Journal of Science and Transport Technology

University of Transport Technology

Forecast of surface chloride concentration of concrete utilizing ensemble decision tree boosted

Anh-Tuan Tran[1], Thanh-Hai Le[1], May Huu Nguyen[1,2,*]

[1]University of Transport Technology, Hanoi 100000, Vietnam

[2]Civil and Environmental Engineering Program, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8527, Japan

Article info

Type of article:
Original research paper

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

Received:
December 08, 2021
Accepted:
March 03, 2022
Published:
March 25, 2022

Abstract: This study proposes the application of Ensemble Decision Tree Boosted (EDT Boosted) model for forecasting the surface chloride concentration of marine concrete C_s. A database of 386 experimental results was collected from 17 different sources covering twelve variables was used to build and verify the predictive power of the EDT model. The input factors considered the changes in eleven variables, including the contents of cement, fly ash, blast furnace slag, silica fume, superplasticizer, water, fine aggregate, coarse aggregate, annual mean temperature, chloride concentration in seawater, and exposure time. The results indicate that EDT Boosted is a good predictor of C_s as verified via good performance evaluation criteria, i.e., R^2, RMSE, MAE, MAPE values were 0.84, 0.16, 0.17, and 17%, respectively. Partial dependence plot (PDP) was then developed to correlate the eleven input variables with the C_s. PDP implied that the strongest factor affecting Cs was the amount of fine aggregate content, chloride concentration, exposure time, amount of cement, and water, which is useful for material engineers in the design of the grade.

Keywords: Machine learning, Ensemble Descion Tree Boosted (EDT Boosted), surface chloride concentration.

1. Introduction

Reinforcement corrosion is the most widespread problem affecting concrete structures' durability, safety, and sustainability exposed to marine environments [1,2]. In marine environments, chlorides can penetrate through concrete cover to break down the protective layer of the reinforcement and cause corrosion [3]. Thus, properly assessing the surface chloride penetration is crucial for controlling the durability of a marine concrete structure.

Theoretically, the chloride ingress into concrete can be assessed by Fick's second law of diffusion [4]. An analytical expression based on Fick's second law to calculate the chloride concentration in concrete is given in Eq. (1), as proposed by Oslakovic et al. [5].

$$C(x,t) = C_0 + (C_s - C_0)\left[1 - erf\left(\frac{x}{2\sqrt{D \times t}}\right)\right] \quad (1)$$

where C (x, t) denotes chloride concentration in the depth x from the exposed surface after exposure time t, C_0 and C_s are the initial and the apparent surface chloride concentrations, respectively in concrete, D is the apparent chloride

diffusion coefficient, and *erf* is the complementary error function. Basically, C_0 is a constant, thus *C (x, t)* depends on the changes in C_s and *D*. Although the effects of chloride diffusion coefficient *D* have been well elucidated in the literature, the debate on how the apparent surface chloride concentration C_s may alter the chloride concentration has been an ongoing debate with inconsistent conclusions [6,7].

Indeed, C_s plays a vital role in reflecting the influence of the surrounding environment and provides the boundary condition for evaluating the durability and service life of a marine concrete structure [8]. Numerous researchers have investigated and proposed numerical models to calculate the C_s value [9–11]. For instance, Yang et al. [11] proposed a computational model for surface chloride concentration of concrete in marine atmosphere zone. Meira et al. [12] evaluated the durability of concrete structures in marine atmosphere zones utilizing chloride deposition rate on the wet candle as an environmental indicator. Nevertheless, these models revealed many shortcomings in predicting C_s due to a lack of considering impact factors, i.e., natural condition, concrete properties, and exposure time. In another approach, time-variant mathematic models were also proposed to calculate C_s utilizing logarithmic, power and exponential functions [13,14]. However, these models cannot reasonably explain the rapid changes in C_s in the early stage and tendency to be steady at the later stage. Besides, the roles of material composition and natural conditions were neglected in the models. Based on materials and environmental conditions, some other models were proposed but they ignored the vital effect of exposure time [11,15]. Consequently, traditional models cannot or partly consider all the influential factors on C_s due to the limitations of the number of experimental databases. Therefore, the development of a new method for predicting C_s that is more comprehensive and accurate is now an inevitable need.

Over the last two decades, machine learning (ML) methods have been broadly applied in civil engineering [16,17]. Artificial neural network (ANN), random forest (RF), support vector machine (SVM), Ensemble Bagged Trees (EDT Bagged), and Ensemble Boosted Trees (EDT Boosted) are some of them. Several ML models have been proposed to predict corrosion of reinforcement bars in concrete [18]. For instance, Ahmad et al. [19] presented the application of novel ML models for predicting the surface chloride concentration in concrete containing waste material. The three models, namely, ANN, DT, and gene expression programming (GEP), were utilized for the investigations. The database consisted of 12 input parameters, including concrete mix proportions, natural conditions, and exposure time. The obtained results indicated that the models were valid for predicting the surface chloride concentration, deprived of the difficulties of laboratory results. Recently, Cai et al. [18] also proposed a novel approach for predicting surface chloride concentration of marine concrete using ensemble machine learning. The five ML models such as linear regression (LR), Gaussian process regression (GPR), SVM, multilayer perceptron artificial neural network (MLP-ANN), and Ramdon forest (RF) model were established based on 642 experimental databases. As expected, this model can easily take into account the influence of twelve factors and has superior prediction performance.

Take part in that flow, the objective of this study is to investigate the application of ensemble decision tree algorithm for predicting surface chloride concentration of marine concrete. To realize this goal, a database of 386 experimental results was used to build a combined ML model based on a decision tree algorithm using a boosting technique called Ensemble Decision Tree (EDT). The cross-validation technique also limits the overfitting phenomenon in the model training process. The predictive performance of the model is evaluated through four criteria: the coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and mean

absolute percentage error (MAPE). The sensitivity analysis using a partial dependency plot (PDP) was then performed to evaluate the effect of each input variable on the surface chloride concentration of marine concrete.

2. Database construction

As mentioned above, in this study, the collected data consists of 386 experimental data collected from 17 experimental research inspired by Cai et al. [18]. The detailed description of the data is described in Table 1. The database consists of eleven parameters of concrete mixture design, exposure time, and environmental conditions provided as inputs for the EDT model, whereas C_s acts as the output. The concrete mixture design parameters included in the study are eight variables representing the contents (units of kg/m^3) of cement (C), fly ash (FA), blast furnace slag (GGBS), silica fume (SF), superplasticizer (SP), water (W), fine aggregate (F. Agg), and coarse aggregate (C. Agg). The changes in environmental conditions are represented via annual mean temperature (T, °C), and chloride concentration in seawater (Cl^-, g/L). Exposure time (E.T, year) is also considered as one input.

Table 1. Database collated in this study

No.	Reference	No. of samples	Proportions (%)
1	Costa and Appleton [20]	8	2.1
2	Chalee et al. [10]	108	28.0
3	Naukuttan et al. [21]	38	9.8
4	Pack et al. [22]	44	11.4
5	Valipour et al. [23]	5	1.3
6	Ghods et al. [24]	6	1.6
7	Markeset et al. [25]	84	21.8
8	Farahani et al. [26]	8	2.1
9	Safehian and Ramezaniapour [27]	9	2.3
10	Xue et al. [28]	5	1.3
11	Mohammed and Hamada [29]	5	1.3
12	Safehian and Ramezaniapour [30]	4	1.0
13	Zhang and Wei [31]	9	2.3
14	Moradllo et al. [32]	34	8.8
15	Zhang et al. [33]	5	1.3
16	Gao et al. [34]	9	2.3
17	Wang et al. [35]	5	1.3
Total		**386**	**100**

Fig. 1 shows the distribution chart and correlation among input and output parameters employed in this study. Here, most of the input variables in the database covered a wide range of values. In detail, C content varies from 100 to 500 (kg/m^3), but the values are mainly in the 200 - 400 (kg/m^3) range. The FA, GGBS, SF, SP, W, F. Agg, C. Agg contents are in the range of 0÷250, 50÷300, 0÷50, 0÷8, 50÷350, 0÷120, 0÷180 (kg/m^3), respectively. The changes in M. Tem and Cl^- vary from 5 to 35°C, and 5 to 30 g/L, respectively. Exposure time changes from 0.5 to 50 years, but most of them range from 0.5 to 15 years. In addition, the Pearson correlation coefficient (r_s) was calculated and noted in each pair of correlations. R_s was employed to evaluate the correlation between used parameters, particularly between output C_s and other parameters. The

results indicated no strong correlation among input and output parameters as observed via low r_s, i.e., r_s <0.49. This observation implied that the employed variables taken in this study were all independent and can be used to build the correlation with C_s. Finally, to minimize the errors generated during simulation by EDT, this dataset is normalized to the range of values 0-1 to limit errors generated by numerical simulations.

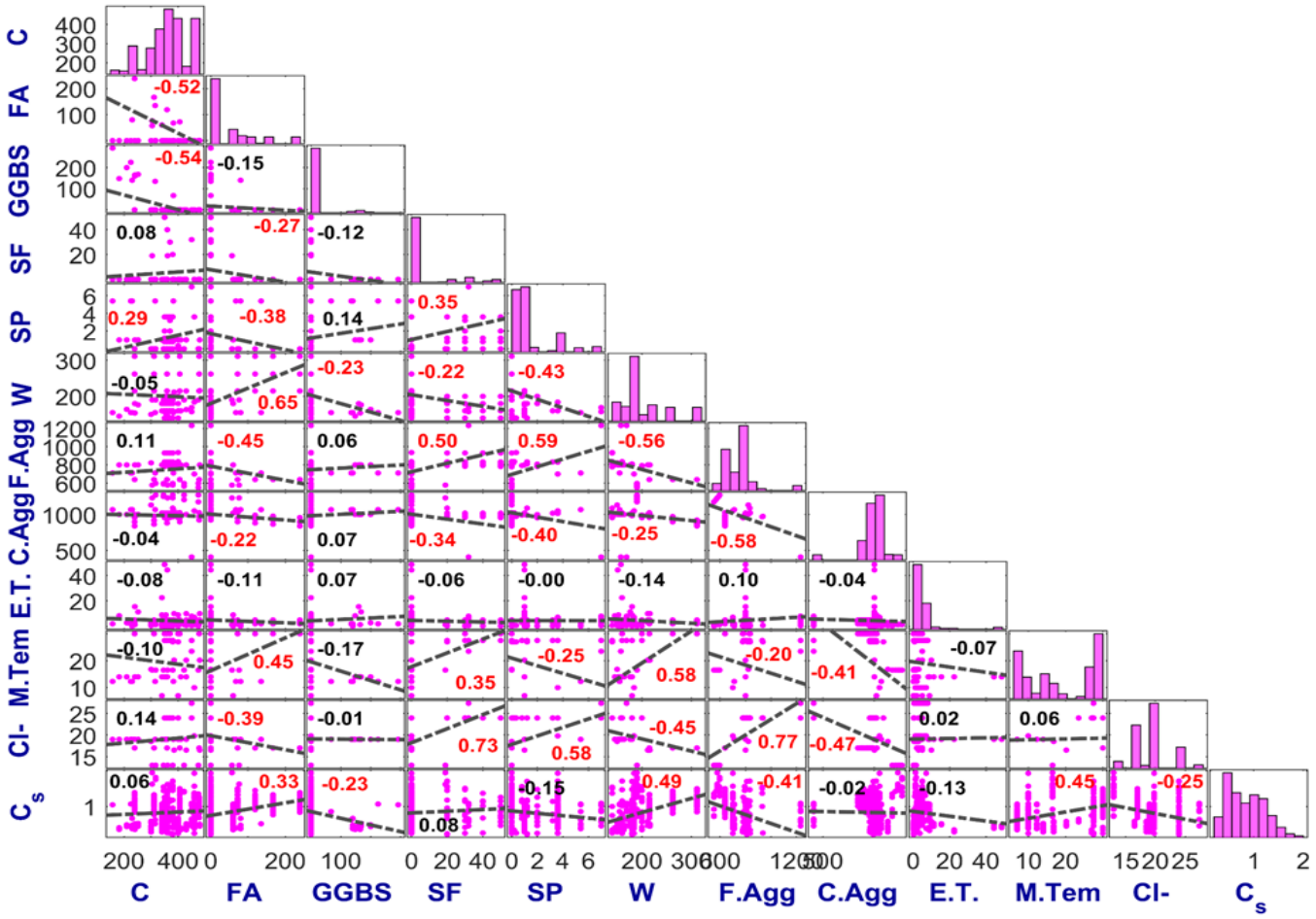

Fig. 1. The distribution chart and correlation between input and output parameters

3. Machine learning methods

3.1. Decision trees

A decision tree algorithm is a nonparametric categorized data configuration applying the divide-and-conquer approach. It solves a complicated regression crisis by diving them into smaller issues and recursively utilizes the same procedure to the sub-problems. Sub-problems can be linked to yield a result of the complicated problem. Advantages of this approach include dividing the instance-space into subspaces wherein each subspace is tailored with various models and addresses many datasets (cases and variables). A typical decision tree includes three types of nodes, namely, decision, chance, and end nodes. Decision nodes work as exam functions with discrete products labeling the branches. Upon input values, exams are utilized, and the corresponding branches are selected [36].

3.2. Ensemble methods

In general, ensemble methods combine several models, wherein each solves the same original task, to improve the generality capacity. Among significant ensemble types methods, Ensemble Bagged Trees and Ensemble Boosted Trees are two commonly used. Ensemble Bagged Trees takes numerous bootstrap random tests from the database to create a new training database. The procedure is repeated up to a

significant subset of the training database is created, and similar results can be withdrawn larger than once [37]. In the Ensemble Boosted Trees method, the dataset is trained in sequence with development from the original to the next model. It uses a data point weight in the training dataset to create various models. The final goal of this method is the weighted mean of the output from the database [37]. In this study, a combination between the decision tree algorithm and Ensemble Boosted Tree, namely, Ensemble Decision Tree Boosted (EDT Boosted), was proposed to predict surface chloride concentration of marine concrete.

3.3. K-Fold cross-validation

In machine learning, cross-validation is a common technique used in training and editing models to overcome overfitting [38]. The database was divided into two parts: training and test datasets. The training dataset will be randomly divided into K equal parts, and each training time will choose 1 part as the validation data and (K-1) the rest as the training data. The model training will be done in K iterations. The final model evaluation result will be the average of the evaluation results of K training times. The choice of K must be appropriate because if K is too large, the training data set will be much larger than the control data set, and the evaluation results will not reflect the true nature of the machine learning method, especially with large data sets. In this study, K=10 was selected to consider a previous work's suggestion [39], and briefly explained in Fig. 2.

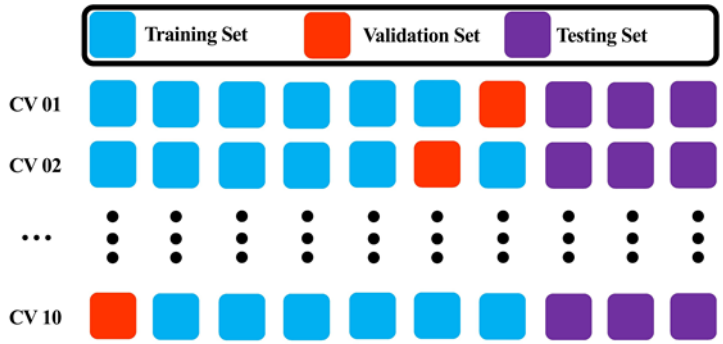

Fig. 2. Cross-validation technique with 10-fold used in this study

3.4. Partial dependence plot

Partial dependence plot (PDP) is a visualization technique that describes the relationships between one or more input variables and the predicted values. PDP allows quantifying how a change in an input variable affects the target variable. The integral of each PDP curve is calculated and serves as an important index to quantify the influence of each input variable [40].

PDP can show the relationship for the linear regression model, and the partial dependence function for regression is calculated by the following formula:

$$f_{x_A}(x_A) = E_{x_M}[f(x_A, x_M)] = \int f(x_A, x_M)\, dP(x_M) \quad (2)$$

where $\mathbf{x_A}$ consists of features for which the partial dependence function will be plotted, $\mathbf{x_M}$ other features used in the machine learning model q. **A** are the features that the user wants to consider and predict (only one or two features inset A). The features in set **A** are those that influence the prediction results, which we want to know. The feature vectors xA and xM makes the total feature space x. PDP does by marginalizing the machine learning model output over the distribution of the features in set **M**; thus, the function presents the relationship between the features in set **A**. By removing the other features, we have a function that only depends on features in A and still interacts with other features.

PDP is a global method because it considers all instances and shows a global relationship of a feature with the predicted outcome. PDP also presents the probability for a specific layer in the different values for features in set A. Thus, PDP can solve the problems of multiple layers by drawing a line or plot for each layer.

3.4. Performance assessment

In this study, the coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and mean absolute percentage error (MAPE) were calculated as statistical measurements to evaluate the accuracy of the EDT model. R^2 is the coefficient of determination, which shows the data's appropriateness for the method, and it ranges from 0 to 1. R^2 values around 0 indicate poor model performance, while R^2 values near 1 indicate strong model accuracy. RMSE is a fundamental criterion for evaluating predictive modeling performance. The RMSE is a standard way to express the error's mean size. RMSE is very sensitive to big error levels. As a result, the model error is more stable when the RMSE is near the MAE. RMSE, like MAE, does not show the difference between the model's output value and the actual value and is in the range of (0; +∞). MAPE is the measure that fulfills the standards of trustworthiness, ease of explanation, and presentation. The following equations reflect these values:

$$R^2 = 1 - \left[\frac{\Sigma_{i=1}^{n}(P_i - E_i)^2}{\Sigma_{i=1}^{n}(P_i)^2}\right] \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(P_i - E_i)^2} \quad (4)$$

$$MAE = \frac{1}{n}\Sigma_{i=1}^{n}|P_i - E_i| \quad (5)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_i - E_i}{P_i}\right| \times 100\% \quad (6)$$

where *E* is the actual experimental value, *P* is the predicted value, calculated according to the model's prediction, *n* is the number of samples in the database.

4. Results and discussion

In this section, the EDT Boosted model building process is performed. In detail, this process consists of two phases: the training phase, which is the process of training the model accompanied by cross-validation (CV) with ten folds EDT Boosed model with default parameters. At a later stage, when the EDT achieves optimal predictive performance on the training dataset, it is used to evaluate the testing dataset. The training dataset (accounting for 70% of the database) was divided into ten parts to conduct a 10-fold CV. To fully evaluate the models' robustness, such a process was repeated ten times in simulating different training and testing data sets. Finally, the testing dataset (the remaining 30% of data) is used to test the model's predictive ability for unknown data. The EDT Boosted prediction performance evaluation results for both data sets are shown in Fig. 4. It is noted that the R^2 CV scores of the proposed model were, in all cases, higher than 0.68.

In general, the obtained results indicated that the changes in training parts lead to the corresponding variation of the prediction power of the EDT Boosted model. The performance evaluation criteria all change within certain intervals, but the amplitude of fluctuations is assessed to be relatively stable. In detail, the obtained R^2 values fluctuate around 0.84 and are almost stable throughout different 10-fold CV runs (see Fig. 3a). The RMSE values vary around 0.15, with the highest and lowest RMSE values being 0.17 and 0.14 at CV5 and CV9, respectively (Fig. 3b), in which CV5 denotes the fifth simulation of random shuffling of training and testing parts. A similar result was also observed in terms of MAE and MAPE, where values range around 0.11 and 16%, respectively (Fig. 3c, d). The results suggested that, in the training parts, the trained

EDT Boosted model has good predictive capacity, which can be selected for evaluation on the testing dataset.

The testing dataset consists of 116 experimental data that were independent during training and validation stages by 10-fold CV. The results indicated that the proposed EDT Boosted model has good predictive power as verified via evaluation criteria. Moreover, there was no overfitting phenomenon due to the capacity of the above training set being better than the control dataset. The results of EDT Boosted model when forecasting new data were quite good as indicated via R^2, RMSE, MAE, MAPE values were around 0.83, 0.16, 0.17, and 17%, respectively. It can be seen that the difference between training and testing parts was insignificant, which suggested the usability of the proposed EDT Boosted model. In particular, based on R^2, RMSE, MAE, MAPE results, CV6 was the best run. Thus, its result was selected to discuss in the following parts.

Fig. 4 compares the target and output of concrete's surface chloride concentration (C_s) utilizing the EDT Boosted model for both training and resting parts. The obtained results indicated that the predicted values were close to the experimental ones for both parts.

The target and output data relationships were given in the form of regression plots in Fig. 5. The obtained results indicated that the linear fit lines were close to the ideal regression line for both training and testing tests. Based on the analysis results, it can be confirmed that the EDT Boosted models successfully predicted the C_S of marine concrete.

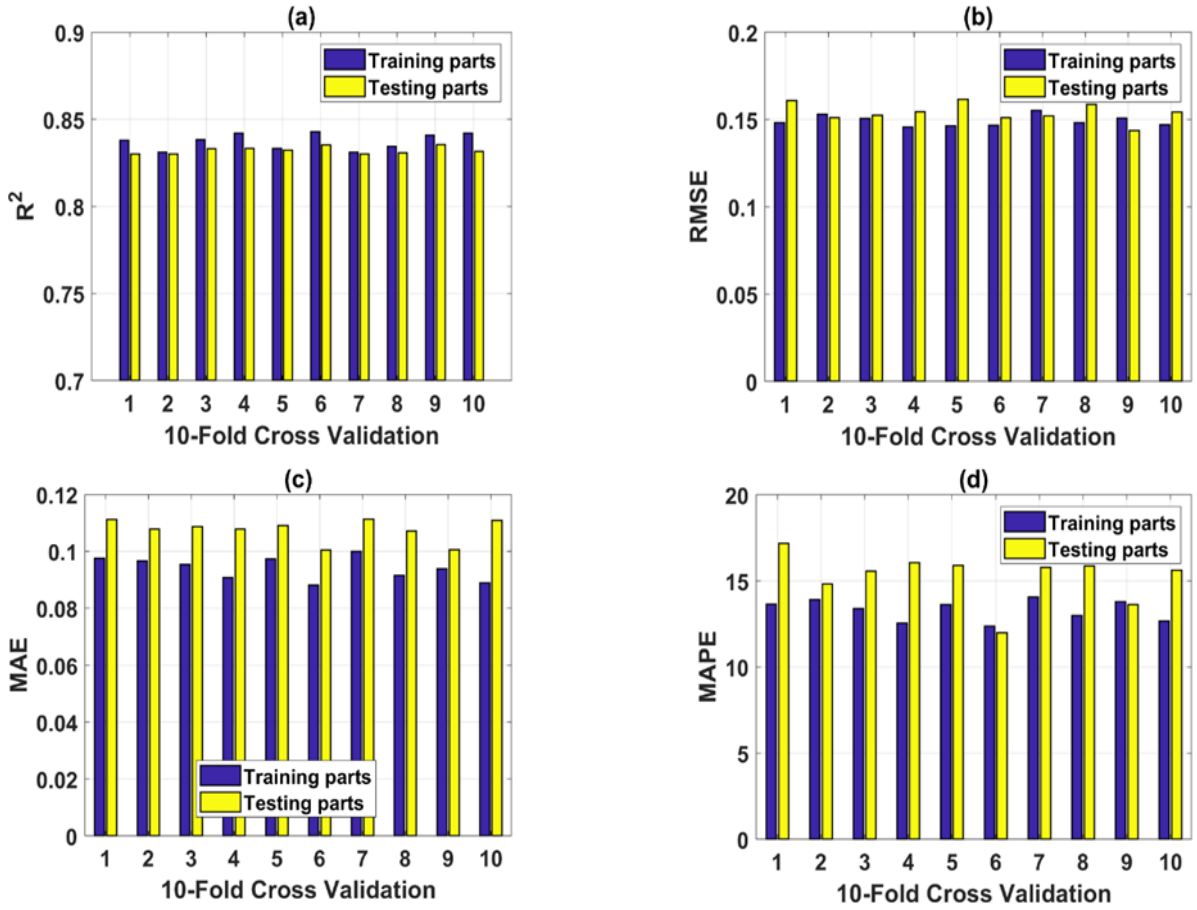

Fig. 3. Results of training and testing of 10-fold cross-validated EDT model based on different performance evaluation criteria: (a) R^2, (b) RMSE, (c) MAE, and (d) MAPE.

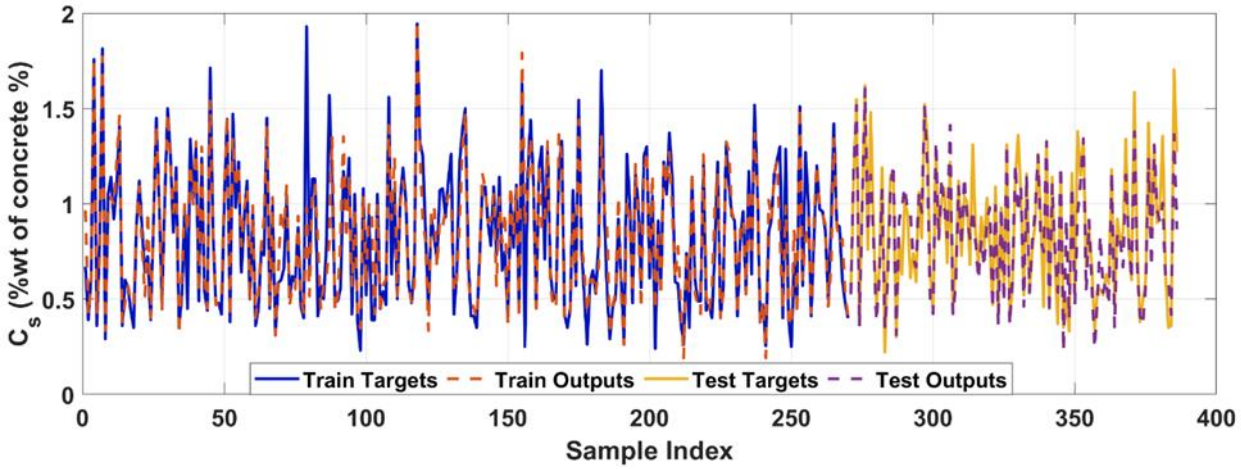

Fig. 4. Comparison between experimental and predicted surface chloride concentration (Cs) of concrete utilizing EDT Boosted model for both training and resting parts

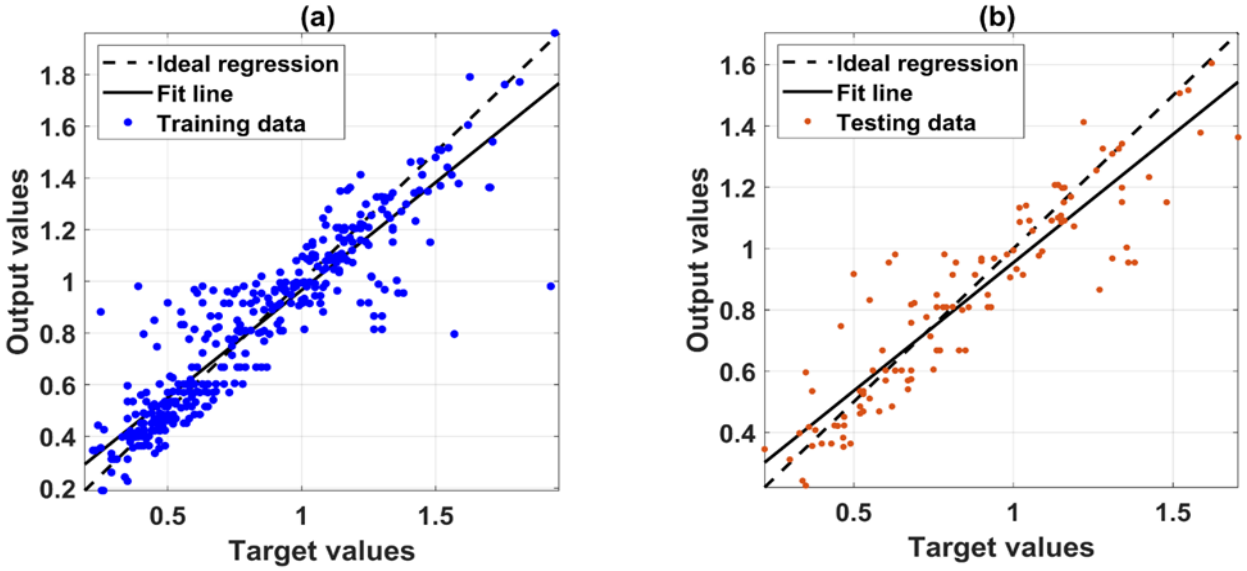

Fig. 5. Correlation analysis between target and output values of C_s for (a) training dataset, (b) testing dataset

As early mentioned, surface chloride concentration (C_s) of concrete is affected by concrete properties, environmental factors, and exposition period. Thus, this part aims to examine the effect of changing the value of an independent variable on a specific dependent variable under a specified set of assumptions.

Based on the optimal EDT Boosted (i.e., at CV9) model in the previous section, PDP analysis is employed to interpret the prediction results in function of the input variables. Here, the effect of each input variable on C_s is computed by varying the value in the corresponding range, while keeping constant all the values of the remaining parameters. The influences of C, FA, GGBS, SF, SP, W, F. Agg, C. Agg, E. T, M. Tem, and Cl^- on C_s are presented in Fig. 6. The PDP of C_s gradually increased with increasing C, W, and Cl^- (see Figs. 6a, f, and k) due to increased porosity and chloride ion concentration [41]. In contrast, The PDP of C_s gradually decreased with increasing GGBS. The influences of FA, SF, SP, F. Agg, and M. Tem on PDP of C_s were insignificantly (Figs. 6b,d, e, g, and j), in agreement with previous works [42]. In particular, the influence of C. Agg on PDP of C_s was significant, similar to the literature finding [43]. The results also implied that the most decisive factor affecting the PDP of C_s was F. Agg; the following orders were Cl^-, E. T, C, and W.

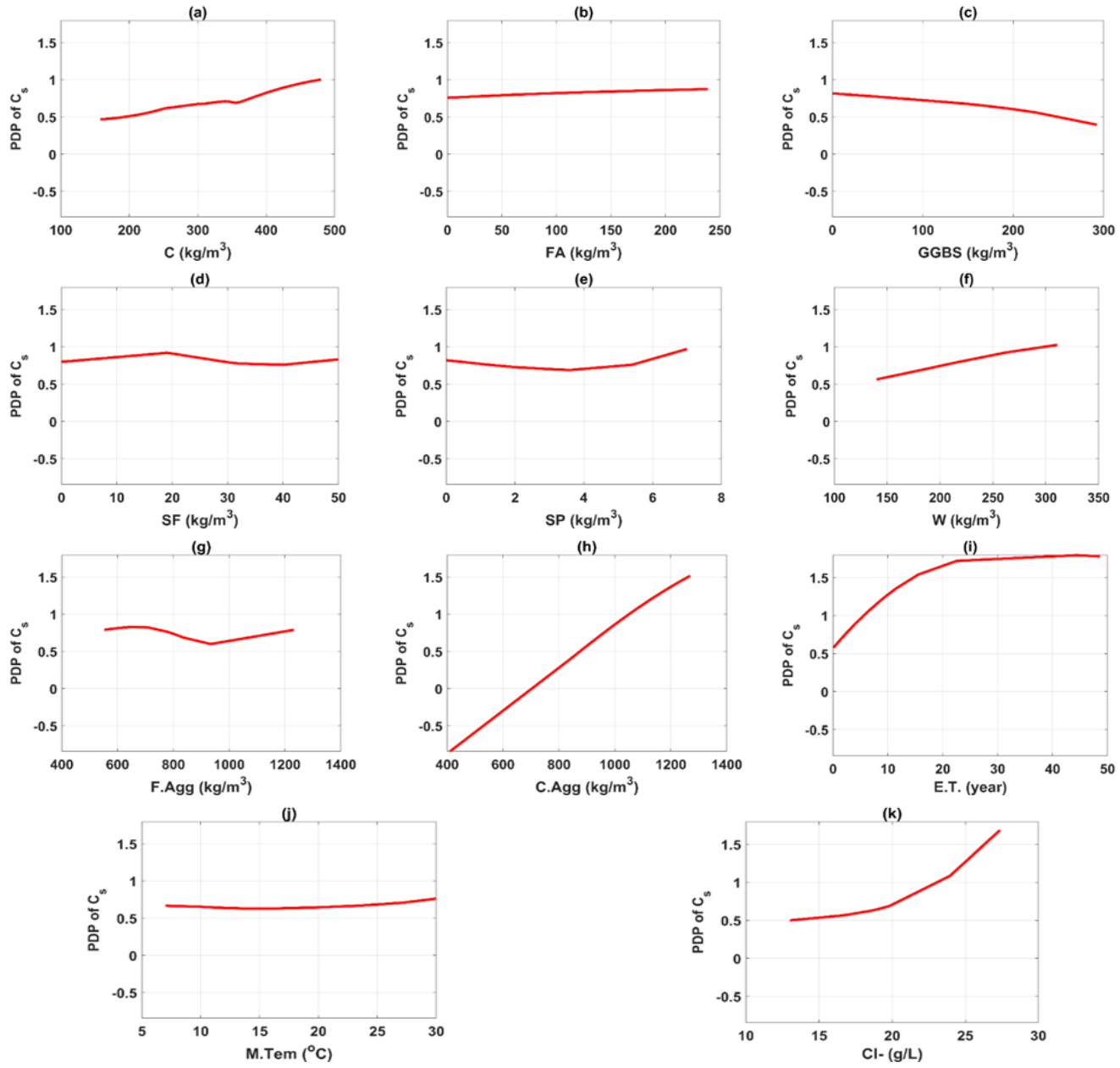

Fig. 6. Feature importance of variables used in this study

6. Conclusions

In this study, to predict surface chloride concentration (Cs) of marine concrete, a combined ML model based on a decision tree algorithm using a boosting technique called Ensemble Decision Tree (EDT) was developed. The EDT Boosted model takes advantage of the decision tree method and overcomes the disadvantages of single tree models to improve performance and predictability. A database of 386 experimental results was collected from 17 different sources covering twelve variables. After evaluating the correlation, these twelve variables can all be considered independent of each other and are selected as input parameters when building the EDT Boosted model. During the training phase of the EBT model, the 10-fold cross-validation technique was applied to limit the overfitting phenomenon. The predictive performance of the proposed model is evaluated through 4 statistical criteria: R^2, RMSE, MAE, and MAPE. The results indicated that EDT Boosted is a good model surface chloride concentration (Cs) of marine concrete as proved via high and stable predictive performance. Moreover, by the partial dependence graph (PDP) technique, the research has analyzed the influence of twelve input

parameters on surface chloride concentration (Cs) of marine concrete. The results also implied that the most vital factor affecting the PDP of Cs was the fine aggregate content, followed by Cl–, E. T, C, and W. The research results help the engineer in the mix design to rely on the above qualitative and quantitative analysis to determine the content of the components in the concrete mix to enhance the chloride resistance of the marine concrete.

References

[1] R.E. Melchers, C.Q. Li, Reinforcement corrosion initiation and activation times in concrete structures exposed to severe marine environments, Cement and Concrete Research. 39 (2009) 1068–1076. https://doi.org/10.1016/j.cemconres.2009.07.003.

[2] A. James, E. Bazarchi, A.A. Chiniforush, P. Panjebashi Aghdam, M.R. Hosseini, A. Akbarnezhad, I. Martek, F. Ghodoosi, Rebar corrosion detection, protection, and rehabilitation of reinforced concrete structures in coastal environments: A review, Construction and Building Materials. 224 (2019) 1026–1039. https://doi.org/10.1016/j.conbuildmat.2019.07.250.

[3] H. Beushausen, R. Torrent, M.G. Alexander, Performance-based approaches for concrete durability: State of the art and future research needs, Cement and Concrete Research. 119 (2019) 11–20. https://doi.org/10.1016/j.cemconres.2019.01.003.

[4] W. Chalee, C. Jaturapitakkul, Effects of W/B ratios and fly ash finenesses on chloride diffusion coefficient of concrete in marine environment, Mater Struct. 42 (2009) 505–514. https://doi.org/10.1617/s11527-008-9398-2.

[5] I. Stipanovic Oslakovic, D. Bjegovic, D. Mikulic, Evaluation of service life design models on concrete structures exposed to marine environment, Mater Struct. 43 (2010) 1397–1412. https://doi.org/10.1617/s11527-010-9590-z.

[6] Y. Yoshizumi, K. Nakarai, M.H. Nguyen, N. Tada, M. Yamamoto, Chloride Penetration Test of Concrete Simulating Deicing Salt Attack, in: J.N. Reddy, C.M. Wang, V.H. Luong, A.T. Le (Eds.), ICSCEA 2019, Springer, Singapore, 2020: pp. 431–436. https://doi.org/10.1007/978-981-15-5144-4_38.

[7] T. Luping, J. Gulikers, On the mathematics of time-dependent apparent chloride diffusion coefficient in concrete, Cement and Concrete Research. 37 (2007) 589–595. https://doi.org/10.1016/j.cemconres.2007.01.006.

[8] L. Yang, R. Cai, B. Yu, Modeling of environmental action for submerged marine concrete in terms of surface chloride concentration, Structural Concrete. 19 (2018) 1512–1520. https://doi.org/10.1002/suco.201800072.

[9] L.F. Yang, R. Cai, B. Yu, Investigation of computational model for surface chloride concentration of concrete in marine atmosphere zone, Ocean Engineering. 138 (2017) 105–111. https://doi.org/10.1016/j.oceaneng.2017.04.024.

[10] W. Chalee, C. Jaturapitakkul, P. Chindaprasirt, Predicting the chloride penetration of fly ash concrete in seawater, Marine Structures. 22 (2009) 341–353. https://doi.org/10.1016/j.marstruc.2008.12.001.

[11] P.F. Marques, A. Costa, F. Lanata, Service life of RC structures: chloride induced corrosion: prescriptive versus performance-based methodologies, Mater Struct. 45 (2012) 277–296. https://doi.org/10.1617/s11527-011-9765-2.

[12] G.R. Meira, C. Andrade, C. Alonso, J.C. Borba, M. Padilha, Durability of concrete structures in marine atmosphere zones – The use of chloride deposition rate on the wet

candle as an environmental indicator, Cement and Concrete Composites. 32 (2010) 427–435. https://doi.org/10.1016/j.cemconcomp.2010.03.002.

[13] H.-W. Song, C.-H. Lee, K.Y. Ann, Factors influencing chloride transport in concrete structures exposed to marine environments, Cement and Concrete Composites. 30 (2008) 113–121. https://doi.org/10.1016/j.cemconcomp.2007.09.005.

[14] M.K. Kassir, M. Ghosn, Chloride-induced corrosion of reinforced concrete bridge decks, Cement and Concrete Research. 32 (2002) 139–143. https://doi.org/10.1016/S0008-8846(01)00644-5.

[15] DuraCrete Final Technical Report-General Guidelines for Durability Design and Redesign, Project Document BE95-1347/R17, Denmark (2000), (n.d.).

[16] T.-H. Le, H.-L. Nguyen, B.T. Pham, M.H. Nguyen, C.-T. Pham, N.-L. Nguyen, T.-T. Le, H.-B. Ly, Artificial Intelligence-Based Model for the Prediction of Dynamic Modulus of Stone Mastic Asphalt, Applied Sciences. 10 (2020) 5242. https://doi.org/10.3390/app10155242.

[17] H.-B. Ly, M.H. Nguyen, B.T. Pham, Metaheuristic optimization of Levenberg–Marquardt-based artificial neural network using particle swarm optimization for prediction of foamed concrete compressive strength, Neural Comput & Applic. (2021). https://doi.org/10.1007/s00521-021-06321-y.

[18] R. Cai, T. Han, W. Liao, J. Huang, D. Li, A. Kumar, H. Ma, Prediction of surface chloride concentration of marine concrete using ensemble machine learning, Cement and Concrete Research. 136 (2020) 106164. https://doi.org/10.1016/j.cemconres.2020.106164.

[19] A. Ahmad, F. Farooq, K.A. Ostrowski, K. Śliwa-Wieczorek, S. Czarnecki, Application of Novel Machine Learning Techniques for Predicting the Surface Chloride Concentration in Concrete Containing Waste Material, Materials. 14 (2021) 2297. https://doi.org/10.3390/ma14092297.

[20] A. Costa, J. Appleton, Chloride penetration into concrete in marine environment—Part I: Main parameters affecting chloride penetration, Mat. Struct. 32 (1999) 252. https://doi.org/10.1007/BF02479594.

[21] Nanukuttan, S. V., Basheer, L., McCarter, W. J., Robinson, D. J., & Basheer, P. M. (2008). Full-Scale Marine Exposure Tests on Treated and Untreated Concretes--Initial 7-Year Results. ACI Materials Journal, 105(1), 81.

[22] S.-W. Pack, M.-S. Jung, H.-W. Song, S.-H. Kim, K.Y. Ann, Prediction of time dependent chloride transport in concrete structures exposed to a marine environment, Cement and Concrete Research. 40 (2010) 302–312. https://doi.org/10.1016/j.cemconres.2009.09.023.

[23] M. Valipour, F. Pargar, M. Shekarchi, S. Khani, M. Moradian, In situ study of chloride ingress in concretes containing natural zeolite, metakaolin and silica fume exposed to various exposure conditions in a harsh marine environment, Construction and Building Materials. 46 (2013) 63–70. https://doi.org/10.1016/j.conbuildmat.2013.03.026.

[24] P. Ghods, M. Chini, R. Alizadeh, M. Hoseini, M. Shekarchi, P. Ghods, The Effect of Different Exposure Conditions on the Chloride Diffusion into Concrete in the Persian Gulf Region.

[25] G. Markeset, O. Skjølsvold, Time dependent chloride diffusion coefficient : field studies of concrete exposed to marine environment in Norway, RILEM Publications, 2010. https://oda.oslomet.no/oda-xmlui/handle/10642/655 (accessed November 28, 2021).

[26] A. Farahani, H. Taghaddos, M. Shekarchi, Prediction of long-term chloride diffusion in silica fume concrete in a marine environment,

Cement and Concrete Composites. 59 (2015) 10–17. https://doi.org/10.1016/j.cemconcomp.2015.03.006.

[27] M. Safehian, A. Ramezanianpour, Prediction of RC structure service life from field long term chloride diffusion, Computers and Concrete. 15 (2015) 589–606. https://doi.org/10.12989/cac.2015.15.4.589.

[28] Chloride ion penetration into concrete exposed to marine environment for a long period. The Ocean Engineering. 2015

[29] T.U. Mohammed, H. Hamada, Relationship between free chloride and total chloride contents in concrete, Cement and Concrete Research. 33 (2003) 1487–1490. https://doi.org/10.1016/S0008-8846(03)00065-6.

[30] M. Safehian, A.A. Ramezanianpour, Assessment of service life models for determination of chloride penetration into silica fume concrete in the severe marine environmental condition, Construction and Building Materials. 48 (2013) 287–294. https://doi.org/10.1016/j.conbuildmat.2013.07.006.

[31] Zhang, B. L., & Wei, S. S. (1998). Decade exposure test of harbor reinforced concrete for 10 years. South China Harbour Eng., 98(2), 32-40.

[32] M. Khanzadeh Moradllo, S. Sadati, M. Shekarchi, Quantifying maximum phenomenon in chloride ion profiles and its influence on service-life prediction of concrete structures exposed to seawater tidal zone – A field oriented study, Construction and Building Materials. 180 (2018) 109–116. https://doi.org/10.1016/j.conbuildmat.2018.05.284.

[33] Y. Zhang, Y. Zhang, J. Huang, H. Zhuang, J. Zhang, Time dependence and similarity analysis of peak value of chloride concentration of concrete under the simulated chloride environment, Construction and Building Materials. 181 (2018) 609–617. https://doi.org/10.1016/j.conbuildmat.2018.06.030.

[34] Y. Gao, J. Zhang, S. Zhang, Y. Zhang, Probability distribution of convection zone depth of chloride in concrete in a marine tidal environment, Construction and Building Materials. 140 (2017) 485–495. https://doi.org/10.1016/j.conbuildmat.2017.02.134.

[35] Y. Wang, L. Wu, Y. Wang, Q. Li, Z. Xiao, Prediction model of long-term chloride diffusion into plain concrete considering the effect of the heterogeneity of materials exposed to marine tidal zone, Construction and Building Materials. 159 (2018) 297–315. https://doi.org/10.1016/j.conbuildmat.2017.10.083.

[36] W.Z. Taffese, E. Sistonen, J. Puttonen, CaPrM: Carbonation prediction model for reinforced concrete using machine learning methods, Construction and Building Materials. 100 (2015) 70–82. https://doi.org/10.1016/j.conbuildmat.2015.09.058.

[37] W.W. Hsieh, Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels, Cambridge University Press, Cambridge, 2009. https://doi.org/10.1017/CBO9780511627217.

[38] Y. Bengio, Y. Grandvalet, No Unbiased Estimator of the Variance of K-Fold Cross-Validation, Journal of Machine Learning Research. 5 (2004) 1089–1105.

[39] Kohavi, R., 1995, August. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai (Vol. 14, No. 2, pp. 1137-1145).

[40] T. Parr, J.D. Wilson, Partial dependence through stratification, Machine Learning with Applications. 6 (2021) 100146. https://doi.org/10.1016/j.mlwa.2021.100146.

[41] M.H. Nguyen, K. Nakarai, Y. Kubori, S. Nishio, Validation of simple nondestructive method for evaluation of cover concrete quality,

Construction and Building Materials. 201 (2019) 430–438. https://doi.org/10.1016/j.conbuildmat.2018.12.109.

[42] J.R. Hino Junior, C.E.T. Balestra, R.A. Medeiros-Junior, Comparison of test methods to determine resistance to chloride penetration in concrete: Sensitivity to the effect of fly ash, Construction and Building Materials. 277 (2021) 122265. https://doi.org/10.1016/j.conbuildmat.2021.122265

[43] J. Bao, S. Li, P. Zhang, X. Ding, S. Xue, Y. Cui, T. Zhao, Influence of the incorporation of recycled coarse aggregate on water absorption and chloride penetration into concrete, Construction and Building Materials. 239 (2020) 117845. https://doi.org/10.1016/j.conbuildmat.2019.117845