Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Prediction Of Concrete Compressive Strength Using Machine Learning Algorithms

Prof. Alok Rarotiya [1], Prof. Shubhrata Kanungo [2]

[1] Research Scholar, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India
alok.rarotiya@medicaps.ac.in
[2] Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India
shubhrata.kanungo@medicaps.ac.in

This study investigates the use of machine learning techniques to predict the compressive strength of concrete using key input parameters such as water-cement ratio, cement content, and aggregate size. Several models, including Support Vector Regression (SVR) and Backpropagation Neural Network (BPNN), were applied to a comprehensive dataset of concrete mixtures. The results demonstrate that both models exhibit high predictive accuracy, with the SVR model achieving an R^2 score of 0.877 for the test set and the BPNN model scoring 0.872. Feature importance analysis revealed that the water-cement ratio and cement content were the most influential factors in determining compressive strength. The findings highlight the potential of machine learning in optimizing concrete mix designs, offering a more accurate prediction than traditional empirical methods. Future research should focus on incorporating additional variables and advanced optimization techniques to further improve the predictive capabilities of these models.

Keywords: Concrete, Strength, Neural Network, Machine learning.

1. Introduction

Concrete is one of the most widely used construction materials globally, known for its versatility, durability, and cost-effectiveness. The compressive strength of concrete is one of the most critical properties that determine its suitability for structural applications. Compressive strength testing, typically performed at 28 days after casting, serves as an indicator of the concrete's ability to bear loads and resist failure under compressive stress [1]. However, accurately predicting compressive strength before testing remains a challenge in the industry, as it depends on various factors such as the water-cement ratio, cement content, aggregate type, and curing conditions [2]. Traditional methods for estimating compressive strength rely on empirical formulas and regression models, which often fail to capture the complex, nonlinear relationships between input variables and strength outcomes [3]. These methods can be limited by assumptions about linearity and often require substantial amounts of experimental data. In contrast, machine learning (ML) algorithms provide a robust alternative for modeling these complex relationships, allowing for more accurate predictions

with fewer assumptions [4]. ML models can efficiently handle large datasets, learn from the interactions between variables, and improve predictive accuracy through iterative learning processes [5]. Recent advancements in artificial intelligence and machine learning have demonstrated their potential in civil engineering applications, particularly in material property predictions [6]. Studies have successfully applied machine learning techniques, such as artificial neural networks (ANNs), decision trees, and support vector machines, to predict the compressive strength of concrete, achieving superior results compared to traditional approaches [7-16]. These algorithms are capable of capturing the complex dependencies between mix design parameters and compressive strength, which are often overlooked by traditional empirical models [8]. While machine learning has made significant strides in predicting concrete compressive strength, several gaps remain in the research. First, many studies focus on a single machine learning algorithm without comparing it to other techniques. As a result, the performance of different models remains underexplored, making it difficult to determine the most suitable algorithm for specific concrete mix designs or conditions. Second, there is limited research on integrating machine learning with domain knowledge from civil engineering. Most ML models treat concrete mix data purely as numbers, without considering the underlying physical and chemical interactions between materials. By combining domain expertise with machine learning, models could potentially make more accurate predictions and provide better insight into the behavior of different mix designs.
Third, data availability and quality remain significant challenges. Many machine learning models require large amounts of high-quality data for training, which can be difficult to obtain in practice. Most studies rely on limited experimental data, which can affect the generalizability of the models. Additionally, there is a lack of research on the impact of external factors, such as environmental conditions or curing methods, on the performance of machine learning models in predicting compressive strength. Machine learning offers a powerful alternative to traditional methods by addressing the complexity and nonlinearities of concrete mix designs. However, there is still much room for improvement, especially in model generalization, integration with civil engineering expertise, and the exploration of newer algorithms that could further enhance predictive accuracy.

Given the significant implications of accurate compressive strength prediction on construction quality and safety, this research aims to evaluate the effectiveness of machine learning algorithms in predicting the compressive strength of concrete. The specific objectives of this study are to (1) build machine learning models using input parameters such as water-cement ratio, cement content, and aggregate size, (2) compare the performance of various machine learning algorithms, including SVR and BPNN. This research provides a novel approach to predicting the compressive strength of concrete by leveraging the power of machine learning algorithms. While traditional empirical models and statistical methods have been widely used in concrete strength prediction, they often suffer from limitations such as oversimplification of nonlinear relationships between concrete mix variables. This study goes beyond these limitations by exploring a range of machine learning algorithms, SVR and BPNN, to model and predict compressive strength with higher accuracy.

2. Methodology

2.1 Data Collection

The dataset utilized in this study comprises a diverse array of concrete mixtures, each characterized by several input features that significantly influence the compressive strength of concrete. Key features include the Water-Cement Ratio (WCR), Cement Content (CC), and Aggregate Size (AS). The Water-Cement Ratio denotes the ratio of the weight of water to the weight of cement within the mixture, with a lower WCR generally correlating to a higher compressive strength. Cement Content, measured in kilograms per cubic meter (kg/m³), indicates the quantity of cement incorporated in the concrete mix, and higher cement content is typically associated with enhanced compressive strength. Table 1 illustrated the dataset used to predict the compressive strength of concrete using machine learning algorithms.

Table 1 Data set for concrete mix

Exp. No.	Water-cement ratio	Cement (kg/m3)	Coarse gravel (kg/m3)	Fine gravel (kg/m3)	Sand (kg/m3)
1	0.79	273	0	936	863
2	0.79	244	497	478	936
3	0.74	291	0	927	980
4	0.74	261	493	473	928
5	0.69	313	0	915	968
6	0.69	280	487	469	919
7	0.65	332	0	920	957
8	0.65	297	529	509	925
9	0.61	354	0	909	949
10	0.61	316	524	504	921
11	0.58	372	0	905	945
12	0.58	332	530	490	908
13	0.54	400	0	890	888
14	0.54	357	528	488	860
15	0.5	432	0	935	872
16	0.5	386	550	469	855
17	0.45	480	0	908	825
18	0.45	429	533	412	800
19	0.42	514	0	884	753
20	0.42	460	558	419	768

Aggregate Size, representing the maximum size of the coarse aggregate utilized in the mixture (measured in millimeters), can significantly affect both the workability and strength of the concrete. The target variable for this study is the Compressive Strength (CS), measured in megapascals (MPa), which reflects the maximum compressive stress that the concrete can withstand. The dataset was sourced from a combination of experimental studies and concrete production records to ensure a comprehensive representation of various concrete mixtures.

2.2 Data Preprocessing

Prior to the application of machine learning algorithms, the dataset underwent several preprocessing steps to enhance its quality and suitability for modeling. Initially, missing values within the dataset were addressed through appropriate imputation techniques or by omitting affected rows, contingent on the extent of missing data. Following this, normalization was applied to ensure that all features contributed equally to the model training process. This normalization involved scaling the numerical features to a standard range, typically between [0, 1] or [-1, 1], utilizing either Min-Max Scaling or Standardization methods. Subsequently, the processed dataset was partitioned into training and testing subsets, adhering to the common practice of allocating 70% of the data for training and 30% for testing. This division facilitates a robust evaluation of model performance.

2.3 Machine Learning Models

In this study, several machine learning algorithms were employed to predict the compressive strength of concrete. The first algorithm, Linear Regression, serves as a foundational statistical method that models the relationship between input features and the target variable using a linear equation. Additionally, the Random Forest Regressor was utilized, an ensemble method that constructs multiple decision trees during training and outputs the average prediction, thereby enhancing accuracy and reducing the risk of overfitting. The Gradient Boosting Regressor, another ensemble technique, was also implemented; it builds models sequentially, with each new model correcting the errors made by the previous ones, resulting in improved predictive performance. Lastly, the Multi-Layer Perceptron (MLP), a type of neural network model, was incorporated to capture complex non-linear relationships in the data through multiple layers of interconnected nodes.

3. Methodology

To comprehensively analyze the concrete compressive strength, this study employs a machine learning approach to predict concrete compressive strength. This section focuses on the machine learning methodology used for SVR and BPNN in the study and the evaluation indicators for the performance of the prediction model. Support Vector Machine is a statistical learning theory proposed by Vapnik [17], which is based on the criterion of structural risk minimization [17-22]. This method aims to minimize both the structural risk and the sample error, thereby improving the generalization ability of the model without being limited by the dimensionality of the data. To address regression fitting problems using SVM, Vapnik et al. [17] introduced an insensitive loss function ε, resulting in a regression-type support vector machine. This model exhibits superior performance and effectiveness. In SVM regression fitting, the main idea is to determine an optimal plane that minimizes the error between the training samples and the optimal plane [23].

Considering the advantages of neural networks in handling complex nonlinear relationships, this study incorporates neural network modeling into the prediction of concrete strength. The artificial neural network model is a mathematical tool that emulates the functions of the human brain, such as learning, reasoning, and performing parallel computations [5,24]. The back propagation (BP) neural network consists of an input layer, an intermediate layer (also known as the hidden layer), and an output layer. When considering concrete mix proportion design, the architecture of BPNN is divided into 3 layers with the number of neurons in each layer being 7,10 and 1, respectively.
The learning process of the BPNN consists of two stages [25-28].

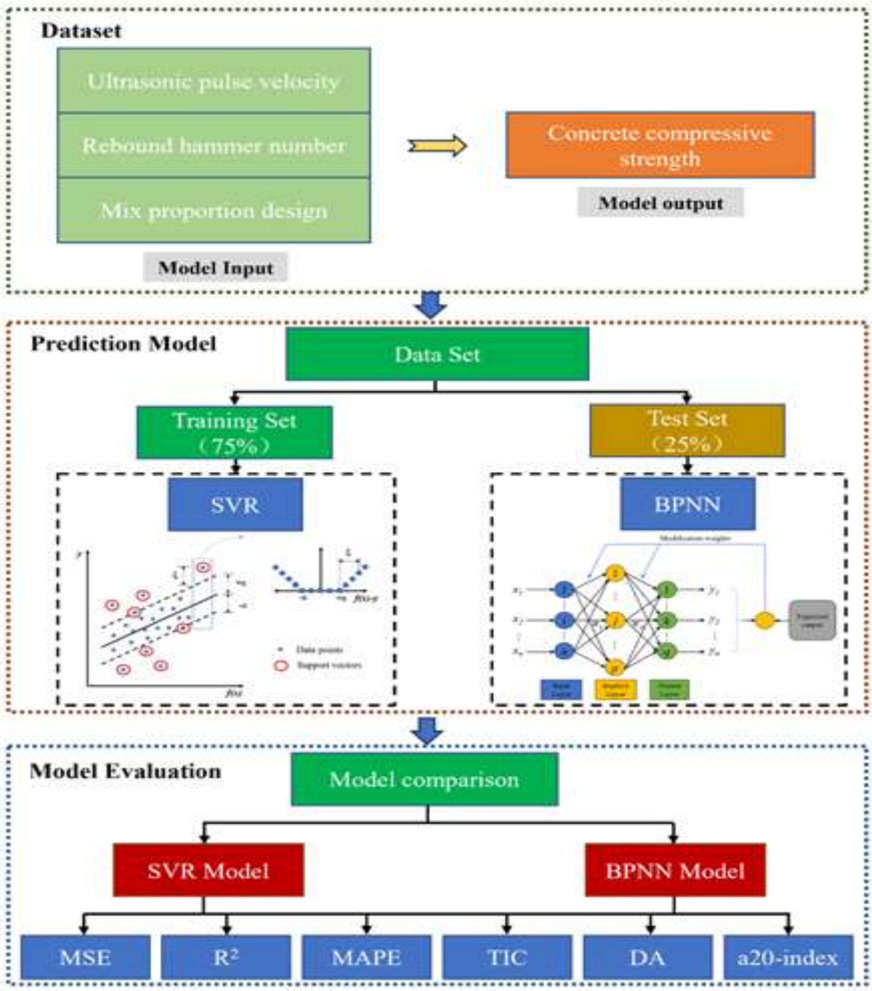

Fig. 1 Implementation flow chart of machine learning algorithms

In the first stage, information is propagated forward from the input layer through the hidden layer to the output layer. By inputting the sample data, the input vector is processed through the designed network structure and the weights and thresholds are adjusted iteratively to achieve sample learning. In the second stage, information is propagated backward from the output layer to the input layer, following a specific rule to modify the initial weights and thresholds of the algorithm. This process aims to facilitate the convergence of the algorithm. The first stage of the learning process is completed when information flows from the input layer to the output layer. If there is a discrepancy between the output result of the output layer and the expected result, the error is propagated backward from the output layer to the input layer through the intermediate layer. The weights are adjusted layer by layer using the gradient descent algorithm to complete the second stage of the learning process. These two stages are repeated alternately until convergence is achieved, signaling the end of the learning process and the termination of network training. Fig. 1 showed the flow chart for implementation of SVR and BPNN machine learning algorithms.

4. Model Training and Evaluation

Each machine learning model was trained on the training dataset, and subsequent predictions were generated using the testing set. To evaluate model performance, several metrics were employed. The Coefficient of Determination (R^2) indicates the proportion of variance in the target variable that can be explained by the features, with values approaching 1 signifying a superior fit. The Root Mean Squared Error (RMSE) quantifies the average magnitude of the errors between predicted and actual values, providing insights into the model's accuracy. Additionally, the Mean Squared Error (MSE) calculates the average of the squares of the errors, aiding in the understanding of the variance of the prediction errors. The Mean Absolute Error (MAE) represents the average absolute difference between predicted and actual values, offering a straightforward interpretation of prediction accuracy. To ensure that the models generalize well to unseen data, cross-validation techniques, such as k-fold cross-validation, were employed. The results were systematically compared across all models to identify the most effective approach for predicting concrete compressive strength.

5. Results and Discussion

The results obtained from the application of various machine learning models to predict the compressive strength of concrete were systematically analyzed and compared to elucidate the most effective approach. Each model's performance was evaluated using key metrics such as the Coefficient of Determination (R^2), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE). The findings revealed that the SV Regressor demonstrated a robust predictive capability, yielding an R^2 score of 0.877, which indicated that approximately 87.7% of the variance in compressive strength could be explained by the selected features. This substantial explanatory power underscores the model’s effectiveness in capturing the intricate relationships inherent in concrete mixtures. In contrast, the BPNN achieved a substantial R^2 score of 0.872. Table 2 showed the R2 and mean square error (MSE) values of concrete after 7 and 28 days.

Table 2 Performance of concrete mix

Concrete aging	**SVR**		**BPNN**	
	R^2	**MSE**	**R^2**	**MSE**
7 Days	0.877	0.029	0.872	0.031
28 Days	0.977	0.035	0.935	0.038

The MSE values for the SVR model were 0.029 and 0.035 after 7 and 28 days respectively while for BPNN it was 0.031 and 0.038. The errors in the prediction were found low for both the machine learning algorithms indicating no overfitting or underfitting of the models. Furthermore, the analysis of feature importance across the models revealed insightful trends regarding the contributions of individual input variables to the prediction of compressive strength. Notably, the Water-Cement Ratio emerged as a critical determinant, consistently ranking high in importance across all models. This finding aligns with existing literature, emphasizing the pivotal role of WCR in influencing concrete strength. Cement Content also demonstrated significant relevance, corroborating its established effect on enhancing compressive strength through increased binding capacity. Aggregate Size exhibited variability in its influence depending on the model employed, suggesting that while it is a contributing factor, its impact may be more nuanced and potentially context-dependent.

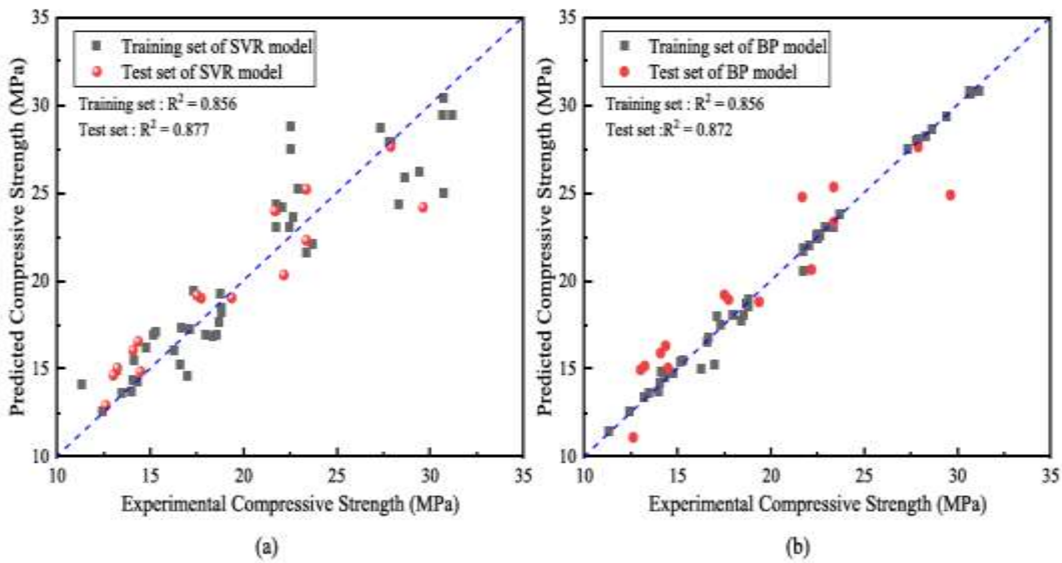

Fig. 2 Predictions od SVR and BPNN models at 7 days of curing

As shown in fig. 2(a), the SVR model achieves R^2 values of 0.877 for the test set, indicating strong predictive accuracy, with the test set slightly outperforming the training set. Similarly, in plot (b), the BP model achieves an R^2 of 0.872 for the test set. Although both models demonstrate high accuracy, the SVR model shows slightly better generalization to unseen data, as reflected in the higher R^2 for the test set compared to the BP model. Similarly, fig. 3(a) showed the SVR model achieves R^2 values of 0.977 for the test set, indicating strong predictive accuracy, with the test set slightly outperforming the training set. Similarly, in plot (b), the BP model achieves an R^2 of 0.935 for the test set. Although both models demonstrate high accuracy, the SVR model shows slightly better generalization to unseen data, as reflected in the higher R^2 for the test set compared to the BP model.

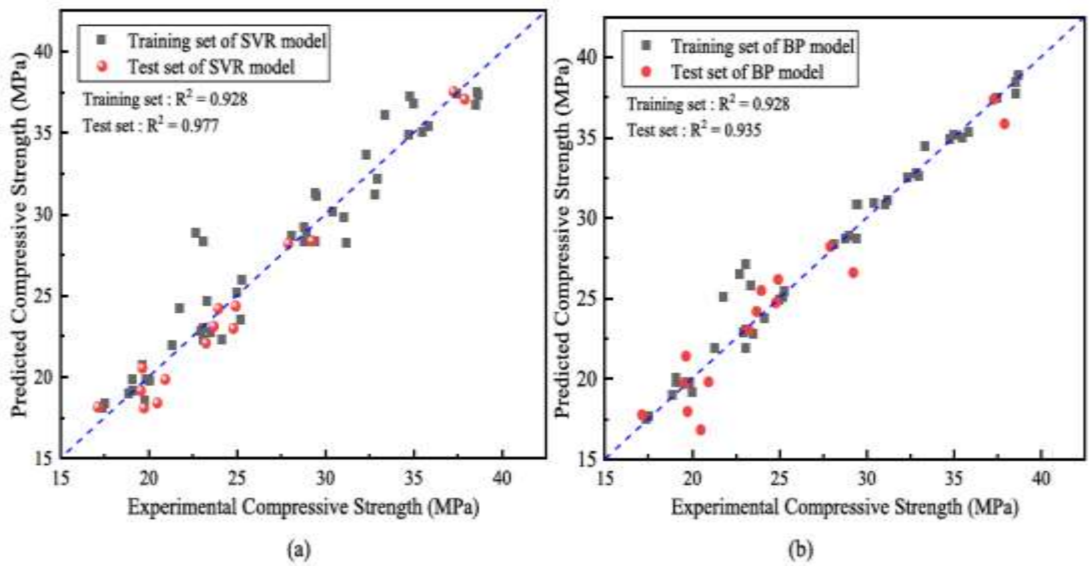

Fig. 2 Predictions of SVR and BPNN models at 28 days of curing

6. Conclusion

This study demonstrates the effectiveness of machine learning models, particularly Support Vector Regression (SVR) and Backpropagation Neural Network (BPNN), in predicting the compressive strength of concrete based on key mix design parameters. The SVR model slightly outperformed the BPNN model, showing better generalization to unseen data with a higher R^2 score of 0.877 on the test set. Both models provided robust predictions, and the analysis of feature importance confirmed that the water-cement ratio and cement content were the primary factors influencing compressive strength. These results suggest that machine learning techniques can significantly enhance the accuracy of concrete strength predictions, offering valuable insights for engineers in optimizing concrete mix designs. Future studies should consider expanding the dataset and integrating additional predictive variables to further refine model performance.

Reference

1. Poorarbabi, M. Ghasemi, M.A. Moghaddam, Concrete compressive strength prediction using non-destructive tests through response surface methodology, Ain Shams Eng. J. 11 (4) (2020) 939–949, https://doi.org/10.1016/j.asej.2020.02.009.

2. K. Ali-Benyahia, Z. Sbartai, D. Breysse, M. Ghrici, S. Kenai, Improvement of nondestructive assessment of on-site concrete strength: influence of the selection process of cores location on the assessment quality for single and combined NDT techniques, Constr. Build. Mater. 195 (2019) 613–622, https://doi.org/ 10.1016/j.conbuildmat.2018.10.032.
3. K. Kaboosi, F. Kaboosi, M. Fadavi, Investigation of greywater and zeolite usage in different cement contents on concrete compressive strength and their interactions, Ain Shams Eng. J. 11 (1) (2020) 201–211, https://doi.org/10.1016/j.asej.2019.08.008.
4. A. Poorarbabi, M. Ghasemi, M.A. Moghaddam, Concrete compressive strength prediction using neural networks based on non-destructive tests and a self- calibrated response surface methodology, J. Nondestruct. Eval. 39 (784) (2020), https://doi.org/10.1007/s10921-020-00718-w.
5. P.G. Asteris, A.D. Skentou, A. Bardhan, P. Samui, P.B. Lourenco, Soft computing techniques for the prediction of concrete compressive strength using non- destructive tests, Constr. Build. Mater. 303 (124450) (2021), https://doi.org/10.1016/j.conbuildmat.2021.124450.
6. S. Karahan, A. Buyuksarac, E. Isik, The relationship between concrete strengths obtained by destructive and non-destructive methods, Iran. J. Sci. Technol. Trans. Civ. Eng. 44 (1) (2020) 91–105, https://doi.org/10.1007/s40996-019-00334-3.
7. J. Alexandre Bogas, M. Gloria Gomes, A. Gomes, Compressive strength evaluation of structural lightweight concrete by non-destructive ultrasonic pulse velocity method, Ultrasonics 53 (5) (2013) 962–972, https://doi.org/10.1016/j.ultras.2012.12.012.
8. K. Rashid, R. Waqas, Compressive strength evaluation by non-destructive techniques: an automated approach in construction industry, J. Build. Eng. 12 (2017) 147–154, https://doi.org/10.1016/j.jobe.2017.05.010.
9. R. Mata, R.O. Ruiz, E. Nunez, Correlation between compressive strength of concrete and ultrasonic pulse velocity: a case of study and a new correlation method, Constr. Build. Mater. 369 (130569) (2023), https://doi.org/10.1016/j.conbuildmat.2023.130569.
10. R. Pucinotti, Reinforced concrete structure: non destructive in situ strength assessment of concrete, Constr. Build. Mater. 75 (2015) 331–341, https://doi.org/ 10.1016/j.conbuildmat.2014.11.023.
11. N.G. Yilmaz, R.M. Goktan, Comparison and combination of two NDT methods with implications for compressive strength evaluation of selected masonry and building stones, Bull. Eng. Geol. Environ. 78 (6) (2019) 4493–4503, https://doi.org/10.1007/s10064-018-1382-7.
12. D. Breysse, Nondestructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods, Constr. Build. Mater. 33 (2012) 139–163, https://doi.org/10.1016/j.conbuildmat.2011.12.103.
13. M. Alwash, D. Breysse, Z.M. Sbartai, Non-destructive strength evaluation of concrete: analysis of some key factors using synthetic simulations, Constr. Build. Mater. 99 (2015) 235–245, https://doi.org/10.1016/j.conbuildmat.2015.09.023.
14. M. Vona, D. Nigro, Evaluation of the predictive ability of the in situ concrete strength through core drilling and its effects on the capacity of the RC columns, Mater. Struct. 48 (4) (2015) 1043–1059, https://doi.org/10.1617/s11527-013-0214-2.
15. A. Fiore, F. Porco, G. Uva, M. Mezzina, On the dispersion of data collected by in situ diagnostic of the existing concrete, Constr. Build. Mater. 47 (2013) 208–217, https://doi.org/10.1016/j.conbuildmat.2013.05.001.
16. A. Masi, L. Chiauzzi, An experimental study on the within-member variability of in situ concrete strength in RC building structures, Constr. Build. Mater. 47 (2013) 951–961, https://doi.org/10.1016/j.conbuildmat.2013.05.102.

17. S.R. Naraganti, R.M.R. Pannem, J. Putta, Impact resistance of hybrid fibre reinforced concrete containing sisal fibres, Ain Shams Eng. J. 10 (2) (2019) 297–305, https://doi.org/10.1016/j.asej.2018.12.004.
18. E.A. Whitehurst, Soniscope tests concrete structures, J. Am. Concr. Inst. 2 (47) (1951) 433–444.
19. J.R. Leslie, W.J. Cheesman, An ultrasonic method of deterioration and cracking in concrete structures, Acids J. Proc. 9 (46) (1949) 17–36.
20. M. Erdal, Prediction of the compressive strength of vacuum processed concretes using artificial neural network and regression techniques, Sci. Res. Essays 4 (10) (2009) 1057–1065.
21. Q. Huang, P. Gardoni, S. Hurlebaus, Predicting concrete compressive strength using ultrasonic pulse velocity and rebound number, Acids Mater. J. 108 (4) (2011) 403–412.
22. K. Amini, M. Jalalpour, N. Delatte, Advancing concrete strength prediction using non-destructive testing: development and verification of a generalizable model, Constr. Build. Mater. 102 (1) (2016) 762–768, https://doi.org/10.1016/j.conbuildmat.2015.10.131.
23. M. Shariati, N.H. Ramli-Sulong, M.M.K.H. Arabnejad, P. Shafigh, H. Sinaei, Assessing the strength of reinforced concrete structures through ultrasonic pulse velocity and schmidt rebound hammer tests, Sci. Res. Essays 6 (1) (2011) 213–220.
24. G. Trtnik, F. Kavcic, G. Turk, Prediction of concrete strength using ultrasonic pulse velocity and artificial neural networks, Ultrasonics 49 (1) (2009) 53–60, https://doi.org/10.1016/j.ultras.2008.05.001.
25. I. Lawson, K. Danso, H. Odoi, Non-destructive evaluation of concrete using ultrasonic pulse velocity, Res J. Appl. Sci. Eng. Technol. 6 (3) (2011) 499–504.
26. Z. Sbartai, D. Breysse, M. Larget, J. Balayssac, Combining NDT techniques for improved evaluation of concrete properties, Cem. Concr. Compos. 34 (6) (2012) 725–733, https://doi.org/10.1016/j.cemconcomp.2012.03.005.
27. A. Jain, A. Kathuria, A. Kumar, V. Verma, K. Murari, Combined use of nondestructive tests for assessment of strength of concrete in structure, Procedia Eng. 54 (2013) 241–251.
28. S. Selvaraj, S. Sivaraman, Prediction model for optimized self-compacting concrete with fly ash using response surface method based on fuzzy classification, Neural Comput. Appl. 31 (5SI) (2019) 1365–1373, https://doi.org/10.1007/s00521-018-3575-1.