE, R^2, and MAPE are the performance metrics used in this study to assess the effectiveness and precision of the LightGBM model in predicting the FC of rubberized concrete. The following formula is used to determine these performance indicators:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Y_{tt,i} - Y_{db,i}\right)^2} \quad (1)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Y_{tt,i} - Y_{db,i}\right| \quad (2)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\left|Y_{tt,i} - Y_{db,i}\right|}{Y_{tt,i}}\right| \times 100\% \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_{tt,i} - Y_{db,i}\right)^2}{\sum_{i=1}^{N}\left(Y_{tt,i} - \bar{Y}_{tt}\right)^2} \quad (4)$$

where $Y_{tt,i}$ is the value of i^{th} test sample, $\bar{Y}_{tt}$ is the average of these values, $Y_{db,i}$ is predicted value of i^{th} sample, calculated by ML model, and N denotes the total numbers. The RMSE, MAE, and MAPE performance measures all achieve their ideal value when equal to 0, however the value of R^2 reaches its optimal value when it is equal to 1.

4. Results and Discussion

4.1. Hyperparameter tuning of the LightGBM model

Hyperparameter tuning is an important aspect of ML that involves adjusting the parameters of the learning algorithm to optimize its performance on a specific task. The choice of hyperparameters significantly impacts the model's performance, and their optimization is critical to achieving the best results. The objective of hyperparameter tuning is to find the best combination of hyperparameters that leads to the optimal ML model performance. The process of hyperparameter tuning involves evaluating a range of hyperparameter values to find the combination that gives the best performance on a validation set. The optimization process can be time-consuming and requires a lot of computational resources, but ensuring that the model has the best possible performance is essential.

The tuning of the hyperparameters of the LightGBM model is performed through an extensive trial and error process. The selection of the hyperparameters to be tuned is based on their significant impact on the LightGBM algorithm. The four hyperparameters that are chosen for tuning are n estimators, max depth, learning rate, and min child samples. The search domains for the hyperparameters associated with the LightGBM model are presented in Table 2. The default values for the remaining hyperparameters in the LightGBM package utilized in Python are utilized.

The utilization of a 5-fold cross-validation (CV) methodology is employed to prevent overfitting and increase the reliability of the training operations. The training dataset is randomly divided into five equal folds, with each fold selected four times to serve as the training component, while the remaining fold was used as the validation component. The performance metrics are calculated by taking the mean of the results obtained after five training and validation iterations have been conducted. The validation dataset created by the 5-fold CV method tests and compares the model's performance throughout the training phase. The hyperparameter tuning

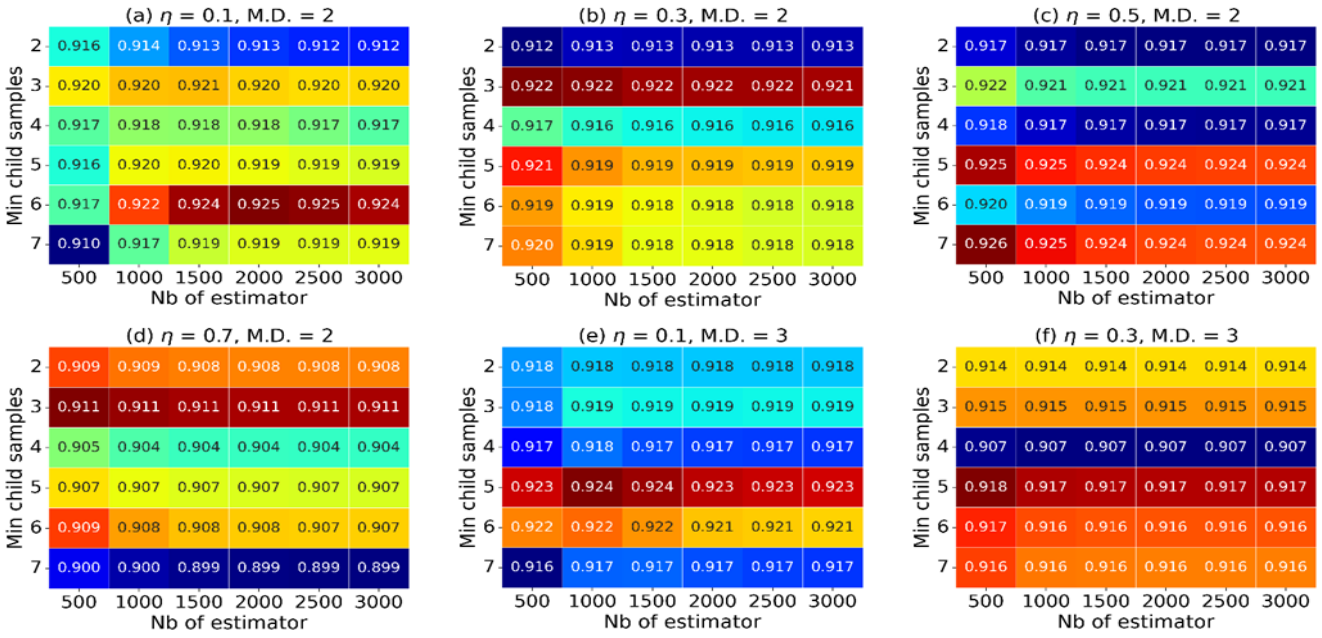
process is assessed using RMSE and R^2 criteria. The optimal hyperparameters for the LightGBM model are selected when the lowest RMSE and the highest R^2 are exhibited.

The optimal values of the hyperparameters are identified by evaluating the validation scores, as indicated in the results. Specifically, the combination of η = 0.5 and M.D. = 0.2 results in the best validation scores, with an R^2 value of 0.926. Furthermore, this combination also yields the lowest value of RMSE, which is 5.253 MPa. These findings suggest that tuning the hyperparameters using the trial and error approach has enabled the LightGBM model to achieve improved performance and greater accuracy in predicting the FC. The results above are obtained with a min child sample value of 7 and a number of estimators set to 500. These hyperparameters reflect the minimum number of samples required in a decision tree leaf, and the number of decision trees created by the LightGBM model, respectively. High validation scores are also observed for other combinations, including η = 0.1 and M.D. = 0.2 with an R^2 of 0.925, η = 0.1 and M.D. = 0.3 with an R^2 of 0.924, and η = 0.1 and M.D. = 0.4 with an R^2 of 0.923. Based on the presented results, it appears that the LightGBM model's hyperparameters significantly impact the model's performance. It is clear that the combination of η = 0.5 and M.D. = 0.2 yields the best performance, as it gives the highest validation R^2 score and the lowest value of RMSE. However, other combinations also result in high validation scores, with R^2 values ranging from 0.925 to 0.923, indicating that these hyperparameter combinations are also feasible.

It is noted that the results are based on the specific dataset and problem being investigated. The optimal hyperparameters for the LightGBM model may vary depending on the specific characteristics of the dataset and problem, the feature number, and relationship between inputs and the target.

Table 2. Search domain of 4 hyperparameters tuned in LightGBM model

Hyperparameter	Notation	Studied range of value
n_estimators (N_e)	estimator	500, 1000, 1500, 2000, 2500, 3000
learning_rate	η	0.1, 0.3, 0.5, 0.7
max_depth	M.D.	2, 3, 4
min_child_samples	Min child samples	2, 3, 4, 5, 6, 7



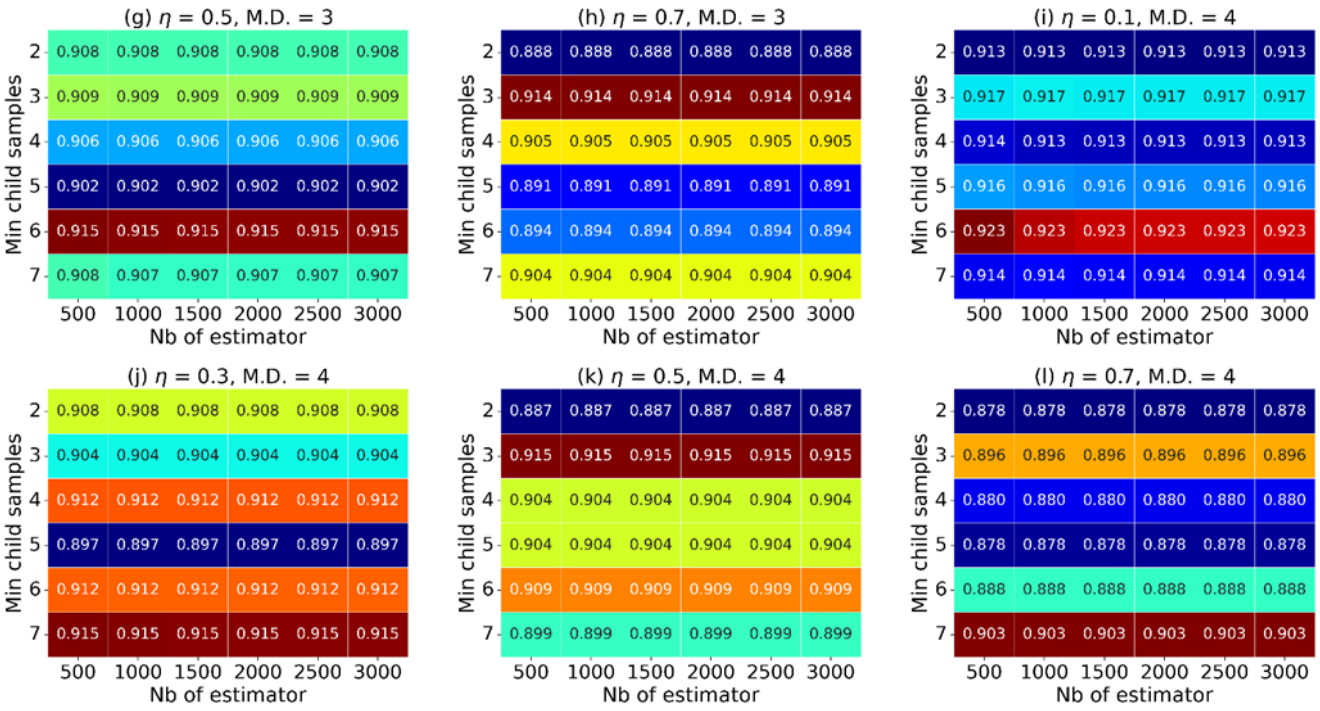
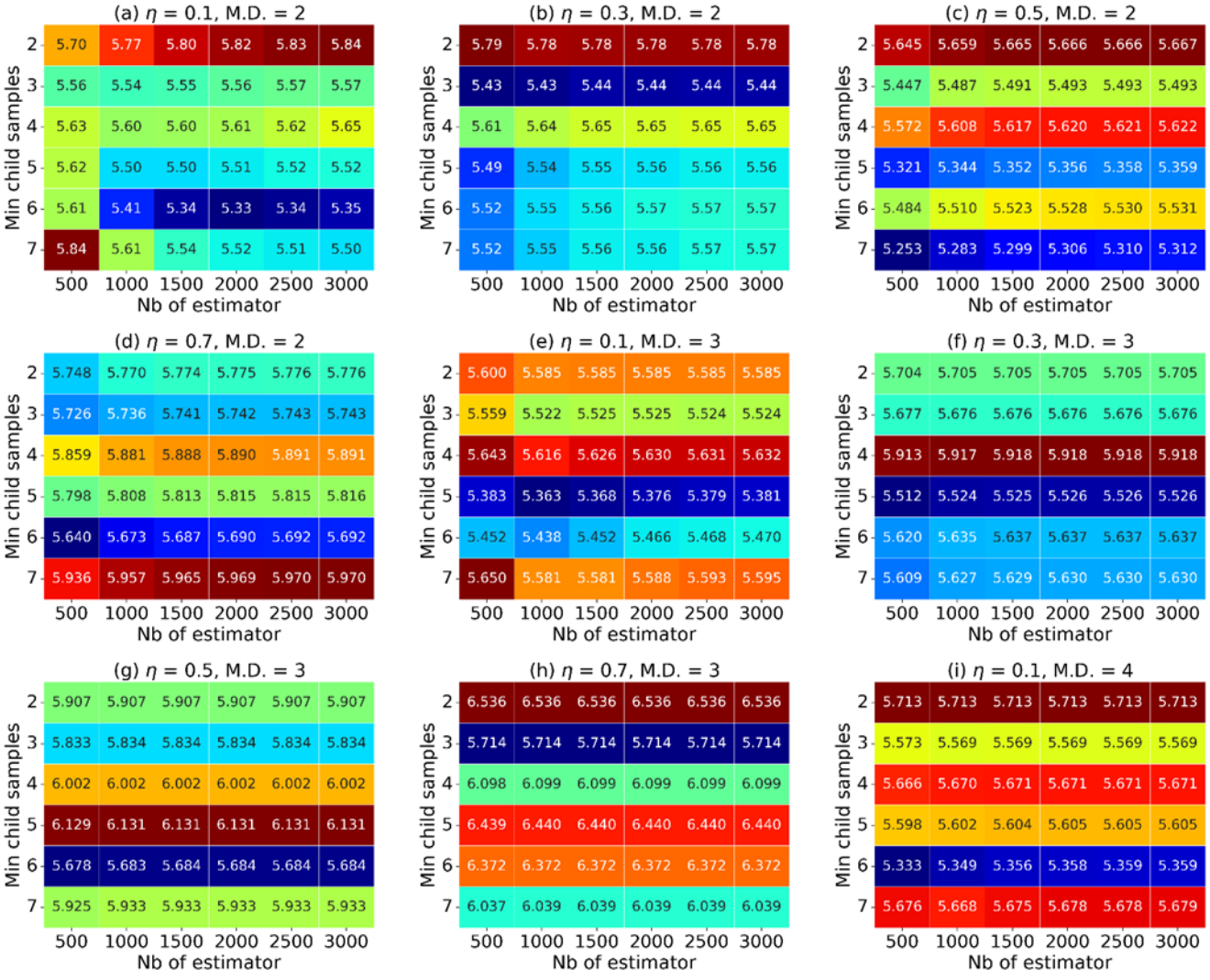


Fig. 3. R^2 validation scores of LightGBM model



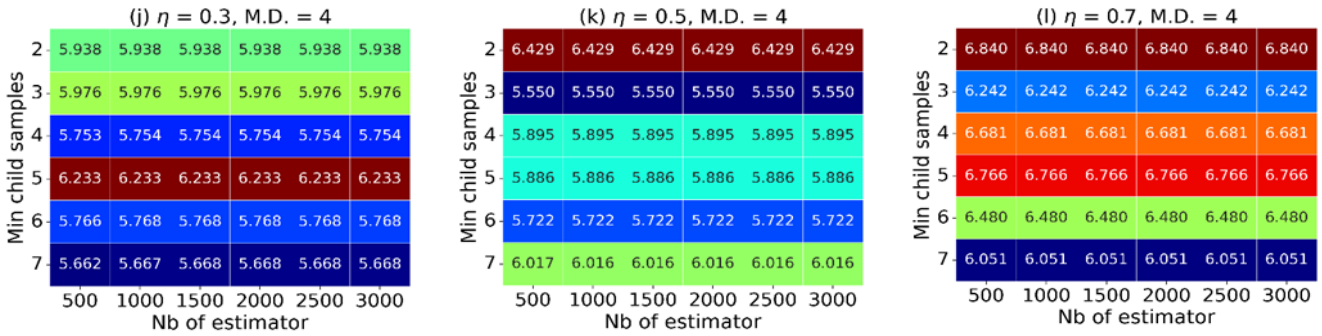


Fig. 4. RMSE validation scores of LightGBM model

4.2. Prediction results

In this section, the finely-tuned LightGBM model's prediction results are presented. The regression analysis of the training, testing, and all data is displayed in Fig. 5. In each case, the two error lines with a deviation of 10% and 20% of the diagonal are shown, providing valuable information regarding the accuracy and precision of the model's predictions. The success of the training phase can be observed from the location of all the data points near the diagonal in Fig. 5. This indicates that the LightGBM model has accurately predicted the FC values and are highly consistent with the experimental FC values. This is a strong indication of the reliability and effectiveness of the model in performing the regression analysis of the data. Regarding the testing data, the model's performance is observed to be relatively lower than that of the training phase, but the results remain precise. Several errors can be noticed with low FC values, particularly those below 50 MPa.

The distribution histogram and cumulative distribution of the error between the predicted and actual FC of the rubberized concrete for both training and testing parts are presented in this section. The distribution histogram of the error, which reflects the frequency distribution of the error range, can help identify the range and distribution of the errors. Among the 193 training and 82 testing data samples, it is observed that a significant error value of approximately 0 exists. It is important to point out that the training dataset includes only 2 samples have errors that lie outside the range of [-2, 2] MPa. In addition, when compared to the overall number of data points, this percentage of samples that include significant errors is negligible. The testing dataset contains a greater error than the training dataset, with a maximum error value of 15 MPa. Furthermore, only 5 samples out of the total number of samples exhibit errors outside of the range [-10, 10] MPa, which is also an insignificant proportion.

Table 3 highlights the quantitative values of the evaluation criteria used to assess the performance of the LightGBM model. It has been noticed that the model has a high degree of predictive accuracy, with an R^2 value of 0.982, an RMSE value of 2.663 MPa, an MAE value of 1.332 MPa, and a MAPE value of 0.109 for all data. In addition, a comparison between the optimized and default LightGBM model is also conducted. The results show that the default model achieved lower prediction accuracy compared to the optimized model, indicating the importance of hyperparameter tuning in achieving optimal performance. The detailed comparison metrics and performance measures in also presented (Table 3). The results demonstrate that the LightGBM model can accurately predict the FC of the rubberized concrete. By doing so, the model is expected to contribute to high cost and time savings in designing and testing rubberized concrete.

4.3. Feature importance

The predictive capability of the ML model developed for forecasting the FC of rubberized concrete is substantial. Nonetheless, interpreting or explaining the model predictions is crucial [47],

given that ML models are typically "black boxes." The interpretation of the model has the potential to aid in directing the development of the model and decision-making procedures, while also promoting user confidence in the trained model. The analysis of feature importance is the most common approach for explaining the models. Python's LightGBM package provides a built-in function that allow for the determination of input parameter sensitivities. The feature_importance() function computes feature importance scores for each input based on its contribution to the model's predictions, in which the selected method for computing is permutation.

The relative importance of features for the FC of rubberized concrete is analyzed using the LightGBM model, and the results are presented in Fig. 7. The analysis shows that SP is the most critical parameter, followed by chipped rubber, crumb rubber, and coarse aggregate. Water (W) and fine aggregate are also important among the other parameters. Conversely, the cement content parameter has minor relevance, followed by SF, BFS, type, and FA. It can be inferred that a significant change in the FC of rubberized concrete can result from changing the proportion of SP. Additionally, using more or less chipped rubber, crumb rubber, and coarse aggregate can also impact the FC of rubberized concrete. The magnitude of the alteration in FC of rubberized concrete would be reduced when modifying the quantity of water and fine aggregate. In contrast, the modification of cement content, SF, BFS, type, and FA would not affect the FC of rubberized concrete.

4.4. Partial dependence analysis

The analysis of feature importance shows whether the features are significant, however, it does not explain the positive or negative influence of the variable of interest. Partial Dependence Plots (PDP) are graphical tools that illustrate a feature of interest and the predicted outcome' relationship, while marginalizing the effect of other features in a ML model. PDPs show the average impact of a particular feature on the predicted outcome across the range of values that it takes in the dataset. The main advantage of PDP over feature importance is that it provides a detailed and more interpretable view of how a feature influences the model's prediction. PDPs can reveal nonlinear relationships between the input and output variables, capture interactions between multiple features, and identify regions of the feature space where the model is most sensitive to changes in the feature values. Therefore, PDPs can offer deeper insights into the model's behavior and help practitioners better understand the model's decision-making process.

The PDP analysis for the 11 input parameters is depicted in Fig. 8. The horizontal axis shows the changes in input parameter values, while the vertical axis displays the PDP output value for each considered feature. The PDP illustrates that the input parameters have a comparable influence on the FC of rubberized concrete to the results obtained by feature importance analysis. By increasing the content of SP, the FC of rubberized concrete can be increased. The PDP shows that an increase in the content of SP from 0 to 14 kg/m3 can result in a significant rise in FC, from 25 to 46 MPa. This illustrates the strong positive correlation between SP content and the FC of rubberized concrete. However, it should be noted that a further increase in the content of SP could have a certain impact on the FC, which needs to be investigated further. In addition, it has been shown that a decrease in the FC of rubberized concrete may occur when there is a greater proportion of chipped rubber and crumb rubber present in the mixture. Specifically, when the content of chipped rubber and crumb rubber is increased, the FC can decrease from 36 to 15 MPa and 36 to 18 MPa, respectively. The FC of rubberized concrete may be made higher by increasing the amount of coarse and fine aggregate that is included in the mixture. In

particular, changing the coarse aggregate content from 600 to 1100 kg/m^3 increases FC from 24 to 35 MPa. Similarly, changing the fine aggregate content from 0 to 800 kg/m^3 increases FC from 22 to 34 MPa. According to the findings, the impact of water on the FC of rubberized concrete has a complex relationship with the variables involved. Increasing the water content from 120 to 180 kg/m3 leads to a reduction in the FC from 37 to 22 MPa. However, beyond this threshold, the change in FC is minor, ranging around 28 MPa. Within the range of the input space, a variation in cement content (from 260 to 680 kg/m3) only leads to a change of 9 MPa in FC of rubberized concrete. It is worth noting that changes in the other input parameters result in minor changes in the FC of rubberized concrete. A change in SF results in a minor change in the FC of rubberized concrete, approximately 4 MPa. Similarly, modifying BFS within the range of input space has resulted in a change of FC of only about 6.5 MPa. The FA feature is found to have a very minimal effect on the FC of rubberized concrete, only leading to a change of approximately 1 MPa. Furthermore, the type feature's effect on FC is insignificant, resulting in a change of only about 0.7 MPa.

The PDP analysis reveals that the influence of the features on the FC of rubberized concrete is comparable to the results of the feature importance analysis. As a result, PDP is an effective instrument for gaining a knowledge of the effect that input factors have on the FC of rubberized concrete. The findings illustrate how critical it is to thoughtfully pick the proportions of SP, chipped rubber, crumb rubber, coarse and fine aggregate, and water content in order to get the required FC. The study also suggests that changes in other input parameters, such as cement content, SF, BFS, FA, and type, have a minor effect on the FC. It is feasible to acquire a better understanding of the influence that input factors have on the target variable by making use of PDP, which can aid in the optimization of rubberized concrete formulations for specific applications.

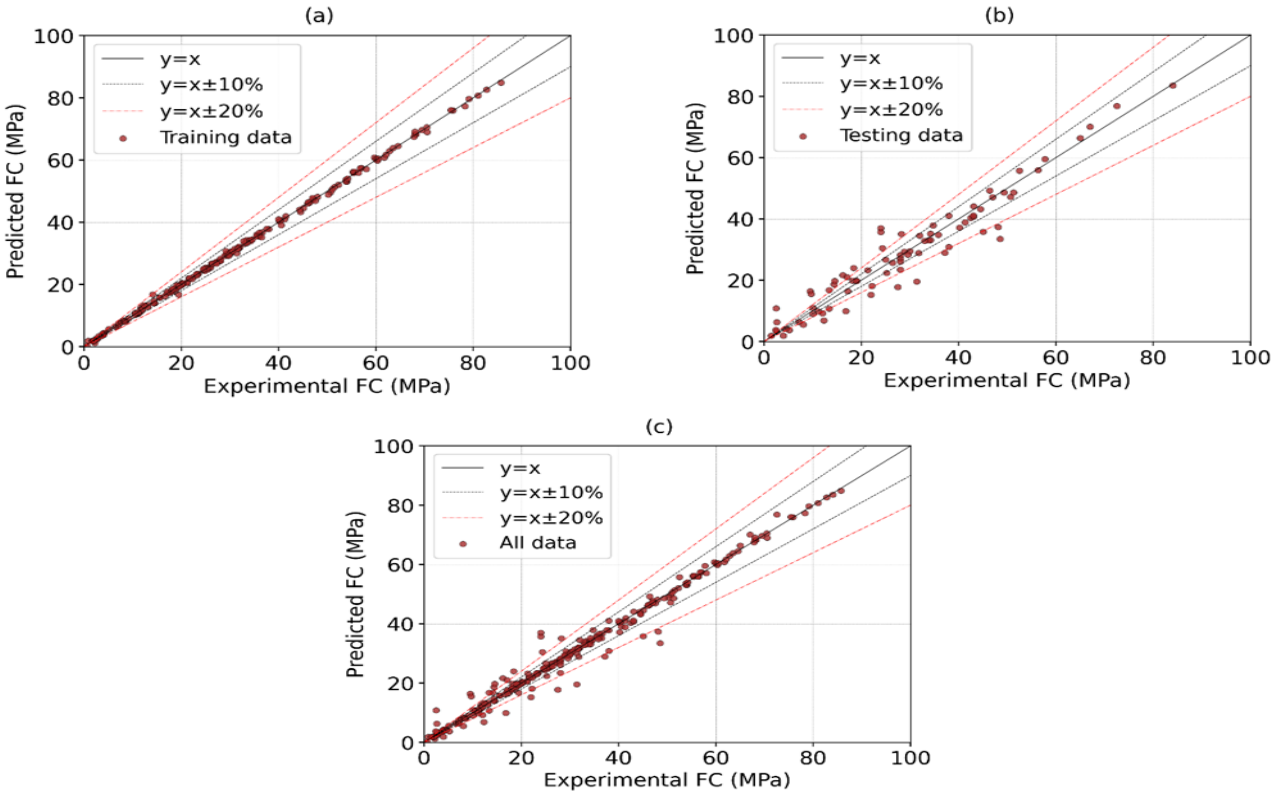


Fig. 5. Regression analysis between experimental and predicted FC of rubberized concrete for the training part; and testing part

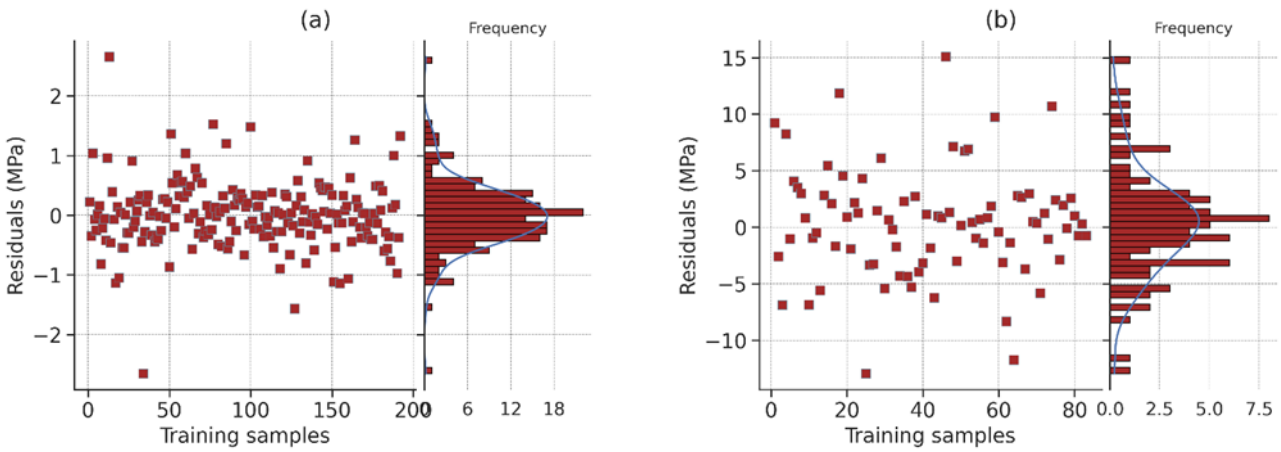


Fig. 6. Error results (a) training set; and (b) testing set

Table 3. Statistical criteria values for typical results of LightGBM model

	RMSE (MPa)	MAE (MPa)	R^2	MAPE (%)
Finely-tuned LightGBM model				
Training part	0.565	0.397	0.999	0.031
Testing part	4.770	3.495	0.930	0.289
All dataset	2.663	1.332	0.982	0.109
Default LightGBM model				
Training part	2.771	2.007	0.981	0.137
Testing part	6.207	4.442	0.881	0.344
All dataset	4.122	2.742	0.956	0.199

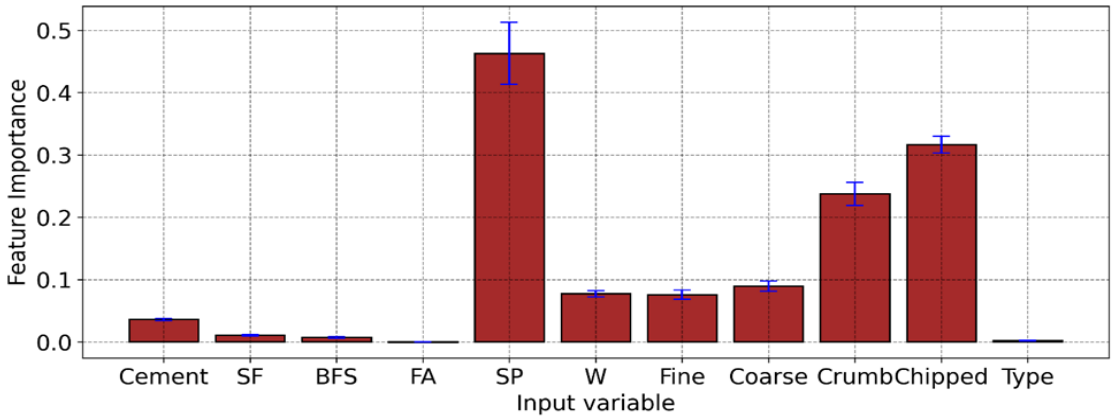
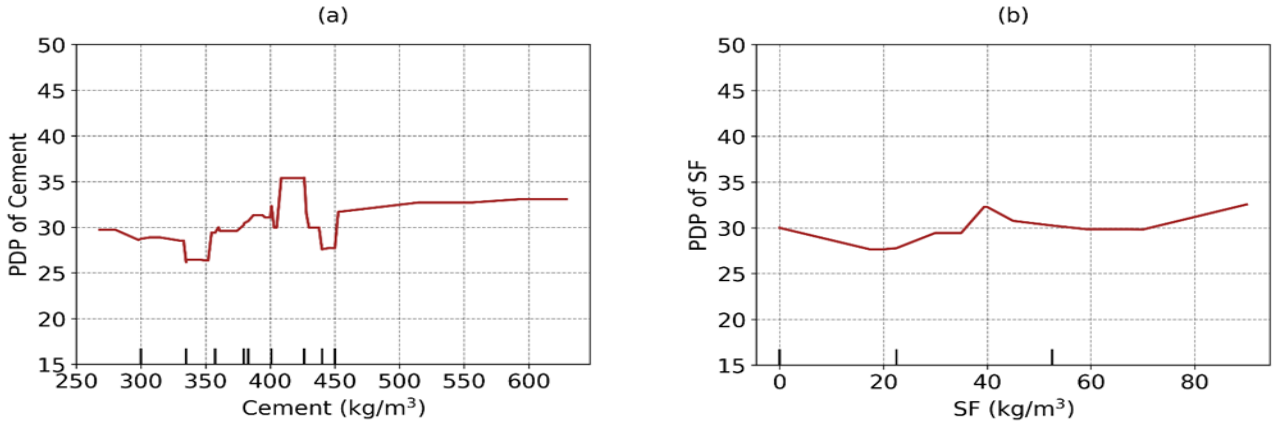
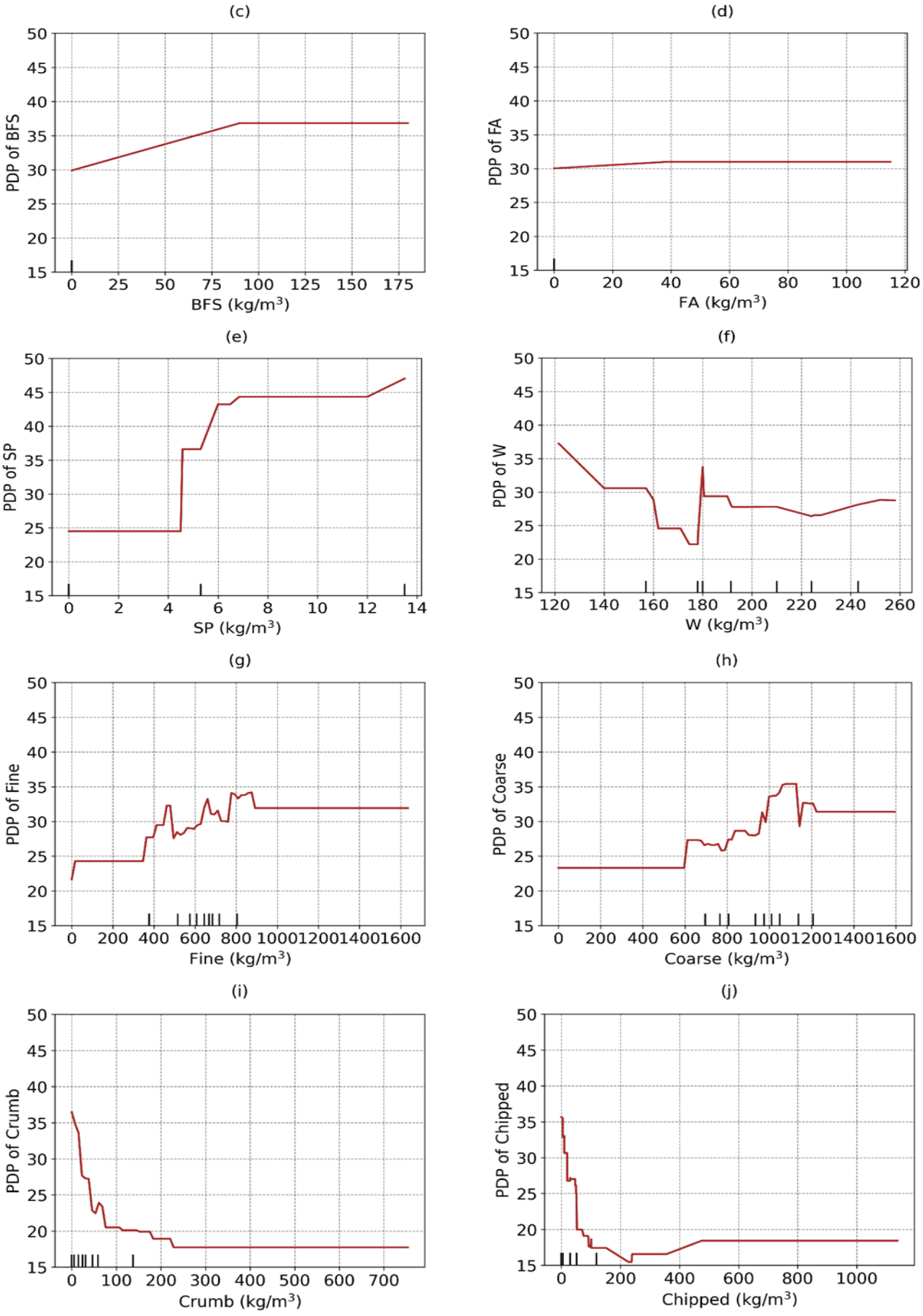


Fig. 7. Feature importance analysis results




(c)
PDP of BFS
BFS (kg/m³)
(d)
PDP of FA
FA (kg/m³)
(e)
PDP of SP
SP (kg/m³)
(f)
PDP of W
W (kg/m³)
(g)
PDP of Fine
Fine (kg/m³)
(h)
PDP of Coarse
Coarse (kg/m³)
(i)
PDP of Crumb
Crumb (kg/m³)
(j)
PDP of Chipped
Chipped (kg/m³)

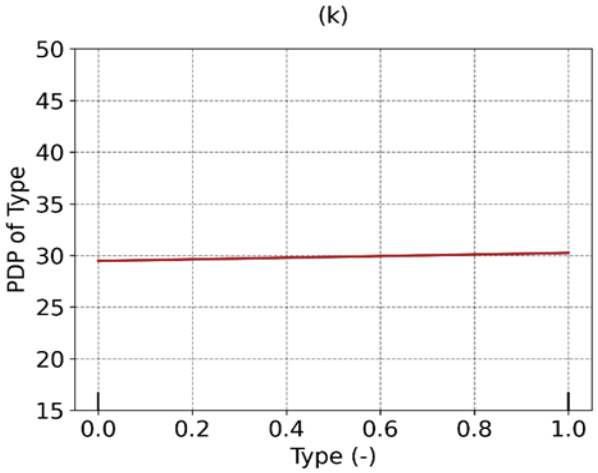


Fig. 8. PDP analysis of the variables considered in this work

5. Conclusions

This work analyzes the possibility of utilizing ML model and statistical approaches in the prediction of the CS of rubberized concrete, also illustrating the usefulness of applying these methods. In order to accomplish this goal, a dataset consisting of 275 samples of rubberized concrete is employed in developing the LightGBM algorithm. This method achieves a high level of prediction accuracy while maintaining a high level of efficiency, as evidenced by an R^2 value of 0.982, RMSE value of 2.663 MPa, MAE value of 1.332 MPa, and MAPE value of 0.109 for all data.

The influence of each input parameter on the CS of rubberized concrete is investigated with the use of PDP. The outcomes of the feature importance analysis are reflected in the fact that the influence of each input parameter on CS is comparable to that of PDP analysis. The study also finds that the most critical input parameter for predicting the CS of rubberized concrete is the superplasticizer (SP) content, followed by chipped rubber, crumb rubber, and coarse and fine aggregates. It is also determined that the amount of water in the rubberized concrete has a complicated influence on the CS of the concrete, and the amount of CS drops as the water content increases.

In general, the findings of this research give useful insights into the effects that different input parameters have on the CS of rubberized concrete. These findings may assist in the creation of concrete mix designs that are more efficient and effective.

Acknowledgments: This research is funded by the University of Transport Technology, Thanh Xuan, Hanoi, Vietnam (UTT), under grant number DTTD 2022-06.

Availability of data and material: Data will be made available on request.

Conflict of Interest: The authors declare that there is no conflict of interest.

References

[1] A. Sofi. (2018). Effect of waste tyre rubber on mechanical and durability properties of concrete – A review. Ain Shams Engineering Journal, 9(4), 2691-2700.

[2] M.K. Batayneh, I. Marie, I. Asi. (2008). Promoting the use of crumb rubber concrete in developing countries. *Waste Management*, 28(11), 2171-2176.

[3] I.B. Topçu. (1995). The properties of rubberized concretes. *Cement and Concrete Research*, 25(2), 304-310.

[4] İ.B. Topçu. (1997). Assessment of the brittleness index of rubberized concretes. *Cement and Concrete Research*, 27(2), 177-183.

[5] G. Li, M.A. Stubblefield, G. Garrick, J. Eggers,

C. Abadie, B. Huang. (2004). Development of waste tire modified concrete. *Cement and Concrete Research*, 34(12), 2283-2289.

[6] F. Hernández-Olivares, G. Barluenga, B. Parga-Landa, M. Bollati, B. Witoszek. (2007). Fatigue behaviour of recycled tyre rubber-filled concrete and its implications in the design of rigid pavements. *Construction and Building Materials*, 21(10), 1918-1927.

[7] F. Pelisser, N. Zavarise, T.A. Longo, A.M. Bernardin. (2011). Concrete made with recycled tire rubber: effect of alkaline activation and silica fume addition. *Journal of Cleaner Production*, 19(6-7), 757-763.

[8] A. Turatsinze, S. Bonnet, J.-L. Granju. (2005). Mechanical characterisation of cement-based mortar incorporating rubber aggregates from recycled worn tyres. *Building and Environment*, 40(2), 221-226.

[9] A.R. Khaloo, M. Dehestani, P. Rahmatabadi. (2008). Mechanical properties of concrete containing a high volume of tire–rubber particles. *Waste Management*, 28(12), 2472-2482.

[10] A. Turatsinze, M. Garros. (2008). On the modulus of elasticity and strain capacity of self-compacting concrete incorporating rubber aggregates. *Resources, Conservation and Recycling*, 52(10), 1209-1215.

[11] S. Sgobba, G.C. Marano, M. Borsa, M. Molfetta. (2010). Use of rubber particles from recycled tires as concrete aggregate for engineering applications. *2nd International Conference on Sustainable Construction Materials and Technologies*.

[12] A.J. Kardos. (2011). Beneficial use of crumb rubber in concrete mixtures. *University of Colorado at Denver*.

[13] A.C. Ho, A. Turatsinze, A. Abou-Chakra, D.C. Vu. (2012). Rubberised concrete for the design of pavement on soil. *International Journal of Materials Engineering Innovation*, 3(2), 101-116.

[14] N.N. Eldin, A.B. Senouci. (1994). Measurement and prediction of the strength of rubberized concrete. *Cement and Concrete Composites*, 16(4), 287-298.

[15] Z.K. Khatib, F.M. Bayomy. (1999). Rubberized Portland cement concrete. *Journal of Materials in Civil Engineering*, 11(3), 206-213.

[16] M.M. Reda Taha, A.S. El-Dieb, M.A. Abd El-Wahab, M.E. Abdel-Hameed. (2008). Mechanical, fracture, and microstructural investigations of rubber concrete. *Journal of Materials in Civil Engineering*, 20(10), 640-649.

[17] M. Gesoglu, E. Güneyisi, O. Hansu, S. İpek, D.S. Asaad. (2015). Influence of waste rubber utilization on the fracture and steel–concrete bond strength properties of concrete. *Construction and Building Materials*, 101(1), 1113-1121.

[18] E. Güneyisi, M. Gesoğlu, T. Özturan. (2004). Properties of rubberized concretes containing silica fume. *Cement and Concrete Research*, 34(12), 2309-2317.

[19] H. Nguyen, H.-B. Ly, T.V. Quan, T.-A. Nguyen, T.-T. Le, B.T. Pham. (2020). Optimization of Artificial Intelligence System by Evolutionary Algorithm for Prediction of Axial Capacity of Rectangular Concrete Filled Steel Tubes under Compression. *Materials*, 13(5), 1205.

[20] P. Zhang, Z.-Y. Yin, Y.-F. Jin. (2022). Machine learning-based modelling of soil properties for geotechnical design: review, tool development and comparison. *Archives of Computational Methods in Engineering*, 29, 1229-1245.

[21] J.-S. Chou, J.P.P. Thedja. (2016). Metaheuristic optimization within machine learning-based classification system for early warnings related to geotechnical problems. *Automation in Construction*, 68, 65-80.

[22] B. Indraratna, D.J. Armaghani, A.G. Correia, H. Hunt, T. Ngo. (2023). Prediction of

resilient modulus of ballast under cyclic loading using machine learning techniques. *Transportation Geotechnics*, 38, 100895.

[23] B. He, D.J. Armaghani, S.H. Lai. (2023). Assessment of tunnel blasting-induced overbreak: A novel metaheuristic-based random forest approach. *Tunnelling and Underground Space Technology*, 133, 104979.

[24] F. Shan, X. He, D.J. Armaghani, P. Zhang, D. Sheng. (2022). Success and challenges in predicting TBM penetration rate using recurrent neural networks. *Tunnelling and Underground Space Technology*, 130, 104728.

[25] A.D. Skentou, A. Bardhan, A. Mamou, M.E. Lemonis, G. Kumar, P. Samui, D.J. Armaghani, P.G. Asteris. (2022). Closed-Form Equation for Estimating Unconfined Compressive Strength of Granite from Three Non-destructive Tests Using Soft Computing Models. *Rock Mechanics and Rock Engineering*, 56, 487-514.

[26] H.-B. Ly, T.-A. Nguyen, V.Q. Tran. (2021). Development of deep neural network model to predict the compressive strength of rubber concrete. *Construction and Building Materials*, 301, 124081.

[27] M. Gesoğlu, E. Güneyisi. (2007). Strength development and chloride penetration in rubberized concretes with and without silica fume. *Materials and Structures*, 40, 953-964.

[28] L. Zheng, X.S. Huo, Y. Yuan. (2008). Strength, modulus of elasticity, and brittleness index of rubberized concrete. *Journal of Materials in Civil Engineering*, 20(11), 692-699.

[29] E. Ganjian, M. Khorami, A.A. Maghsoudi. (2009). Scrap-tyre-rubber replacement for aggregate and filler in concrete. *Construction and Building Materials*, 23(5), 1828-1836.

[30] E. Ozbay, M. Lachemi, U.K. Sevim. (2011). Compressive strength, abrasion resistance and energy absorption capacity of rubberized concretes with and without slag. *Materials and Structures*, 44, 1297-1307.

[31] K.A. Paine, R.K. Dhir, R. Moroney, K. Kopasakis. (2002). Use of Crumb Rubber to Achieve Freeze/Thaw Resisting Concrete. *Challenges of Concrete Construction: Vol 6, Concrete for Extreme Conditions*, 485–498.

[32] R.H. Ghedan, D.M. Hamza. (2011). Effect of rubber treatment on compressive strength and thermal conductivity of modified rubberized concrete. *Journal of Engineering and Development*, 15(4), 21-29.

[33] A. Bala, V.K. Sehgal, B. Saini. (2014). Effect of Fly ash and Waste Rubber on Properties of Concrete composite. *ISSR Journal*, 5(3), 842-857.

[34] A. Fiore, G.C. Marano, C. Marti, M. Molfetta. (2014). On the fresh/hardened properties of cement composites incorporating rubber particles from recycled tires. *Advances in Civil Engineering*, 4, 876158, 1-12.

[35] M. Gesoğlu, E. Güneyisi, G. Khoshnaw, S. İpek. (2014). Investigating properties of pervious concretes containing waste tire rubbers. *Construction and Building Materials*, 63, 206-213.

[36] G.N. Kumar, V. Sandeep, C. Sudharani. (2014). Using tyres wastes as aggregates in concrete to form rubcrete-mix for engineering applications. *International Journal of Research in Engineering and Technology*, 3(11), 500-509.

[37] B. Mohammed, N. Azmi. (2014). Strength reduction factors for structural rubbercrete. *Frontiers of Structural and Civil Engineering*, 8 270-281.

[38] A.M. Almaleeh, S.M. Shitote, T. Nyomboi. (2017). Use of waste rubber tyres as aggregate in concrete. *Journal of Civil Engineering and Construction Technology*, 8(2), 11-19.

[39] R. Bharathi Murugan, E. Sai, C. Natarajan, S. Chen. (2017). Flexural fatigue performance and mechanical properties of rubberized concrete. *Gradjevinar*, 69(11), 983-990.

[40] S.W. Abusharar. (2015). Effect of particle sizes on mechanical properties of concrete

containing crumb rubber. *Innovative Systems Design and Engineering*, 6(2), 114-125.

[41] T. Ishwariya. (2016). An Experimental study on partial replacement of coarse aggregate by crumb rubber. *International Research Journal of Engineering and Technology*, 3(6), 1047-1050.

[42] H. Liu, X. Wang, Y. Jiao, T. Sha. (2016). Experimental Investigation of the Mechanical and Durability Properties of Crumb Rubber Concrete. *Materials*, 9(3), 172.

[43] P. Asutkar, S.B. Shinde, R. Patel. (2017). Study on the behaviour of rubber aggregates concrete beams using analytical approach. *Engineering Science and Technology, an International Journal*, 20(1), 151-159.

[44] B.G. Tabachnick, L.S. Fidell, J.B. Ullman. (2007). Using multivariate statistics. Pearson Boston, MA.

[45] Q.H. Nguyen, H.-B. Ly, L.S. Ho, N. Al-Ansari, H.V. Le, V.Q. Tran, I. Prakash, B.T. Pham. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering,* 2021(1), 4832864.

[46] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30.

[47] S. Mangalathu, S.-H. Hwang, J.-S. Jeon. (2020). Failure mode and effects analysis of RC members based on machine-learning-based SHapley Additive exPlanations (SHAP) approach. *Engineering Structures*, 219, 110927.