Research Article

Omar Alomari and Mohammad Al-Rawashdeh*

Application of soft computing in estimating primary crack spacing of reinforced concrete structures

https://doi.org/10.1515/cls-2022-0194
received October 29, 2022; accepted April 26, 2023

Abstract: The investigation related to the serviceability analysis, particularly in terms of crack spacing prediction, has remarkably increased recently. In addition, the prediction of serviceability analysis is highly dependent and influenced by different physical and material factors that contribute to the crack spacing of reinforced concrete (RC) structures. As a result, the cracking phenomenon has not been fully grasped due to these factors' wide variety and complexity. Recently, soft computing techniques have gained considerable popularity due to their capability of learning and producing generalized solutions and exhibiting desirable performance in terms of time, effort, and cost. However, the literature on crack spacing prediction using various machine learning approaches is limited and insufficient. Therefore, this article is dedicated to estimating the primary crack spacing of RC structures using different machine learning methods. As a part of the study, the findings of these approaches will be computed and compared to the benchmark experimental results. Besides, the results of the developed models will be compared against that of available approaches in the literature to highlight their reliability. Furthermore, a parametric assessment will be conducted to emphasize the most influencing input parameter on the primary crack spacing of RC structures.

Keywords: reinforced concrete, serviceability analysis, crack spacing prediction, machine learning techniques

*** Corresponding author: Mohammad Al-Rawashdeh,** Department of Civil Engineering, Faculty of Engineering, Al-Balqa Applied University, As-Salt, Jordan, e-mail: dr_alrawashdeh@bau.edu.jo
Omar Alomari: Department of Architecture, Faculty of Engineering, Al-Balqa Applied University, As-Salt, Jordan
ORCID: Mohammad Al-Rawashdeh 0000-0002-4239-9166

1 Introduction

Concrete is considered a globally important structural material in the field of construction due to its advantages in tolerating intense compressive loads, adequately resisting hostile conditions, and possessing the capability to be molded into various shapes and forms [1]. However, the tensile strength of concrete poses a critical detriment since the tensile strength is much smaller than the compressive strength of concrete. Accordingly, the considerable increase in the tensile stresses is accompanied by the development of cracks in the concrete where these cracks mainly depend on different factors, including the bond properties between the concrete and reinforcement, the rebar diameter and spacing, the amount of steel in the tensile zone, and the concrete cover relative to the reinforcing steel. Moreover, the width and spacing of cracks are directly associated with the discretization mechanism, which can be described as separating material into small-scale particles forming small-scale cracks with random spacings [2]. The serviceability design of concrete structures is inevitably influenced by the crack width and spacing evaluation. Hence, reliable and coherent modeling of crack width and crack spacing of reinforced concrete (RC) structures is highly essential in producing precise outcomes for the prediction of crack spacing of RC to assure sustainable serviceability design [3]. Nonetheless, the accuracy of serviceability analysis estimation based on the several design codes and cracking methods showed low values, where the error of these values was recorded to occasionally be 100% [4–6]. In general, the majority of analytic models utilized for estimating the behavior of cracks in RC structures are developed according to the typical concrete elements leading to limitations in accurately predicting oversized concrete slabs implemented in special structures such as power plants and offshore oil platforms. These drawbacks and limitations can be particularly attributed to the RC's mechanical characteristics, including the disparity of values in the mechanical characteristics and the bond-slip behavior of concrete and steel. In addition, the complexity of establishing reliable analytic approaches is related to the complicacy of the material, which is the result of the arbitrary behavior of the composite structure in the concrete and the arbitrary development of cracks. In fact, the development of crack widths and spacing in the RC structure can be investigated based on three types of crack models: rate-dependent, gradient, and fracture-based energy crack models [7]. Thus, the conduction of the sensitivity analysis is highly necessary for assessing contributive variables of cracks, such as the exfoliation of concrete, steel reinforcement exposure to various damage sources and levels, and the maximum width of cracks [8].

The development of cracks in the concrete occurs when an extensive amount of tensile stress is transferred using the bond action where the resistance to these loadings is performed *via* the reinforcement, and the concrete cannot endure the tensile stress. As a result, the tensile stress is transmitted through the reinforcement to the concrete, where the reinforcement exhibits a reduction in the stress, and an equivalent increase in the stress is expressed by the concrete, which is accompanied by an increase in the distance from the crack, leading to a spatial difference of stress up until the tensile strength of concrete is achieved, and hence, a crack is formed. Various cracking models have been developed over the previous decades as attempts to precisely evaluate and predict the most crucial factors in forming cracks in RC structures. Generally, the experimental data represent the mainstay for most developed cracking models based on bond action between concrete and reinforcement bar or empirical/semi-empirical methods [9–11]. The investigation of the bond action method was originally presented in the work of Saliger [12] and Noakowski [13], where the suggested mathematical expressions comprised factors such as the compressive strength of concrete, the bond stress, mean crack spacing, and the reinforcement ratio. Another study by Broms [9] investigated the influence of the concrete cover on the development of cracks, where he found that the relationship between the thickness of the concrete cover and the mean crack spacing is considered directly proportional. Borges [14] and Farra and Jaccoud [15] evaluated the linearly proportional correlation between the crack spacing with concrete cover and the ratio of bar diameter-to-effective reinforcement ratio and suggested experimental expressions based on these independent parameters. The effective reinforcement is directly associated with the effective concrete area, demonstrating the ultimate empirical nature of the relationship. Despite the fact that there is no definite agreement regarding the impact of the bar diameter-to-effective reinforcement ratio on the crack spacing was reached, some prominent design codes such as Eurocode 2 [16] and CEB-FIP [17] consider this variable. Beeby [18] concluded the crucial impact of the spacing between reinforcing bars on the magnitude of crack spacing. Moreover, the dominance of concrete cover thickness on the cracking behavior of concrete compared to the bar diameter-to-effective reinforcement ratio was shown. Another study conducted by Oh and Kang [19] proposed a semi-experimental expression for the crack spacing in reference to the axial strain and bar diameter. A comprehensive assessment concerning the most influential independent parameters, such as concrete cover thickness, bar diameter to effective reinforcement ratio, and spacing of bar reinforcements, on the crack spacing presented in the majority of international codes, was carried out by Lapi *et al.* [20]. A distinctive group of crack spacing models that depends on the fracture mechanics was introduced by Oh and Kang [19] and Wang *et al.* [21]. The direct and indirect investigation of the cracking phenomenon is still considered a complex process nowadays due to the variety of physical and material variables where only partial comprehension of the cracking phenomenon is achieved [22,23]. This can be partially attributed to the limited compatibility between the physical/mechanical variables, such as concrete compressive strength, concrete cover thickness, or spacing of bar reinforcement and the empirical ones such as bar diameter-to-effective reinforcement ratio, resulting in the lack of harmonization between the deformed RC element and the cracking and hence, incapability to conclude the dispersed outcomes of the cracking analysis [20,24]. Nonetheless, a novel approach based on the strain compliance of the reinforcing bars, which demonstrates the equality of the mean strains *via* mean strain and stress transfer approaches, was suggested by ref. [25] and later improved by refs. [26,27]. This innovative technique showed an adequate and precise embodiment of RC elements' crack spacing and mean strain behavior. Accordingly, the application of this approach can be extended to include the serviceability analysis of the cracking of the concrete as far as strain compliance is guaranteed since it presents a uniquely versatile framework composed of diverse sets with various principles. However, further investigations and comparisons with other methods *via* experiments and examinations are required to expand this approach's applications and reduce its drawbacks. Currently, multiple methods to estimate the primary crack spacing exist, including code-based methods such as the Eurocode 2 [16] and CEB-FIP [17], strain compliance approach (Figure 1), and artificial neural network (ANN)-based method.

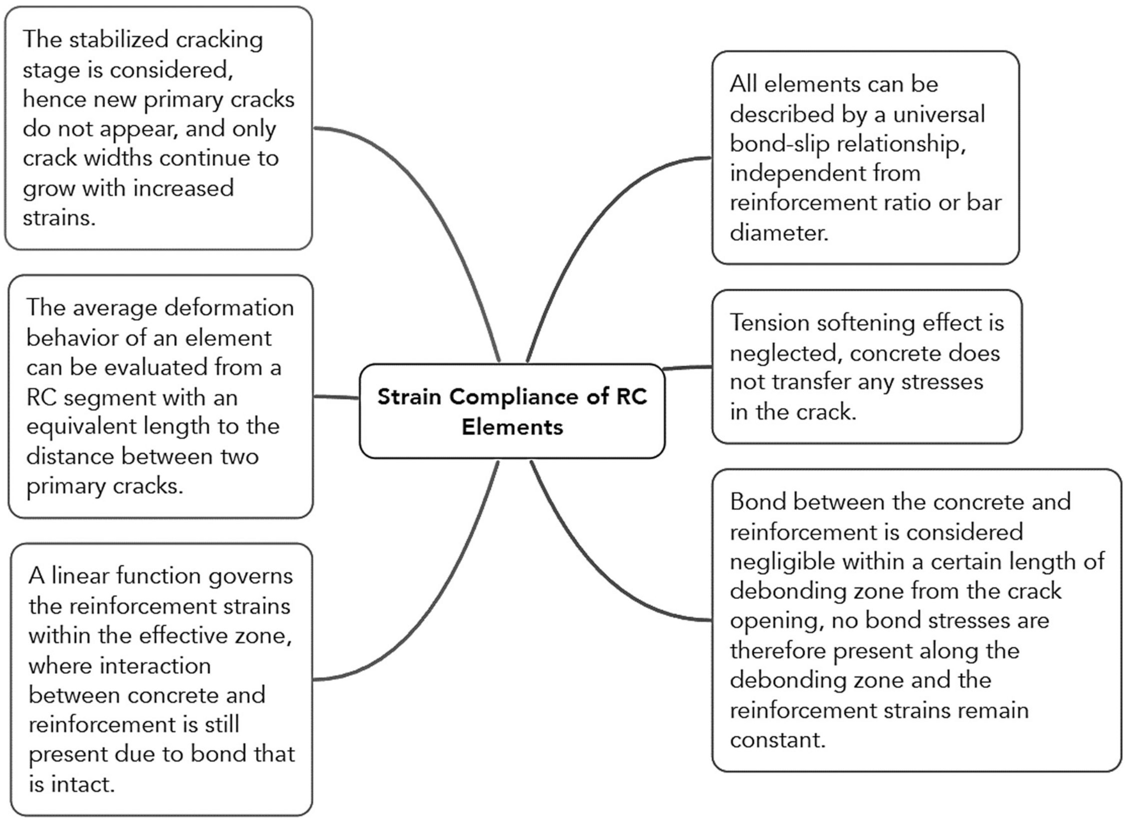

Figure 1: The principles of the strain compliance approach.

Indeed, the numerical approaches reflected their substantial capability in accurately forecasting the crack spacing of RC elements in case the proper understanding of the reinforcement and concrete behavior is ensured regardless of the complexity and time consumption for applying these techniques [28]. Implementing numerical models can be expanded to include the cracking models associated with fracture mechanics [29] and the complex, large-scale RC structures such as power plants and dams [30]. Recently, the utilization of ANNs as one of the numerical approaches for predicting crack spacing has slightly increased. ANN can be described as an adaptive mathematical model composed of many interlinked artificial neurons that simulate the processes of biological neural networks in the brain to model complicated relationships, detect existing patterns in the data, or calculate outputs [31]. The early application of ANN and machine learning targeted the estimation of concrete mechanical characteristics, including the strength of the concrete mix [32]. The superiority of ANN stems from its capability to develop and determine the relationship between the evaluated factors based on the investigated data, hence allowing for generalizations of the estimations for new cases. Furthermore, these models can improve over time by constantly feeding data to the models to establish relationships independently from users and clear of any preliminary presumptions. Some research articles aimed to derive mathematical equations and formulas for the fracture energy of concrete using ANN [33]. The current orientation of studies is toward deploying deep convolution networks for identifying cracks *via* image sources compared to serviceability analysis, such as forecasting crack spacing and crack width of concrete [34–38]. Even though numerous studies investigated the performance of ANN in predicting crack width [39,40], the incorporation of ANN in the evaluation of crack spacing is excessively scarce. Elshafey *et al.* [31] performed a study to estimate the mean crack spacing of concrete using ANN and concluded to be limited due to the insufficiency of adequate network calibration and reliable data. The behavior of ANN is highly sensitive to the number of data points where alternation of the initial weights is considerably impactful [41]. On the other hand, new soft computing techniques, known as ensemble machine learning models, have been extensively applied over the past few years. On the other hand, the capability of these models in estimating the primary crack spacing is unclear and must be highlighted to understand the various methods that can be used in such tasks. Thus, this study extensively investigates the performance of various machine learning models to predict the primary crack spacing of RC elements. Besides, it compares the estimation results to available alternatives to report the reliability of the developed approaches. Ultimately, the study conducts a parametric assessment that investigates the impact of each input variable on the output of the models.

2 Research significance

The significance of this research lies in the application of soft computing techniques to estimate the primary crack spacing in RC structures, which is a critical aspect of serviceability analysis. Accurate predictions of crack spacing are essential for ensuring the structural integrity, durability, and safety of such structures. The complex interplay of various physical and material factors affecting crack spacing has made it challenging to develop a comprehensive understanding of the cracking phenomenon. This study addresses this gap by employing different machine learning methods for predicting crack spacing in RC structures. By comparing the results of these methods with benchmark experimental data, the research provides valuable insights into the effectiveness and reliability of machine learning-based models in this context. Furthermore, the comparison of the developed models with existing approaches in the literature will showcase the advantages and potential improvements offered by the proposed methods. The parametric assessment conducted as part of this study also contributes to a deeper understanding of the most influential input parameters on primary crack spacing. This knowledge can inform the design and construction of RC structures, leading to better control of cracking behavior and enhanced serviceability performance. Finally, while this study offers more theoretical rather than application information, its findings have potential implications for curved RC structures, which are often used in the construction of bridges, tunnels, and other infrastructure projects. These structures have unique geometries that can lead to complex cracking patterns, making crack spacing prediction particularly challenging. Therefore, the use of the best soft computing model that will be suggested at the end of this study can learn and produce generalized solutions and exhibit desirable performance in terms of time, effort, and cost, which can be particularly beneficial for such structures. Hence, the ability to accurately predict crack spacing in such structures can contribute to their long-term durability and safety, ultimately benefiting the broader civil engineering community and society as a whole. By demonstrating the applicability and potential advantages of soft computing techniques in estimating primary crack spacing, this study paves the way for further research and development in this crucial area of structural engineering.

3 Research methodology

In this section, a comprehensive discussion of different machine learning models utilized for estimating crack spacing of concrete is conducted. This article aims to investigate the performance and adequacy of these machine learning models in predicting the crack spacing of concrete. In addition, a comparison between the tested machine learning models' results is carried out to determine the best model. The general methodology adopted in this study is shown in Figure 2.

3.1 Database selection

This study utilizes the large dataset prepared and collected previously by Kaklauskas [1,27] to develop machine learning models that can forecast the crack spacing of concrete. The dataset comprises 96 rectangular sections of RC beam and slab samples evaluated experimentally. The descriptive statistics of the database are shown in Table 1.

3.2 Machine learning approaches

A decision tree (DT) can be considered one of the most important and powerful machine learning approaches

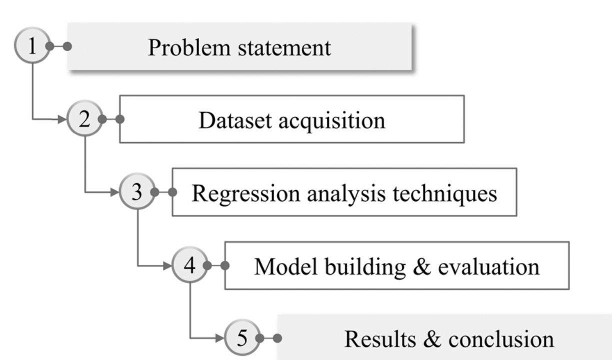

Figure 2: Demonstration of the implemented research methodology.

Table 1: Descriptive statistics of the database

Variable	Sample size	Mean	Standard deviation	Coefficient of variation	Minimum	Q1	Median	Q3	Maximum
Height (m)	96	0.4574	0.1882	41.14	0.14	0.3367	0.46	0.625	1.2
Section width (m)	96	0.3332	0.2226	66.79	0.12	0.203	0.2755	0.3	0.914
Top cover (m)	96	0.0132	0.01725	130.33	0	0	0	0.034	0.043
Bottom cover (m)	96	0.035	0.01208	34.55	0.019	0.028	0.035	0.038	0.111
Area of tensile bars (m^2)	96	0.0009	0.00062	64.99	0.00016	0.00052	0.0008	0.00106	0.00265
Area of compressive bars (m^2)	96	9×10^{-5}	0.00024	264.91	0	0	0	0.0001	0.00126
Concrete elasticity modulus (MPa)	96	30,233	3,859	12.76	23,843	27,868	29,754	33,042	46,166
Steel bar elasticity modulus (MPa)	96	200,059	3,888	1.94	186,000	200,000	200,000	200,000	210,505
Concrete compressive strength (MPa)	96	30.72	14.23	46.33	13.1	22	27.35	38.8	118.3
Concrete tensile strength (MPa)	96	2.301	0.8685	37.75	0.89	1.74	2.165	2.95	5.41
Primary mean crack spacing (m)	96	0.1981	0.07333	37.02	0.091	0.14075	0.1875	0.23975	0.503

used for classification and prediction. The learning of the DT can be achieved by dividing source data into subsets depending on the attribute value test, and recursion of this splitting process is performed for each resulting subset, known as recursive partitioning. The recursion process is terminated when the division of the subsets contributes to no additional value or when the same value is observed in each node. One of the critical merits of the DT method is the evaluation of all potential outcomes and tracing each path to a conclusion by performing an extensive investigation of the results for all branches to specify the decision nodes that demand more analysis. The most fundamental advantages of the DT approach are the classification without demanding much computation time and the capability of dealing with continuous and categorical variables. The primary disadvantages of this method are the possibility of producing errors in the classification problems and the expensive computations for training the DT model.

Random forest (RF) is the combination of tree predictors where the selection of a random vector is conducted individually for each tree. This method is based on the simultaneous training of multiple DTs with bootstrapping, where different subsets of the training dataset using various subsets of existing features are maintained. The overall variance of RF is typically minimal due to the uniqueness of each single DT. Accordingly, the aggregation of the individual DTs aims to generate adequate generalizations with high accuracy and devoid of any overfitting issues. The intersection between DT and RF classifiers is the lack of feature scaling and the ease of tuning the hyperparameters. In contrast, the difference between these two approaches is that the RF classifier is considered more robust in terms of training samples and noise in the training dataset, and the RF classifier is more complex in interpreting the results.

Extremely randomized trees (ERT), also known as extra trees, are an approach for randomly selecting the splits inside the tree's nodes. This method develops numerous trees with several feature subsets where bootstrapping can later be incorporated into the ERT structure. The most paramount advantage of ERT is the minimization of different biases, which are produced in various subsets of the data during the sampling of the whole dataset during the building of the trees. Moreover, another advantage of the ERT method is the minimization of variance produced during the splitting of the nodes within the DTs, indicating no significant impact by particular features or patterns in the dataset.

Adaptive boosting (Ada) is one of the most prevalent machine learning models that is implemented as an ensemble method. This approach adopts the DTs algorithm with only one split, which is known as decision stumps. This approach aims to combine numerous weak classifiers to generate one strong classifier. Even though estimating by a single tree is usually low, the incremental learning performed within different weak classifiers can lead to constructing one strong model. The fitting of the Ada classifier is conducted over the entire dataset, followed by additional classifier fitting where the wrong weights of the examples are adjusted.

Stochastic gradient boosting (GB) is a well-known repeated version used for regression and classification. This method was originally developed to reduce the loss function, where the main difference between this method and the Ada approach is that the objective of GB is to minimize the loss function of the learner *via* the addition of weak learners using gradient descent. The primary merits of this approach are the high flexibility in optimizing several loss functions and providing various hyperparameter options, as well as the suitability in dealing with numerical and categorical values. On the contrary, the main disadvantages of this method are the possibility of causing overfitting due to the continuous reduction in errors, which demand cross-validation to neutralize this, and the requirement for many trees exceeding 1,000, leading to expensive computation in time and memory.

Extreme gradient boosting (XGB) is a machine learning approach that implements DTs as base learners, where any individual learning algorithm can be utilized in an ensemble method. This approach assigns a particular weight to each independent parameter, which is later placed in the DT for predicting the outcomes. The wrongly predicted parameters are assigned more weight before the second placement in the DT. A combination of these unique classifiers and predictors is performed to generate a robust model, indicating the significant influence of weights in the algorithm. Thus, XGB utilizes a normalized model formalization to handle the overfitting issues and produce far higher accuracy.

3.3 Model development and hyperparameters tunning

This section discusses the method adopted in this study for developing the machine learning model. In general, the grid search approach with k-fold cross-validation was adopted for optimizing the hyperparameters during the training stage, where a set of 10 folds and many parameters were implemented for each machine learning model. Afterward, the optimal results were used for testing the model, in which 80% of the database was used for training each model, and the remaining 20% was used for testing them. The model development and estimation were entirely handled in python. Thereafter, the optimal hyperparameters is selected and the model performance is controlled using the testing dataset. Finally, the model's performance is reported in this study by comparison with other machine learning approaches. The flowchart of the approach implemented for constructing these models is summarized in Figure 3.

3.4 Models' performance assessment

The goodness of fit of the linear regression model was determined using the coefficient of determination (R^2) where the fraction's numerator is related to the undetermined differences by the response-independent parameters and the fraction's denominators are related to the differences in the response, Eq. (1) [42]. The coefficient values range varies between 0 and 1, where 1 represents the strongest linear relationship. Generally, the root-mean-square error (RMSE) is widely utilized for computing the difference between the observed and predicted values, Eq. (2). Another error analysis used in this study is the mean absolute error (MAE), which measures the variance between the absolute estimated and measured values, Eq. (3).

$$R^2 = 1 - \frac{\sum(x_i - y_i)^2}{\sum(x_i - \bar{x}_i)^2}, \quad (1)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}}, \quad (2)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \bar{y}_i|, \quad (3)$$

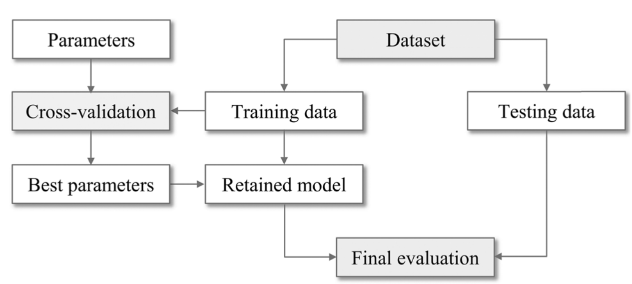

Figure 3: The methodology used for developing the machine learning model.

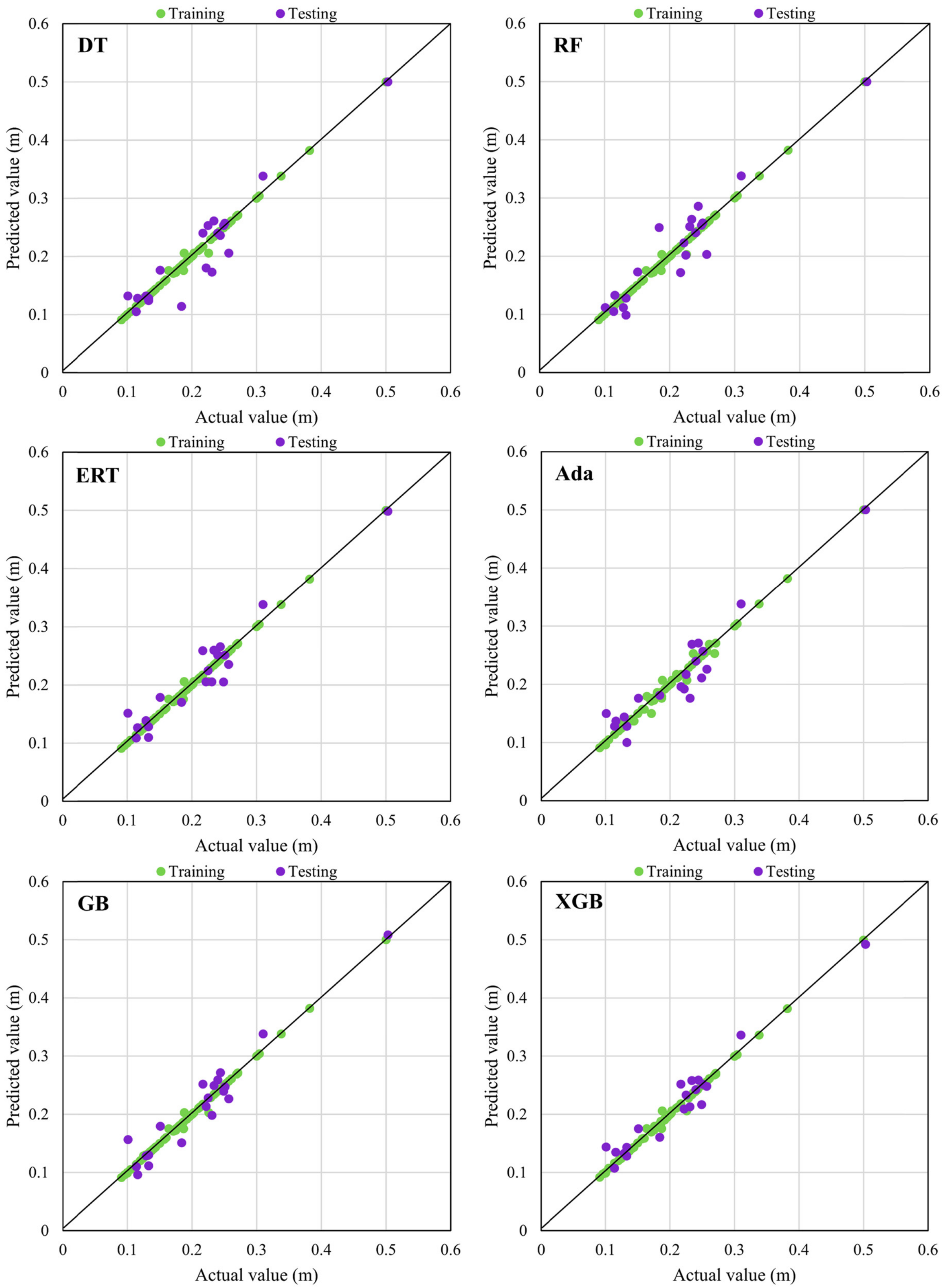

Figure 4: Performance of the selected results.

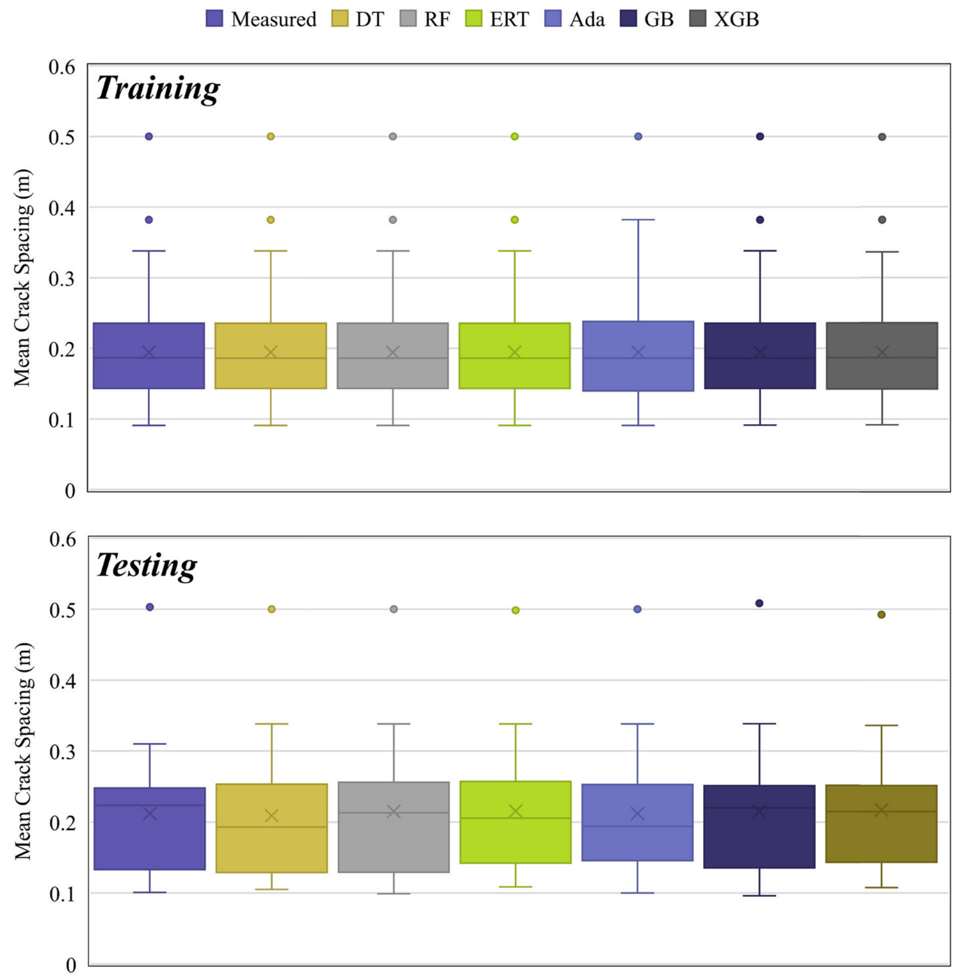

Figure 5: Comparative of the machine learning models' outcomes.

where x_i is the measured value, $\bar{x}_i$ is the mean of the measured values, y_i is the predicted value, $\bar{y}_i$ is the mean of the predicted values, and n is the number of observations.

3.5 Feature importance and partial dependence analyses

Indeed, understanding the impact of each input parameter on the prediction is a vital task to achieve a good estimation model. There are several innovations in machine learning models, such as the ability to do parametric assessments by evaluating the importance of each input parameter. Numerous articles have applied feature importance and partial dependence assessments in machine learning parametric analysis. The idea of partial dependence analysis is the calculation of the model's loss in terms of the degree of fitting caused by variations in a particular input. On the other hand, the relevance of feature importance analysis stems from its capacity to determine the impact of deleting a particular input variable on the quality of the model and, consequently, the value of the prediction output. To analyze the effect of each input parameter on primary crack spacing, feature importance and partial dependence analysis were performed in this work.

4 Results and discussion

The performance and efficiency of various machine learning models for predicting the primary crack spacing of RC structural elements were comprehensively discussed and investigated, and a comparative evaluation of these models was conducted to determine the best model. In general, the efficacy of machine learning models in the training dataset reflected higher results and better performance than the testing dataset indicated by the concentration of the values

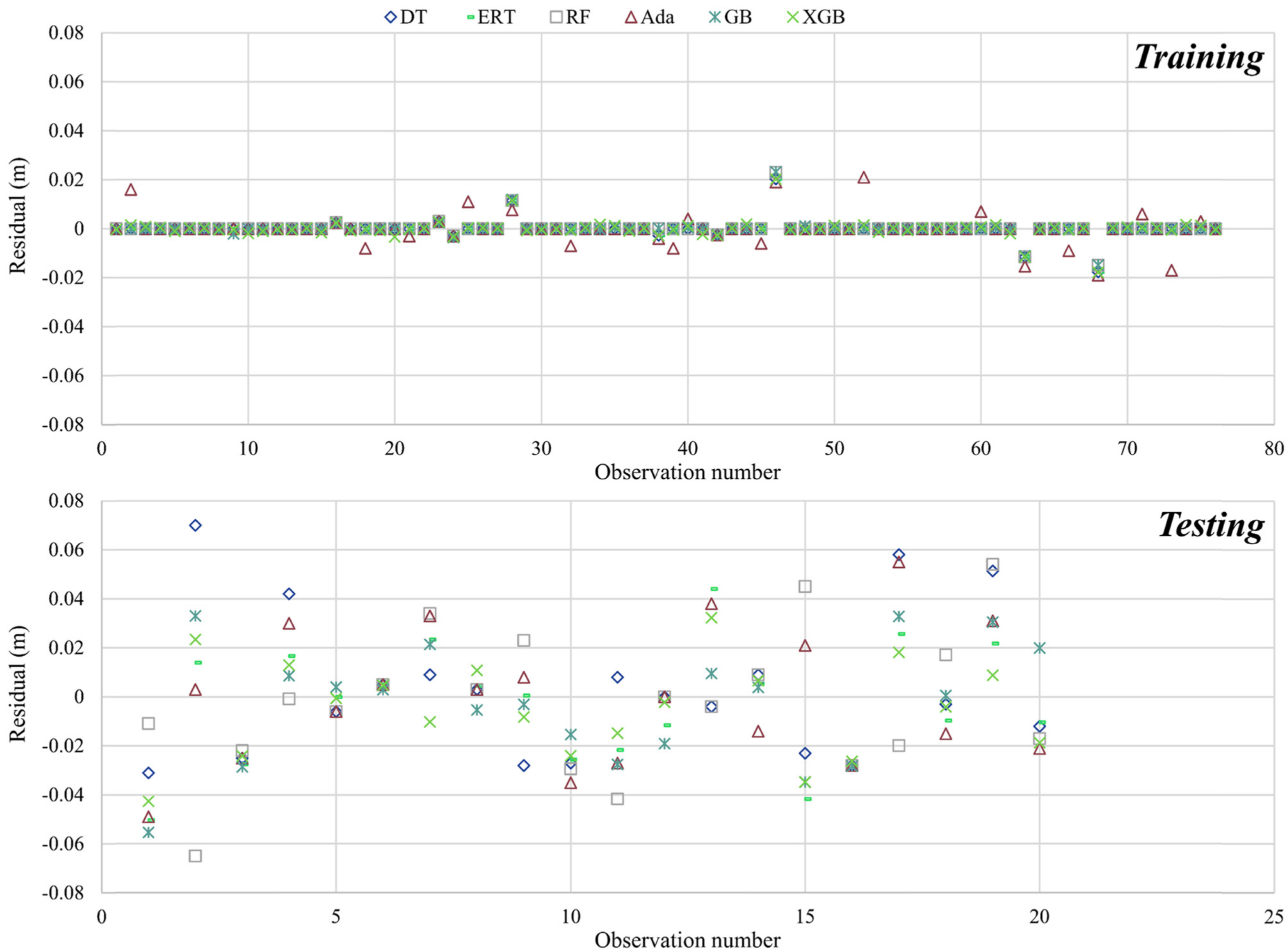

Figure 6: Residuals of the machine learning models for the training and testing datasets.

Table 2: Performance of the selected machine learning models

	R^2		MAE		RMSE	
	Training	**Testing**	**Training**	**Testing**	**Training**	**Testing**
DT	0.997	0.888	0.001	0.022	0.004	0.030
RF	0.997	0.898	0.001	0.022	0.004	0.028
ERT	0.997	0.927	0.001	0.019	0.004	0.024
Ada	0.993	0.907	0.003	0.022	0.006	0.027
GB	0.997	0.927	0.001	0.019	0.004	0.024
XGB	0.997	0.949	0.002	0.016	0.004	0.020

near the equity line, as shown in Figure 4. The performance of the XGB machine learning models showed the best results among all models, whereas the DT machine learning performance demonstrated the worst case.

A comparison of the estimated outcomes of the machine learning models with the observed outcomes for the training and testing datasets was conducted, as illustrated in Figure 5. In fact, the highest mean crack spacing in the training dataset was exhibited by the Ada model reaching the value of 0.387 m followed by DT, RF, ERT, and GB models as well as the measured case, which all marked the same mean crack spacing value of 0.35 m. Furthermore, the lowest mean crack spacing value was recorded in the case of the XGB model, which displayed a slightly smaller value than the remaining models, with a value of 0.349 m. The highest mean crack spacing value in the testing dataset was represented by all machine learning models with an approximate value of 0.342 m, whereas the lowest mean crack spacing was expressed by the measured case with a value of 0.31 m.

The difference between the observed and predicted values was computed using the residual analysis approach for validating the unspecified outcomes, Eq. (4).

$$e_i = (y_i - \hat{y}_i). \tag{4}$$

Regarding the training dataset, the residual of the several machine learning models was demonstrated over the entire observation numbers, as shown in Figure 6. In fact,

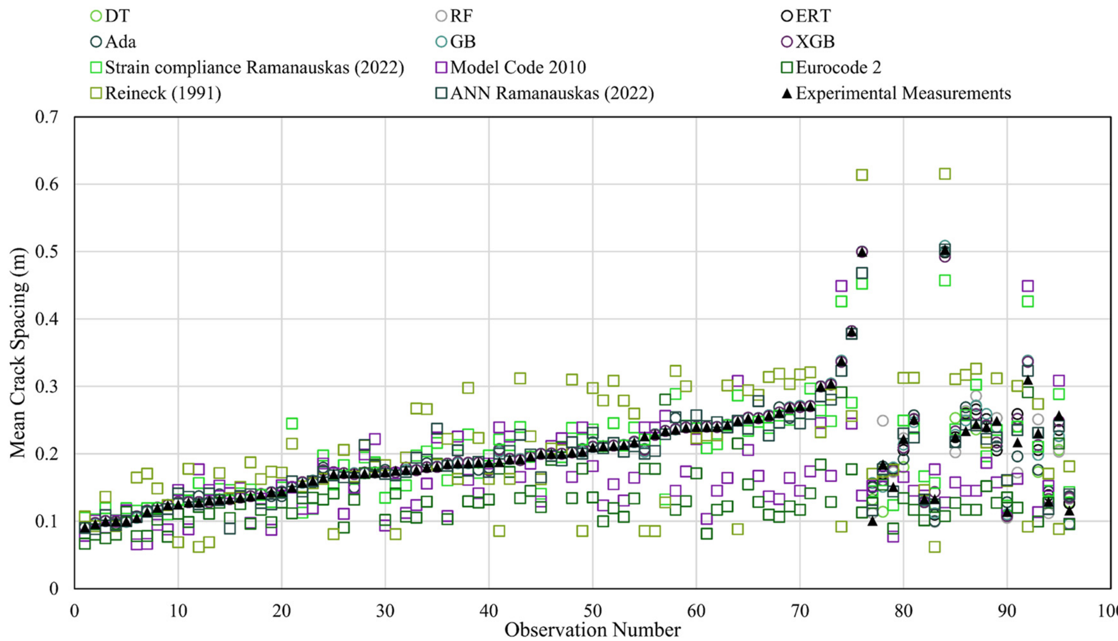

Figure 7: Performance of the machine learning models against other techniques.

the concentrated values of the tested machine learning models were represented over the entire observation number except for the Ada model. Accordingly, the RF model observed the highest residual result roughly at 0.021 m, while the lowest residual result was recorded in the Ada model exactly at −0.02 m.

As can be seen in Table 2, R^2 was determined for the training and testing datasets of the examined machine learning models. In general, the values of R^2 for the training and testing datasets are, to a certain extent, similar. All tested machine learning models scored a value of 1 during the training stage except for the Ada model, which reached a value of 0.99. Regarding the testing dataset, the values of R^2 varied considerably. The XGB model achieved the highest R^2 value of 0.95, while the lowest R^2 value of 0.89 was observed in the

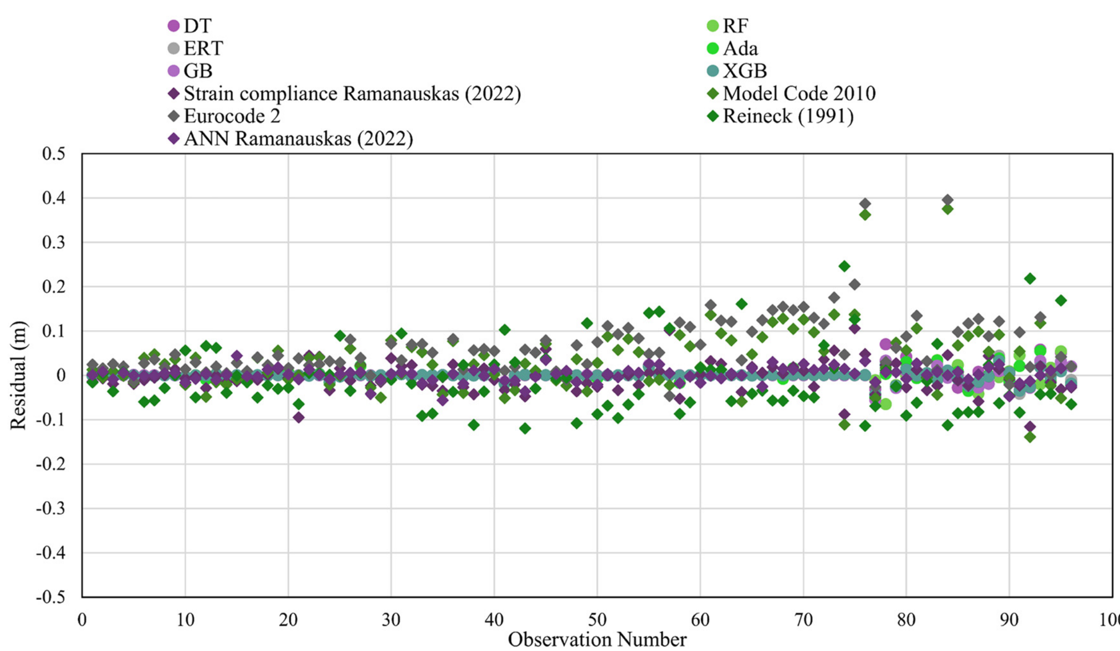

Figure 8: Benchmarking the residuals of the machine learning models against other techniques.

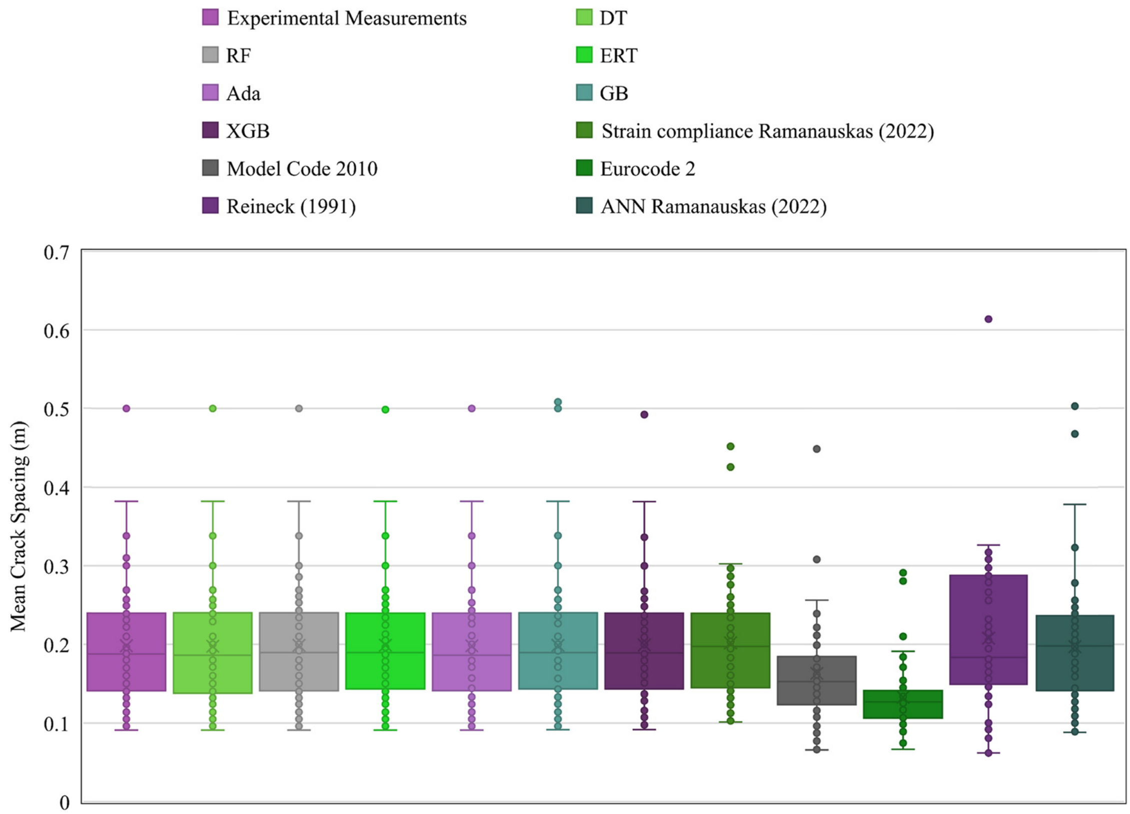

Figure 9: Statistical visualizations of the machine learning models against other techniques.

case of the DT model. The MAE results of the evaluated machine learning approaches for the training and testing datasets were represented. The MAE of the machine learning models for the training dataset reflected excellent values, where the MAE value for all machine approaches was 0.001–0.003 m. On the contrary, the lowest MAE value was recorded in the XGB model with a value of 0.016–0.02 m, while the highest was recorded in the DT, RF, and Ada models. Another error analysis method (RMSE) was computed to evaluate the efficiency of all machine learning models for the training and testing datasets, as presented in Table 2. The RMSE of the training dataset showed a perfect value for all machine learning models with a result of 0.004 m, except for the Ada model with a result of 0.006 m. On the other hand, the highest RMSE value during the testing stage was

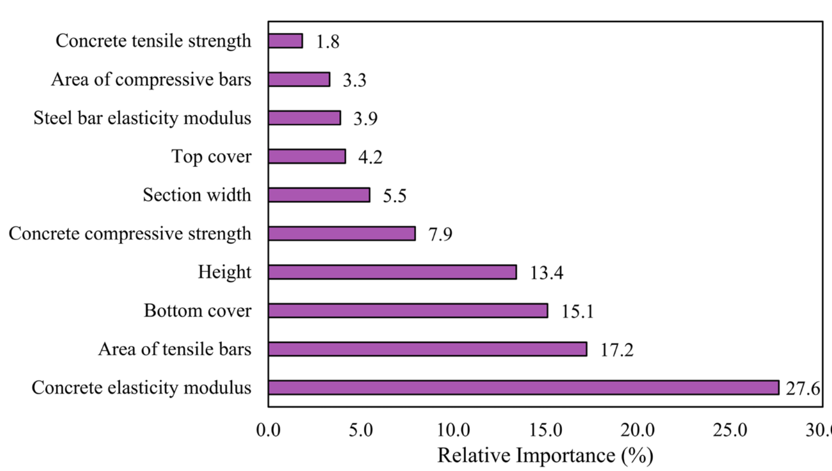

Figure 10: Parametric assessment using feature importance of the input parameters in the XGB model.

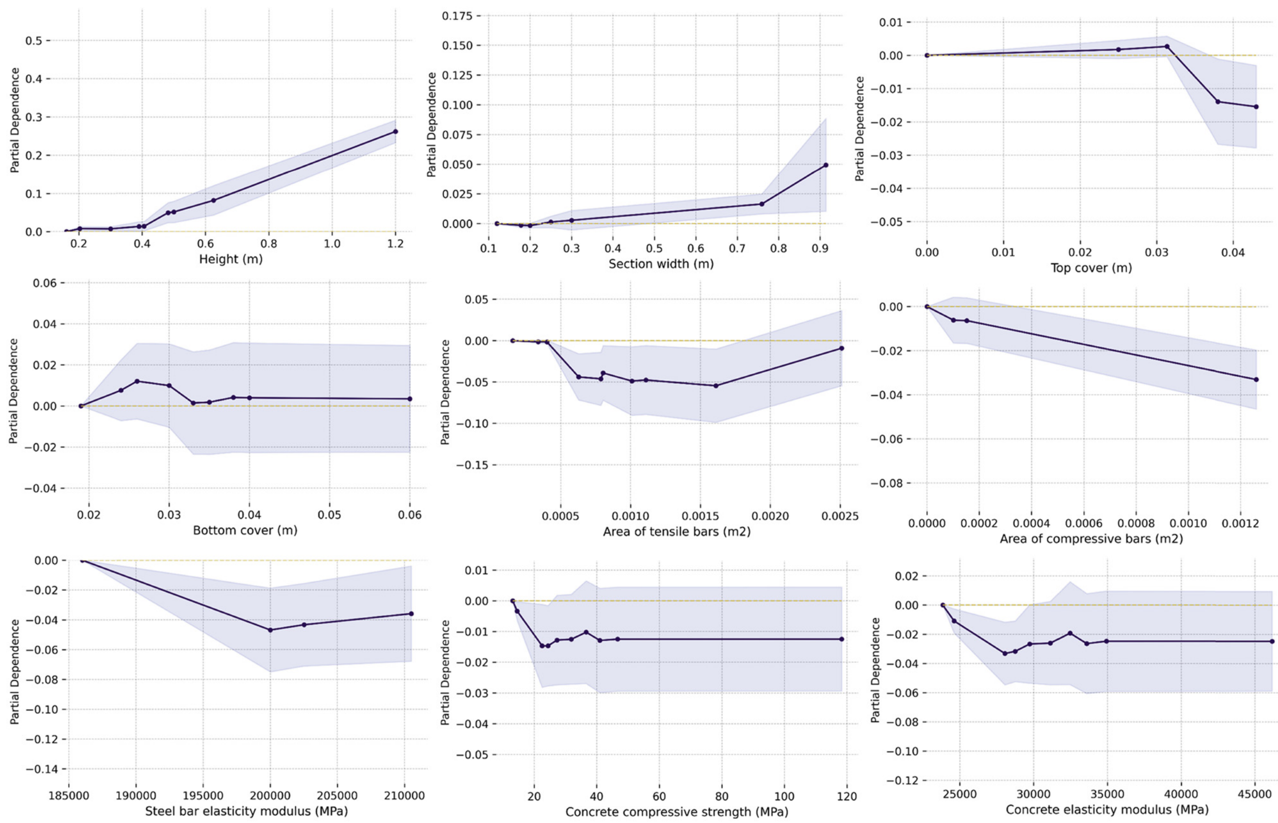

Figure 11: Influence of input parameters using partial dependence curves through the XGB model.

expressed at 0.03 m in the case of the DT model, whereas the lowest RMSE value was reached at 0.02 m in the case of the XGB model.

Figures 7 and 8 compare the performance of the machine learning models against five prediction methods presented by Ramanauskas *et al.* [3], which are Reineck, strain compliance, Eurocode 2, CEB-FIP (model code 2010), and ANN approaches. The first approach was proposed by Reineck [43], where he discussed the structural performance of the prestressed concrete slabs from cracking where the shear force is fundamentally transferred in the web. He derived a mathematical equation for detecting the crack width in the web by determining the ultimate shear force within the web. The second method is strain compliance based on the average strains' equality *via* average strain and stress transfer techniques. In addition, the third prediction method is Eurocode 2 [16], which provides a mathematical expression for computing the average crack spacing considering the effect of the concrete cover. The fourth prediction method is model code 2010 by the CEB-FIP [17], which calculates the crack spacing based on the bond-slip length. Finally, the last prediction method is a feed-forward back-propagation ANN with the Levenberg–Marquardt training function. The highest residual values representing the lowest accuracy and worst efficiency were recorded precisely at 0.4 m in the case of Eurocode 2, followed directly by the model code 2010 and Reineck, where the results of these prediction methods were scattered and dispersed. On the other hand, the examined machine learning models exhibited the lowest residual values representing the highest accuracy and best efficiency, where the results were concentrated and condensed. Finally, the results of machine learning models act as a benchmark for other prediction techniques.

A comparative assessment of the performance of the investigated machine learning models with the prediction methods provided in ref. [3] is illustrated in Figure 9. The tested machine learning models were observed to mark the exact mean crack spacing of the experimental measurements with a value of 0.394 m. The lowest mean crack spacing values were expressed in the case of Eurocode 2, followed by model code 2010 and strain compliance with values of 0.199, 0.265, and 0.3 m.

Finally, a parameter assessment of the impact of the input variables on the prediction of the machine learning models is given in Figures 10 and 11. It can be seen that the concrete elastic modulus holds the highest impact on the prediction decision of the machine learning model, whereas the tensile strength of concrete holds the lowest impact. In addition, it can be seen from the partial dependence curves that there is a positive relationship between the section height and the crack spacing, whereas there is a negative relationship between the area of compression bars and the crack spacing. Conversely, the other variables have a flaunting trend with respect to the crack spacing.

5 Conclusion

This article aims to study and examine the effectiveness and performance of different machine learning models in predicting the primary crack spacing of RC structures. In addition, a comprehensive comparison between the outcomes of these machine learning models with the other prediction techniques and the experimental measurements was performed.

On the basis of the aforementioned statement, the following conclusions are drawn:
- The tested machine learning models showed satisfactory capability in accurately estimating the primary crack spacing.
- The DT, RF, and Ada models provide the highest MAE errors, while the XGB model achieves the best performance.
- The residuals of the evaluated machine learning models demonstrated concentrated values about zero over the entire observation numbers in the training dataset in contrast to the testing dataset, which displayed scattered values, with DT and RF having the worst results.
- Compared to other techniques available in the literature, the highest residual value, and hence the lowest accuracy of the mean crack spacing, was produced in the case of the Eurocode 2 followed by model code 2010 and Reineck methods. On the contrary, the lowest residual, and hence the highest accuracy of the mean crack spacing, was seen in the case of the XGB model, which represents the reliability of this technique.
- The concrete elastic modulus holds the highest impact on the prediction decision of the machine learning model, whereas the tensile strength of concrete holds the lowest impact on the decision of the machine learning models.

Funding information: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author contributions: OA and MA designed the study, analyzed the data, and wrote the manuscript.

Conflict of interest: The authors declare no conflict of interest.

References

[1] Kaklauskas G, Ramanauskas R, Ng PL. Predicting cracks pacing of reinforced concrete tension members using strain compliance approach with debonding. J Civ Eng Manag. 2019;25:422–30.
[2] Carino NJ, Carino NJ. Prediction of cracking in reinforced concrete structures. Gaithersburg (MD), USA: US Department of Commerce, National Institute of Standards and Technology; 1995.
[3] Ramanauskas R, Kaklauskas G, Sokolov A. Estimating the primary crack spacing of reinforced concrete structures: Predictions by neural network versus the innovative strain compliance approach. Mech Adv Mater Struct. 2022;29:53–69.
[4] Kaklauskas G. Flexural layered deformational model of reinforced concrete members. Mag Concr Res. 2004;56:575–84.
[5] Kaklauskas G, Gribniak V, Bacinskas D, Vainiunas P. Shrinkage influence on tension stiffening in concrete members. Eng Struct. 2009;31:1305–12.
[6] Gribniak V, Kaklauskas G, Bacinskas D. Experimental investigation of shrinkage influence on tension stiffening of RC beams. In: Tanabe T, Sakata K, Mihashi H, Sato R, Maekawa K, Nakamura H, editors. Proceedings of Creep, Shrinkage and Durability of Concrete and Concrete Structures (ConCreep 8), 2008 30 Sep–2 Oct; Ise-Shima, Japan: CRC Press, 2009.
[7] Sluys LJ, Borst RDe. Failure in plain and reinforced concrete—an analysis of crack width and crack spacing. Int J Solids Struct. 1996;33:3257–76.
[8] Yasuda N, Tsutsumi T, Kawamura T, Matsuho S, Shiraki W. Assessment of deteriorating reinforced concrete structures using artificial neural networks. 1993 (2nd) International Symposium on Uncertainty Modeling and Analysis; 1993 Apr 23–25; College Park (MD), USA. IEEE, 2002. p. 581–6.
[9] Broms BB. Crack width and crack spacing in reinforced concrete members. J Proc. 1965;62:1237–56.
[10] Broms BB, Lutz LA. Effects of arrangement of reinforcement on crack width and spacing of reinforced concrete members. J Proc. 1965;62:1395–410.
[11] Gergely P, Lutz LA. Maximum crack width in reinforced concrete flexural members. Am Concr Inst. 1968;20:87–117.
[12] Saliger R. High grade steel in reinforced concrete. IABSE Congress Report; 1936.
[13] Noakowski P. Verbundorientierte, kontinuierliche Theorie zur Ermittlung der Rißbreite. Wirklichkeitsnaher und einfacher Nachweis unter Berücksichtigung der Verbundgesetze und der Betonzugfestigkeit sowie unter Verknüpfung des Erstriß-und Endrißzustands (Teil 1). Beton-und Stahlbetonbau. 1985;80:185–90.

[14] Borges JF. Cracking and deformability of reinforced concrete beams. Lisbon, Portugal: Laboratório Nacional de Engenharia Civil; 1965.
[15] Farra B, Jaccoud JP. Bond behaviour, tension stiffening and crack prediction of high strength concrete. Proceedings of the International Conference of Bond in concrete. Concrete; 1992 Oct 15–17; Riga, Latvia; 1992.
[16] Eurocode 2. Design of concrete structures—Part 1-1: General rules and rules for buildings. London, UK: British Standard Institution; 2004.
[17] CEB-FIP. fib Model Code for Concrete Structures 2010. Berlin, Germany: Ernst & Sohn; 2010.
[18] Beeby AW. The influence of the parameter ϕ/ρ eff on crack widths. Struct Concr. 2005;6:155–65.
[19] Oh BH, Kang YJ. New formulas for maximum crack width and crack spacing in reinforced concrete flexural members. Struct J. 1987;84:103–12.
[20] Lapi M, Orlando M, Spinelli P. A review of literature and code formulations for cracking in R/C members. Struct Concr. 2018;19:1481–503.
[21] Wang JJ, Tao MX, Nie X. Fracture energy-based model for average crack spacing of reinforced concrete considering size effect and concrete strength variation. Constr Build Mater. 2017;148:398–410.
[22] Sanal I. Effect of shear span-to-depth ratio on mechanical performance and cracking behavior of high strength steel fiber-reinforced concrete beams without conventional reinforcement. Mech Adv Mater Struct. 2020;27:1849–64.
[23] Pozharskii DA, Sobol BV, Vasiliev PV. Periodic crack system in a layered elastic wedge. Mech Adv Mater Struct. 2020;27:318–24.
[24] Borosnyói A, Balázs GL. Models for flexural cracking in concrete: the state of the art. Struct Concr. 2005;6:53–62.
[25] Kaklauskas G, Ramanauskas R. A new approach in predicting the mean crack spacing of flexural reinforced concrete elements. In: Beushausen H, editor. fib Symposium 2016, Performance-Based Approaches for Concrete Structures; 2016 Nov 21–23; Cape Town, South Africa. fib, 2016. 2016.
[26] Kaklauskas G, Ramanauskas R, Jakubovskis R. Mean crack spacing modelling for RC tension elements. Eng Struct. 2017;150:843–51.
[27] Kaklauskas G. Crack model for RC members based on compatibility of stress-transfer and mean-strain approaches. J Struct Eng. 2017;143:04017105.
[28] Ng PL, Ma FJ, Kwan AKH. Crack analysis of concrete beams based on pseudo-discrete crack model. Proceedings of the Second International Conference on Performance-Based and Life-Cycle Structural Engineering; 2015 Dec 9–11; Brisbane, Australia. School of Engineering, The University of Queensland, 2015. p. 669–78.
[29] Kurumatani M, Soma Y, Terada K. Simulations of cohesive fracture behavior of reinforced concrete by a fracture-mechanics-based damage model. Eng Fract Mech. 2019;206:392–407.
[30] Nader C, Rossi P, Tailhan JL. Multi-scale strategy for modeling macrocracks propagation in reinforced concrete structures. Cem Concr Compos. 2019;99:262–74.
[31] Elshafey AA, Dawood N, Marzouk H, Haddara M. Predicting of crack spacing for concrete by using neural networks. Eng Fail Anal. 2013;31:344–59.
[32] Vanluchene RD, Sun R. Neural networks in structural engineering. Comput-Aided Civ Infrastruct Eng. 1990;5:207–15.
[33] Nikbin IM, Rahimi S, Allahyari H. A new empirical formula for prediction of fracture energy of concrete based on the artificial neural network. Eng Fract Mech. 2017;186:66–482.
[34] Moon HG, Kim JH. Inteligent crack detecting algorithm on the concrete crack image using neural network. 28th International Symposium on Automation and Robotics in Construction (ISARC); 2011 29 Jun–2 Jul; Seoul, Republic of Korea. p. 1461–7.
[35] Cha YJ, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. Comput-Aided Civ Infrastruct Eng. 2017;32:361–78.
[36] Dorafshan S, Thomas RJ, Maguire M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. Constr Build Mater. 2018;186:1031–45.
[37] Dung CV. Autonomous concrete crack detection using deep fully convolutional neural network. Autom Constr. 2019;99:52–8.
[38] Bayar G, Bilir T. A novel study for the estimation of crack propagation in concrete using machine learning algorithms. Constr Build Mater. 2019;215:670–85.
[39] Ince R. Prediction of fracture parameters of concrete by artificial neural networks. Eng Fract Mech. 2004;71:2143–59.
[40] Elshafey AA, Dawood N, Marzouk H, Haddara M. Crack width in concrete using artificial neural networks. Eng Struct. 2013;52:676–86.
[41] Shaikhina T, Khovanova NA. Handling limited datasets with neural networks in medical applications: A small-data approach. Artif Intell Med. 2017;75:51–63.
[42] Renaud O, Victoria-Feser MP. A robust coefficient of determination for regression. J Stat Plan Inference. 2010;140:1852–62.
[43] Reineck KH. Ultimate shear force of structural concrete members without transverse reinforcement derived from a mechanical model (SP-885). Struct J. 1991;88:592–602.