



Statistical analysis for evaluating concrete strength of existing structure using non-destructive and destructive test

Hemraj Ramdas Kumavat¹ · Narayan R. Chandak²

Received: 30 November 2023 / Accepted: 11 April 2024 / Published online: 1 May 2024
© Springer Nature Switzerland AG 2024

Abstract

Destructive tests for evaluating concrete compressive strength are costly and challenging in certain instances. Using a rebound hammer (RH) and ultrasonic pulse velocity (UPV), i.e., non-destructive methods for strength evaluation, proved more beneficial in all senses. However, calibrating the model between non-destructive testing (NDT) and compressive strength is essential for estimating strength. The reliability of this calibration is a crucial task that leads to selecting a minimal number of cores to be taken out (core) from a structure. The present study aims to identify and optimize the on-site reliability model. Extensive data from 275 core samples were obtained from the Construction Diagnostic Centre, Pune (India), which RH and UPV examined. The cores are taken from thirty existing RCC structures built between 1975 and 2005. The Root Mean Square Error (RMSE) and the coefficient of determination (R^2) for single method (SM) and combined method (CM) are used to investigate the total number of cores needed for calibration. According to $RMSE_{pred}$ and R^2_{pred} , at least 6–8 cores are required to achieve the correct prediction phase with a CM rather than using SM. The CM leads to more reliable results than an SM with the least RMSE and higher R^2 values by analyzing 100 iterations for each number of cores (NC). Also, the CM shows more reliable results than the SM in the fitting and prediction phase. As a reasonable number of samples, 9 cores must be considered to converge for an SM, compared to 6–8 cores necessary for CM to estimate the strength precisely.

Keywords Concrete strength · Calibration model · Non-destructive testing methods · On-site evaluation · Statistical analysis

Introduction

The structural members must be compacted at the construction stage to achieve durability and performance. For many reasons, some parts of structural members are compacted manually, or some parts may remain without compaction, and these also vary with the type of structure and location. When structural concrete is exposed to heat load, its hardness and microstructural behavior are affected. It is well-known that reducing the water-to-cement ratio enhances the strength of concrete and affects its hydration and hardness

properties. The strength of the existing concrete structure degrades over time due to environmental conditions and carbonation. In the construction phase, concrete shows higher alkalinity (pH-12 to 13) due to the $Ca(OH)_2$ in its interfacial transition zone. Reduction of voids and increased volume of reactants in the concrete due to lower pH and carbonation progress from surface to the internal structure.

The existing structures often need to be assessed for stability in various circumstances, such as rehabilitation, seismic behavior forecasting, and alteration following severe damage. It is essential to test the characteristics of the concrete to evaluate structural performance more precisely [1]. Even so, destructive testing of concrete allows only a tiny range of experiments, attributable to the explanation that, in some cases, it is costly, inefficient, and challenging in many old structures. The rebound hammer (RH) and ultrasonic pulse velocity (UPV) results for the same structures are used for reliability assessment with destructive testing results (DT). This will help reduce the need to take many cores to assess the performance of whole structural concrete

✉ Hemraj Ramdas Kumavat
kumavathr1981@gmail.com

Narayan R. Chandak
narayan.chandak@nmims.edu

¹ Department of Civil Engineering, R C Patel Institute of Technology, Shirpur, Maharashtra, India

² Department of Civil Engineering, SVKM Institute of Technology, Dhule, Maharashtra, India

behavior using a non-destructive technique [2]. The selection of core location is based on the average core results that can provide a representative picture of the complete dataset. Measurement uncertainty arises because the values obtained from destructive tests do not perfectly match non-destructive test values at the exact test location. Consideration of the position of the test is not represented as a specific point but as a restricted region, which makes it possible to consider the repeatability of the rebound test results.

The average values of root mean square error (RMSE) of the fitted and predicted model dataset could be more efficient, and the resulting reliability of the model needs to be addressed during the prediction phase. However, the reliability of the predicted model could be less than that of the fitted model. The difference between fitted and predicted error in the model generally evolves due to the generalization of the model approach. Benyahia [3] pointed out the technological challenge of testing the model identification approach's real dataset at the prediction stage.

In some instances, the use of a single NDT method can be less precise; hence, combining the NDT (RH and UPV) methods is always better to ensure higher accuracy and reliability of the concrete strength assessment. More emphasis is placed on the trustworthy conversion model, which depends on factors such as the number of core and heterogeneous datasets. [4–6].

NDT incorporated with cores provides a better option for structural strength evaluation [7, 8]. An empirical relation must be formed between the findings of the NDT and the strength calculated on cores taken from the exact locations. EN (13,791–2019) [9] suggested that at least eight cores for > 75 mm diameter, where it is recommended to core at least ten test locations, to allow for possible outliers. The minimum number of cores > 75 mm diameter of cores for use in combination with indirect testing is three, where at least four test locations should be cored to allow for a possible outlier. ACI 228.1R [2] standard suggested that at least 6–9 data sets are needed for the non-destructive strength assessment. Even so, consulting practices are typically based on a minimum of 3 or 4 cores, although the accuracy of this measurement is still being determined and needs to be addressed [4].

Cristoforo [10] offered a wide range of approaches for describing concrete strengths, which made use of a broad dataset to verify the efficiency of established prediction models and suggested new relationships that are successful in predicting the substantial strength of the Italian reinforced cement concrete (RCC) structures constructed in the twentieth century.

Careful consideration has been paid to the reliability of the model based on the following parameters: a large number of cores data, NDT methods for quality control, range of statistical parameters (Range, Standard Deviation (SD),

coefficient of variation (COV)), reliability of the model, the influence of uncontrolled parameters (cracks, carbonation, moisture) [6, 11].

Due to variations in the mix proportion, the core stresses in normal-strength concrete are double that of those in low-strength concrete if the transverse and longitudinal bars are identical [12–14]. A machine learning algorithm was employed to establish a correlation between variables and outcomes and forecast the 28-day compressive strength while modifying the water-to-cement ratio and superplasticizer. The various models are used for predicting future results [15, 16]. An acceptable level of precision is additionally appreciated for concrete strength estimation. Therefore, the resulting equations for strength prediction could be used securely for in situ concrete strength assessment [17].

Implication of the research

1. To address the influence of the number of cores used for calibration using statistical parameters: the root mean square error (RMSE) and the coefficient of determination (R^2).
2. To conduct statistical analysis for a dataset collected from various structures constructed between 1975 and 2005, and to compare with two indicators RMSE, R^2 .
3. To access the reliability precision of the model in the fitting and prediction phase.
4. To compare model identification approach for the effectiveness and performance of single-method versus combined-method strategies.
5. To scrutinize the practical performance of the model in real-world scenario, with a specific emphasis on evaluating the effectiveness of the calibration approach.

Dataset collection

The dataset collected from the Construction Diagnostic Centre, Pune, shown in this research study, relates to thirty (five structures in each group) RCC structures located in Pune, which comes under the metropolitan area, where the humidity and carbonation affect the atmosphere severely. The structural concrete elements (column and beam) underwent NDT (RH and UPV), and DT tests of cores were taken. Coring and NDT measurements were conducted perpendicularly to the direction of the concrete casting. The UPV test is performed at the propagation rate of the ultrasound velocity determined on the structural elements under study. In contrast, the rebound hammer test is defined by the rebound index. A total of 275 elements, consisting of columns and beams, underwent testing through a combination of NDT and coring. For each element, two NDT measurements were conducted before coring, and one to three cores were

extracted at the NDT measurement location. In total, 275 core samples with a 61–105 mm dia. were extracted, as shown in Table 1 and subjected to compression testing until ultimate failure.

The core diameter drops for a specific height/diameter (H/D) ratio, reducing the specimen volume significantly. As a result, the diameter of the core increases with decreasing strength. The core length generally depends upon the length/diameter (L/D) ratio; it should be 1.0 for the cube strength value and 2.0 for the cylinder strength value. It is required to apply a correction factor if the ratio is less than 2.0. The reason for using a correction factor is that long cores tend to give lower test values than short cores of the same cross-sectional area. A correction factor is applied to the calculated strength to convert all core specimens to approximately the same strength as the standard test cylinder. It is experienced that the core diameter gets affected by factors like maximum aggregate size, member dimension, and steel reinforcement in the member. EN 12504–1 [18] suggested that the core diameter/aggregate (D/A) ratio must be more than 3.0, while ASTM C42 [19] suggested its value as 2.0. Several standards consider the diameter of the core to be 100 mm,

although, in Australia, a diameter of 75 mm is considered acceptable [20]. Table 1 shows the dataset for different years of construction grouped into six categories (D1 to D6) since it reflects the variation in strength in structural members caused by environmental conditions and voids, cracks, and carbonation, with parameters such as number of cores (diameter and height), ranges, standard deviation (SD), and coefficient of variation (COV). The core strength mentioned in this study has been calculated after conversion to an equivalent in-situ cube strength (equivalent cube strength as core strength/0.8). The value associated with the COV of the core strength, which is in the range of 13–35%, demonstrates that the dataset presented here is heterogeneous.

Assessment methodology

The assessment methodology is based on and influenced by controlled parameters such as the number of test locations for cores-NT, within-test variability, and methods of NDT (RH, UPV, and Combined Method (CM)). Uncontrolled parameters considered are concrete mix characteristics, i.e.,

Table 1 Statistical database parameters based on year of construction

Year	Statistical parameters	Diameter of core (mm)	Height of core (mm)	UPV (Km/Sec)	Rebound No	Core Equi. Cube Strength (MPa)
1975–1980	No of cores	50	50	50	50	50
	Range	73.88–105.74	85.64–204.25	0.768–4.368	10.07–48.37	7.65–23.75
	SD	13.42	30.10	0.84	8.40	4.14
	COV	0.16	0.21	0.27	0.28	0.26
1981–1985	No of cores	50	50	50	50	50
	Range	73.64–105.50	85.27–204.25	0.976–4.078	15.08–50	6.16–25.83
	SD	15.73	29.92	0.82	9.38	5.33
	COV	0.19	0.21	0.29	0.28	0.35
1986–1990	No of cores	50	50	50	50	50
	Range	61.08–105.47	80.23–209.07	2.028–4.503	17.50–47.37	5.94–28.28
	SD	12.98	28.32	0.64	8.94	5.90
	COV	0.16	0.20	0.21	0.27	0.35
1991–1995	No of cores	25	25	25	25	25
	Range	73.53–105	74.38–200.30	1.516–4.409	12.37–50.9	9.37–21.28
	SD	0.36	22.82	0.66	10.58	2.93
	COV	0.00	0.17	0.21	0.29	0.18
1996–2000	No of cores	50	50	50	50	50
	Range	73.49–105.71	80.21–210.99	2.610–4.166	22.12–50.28	12–38.09
	SD	15.82	37.20	0.41	6.25	6.41
	COV	0.18	0.22	0.12	0.17	0.25
2000–2005	No of cores	50	50	50	50	50
	Range	73.04–74.94	87.26–147.87	0.875–4.098	17.05–47	5.96–24.98
	SD	0.47	18.02	0.69	8.07	4.72
	COV	0.01	0.13	0.23	0.23	0.29

type of cement, aggregate, compaction methods, curing conditions, water-to-cement ratio, elevated temperature, and carbonation. Fig. 1 shows a flow chart of the strength evaluation using reliability techniques. The evaluation approach for model accuracy is implemented and tested on a homogeneous data set of 275 core results collected from the in-situ structure categorized in D1 to D6, following normal distribution. Values obtained from the difference between estimated and observed results are strength values achieved by applying the same technique to a small set of cores. The core strength results selected for each number of cores (NC) between 2 and 20 are randomly chosen from the 275 core samples.

The statistical regression is used to validate the relationship between core strength and strength estimated by the NDT (RH, UPV) method. A set of NC (x and f_c) pairs is obtained, where x represents the NDT measurement (RH, UPV- Independent variable) and f_c (Dependent variable) is the equivalent cube strength of the core corresponding to the test location. The fitted error has been estimated from the NC pairs, and the predicated error has been estimated from the (NT-NC) pairs in terms of RMSE and the coefficient R^2 . The compressive strength values are available at all core points in this study. This condition is often impractical in common scenarios due to the constraints on the dataset size. Since the findings may have some randomness, every collection of NC data showed a different outcome. Because

of this, the same process is repeated 100 times for each NC element. The mean and standard deviation of the strength values over 100 repetitions shall be calculated for the NC reliability assessment.

Estimation methodology

Coefficient of determination (R^2) estimation

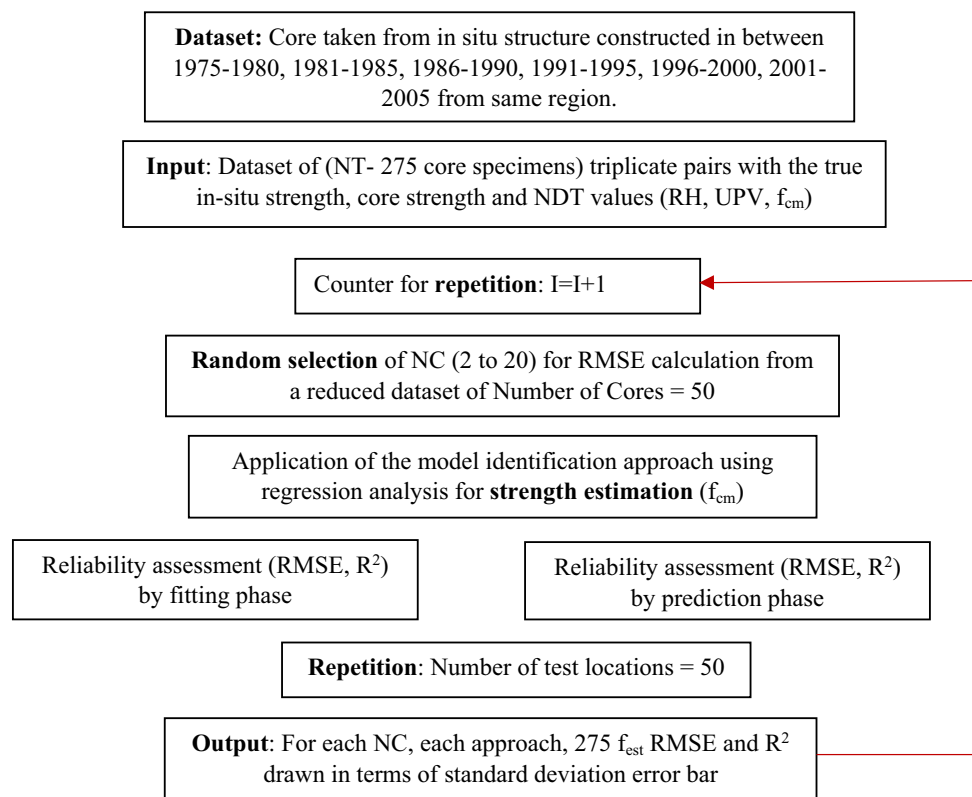
The coefficient of determination (R^2) is commonly used by researchers as a measure of the effectiveness of the fitted regression model, as indicated in Eq. (1).

$$R^2 = 1 - \frac{SSE}{SST} \tag{1}$$

where R^2 —coefficient of determination, SST- is the total sum of squares (total variation in values of the response variable) and SSE- is the sum of squares of residuals (amount of variation unexplained by the fitted regression model)

To analyze the relationship between core strength and NDT parameters (RH, UPV). We used the R^2 coefficient as stated in Eq. (2). NDT tests were conducted before extracting the core from the in-situ members. The correlation between the two measures is used to evaluate concrete compressive strength. Ministry of railways-India [21] suggested that the probable accuracy of estimating

Fig. 1 Flow chart of the strength evaluation using reliability techniques



concrete strength in a structure is $\pm 25\%$. The coefficient R^2 assesses the reliability of goodness of fit but does not warn about the model's reliability when applied to the prediction. The R^2 value for single and multiple variables with two regression equations (Power and Exponential) with nonlinear mathematical models are used in this analysis, as shown in Table 2. Equations of one variable (RH or UPV) provide best-fit correlations $R^2 > 0.70$, which has improved $R^2 > 0.80$ in the case of two variables (combined) equations.

$$R^2 = \frac{\left[\sum_{i=1}^{NC} (f_m - \bar{f})^2 - \sum_{i=1}^{NC} (f_m - f_e)^2 \right]}{\sum_{i=1}^{NC} (f_m - \bar{f})^2} \tag{2}$$

where f_m , f_e , and f are measured, estimated, and average strength respectively, NC- number of cores

According to the EN (13,791–2019) [9], at least eight cores for > 75 mm diameter of cores, where it is recommended to take at least ten test locations (core) to allow for possible outliers. According to the ACI 228.1R-03 [2], a minimum of twelve cores is required for the regression approach. RILEM TC 43-CND [22] committee

Table 2 Destructive and Non-destructive parameters

Year	Method	Equations	Type	R^2
1975–1980 (D1)	UPV	$f_{cm} = 6.8454 * V^{0.7359}$	Power	0.75
		$f_{cm} = 6.2165 * e^{0.2885V}$	Expo	0.73
	RH	$f_{cm} = 1.4347 * R^{0.7108}$	Power	0.72
		$f_{cm} = 6.7881 * e^{0.0273R}$	Expo	0.65
		$f_{cm} = 2.317 * (R^{0.387}) * (V^{0.546})$	Power	0.82
1981–1985 (D2)	UPV	$f_{cm} = 5.6741 * V^{0.9236}$	Power	0.75
		$f_{cm} = 4.4843 * e^{0.4039V}$	Expo	0.79
	RH	$f_{cm} = 0.4104 * R^{1.0205}$	Power	0.74
		$f_{cm} = 4.3936 * e^{0.0348R}$	Expo	0.76
		$f_{cm} = 0.741 * (R^{0.658}) * (V^{0.655})$	Power	0.85
1986–1990 (D3)	UPV	$f_{cm} = 3.0186 * V^{1.5306}$	Power	0.75
		$f_{cm} = 3.4964 * e^{0.4999V}$	Expo	0.71
	RH	$f_{cm} = 0.3306 * R^{1.121}$	Power	0.72
		$f_{cm} = 4.8074 * e^{0.0363R}$	Expo	0.73
		$f_{cm} = 0.569 * (R^{0.755}) * (V^{0.677})$	Power	0.86
1991–1995 (D4)	UPV	$f_{cm} = 4.104 * e^{(0.025R)} * e^{(0.187V)}$	Expo	0.85
		$f_{cm} = 7.4362 * V^{0.6951}$	Power	0.74
	RH	$f_{cm} = 7.4854 * e^{0.2457V}$	Expo	0.74
		$f_{cm} = 3.5129 * R^{0.4308}$	Power	0.75
		$f_{cm} = 9.1285 * e^{0.0156R}$	Expo	0.78
1996–2000 (D5)	UPV	$f_{cm} = 4.032 * (R^{0.254}) * (V^{0.437})$	Power	0.82
		$f_{cm} = 7.398 * e^{(0.010R)} * e^{(0.136V)}$	Expo	0.86
	RH	$f_{cm} = 2.2338 * V^{1.939}$	Power	0.74
		$f_{cm} = 3.3974 * e^{0.5703V}$	Expo	0.74
		$f_{cm} = 0.2604 * R^{1.2758}$	Power	0.74
2000–2005 (D6)	UPV	$f_{cm} = 6.4351 * e^{0.0373R}$	Expo	0.72
		$f_{cm} = 0.814 * (R^{0.506}) * (V^{1.304})$	Power	0.82
	RH	$f_{cm} = 4.122 * e^{(10.015R)} * e^{(0.363V)}$	Expo	0.81
		$f_{cm} = 5.4942 * V^{0.9874}$	Power	0.78
		$f_{cm} = 4.5742 * e^{0.4112V}$	Expo	0.80
2000–2005 (D6)	UPV	$f_{cm} = 0.3415 * R^{1.091}$	Power	0.78
		$f_{cm} = 4.8052 * e^{0.0344R}$	Expo	0.76
	RH	$f_{cm} = 0.879 * (R^{0.633}) * (V^{0.621})$	Power	0.86
		$f_{cm} = 4.305 * e^{(10.018R)} * e^{(0.22V)}$	Expo	0.87
		$f_{cm} = 4.305 * e^{(10.018R)} * e^{(0.22V)}$	Expo	0.87

recommended a process to develop iso-strength curves for concrete (materials and mix from different countries for which the curves are developed). If the mix is known, a correction factor should be applied, and if it is unknown,

the correction factor should be estimated using cores extracted from the structure under investigation.

Figure 2 illustrates the combined approach's experimental vs. predicted $FCM \approx f_{cm}$ results with the regression model. Less variance is reported in D1, D4, D5, and D6, while more

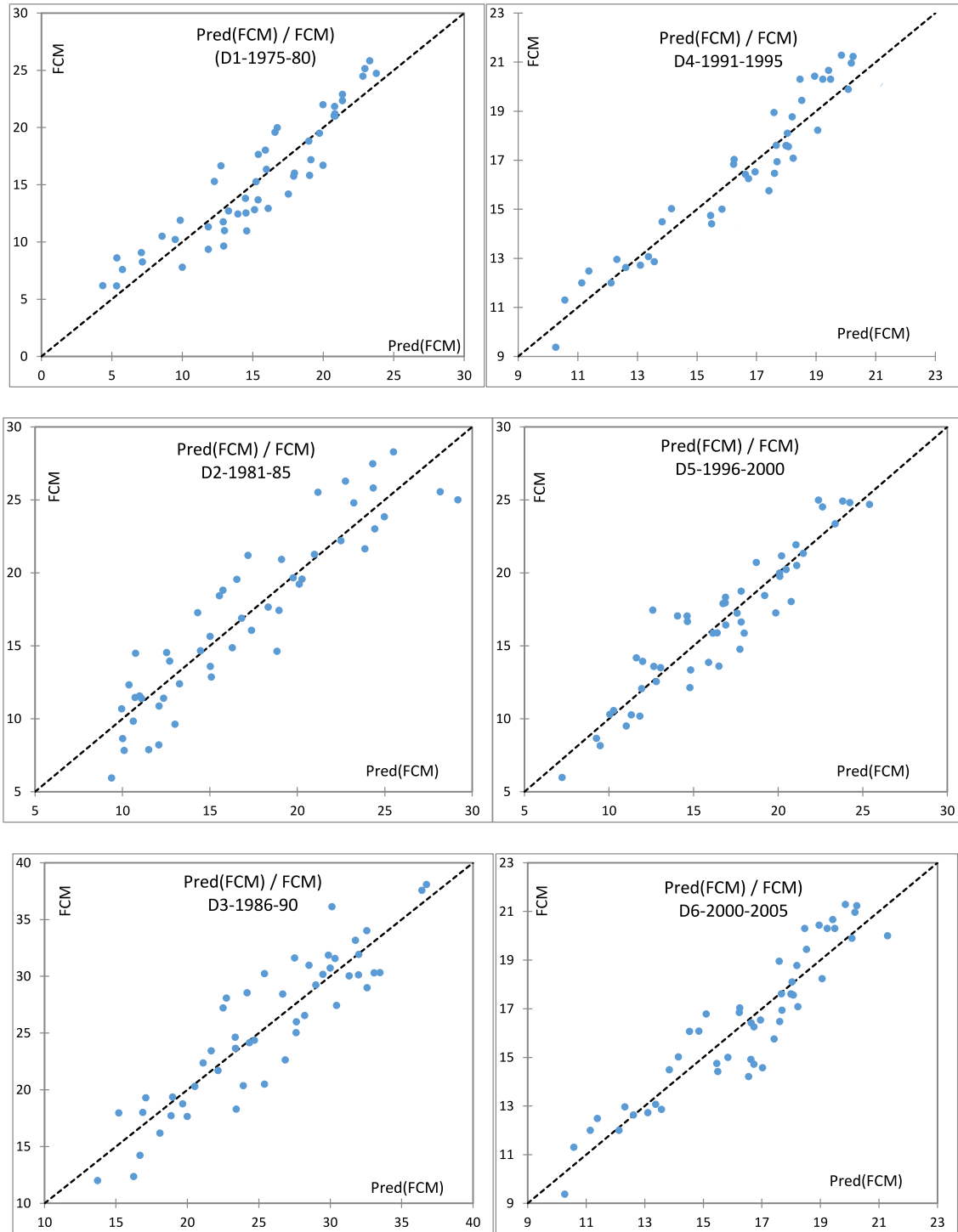


Fig. 2 Experimental vs. predicted FCM results of combined method

significant variance is observed in D2 and D3. Hence, it is essential for correct model validation to use the regression approach. All outcomes are derived using a power function with a standard deviation ranging from 2.93 to 6.41 and a coefficient of determination (R^2) ranging from 0.82 to 0.86.

Variability

This approach is used to identify “regions” that can be considered homogeneous in terms of the COV. All data sets appear to have relatively closer results according to mean, SD, and COV. Results from 1986 to 1990 present slightly better statistical values than others, as shown in Table 2. The variability of the NDT parameters measured between the mean values for each database is approximately 5 percent, which is just one-third of the within-subset variability. Kumavat and Chandak [23] pointed out that uncertainties affect assessment, and the concrete strength variability was induced in the structure due to batch-to-batch (mix composition, compaction deficiency, different curing conditions), member-to-member (environmental conditions), within member (voids and cracks) conditions.

RMSE estimation

Another measure that Liu et al. [24] can be used for this purpose is the Root Mean Square Error (RMSE) as given in Eq. (3). Its functional interest is that it gives the amount of error linked to the predicted strengths more explicitly.

$$RMSE_{(fitted)} = \sqrt{\frac{\sum_{i=1}^{NC} (f_m - f_e)^2}{NC}} \tag{3}$$

The R^2 values, indicating the correlation degree between predicted and experimental values, vary between the lowest and highest across all models. The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) values demonstrate minimal error levels when compared to individual models [25, 26]. The model established for core test locations is used in actual practice to predict strength at test locations where only NDT measurements exist. The predictability of the model is, therefore, a crucial issue. According to Shcherbakov [27], statistics carried out a range of prediction indicators. Between these indicators, RMSE is used by NDT experts to test the error of prediction Nobile [28]. The prediction error can be determined for (NT-NC) test locations using the following adjusted form of Eq. (4).

$$RMSE_{(predicated)} = \sqrt{\frac{\sum_{i=1}^{NT-NC} (f_m - f_e)^2}{NT - NC}} \tag{4}$$

where NC- number of cores, NT- total number of test locations for NDT measurements

Results and discussion

Influence of the number of cores on the evaluation accuracy

The values of RMSE (RH, UPV, and CM (combined method)) calculated for fitted and predicated error are shown in Fig. 3 as a function of the number of cores. The average value of RMSE at the prediction stage is always higher than at the fitting stage, as shown in Fig. 3. When NC increases, their standard deviation decreases. The difference between the fitting and prediction error is due to extrapolation. This precision difference is very significant, mainly when NC is small, but in actual practice, only the precision of the fitting model is estimated. If fitting errors are significantly low, it may indicate overfitting, where the model performs exceptionally well on the training data but fails to generalize to new, unseen data.

On the other hand, understanding prediction errors assists in refining the model to enhance its generalization ability. Recognizing the precision difference allows practitioners to address overfitting and create robust and applicable models to diverse situations. Kumavat and Chandak [29] pointed out that the relationship between CV and rebound values varies with the structural age (18–32 yrs.) and its type. It is helpful to be conscious that a fitting model could have a good quality of precision and a deficient predictive capacity. In the case of minimal NC numbers, this problem is essential because the graphs show that the predicted RMSE can be doubled that of the fitted RMSE. The consistent values of RMSE are seen when the number of cores ranges between 9 and 20; this indicates that a minimal number of nine cores is required for assessment precision. Regarding the number of cores from 9 to 20, prediction and fitting errors are stabilized at 1.8 MPa for RMSE with RH and 1.9 MPa for RMSE with UPV for groups D1, D2, and D6.

Similarly, prediction and fitting errors are stabilized at 2 MPa for RMSE with RH and UPV for groups D3 and D5 and 1.0 MPa for RMSE with RH and UPV for group D4. These errors occur due to measurement uncertainties and the influence of uncontrolled factors like carbonation, moisture condition, and temperature variation Kumavat et al. [30]. When comparing results from different experiments or instruments, knowing uncertainties is essential. It provides a basis for evaluating the reliability of the data and determining whether observed differences are statistically significant or within the expected margin of error. A clear understanding of uncertainties ensures that researchers stay within the precision of their findings. Uncontrolled factors refer to the

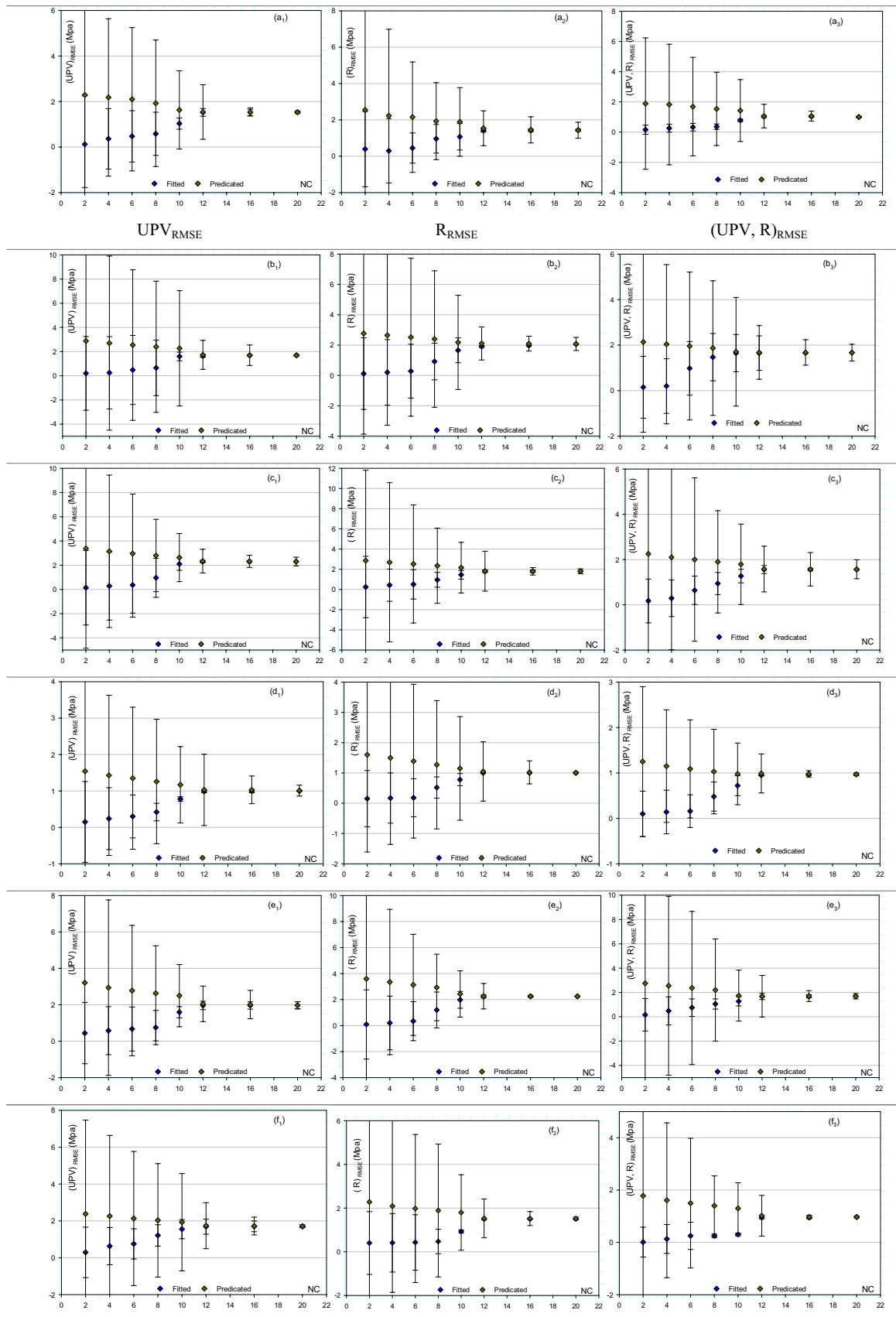


Fig. 3 RMSE (Mean and SD) of fitted and predicted models with construction age between 1975 and 2005

variables that can affect measurements but are not under direct observation or control. Error mitigations may involve refining experimental setups, conducting additional analyses, generalizing results, and improving reproducibility.

Table 3 illustrates the comparative analysis of RMSE_{pred} and R₂_{pred} between single and combined NDT methods. When NC increases, the RMSE_{pred} of UPV is greater than that of RH, and the R₂_{pred} of RH is greater than that of UPV with a consistent variance. According to RMSE_{pred} and R₂_{pred}, at least 6–8 cores are required to achieve a better prediction phase with a combined approach compared to a single method. The combined method yields more reliable results than a single method regarding the least RMSE and more excellent R₂ value by analyzing 100 iterations for each NC.

Generalization of results

Alwash et al. [4] investigated the efficacy of RMSE_{pred} for a combination of NDT measurements (UPV and R) according to NC. Based on average RMSE, the conclusion indicates the effectiveness is observed beyond a minimal number of cores (9 or 5). It is observed when the NDT measurement quality of the first technique is high (low within-test variability) and the second technique is low or average. Benyahia et al. [3] incorporating a second technique proves effective with at least 5 cores if the first technique is UPV and with at least 6 cores if the first technique is R. This difference arises from the repeatability of rebound measurements. The minimal number of cores specified by both authors may be specific to the case. In the present work, a second technique proves to be effective with at least 8 cores if the first technique is UPV and with at least 6 cores if the first technique is R. In addition, both techniques proved effective for earlier years (2000–2005) results with at least 6 cores illustrated in Table 3.

Influence of number of cores on the efficacy of the CM

The combined method (CM) trends for the same database are almost similar to those of a single approach. A modest improvement in RMSE, due to the combined method, has also stabilized a slightly lower RMSE trend than a single method. By observing Fig. 3, the fitted error is lower for the combined approach (CM) than for the single method (SM) when the number of cores (NC) differs. The efficacy of the combined technique during the fitting phase is not affected by the NC.

When NC varies, the RMSE_{fit} of UPV is more significant than RH and equivalent to the constant difference, as shown

in Fig. 4. Also, when NC varies, the R²_{fit} of rebound is more effective than that with UPV which is equivalent to the constant difference as shown in Table 3. Comparative analysis of the fitted R² value in terms of the construction age (D1 to D6) with regression equation (Power and Exponential function) as shown in Fig. 5. It should be stressed here that these findings are specific to this present study and rely primarily on the respective precision of both evaluation methods in this situation. From previous case studies [3], identical or opposite circumstances seem to be based on the structural condition and measuring system.

However, the model of the CM at the fitted phase tends to be less dispersive than the models with an SM, where the NC < 6, as shown in Fig. 4. This is because the model parameters have to be identified from the measurement error database [11]. It demonstrates that using a CM with limited cores at the prediction phase is risky and can result in enormous uncertainties. Hence, attention must be paid to the calibration of the CM since the minimum number of NCs often used is a common practice in consultancy work [4].

The present study shows that at least 6–8 cores (concerning RMSE and coefficient of determination (R²)) are necessary to achieve a better prediction phase with a combined method than a single method. The number of cores should be increased to ensure the CM is extremely strong and significantly more reliable than an SM. Consequently, it is also essential to analyze the statistical interpretation of the data to provide a more detailed understanding of the effectiveness of the CM during the prediction phase. Therefore, it is necessary to check and quantify every 100 iterations to know how the CM contributes to the lower RMSE and a higher R² than SM.

The absence of unanimous agreement among practitioners and researchers based on the study findings on core numbers can be attributed to the choice of the success rate (tolerance), which is considered a genuine benefit when utilizing CM. This success rate is based on consideration of the practical quality of each non-destructive test; the minimum number of cores needed will vary. The authors recommended that CM measurements require a minimum of 6 to 8 cores for an expected outcome.

Conclusion

After conducting tests on cores obtained from structures built between 1975 and 2005 and performing statistical analysis, the following inferences can be made:

The results were utilized for statistical analysis in assessing strength, emphasizing the influence of the number of cores during the model identification approach using SM and CM methods. In the fitting and prediction stage, the variation in error depends on the number of cores used for calibration,

Table 3 Comparative analysis of $RMSE_{pred}$ and R^2_{pred} between SM and CM (NDT)

NC	UPV			RH			CM		
	RMSE	SD	R ²	RMSE	SD	R ²	RMSE	SD	R ²
1975–1980 (D ₁)									
2	2.29	4.07	0.72	2.55	5.45	0.72	1.89	4.35	0.82
4	2.18	3.46	0.73	2.23	4.76	0.74	1.83	3.99	0.85
6	2.10	3.15	0.76	2.15	3.04	0.76	1.69	3.26	0.88
8	1.92	2.78	0.77	1.93	2.12	0.82	1.54	2.43	0.90
10	1.96	1.72	0.78	1.89	1.89	0.83	1.43	2.05	0.91
12	1.53	1.20	0.79	1.54	0.96	0.87	1.06	1.17	0.92
16	1.53	0.18	0.80	1.45	0.72	0.87	1.06	0.85	0.92
20	1.53	0.06	0.82	1.43	0.44	0.88	1.00	0.08	0.93
1981–1985 (D ₂)									
2	2.88	8.51	0.75	2.77	6.64	0.74	2.14	3.97	0.85
4	2.70	7.20	0.76	2.65	5.93	0.76	2.04	3.50	0.87
6	2.54	6.23	0.76	2.52	5.21	0.80	1.96	3.25	0.9
8	2.40	5.43	0.76	2.40	4.50	0.80	1.87	2.96	0.93
10	2.27	4.77	0.78	2.30	3.11	0.82	1.71	2.39	0.94
12	1.73	1.20	0.82	2.11	1.09	0.84	1.68	1.18	0.95
16	1.70	0.85	0.87	2.10	0.49	0.85	1.68	0.56	0.95
20	1.69	0.07	0.87	2.08	0.43	0.87	1.67	0.37	0.96
1986–1990 (D ₃)									
2	3.38	8.24	0.75	2.86	8.95	0.70	2.25	5.08	0.84
4	3.15	6.29	0.76	2.69	7.89	0.76	2.10	4.08	0.86
6	2.96	4.91	0.76	2.52	5.85	0.79	2.00	3.61	0.87
8	2.80	2.99	0.76	2.35	3.73	0.83	1.90	2.26	0.89
10	2.63	1.99	0.76	2.16	2.51	0.86	1.79	1.78	0.89
12	2.34	0.98	0.77	1.81	1.98	0.87	1.58	1.01	0.90
16	2.32	0.50	0.78	1.80	0.37	0.88	1.57	0.74	0.91
20	2.30	0.36	0.78	1.80	0.24	0.90	1.57	0.42	0.91
1991–1995 (D ₄)									
2	1.54	2.70	0.73	1.60	3.21	0.74	1.25	1.65	0.82
4	1.43	2.20	0.75	1.50	2.86	0.77	1.15	1.24	0.83
6	1.35	1.95	0.77	1.39	2.54	0.79	1.09	1.08	0.85
8	1.26	1.71	0.80	1.27	2.12	0.81	1.03	0.93	0.87
10	1.17	1.05	0.82	1.15	1.71	0.84	0.98	0.68	0.87
12	1.03	0.98	0.86	1.05	0.98	0.86	0.97	0.43	0.88
16	1.03	0.38	0.87	1.02	0.38	0.87	0.97	0.07	0.90
20	1.01	0.15	0.87	1.01	0.03	0.87	0.97	0.03	0.90
1996–2000 (D ₅)									
2	3.21	6.98	0.74	3.60	8.94	0.74	2.75	8.87	0.82
4	2.94	4.82	0.76	3.35	5.59	0.77	2.55	7.36	0.84
6	2.78	3.58	0.79	3.13	3.89	0.80	2.37	6.30	0.87
8	2.63	2.61	0.81	2.93	2.56	0.82	2.20	4.20	0.89
10	2.50	1.71	0.83	2.43	1.78	0.87	1.74	2.10	0.89
12	2.05	0.98	0.84	2.26	0.98	0.89	1.69	1.71	0.90
16	2.01	0.78	0.88	2.25	0.08	0.90	1.70	0.45	0.92
20	1.97	0.20	0.88	2.25	0.02	0.91	1.69	0.22	0.93
2000–2005 (D ₆)									
2	2.38	5.09	0.77	2.28	5.63	0.76	1.78	4.11	0.86
4	2.26	4.38	0.77	2.09	3.95	0.80	1.61	2.96	0.87
6	2.13	3.64	0.80	1.98	3.39	0.81	1.50	2.48	0.89

Table 3 (continued)

NC	UPV			RH			CM		
	RMSE	SD	R ²	RMSE	SD	R ²	RMSE	SD	R ²
8	2.03	3.08	0.82	1.89	3.04	0.81	1.40	1.14	0.91
10	1.93	2.64	0.83	1.80	2.73	0.83	1.30	0.98	0.92
12	1.74	1.25	0.85	1.53	1.07	0.87	1.02	0.78	0.92
16	1.72	0.48	0.85	1.52	0.78	0.88	0.97	0.07	0.95
20	1.71	0.08	0.85	1.52	0.07	0.88	0.97	0.03	0.95

Significance of bold values illustrate minimal NC values for RMSE, R2, and SD for each NDT (UPV, RH, CM)

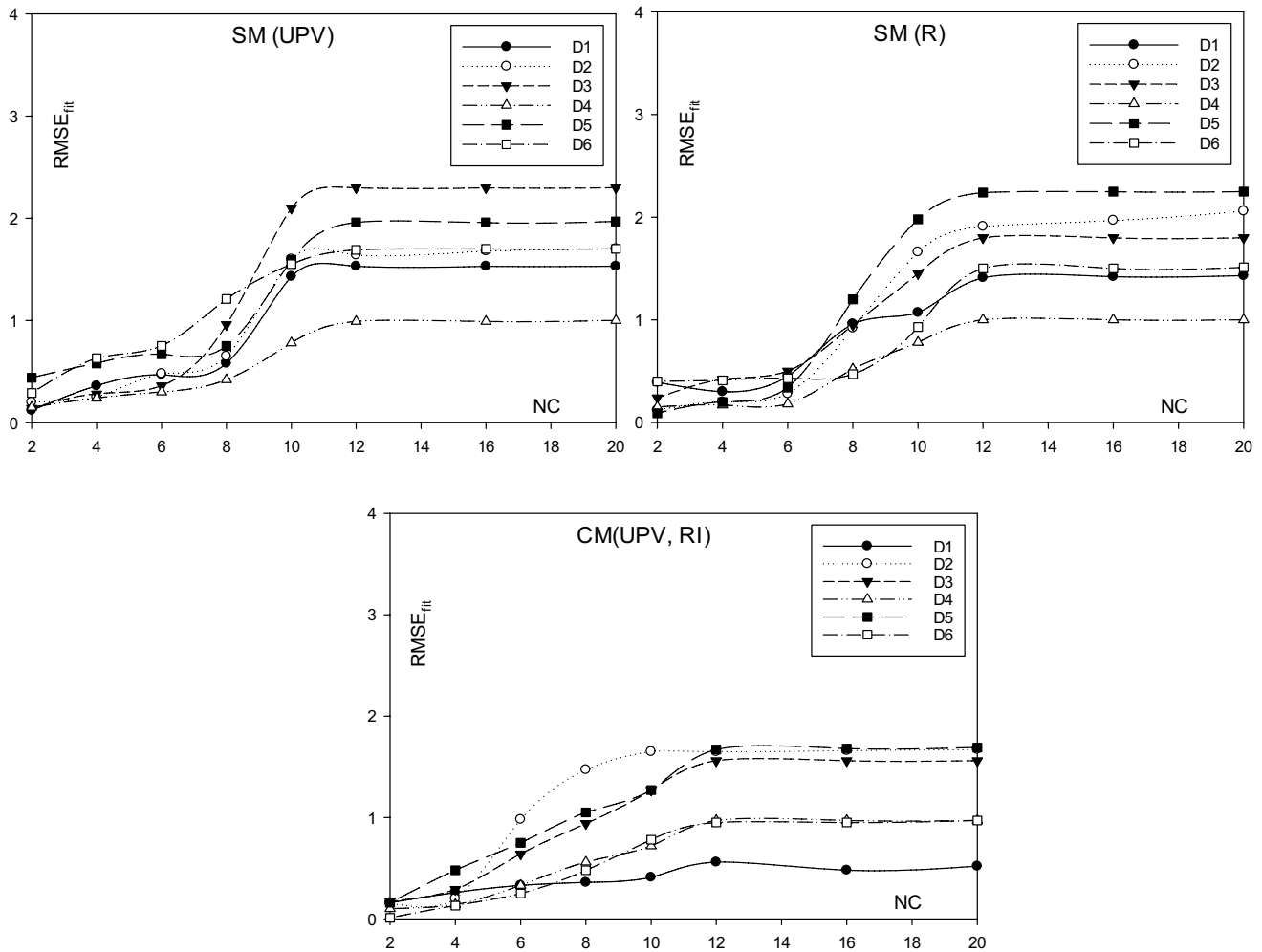


Fig. 4 Comparative analysis of the fitted RMSE in terms of the number NC with construction ages (D1 to D6). **a** NDT method (UPV value) **b** NDT method (R-value) **c** CM (R, UPV value)

expressed in the root mean square error (RMSE) and coefficient of determination (R²), which will logically reduce as NC increases. The present study concluded that nine cores must be considered a reasonable number of samples to converge for a single NDT method; additional increases yielded little improvements in the assessment.

It is, therefore, necessary to check and quantify each of the 100 iterations to know how the combined method contributes to a lower RMSE and higher R² than the single method. The average value of RMSE at the prediction stage is always higher than the fitting stage, but its standard deviation decreases when the number of cores increases.

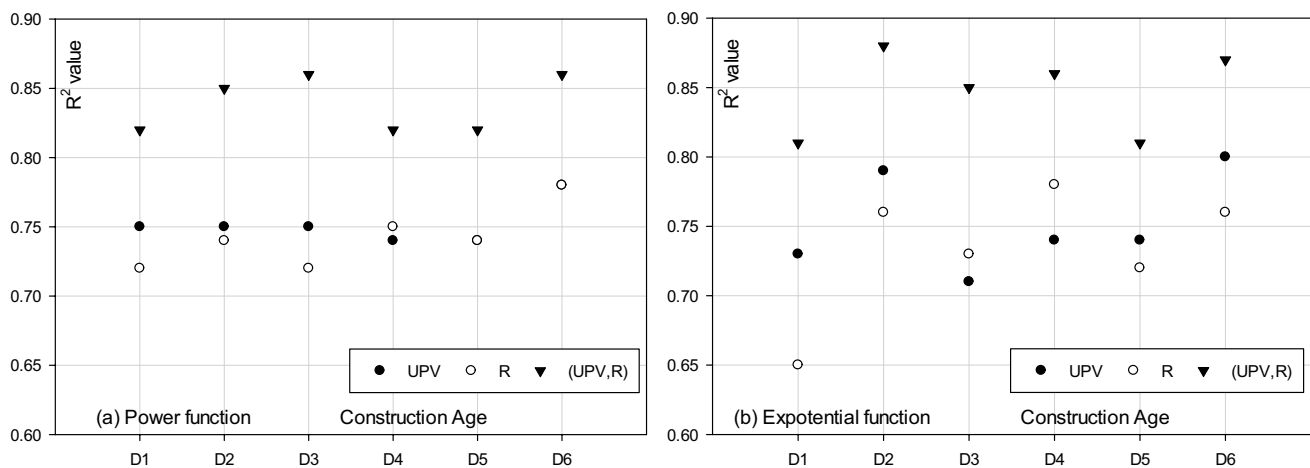


Fig. 5 Comparative analysis of the fitted R^2 value in terms of the construction age (D1 to D6) with regression equation **a** Power function, **b** Exponential function

Compared with a single method, 6–8 cores are necessary for the combined method to estimate strength precisely.

We believe qualitative declarations in this study are expected to hold true across various situations, but quantitative assertions may require further analysis. Examining the model's performance in actual practice is crucial for evaluating the calibration process or increasing the number of testing samples.

Future recommendation

Additional research is needed to determine variations in the number of cores, while keeping in mind that combining two NDT methods enhances the evaluation.

Research contribution

In present study, it is proposed to take 9 cores as a reasonable number of samples to converge for a single NDT method. Additional increase did not yield any significant improvements in the results. In comparison with single method, 6–8 cores are necessary for the combined method to estimate strength precisely. Examining the model's performance in actual practice is crucial for evaluating the calibration process or increasing the number of testing samples.

Acknowledgements The authors declare that there is no funding provided for this research from any institution or organization. Authors of the paper would like to express special thanks to Er. Ravi Ranade, MD of Construction Diagnostic Centre, Pune who has provided testing results to assist this research work.

Declarations

Conflict of interest The authors declare that there is no conflict of interest if there are potential conflicts of interest, we highly encourage each other to identify and declare clearly to avoid any future investigation by the publisher.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study formal consent is not required.

References

- Pucinotti R (2013) Assessment of in situ characteristic concrete strength. *Constr Build Mater* 44:63–73
- ACI 228.1R-03 (2003) In-place methods to estimate concrete strength. American Concrete Institute, Michigan
- Benyahia KA, Sbartai ZM, Breysse D, Kenai S, Kenai GM (2017) Analysis of the single and combined non-destructive test approaches for on-site concrete strength assessment: general statements based on a real case study. *Case Stud in Constr Mater* 6:109–119
- Alwash M, Breysse D, Sbartai ZM (2015) Non-destructive strength evaluation of concrete: analysis of some key factors using synthetic simulations. *Constr Build Mater* 99:235–245
- Pucinotti R (2015) Reinforced concrete structure: non-destructive in situ strength assessment of concrete. *Constr Build Mater* 75:331–341
- Breysse D (2012) Nondestructive evaluation of concrete strength: a historical review and a new perspective by combining NDT methods. *Constr Build Mater* 33:139–163
- Vona M, Nigro D (2015) 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* 48:1043–1059
- Fiore A, Porco F, Uva G, Mezzina M (2013) On the dispersion of data collected by in situ diagnostics of the existing concrete. *Constr Build Mater* 47:208–217

9. EN 13791 (2019) Assessment of in-situ compressive strength in structures and precast concrete components, Brussels: CEN.
10. Cristofaro MT, Viti S, Tanganelli M (2020) New predictive models to evaluate concrete compressive strength using the SonReb method. *J Build Eng* 27:100962
11. Breysse D, Martinez-Fernandez JL (2014) Assessing concrete strength with rebound hammer: a review of key issues and ideas for more reliable conclusions. *Mater Struct* 47:1589–1604
12. Mujahid A, Sheraz A, Azevedo AR, Markssuel TM, Khan MI, Rafiq W (2022) Experimental and analytical investigation on the confinement behavior of low strength concrete under axial compression. *Structures* 36:303–313
13. Raheel A, Mohsin AK, Rayed A, Muhammad FJ, Mujahid A (2023) Promoting the green construction: scientometric review on the mechanical and structural performance of geopolymers concrete. *Constr Build Mater* 368:130502
14. Suleman AK, Mohammad AK, Muhammad NA, Mujahid A, Al F, Fahad A (2023) Sustainable alternate binding material for concrete using waste materials: a testing and computational study for the strength evaluation. *J Build Eng* 80:107932
15. Guoliang L, Hongzhi Z, Muhammad NA, Athar Z, Ahmed MH, Mujahid A, Muhammad FR (2023) Strength predictive models of cementitious matrix by hybrid intrusion of nano and micro silica: hyper-tuning with ensemble approaches. *J Market Res* 26:1808–1832
16. Sohaib N, Jian Y, Xing-Er W, Kaffayatullah K, Muhammad NA, Mohammad FJ, Fadi A, Mujahid A (2023) Estimation of strength, rheological parameters, and impact of raw constituents of alkali-activated mortar using machine learning and shapely additive explanations (SHAP). *Constr Build Mater* 377:131014
17. Baoping Z, Yanbing W, Muhammad NA, Bawar I, Kaffayatullah K, Mujahid A, Fadi A (2023) Artificial intelligence-based optimized models for predicting the slump and compressive strength of sustainable alkali-derived concrete. *Constr Build Mater* 409:134092
18. EN 12504-1, (2000) Testing concrete in structures-Part 1: cored specimens-taking, examining, and testing in compression, Brussels: CEN
19. ASTM C42 (2012) Standard test method for obtaining and testing drilled cores and sawed beams of concrete. ASTM International, Philadelphia
20. Bungey JH, Millard SG, Grantham MG (2006) Testing of concrete in structures, 4th edn. Taylor and Francis, London
21. Ministry of railways-India, (2009) Guidelines on Non-destructive testing of bridges BS-103, Lucknow: B&S Directorate, India.
22. RILEM TC 43-CND (1993) Draft recommendation for in situ concrete strength determination by combined non-destructive methods. *Mater Struct* 26(155):43–49
23. Kumavat HR, Chandak NR (2021) Experimental investigations on effect of compaction, curing, water to cement ratio, cement type, and temperature variation on the rebound hardness of concrete. *KSCE J Civ Eng* 25:1857–1864
24. Liu JC, Sue ML, Kou CH (2009) Estimating the strength of concrete using surface rebound value and design parameters of concrete material. *Tamkang J Sci Eng* 12(1):1–7
25. Khan MI, Nasir K, Hashmi SRZ, Yazid MRM, Md NI, Yusoff RWA, Mujahid A, Roman F (2023) Prediction of compressive strength of cementitious grouts for semi-flexible pavement application using machine learning approach. *Case Stud Constr Mater* 19:e02370
26. Shcherbakov MV (2013) A survey of forecast error measures. *World Appl Sci J* 24:171–176
27. Nobile L (2015) Prediction of concrete compressive strength by combined non-destructive methods. *Meccanica* 50(2):411–417
28. Kumavat HR, Chandak NR (2023) Repeatability of concrete rebound hardness using comprehensive statistical analysis. *J King Saud Univ-Eng Sci* 35(3):167–173
29. Kumavat HR, Chandak NR, Patil IT (2021) Factors influencing the performance of rebound hammer used for non-destructive testing of concrete members: a review. *Case Stud Constr Mater* 14(1–12):e00491
30. Mariam FG, Metwally AE (2013) Implementation of nondestructive combined test methods in prediction of concrete strength in Egypt. *Eng Res J* 36(2):195–204

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.