

Advancing non-destructive concrete compressive strength estimation: Large-Scale datasets and machine learning framework

Benjamin Matthews^{a,*}, Diego Allaix^b, Simon Wijte^c, Marcel Vullings^d

^a Eindhoven University of Technology, Department of the Built Environment, P.O. box 513, 5600 MB, Eindhoven, the Netherlands

^b TNO, Department of Reliable Structures, Delft, the Netherlands

^c Eindhoven University of Technology, Department of the Built Environment, Eindhoven, the Netherlands

^d TNO, Department of Building Materials & Structures, Delft, the Netherlands

ARTICLE INFO

Keywords:

Non-destructive test
UPV
Rebound hammer
SonReb
Compressive strength
Machine learning
CatBoost

ABSTRACT

Non-destructive test (NDT) methods provide an indirect assessment of the compressive strength of in-situ concrete structures. While traditional static models effectively capture the behaviour of small-scale localised datasets, their accuracy diminishes when applied to larger, aggregated datasets, where increased variability in NDT measurements introduces greater uncertainty in predicting concrete compressive strength. This paper presents three exhaustive, largest-to-date NDT databases on the ultrasonic pulse velocity (UPV), rebound hammer (RH), and SonReb methods, comprising 16,531 test results from 115 studies. First, existing empirical models are evaluated against global dataset trends. New relationships are fitted to reflect the global behaviour of each NDT method, highlighting their innate limitations in capturing large-scale variability. A comprehensive three-phase machine learning (ML) program is then introduced, studying the effects of incomplete features with varying levels of missing data on model performance. Seven diverse ML models are included in Phase 1, while Phase 2 assesses different imputation strategies. Phase 3 integrates the top-performers with a Tree-Structured Parzen estimator (TPE) optimisation algorithm to refine hyperparameters and maximise performance. Across all phases, CatBoost regression emerged as the most robust predictive model due to the high proportion of categorical variables included within the databases. The TPE-CatBoost models achieved final R^2 values of 0.928, 0.896, and 0.947 for UPV, RH, and SonReb, respectively. Finally, a Django-based web application was deployed on a cloud server (<https://recreate-ndt.onrender.com/>), allowing practitioners to generate real-time compressive strength predictions for new NDT results. These novel datasets and ML tools can power future innovation through more advanced data-driven modelling.

1. Introduction

Construction and demolition activities contribute 11 % of global energy-related CO₂ emissions, while the European building sector is responsible for 36 % of their solid waste generation [1]. In the late 1990s, demolition activities in Britain produced approximately 30 million tonnes of industrial waste, with only 2 million tonnes of construction materials reclaimed through reuse or recycling [2]. As the deadlines for achieving CO₂ emission targets approach, alternative methodologies that improve material circularity and extend the lifespan of structural components are urgently needed [3]. The effective structural reuse of existing concrete elements is a promising option for substantially reducing emissions and environmental impacts [4–6]. For

example, Küpfer et al. [6] conducted a systematic review identifying 77 completed projects involving the structural reuse of precast and cast-in-place concrete elements across Europe and the USA between 1967 and 2022.

To ensure that reclaimed elements are suitable for additional service lives, their quality and properties must be thoroughly assessed and tested. Non-destructive test (NDT) methods have been employed for decades to evaluate the quality and properties of structural materials [7–10]. Techniques for non-destructive estimation of the concrete compressive strength include ultrasonic methods, impact hammer, penetration testing, radiography, and acoustic technologies [11–14]. A tremendous amount of research has been devoted to the non-destructive study of concrete compressive strength at various life-cycle stages, from

* Corresponding author.

E-mail address: b.j.m.matthews@tue.nl (B. Matthews).

<https://doi.org/10.1016/j.ndteint.2025.103549>

Received 9 June 2025; Received in revised form 3 September 2025; Accepted 11 September 2025

Available online 17 September 2025

0963-8695/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

less than one day to structures over 80 years old [15,16]. Among these methods, the ultrasonic pulse velocity (UPV) and rebound hammer (RH) techniques have been most widely adopted within academic and industry communities. More recently, the combined UPV and RH approach (SonReb) has attracted significant attention, showing improved predictive capability compared to other independent methods [17–19]. Despite their broad use, each method indirectly estimates strength by relying on the establishment of univariate or multivariate correlations that require additional modelling [11,20,21].

The research presented herein is part of the EU Horizon 2020-funded ReCreate project, exploring the complete structural reuse of precast concrete elements in new building constructions, including the assessment of material properties and quality prior to reconstruction. This study comprises several progressive phases. Initially, three databases are established, each representing, to the best of the authors' knowledge, the largest aggregation of experimental data on UPV, RH, and SonReb methods. From a total of 115 published studies, 16,531 test results are compiled. Existing empirical models from the literature are then compared against the global performance of each NDT. Subsequently, new models are developed for the conventional typologies (linear, polynomial, power, and exponential) based on the final empirical distributions. To further advance predictive performance, several machine learning (ML) models are competitively trained and tested to determine the optimal methodology for each NDT. Two modelling phases are undertaken, implementing different feature lists and imputation methods based on the volume of missing data. The most effective predictor for each NDT is then selected, and a rigorous Tree-Structured Parzen Estimator (TPE) optimisation algorithm is integrated into the training process to enhance performance further. Finally, a web-based graphical user interface (GUI) has been developed, which runs the optimised ML models in the backend, thereby providing engineers with a practical and reliable tool for estimating concrete compressive strength from on-site or laboratory NDT results.

1.1. Ultrasonic pulse velocity

The ultrasonic pulse velocity (UPV) method is widely used to assess the properties, homogeneity, structural health, and the presence of defects in concrete. It measures the velocity of an ultrasonic wave transmitted between two piezo transducers at a known distance. Because wave velocity depends on density and elastic properties, it can be correlated to compressive strength and material quality [22,23]. An electrical pulse induces vibrations in the concrete at its resonance frequency, where a receiving transducer records travel time, and velocity is calculated from the known distance [16]. Strong acoustic coupling between the transducer and concrete surface via grease or petroleum jelly is required for accurate measurements [24]. Fig. 1 schematically

illustrates a typical UPV apparatus in a direct arrangement. For longitudinal (P-wave) propagation in elastic concrete:

$$V_p = \sqrt{\frac{E_c(1 - \nu_c)}{\rho_c(1 + \nu_c)(1 - 2\nu_c)}} \quad (1)$$

Where E_c is the concrete elastic modulus, ν_c is the Poisson ratio, and ρ_c is the concrete density. Factors such as moisture content, mixture details, and embedded reinforcing steel also influence V_p , where correction factors of 0.9 and 0.8 have been proposed for reinforcing steel perpendicular and parallel to the wave path, respectively [16]. UPV can monitor cracking and degradation; however, due to concrete heterogeneity and variable pore structure, correlations with compressive strength can be inconsistent in situ [25,26]. UPV tests may be direct, semi-direct, or indirect, depending on the relative placement of the transducers. Direct transmission (Fig. 1) yields the most reliable strength correlation, but is often less practical for in-situ assessments due to accessibility challenges, where semi-direct or indirect arrangements are adopted. Numerous standards, such as EN 12504-4 [27], ASTM C597 [28], and BS 1881-203 [29], delineate the procedural requirements, limitations, and interpretations of UPV tests.

1.2. Rebound hammer

The Schmidt rebound hammer (RH) relates surface rebound hardness to the compressive strength, through a spring-loaded steel mass impacting the concrete surface (Fig. 2). The rebound number (RN) represents the kinetic energy and distance travelled by the rebounding mass, expressed as a percentage of the initial spring energy [30,31]. Velocity is measured before and after impact, and is affected by gravity and the inclination of the hammer [32]. Horizontal applications are therefore preferred (Fig. 2), while vertical readings must be recalibrated to horizontal benchmarks. Since RN indicates surface hardness, local variations in the concrete, such as coarse aggregate beneath the plunger or cracks, can cause over- or under-estimations of strength, respectively. Moreover, moisture content, surface texture, concrete maturity, stress state, degradation, mix design and aggregate properties all further influence RH accuracy [30,33,34]. The standards ASTM C805 [35] and EN 12504-2 [36] recommend a minimum of nine measurements per test location to establish a reliable statistical distribution, using the median value. If more than 20 % of readings deviate by more than 30 % from the median, the entire set should be discarded and repeated [36]. EN 12504-2 [36] further requires rigid support of concrete elements to eliminate displacement upon impact during testing.

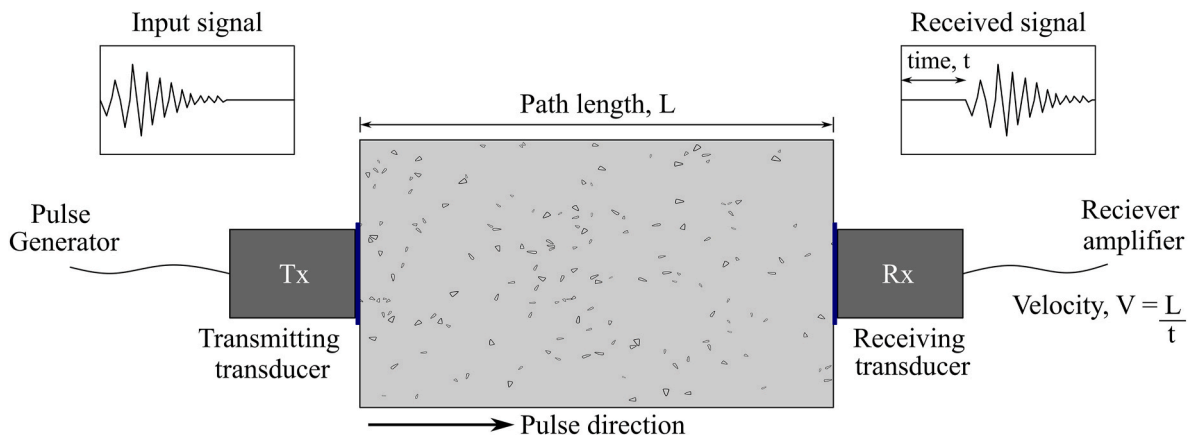


Fig. 1. Schematic of a typical UPV apparatus and test arrangement.

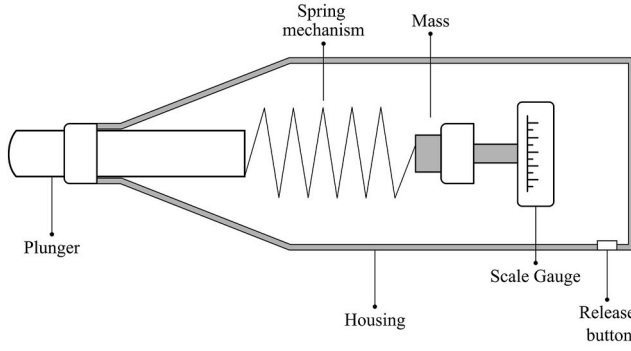


Fig. 2. Rebound hammer device (left) and application on a precast concrete element.

1.3. SonReb

The SonReb method combines UPV and RH tests at the same location to mitigate individual limitations and improve strength estimations. For example, increasing relative humidity reduces RN while inversely raising V_p [18]. Furthermore, UPV probes internal properties, while RH characterises surface hardness, providing a more holistic understanding of the concrete member's condition when used in combination. Since the 1970s, SonReb has generally outperformed the standalone methods, though some studies report only marginal gains [18,19,37–42]. Carbonation in older existing structures increases elastic modulus and density, slightly increasing the ultrasonic velocity (Eq. (1)), while markedly increasing RN, potentially biasing strength estimates [32,41].

Multivariate SonReb relationships, $f_c = f(V_p, RN)$, are commonly presented as 3D surfaces or nomograms, and calibrated by fitting iso-strength curves to strength-stratified datasets at regular intervals [18]. Breyse [18] and Cristofaro et al. [43] synthesised 59 and 17 SonReb models, respectively, from the literature, proposing strength-UPV-RH curves calibrated to local or aggregated datasets. The most common functional forms include double power and bilinear models, although double exponential and higher-order polynomials are also used. Although the increased complexity of SonReb appears to improve predictive ability in most local calibration cases, the additional testing time and cost may not always justify its use.

1.4. Destructive testing

Destructive testing (DT) on cast or drilled specimens remains the standard method for verifying the compressive strength of concrete mix designs. Compression tests performed on poured cylinders or cubes (for new elements) and drilled cylinders (for existing elements) provide direct measurements of concrete compressive strength that can be correlated with NDT results [7,18]. Test outcomes depend heavily on specimen geometry, compaction quality, curing, aggregate characteristics, age, homogeneity, and air-entrainment [8,44]. To standardise DT results across different geometries, normalisation methods are required. EN 1992-1-1 [45] defines concrete strength using a standard reference cylinder geometry of 150×300 mm. Reineck et al. [46] propose a two-step normalisation procedure: first, specimens are standardised to a reference geometry ($150 \times 150 \times 150$ mm for cubes and 150×300 mm for cylinders) using conversion factors [46]. These conversion factors are also adopted in the German code DIN 1045-1 [47], *fib* Bulletin 12 [48], and ASCE-ACI Committee 445 [49] and are therefore considered internationally recognised [46]. Second, reference cubes are converted to cylinders using a bilinear approximation for normal- and high-strength concrete, as given in Eq. (2) [46]:

$$f_{c,cyl} = \begin{cases} 0.83 \cdot f_{c,cu150}, & \text{if } f_{c,cyl} \leq 54 \text{ MPa} \\ 1.0 \cdot f_{c,cu150} - 11.1, & \text{if } f_{c,cyl} > 54 \text{ MPa} \end{cases} \quad (2)$$

Drilling effects also impact the compressive strength of in-situ core

samples [50]. Drilled cores often exhibit cut aggregate and micro-cracking, leading to lower compressive strength than laboratory-cast or on-site cylinders [50]. This variability is not explicitly addressed in the normalisation procedure and may contribute to residual variance observed among experimental results.

2. Experimental databases

Extensive databases of experimental test results for the non-destructive evaluation of the compressive strength of concrete specimens have been compiled for the UPV, RH, and SonReb methods. A total of 16,531 unique sets of NDT-compression strength results were gathered from 115 publications, comprising 6103 UPV, 10,428 RH, and 3299 SonReb results [7–10], [13–17], [19], [22], [23], [30–34], [38–41], [43], [51–66], [67–82], [83–98], [99–114], [115–130], [131–142]. The literature sources ‘Ramboll Finland OY (2023)’ and ‘Matthews et al. (2025)’ are an unpublished industry report and internal ReCreate testing on in-situ precast concrete elements. For brevity, the original experimental tests performed as part of the ReCreate project are not described in detail here. The UPV database includes 20 input variables, while the RH database includes 18, and the SonReb database contains 25. These variables encompass material age, location, composition, geometry, NDT parameters and adopted procedures. The response variable of each database is the compressive strength determined through destructive testing of cubic, cylindrical, or drilled core specimens, normalised through the procedure described in Section 1.4. Each database consists of approximately equal proportions of numerical and categorical variables. Fig. 3 visualises the database structures and shared characteristics between each NDT.

Article discovery was achieved through a programmatic API-based web-scraping algorithm of leading scholarly databases, including Google Scholar, Scopus, CrossRef, and a custom ResearchGate scraper. Keyword-based query limits were set to the first 200 successful matches for Google Scholar and Scopus, 100 for CrossRef, and 150 for ResearchGate. After deduplication, over 250 relevant article metadata sets were exported for both UPV and RH. Following a screening process based on the exclusion criteria mentioned below, data were collected manually for each study to ensure relevance, correctness, and fidelity. A total of 1328 UPV tests, 3042 RH tests, and 544 SonReb tests were performed on in-situ structures, with the remaining number conducted on laboratory-cast specimens. Material ages range from structures built in the 1930s to laboratory specimens tested one day after casting. Tests conducted on concrete less than one day old were excluded. An open-source computer vision-assisted software was used to extract the numerical data from plots to high precision for publications that reported their results visually. Since the drilling effect is not explicitly accounted for in the normalisation approach, the input features ‘NDT Specimen Type’ and ‘Compression Specimen’ are used to qualitatively differentiate between poured or drilled cylinders, enabling the machine learning models developed in later sections to learn underlying patterns

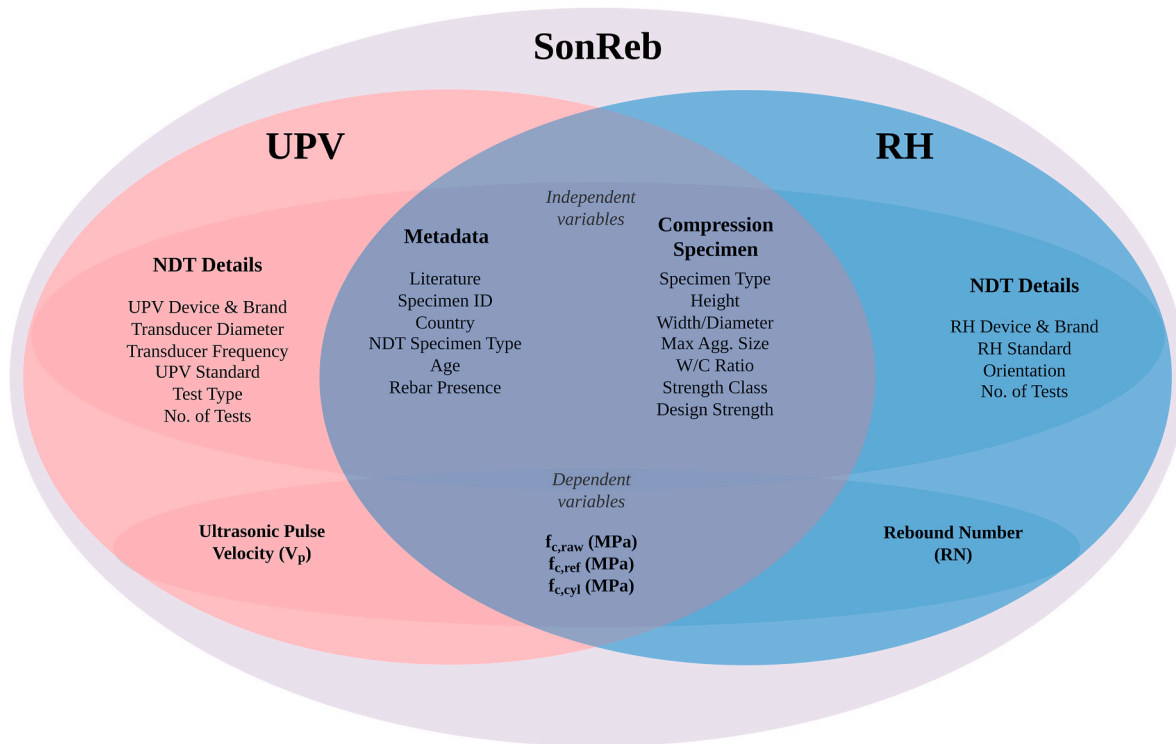


Fig. 3. Database structure and variables recorded for each NDT.

associated with core extraction effects. Fig. 4 summarises the distributions of UPV and RH test counts (individual test records captured within the databases) by year and country. UPV and RH show a clear absolute increase in the number of tests performed over the past five decades, likely due to the growing need for assessment of existing structures and reliable in-situ strength estimations. While other NDT modalities may have grown as well, Fig. 4 reports absolute counts rather than cross-modality shares, which is considered outside of the scope of this paper.

Data filtering criteria were systematically applied throughout the acquisition phase [143]. The primary objective of the raw databases was to gather the maximum amount of experimental data, irrespective of its potential impact on model performance. In this unprocessed state, the databases serve as a comprehensive representation of the overall

predictive capacity of each NDT method for estimating the concrete compressive strength, considered herein as the ‘global’ performance. During data collection, all observations were determined to be statistically independent, with each entry corresponding to a distinct pairing of an NDT measurement and compression strength value. Tests presenting evidence of significant corrosion, deterioration, fire damage, or other physical damage were omitted from the databases, as it is assumed that such elements will undergo substantial rehabilitation or will be disregarded from a reuse perspective. For UPV measurements, only compression P-waves were considered, while shear (S-waves) and Rayleigh (R-waves) were excluded. Other exclusions include unique concrete materials, admixtures, superplasticisers, and aggregates, which may influence compressive strength in niche ways. Table 1 summarises

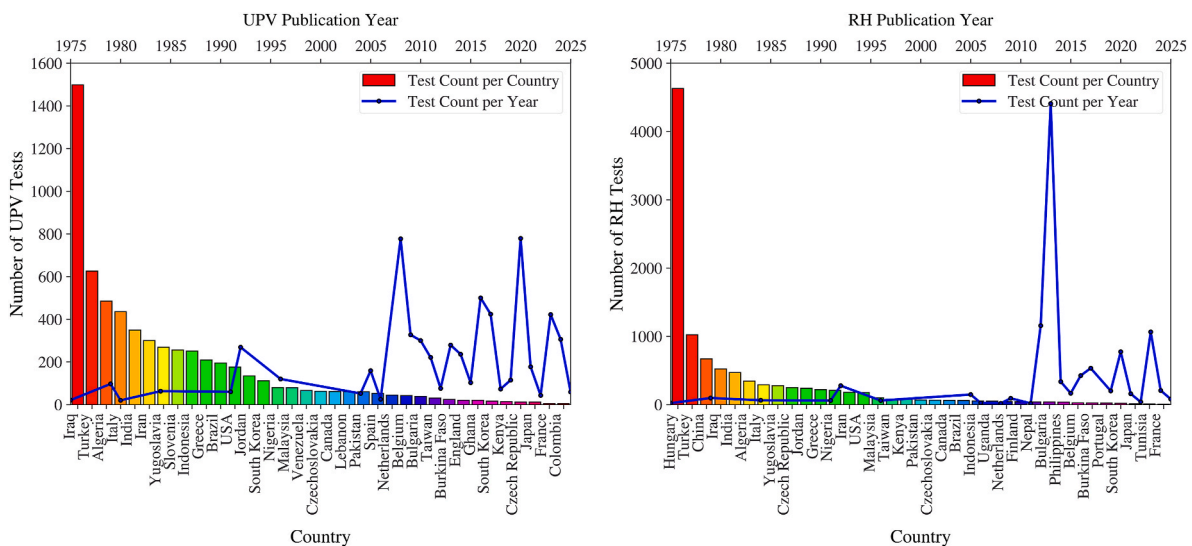


Fig. 4. Annual and geographical test counts for UPV and RH within the compiled databases. “Test count” denotes the number of individual test records extracted from the literature.

Table 1
Statistical domains of key features and response variables for each database.

Variable	UPV		RH		SonReb		
	V_p (m/s)	$f_{c,cyl}$ (MPa)	RN	$f_{c,cyl}$ (MPa)	V_p (m/s)	RN	$f_{c,cyl}$ (MPa)
n	6103	6103	10428	10428	3299	3299	3299
μ	4198	27.78	38.56	42.17	4171	32.97	28.86
σ	637.8	15.43	10.90	23.65	631.6	9.136	14.58
Min	981.2	0.840	7.955	1.343	1510	7.955	1.343
25 %	3880	17.37	30.70	22.71	3840	26.08	18.57
50 %	4300	25.09	38.20	37.08	4272	32.00	26.68
75 %	4638	34.11	45.90	59.15	4605	38.97	35.08
Max	5779	109.2	79.30	128.1	5660	63.50	109.2

Note: n is the sample size, μ is the mean, and σ is the standard deviation.

the statistical domains of the normalised compression strength and key NDT results for each database. The complete open-source databases can be accessed at <https://doi.org/10.5281/zenodo.14921019>. The supplementary documentation in the repository provides an account of all assumptions, nomenclature, and abbreviations applied throughout the databases.

Fig. 5 illustrates the correlation between ultrasonic velocity and RN measurements with normalised compressive strength alongside density distributions and marginal histograms. No data filtering or outlier removal is applied at this stage, ensuring that the global variability and sensitivity of each technique are fully represented. When simplified to univariate relationships, neither NDT method yields a reliable estimation of compressive strength. The UPV distribution encompasses tests conducted across all transmission modes – direct, semi-direct, and indirect – which increases the variance, particularly at lower velocities. The rebound hammer is more heavily influenced by the specific device and orientation (horizontal or vertical) applied during testing. For higher-strength concretes, devices with larger nominal kinetic energies provide more reliable results, as demonstrated by Chen et al. [128].

Furthermore, the large cluster of high RN results, which deviates from the global trends in Fig. 5b, was predominantly produced by one specific device with a 4.5 J nominal kinetic energy, applied on high-strength concrete in Ref. [128]. It is evident that a linear approximation does not adequately capture the global ability of any individual NDT technique. However, on a local scale, linear models often perform relatively well on small subsets of data [13,64,66,144]. The interaction

between UPV and RH measurements within the SonReb database is depicted in Fig. 6. In general, a proportional increase among all three

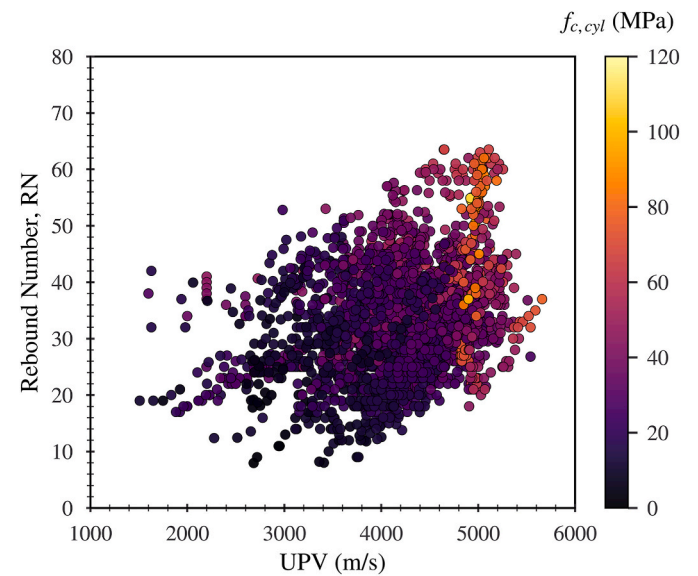


Fig. 6. Multivariate relationship between velocity, RN, and strength (SonReb).

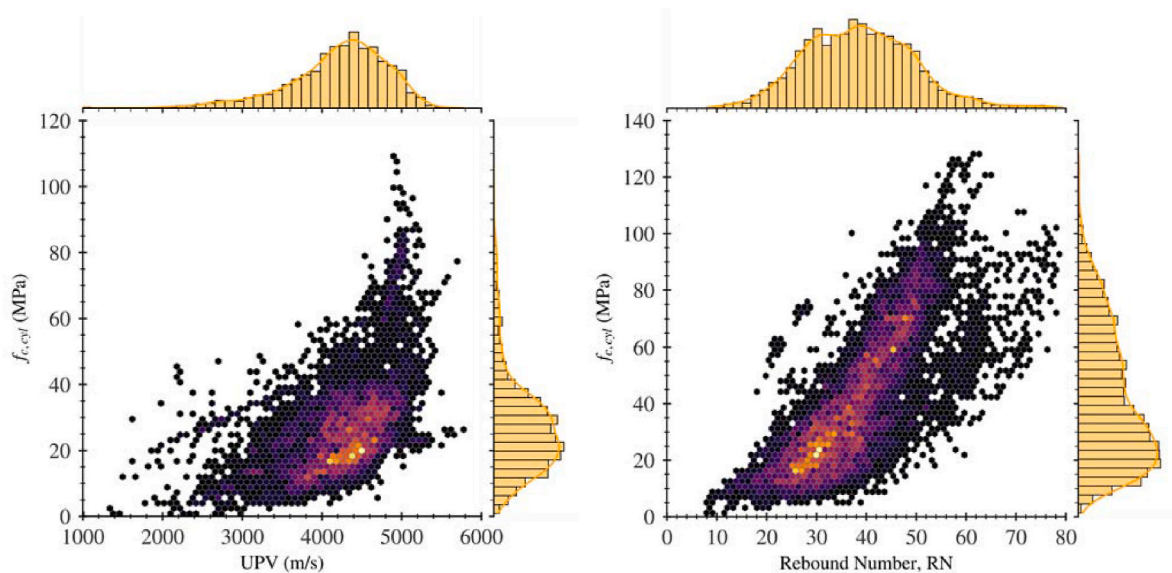


Fig. 5. Density distributions of the UPV (a); Rebound Hammer (b) databases along with the marginal distributions of each variable. The colour gradient represents data density.

variables is observed. However, a significant amount of uncertainty will remain in an empirically fitted multivariate model unless the strength iso-curves are calibrated to a very fine resolution.

3. Empirical modelling

3.1. Existing models

A substantial body of research has been conducted within the scientific community to assess the capability of UPV, RH, and SonReb techniques in estimating the concrete compressive strength and the derivation of empirical laws to capture these relations. Breyse [18] collated 89 published UPV models, 70 RH models, and 59 SonReb models through an extensive review of the literature. More recently, Debroy and Sil [145] performed a comprehensive review of 102 models, proposing their own formulations after evaluating the global uncertainty associated with each technique.

When fitted to localised datasets derived from individual structures, small groups of structures, or controlled laboratory experiments, many empirical models exhibit exceptionally high performance, frequently achieving correlation coefficients (R^2) exceeding 0.90 [71,146–152]. However, as datasets increase in size, the accuracy of these locally calibrated models invariably declines. Furthermore, research proposing new empirical models does not converge on a single optimal typology, accentuating the variability in NDT performance at the local scale. The selection of model typology typically includes exponential, power, polynomial (second order), and linear functions, with additional

adaptations for multivariate approaches in SonReb-based models. EN 13791 [144] advocates a locally calibrated linear method for estimating the in-situ concrete compressive strength by integrating NDT measurements with destructive core tests, generally requiring a minimum of eight core-NDT couples per test region. From this perspective, further research exploring local calibration becomes inherently limited, as newly developed models are site- or program-specific, preventing their applicability to other locations and consequently limiting their contribution to the broader scientific understanding of NDT-based strength estimation. Conversely, Breyse [18] demonstrated that local calibrations consistently outperform global (collective) calibrations for both UPV and RH due to the substantial variability between datasets.

Since the databases compiled as part of this research are, to the best of the author's knowledge, the most extensive of their kind currently available, it is assumed that each database provides a representative measure of the 'global' ability of each NDT to explain the compressive strength. The objective of the present study is, therefore, to evaluate the predictive capability of each NDT on a global scale through various modelling modalities of increasing complexity. As a preliminary phase to the machine learning discussed in Section 4, a selection of existing empirical models is extracted from the literature and benchmarked against the global performance of each NDT technique. Twelve models per method are selected and summarised in Tables A1 to A3 in Appendix A, derived from Refs. [9,11,20,21,31–33,61,71,84,105,126], and [146–155], along with their corresponding local and global performances. Erroneous data points exhibiting excessive deviations from the global trend were removed as outliers, resulting in the cleaned

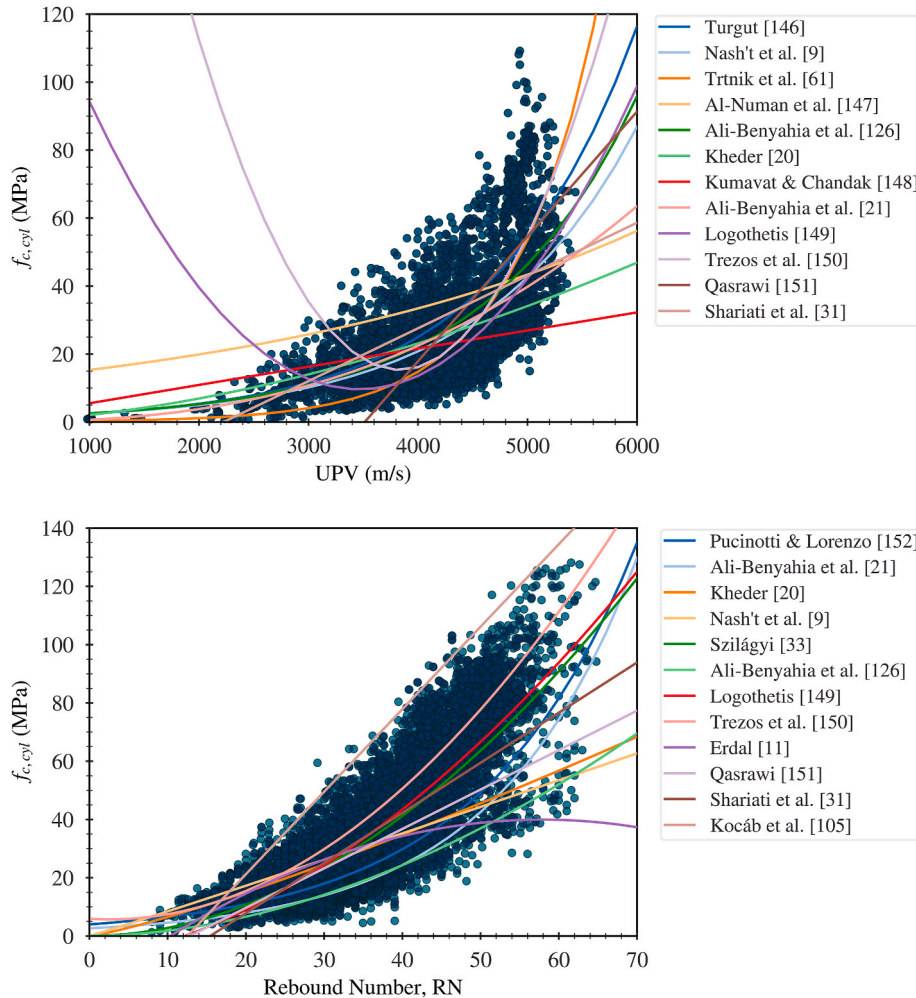


Fig. 7. Performance of existing empirical models for UPV and RH relative to the global datasets.

distributions given in Fig. 7. The final cleaned databases consisted of 5680 test results for UPV, 9824 for RH, and 2980 for SonReb.

Fig. 7 presents the selected models against their respective databases. To ensure appropriate correlation, only UPV tests conducted using the direct transmission method and RH tests performed in the horizontal orientation are considered in this phase. Most manufacturers of Schmidt hammers provide proprietary conversion factors to adjust vertical measurements to horizontal values. However, many of the reviewed articles that employed vertical testing did not specify whether a conversion was applied to the reported results. Therefore, in the absence of explicit conversion details, it is assumed that no adjustments were made in these cases.

There is considerable diversity across different model typologies in their effectiveness at capturing the global behaviour of the UPV and RH tests. Due to the substantial variance within each distribution, the global predictive performance is significantly lower than the originally reported local fits. Several models also yielded negative R^2 results, shown in Appendix A. A negative correlation coefficient indicates that the model predictions are less accurate than a simple mean-based estimation of the dependent variable. If the sum of the squared residuals exceeds the total sum of squares, then the model would perform better by simply predicting the mean value, $\bar{y}$, for each data point, signifying that, despite following the correct trajectory, it fails to capture the variance in the dataset adequately.

Among the NDT methods, UPV proves to be the most challenging to model accurately, with a maximum R^2 of 0.421, as reported by Ali-Benyahia et al. [126]. For RH, the highest achieved R^2 was 0.713, obtained by the second-order polynomial function proposed by Logothetis [149]. Despite slightly lower correlation coefficients, SonReb models demonstrate the most consistency, with a maximum R^2 of 0.631 by Fawzi et al. [71] and a minimum of 0.208 from Nikhil et al. [84]. It is important to note that many of these models were likely fitted to concrete specimens of varying geometries without normalisation to a standard reference shape, further impeding their adaptability to larger datasets.

3.2. Proposed models

New empirical models are proposed based on the global distribution of each NDT method, covering the most commonly used mathematical typologies. Eqs. (3)–(6) define the general forms of the exponential, power, polynomial, and linear models, respectively:

$$y = a \cdot e^{(b \cdot x)} \quad (3)$$

$$y = a \cdot x^b \quad (4)$$

$$y = a \cdot x^2 + b \cdot x + c \quad (5)$$

$$y = a \cdot x + b \quad (6)$$

Model coefficients are derived through a non-linear least-squares optimisation, implementing the Levenberg-Marquardt (LM) algorithm. An additional combined power-exponential model is fitted against the SonReb dataset. Table 2 presents the newly proposed models for each NDT method, along with their corresponding performance metrics. Fig. 8 presents the nomogram iso-curves for the SonReb power model given by Eq. (16).

Among the fitted functions, exponential provides the best fit for the UPV and, less convincingly, SonReb datasets, where the double-power and combined power-exponential functions demonstrate equivalent performance. The power and second-order polynomial functions show comparable performance for the RH. Although the RH models yield the highest correlation coefficients, the mean absolute error (MAE) and root-mean-squared error (RMSE) are similarly the highest, reflecting the considerable variance in the dataset. SonReb achieves the most balanced trade-off between reduced variance and absolute error, with an optimal

Table 2

Proposed empirical models for each NDT based on global fitting to the cylinder compressive strength.

NDT	Equation	Eq.	R^2	RMSE (MPa)	MAE (MPa)
UPV	$f_c(V_p) = 0.9231 \cdot e^{(0.7804V_p)}$	(7)	0.427	12.0	9.13
	$f_c(V_p) = 0.1887 \cdot V_p^{3.4009}$	(8)	0.417	12.1	9.22
	$f_c(V_p) = 5.8184 \cdot V_p^2 - 29.1268 \cdot V_p + 44.6680$	(9)	0.415	12.1	9.24
	$f_c(V_p) = 16.4909 \cdot V_p - 42.1610$	(10)	0.366	12.6	9.56
	$f_c(RN) = 2.1610 \cdot e^{(0.0459RN+1.1575)}$	(11)	0.690	13.1	10.1
RH	$f_c(RN) = 0.0326 \cdot RN^{1.9584}$	(12)	0.708	12.8	9.79
	$f_c(RN) = 0.0287 \cdot RN^2 - 0.0984 \cdot RN + 2.5969$	(13)	0.708	12.8	9.77
	$f_c(RN) = 2.0441 \cdot RN - 34.7406$	(14)	0.688	13.2	10.4
SonReb	$f_c(V_p, RN) = 1.3770 \cdot e^{(0.4101V_p+0.0292RN-0.2423)}$	(15)	0.683	8.06	6.05
	$f_c(V_p, RN) = 0.0372 \cdot V_p^{1.7766} \cdot RN^{1.1495}$	(16)	0.682	8.07	6.11
	$f_c(V_p, RN) = 0.0357 \cdot RN^{1.1439} \cdot e^{(0.6672V_p)}$	(17)	0.683	8.07	6.11
	$f_c(V_p, RN) = 3.6069 \cdot V_p^2 + 0.0188 \cdot RN^2 - 18.5577 \cdot V_p - 0.3917 \cdot RN + 31.9895$	(18)	0.680	8.10	6.03
	$f_c(V_p, RN) = 10.5888 \cdot V_p + 0.9383 \cdot RN - 47.6760$	(19)	0.644	8.54	6.43

Note V_p is in km/s.

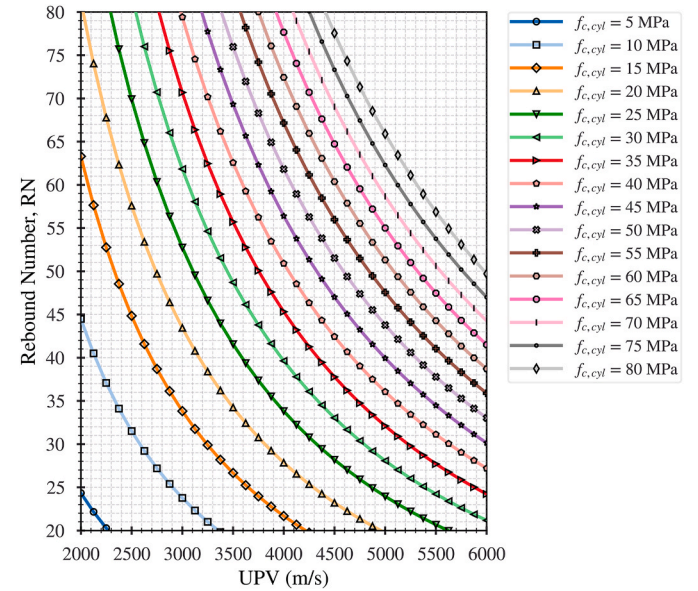


Fig. 8. Iso-curves for SonReb power model Eq. (16).

MAE of 6.05 MPa. The top-performing model for each NDT in Table 2 also surpasses the best local models presented in Tables A1 to A3 (Appendix A), although the improvement margins are minor at only 2.8 % for UPV and 2.9 % for RH.

Despite the abundance of locally calibrated empirical models, numbering in the hundreds [18,145], these global comparisons demonstrate that no empirical models reliably capture the true variability of UPV, RH, or SonReb measurements when applied beyond their original, small-scale datasets. Even when refitted to the new aggregated databases, static univariate and bivariate equations exhibit minor improvement, large predictive errors, and fail to serve as adaptive baselines. Therefore, these new global empirical fits are established as

rigorous null models, against which more advanced ML approaches must demonstrate clear performance gains.

4. Machine learning

To further extend the modelling capabilities of the NDTs, three machine learning phases are introduced to examine the impact of increasing complexity on the predictive performance. The primary distinction between Phases 1 and 2 lies in the extent of feature reporting within the datasets. Variables such as transducer diameter and frequency (for UPV), maximum aggregate size, water/cement (W/C) ratio, and target strength exhibit significant reporting gaps. For instance, within the RH database, the W/C has a 74.1 % non-reporting rate, largely due to the substantial contribution by Ref. [33]. To impute incomplete features of this scale would be impractical.

Accordingly, Phase 1 involves machine learning using a refined list of 'core features', selecting only those with complete or near-complete reporting, where logical imputation can be reasonably justified. Phase 2 then expands the feature set to include all available features within the databases, leaving heavily incomplete features unaltered, and applies machine learning models that can handle missing values. This comparison aims to identify whether incorporating additional features, even with minimal reporting, meaningfully impacts the predictive power of the models.

Seven models from different algorithmic typologies are trained and tested in Phase 1 to determine the most appropriate model for each NDT. Likewise, Phase 2 includes training and testing seven models for the same purposes. Phase 3 then integrates an advanced optimisation algorithm into the strongest-performing model from each NDT, further refining predictive performance. The final optimised models are then deployed within a Django-based web GUI application. Throughout all phases, the response variable is the normalised compressive strength, calibrated to a reference cylinder of 150×300 mm ($f_{c,cyl}$) using the procedure described in Section 1.4.

4.1. Phase 1: core features

4.1.1. Exploratory data analysis

Due to insufficient reporting, several features were excluded from this phase, including transducer diameter and frequency (for UPV and SonReb), maximum aggregate size, W/C ratio, concrete strength class, and target strength. Features with minimal missing data were imputed based on logical inferences and supplementary information given within the publication. The following assumptions were applied for imputation:

- Specimen age was inferred from the reported information and element type; otherwise, 28 days was assumed. In cases where a range of ages was reported, the average value was taken.
- Compression specimen dimensions were estimated based on the type of experiment. Laboratory-cast cylinders were assumed to be 150×300 mm, whereas in-situ cores were more likely to be 100×200 mm or 100×100 mm due to reinforcement constraints.
- Number of tests per location: The number of UPV tests performed at a single location was imputed as one for all missing values. For RH tests, a default value of nine was assigned if not reported, assuming the usual standard recommendations were followed.
- The UPV device was imputed as 'PUNDIT', while the RH device was imputed as 'Original Schmidt Type N' to generalise for all missing values.
- UPV and RH standards were inferred based on similar countries and the timeframe of the experimental program.

The reduced feature set included five numerical and seven categorical features for the UPV and RH datasets, while the SonReb dataset contained seven numerical and ten categorical features. Before modelling, a feature evaluation was conducted to assess the suitability of

features as predictors for the concrete compressive strength. Categorical variables were evaluated to determine whether the inter-group variance was statistically significant. If statistical significance is not observed, all groups have approximately the same effect on the response (null hypothesis). Therefore, the covariate is deemed detrimental to training quality and removal from modelling.

Independent t-testing can be used on features with only two groups (e.g., the presence of reinforcement) [156]. The t-value quantifies the relative difference between groups considering variance, where higher t-values indicate greater statistical significance. A p-value ≤ 0.05 is a standard threshold for statistical significance. Analysis of variance (ANOVA) tests can be used for features with three or more groups. ANOVA tests measure group sensitivity through the f-value and p-value. If the computed f-value is less than the f-critical value, the null hypothesis cannot be rejected, and there is no significant variance between groups [157]. Following independent t-tests and ANOVA tests across all categorical features in each database, only one variable – the compression specimen type (cube, cylinder, or core) in the RH and SonReb databases – failed to demonstrate statistical significance, returning p-values of 0.761 and 0.571, respectively. This outcome implies that the normalisation methodology discussed in Section 1.4 effectively accounts for the variance in compressive strength associated with compression specimen geometries. Since most statistical models require numerical inputs, the validated categorical features were encoded to ordinal integers in the range $[0, n - 1]$.

As the primary purpose of this study is to develop optimal predictive performance, rather than assessing feature importance, multicollinearity among numerical and encoded features was not analysed [158]. Furthermore, due to the wide variety of experimental programs captured within each database, the response variable ($f_{c,cyl}$) displays mild skewness and heavy-tailed distributions. For optimal model performance, normally distributed response data is typically preferred. A logarithmic transformation appeared to be the most appropriate option for the UPV and SonReb distributions, whereas a square-root transformation produced the closest near-Gaussian state for the RH compressive strength. Fig. 9 displays the probability distribution functions of the transformed response variables used for model training. Observations exhibiting anomalous behaviour (outliers) were identified and excluded based on their statistical deviation from the global trends.

4.1.2. Model selection

Seven machine learning models were trained, optimised, tested, and evaluated for Phase 1, covering a range of typologies. The selected models include artificial neural networks (ANN), gradient boosting regression trees (GBRT), XGBoost, LightGBM, CatBoost regression, support vector regression (SVR), and k-nearest neighbour (k-NN). Past research by the authors has indicated that ensemble tree-based models perform particularly well in structural engineering applications, especially when the datasets contain a large proportion of categorical features [159]. Thus, tree-based models are given specific attention in the present study. The models empirically fitted to the databases in Section 3.2 and Table 2 serve as the benchmark null models for comparative analysis. The same cleaned datasets presented in Fig. 7 were used for all machine learning phases. For brevity, theoretical descriptions of the selected models are not given here. Table 3 summarises the models applied in both Phases 1 and 2.

4.1.3. Model training

All models were trained and validated using the k-fold cross-validation technique. This technique partitions the whole dataset into k subsets, where the model is trained on $[k - 1]$ subsets and tested on the final hold-out set. The process is then iteratively repeated until each subset has been used for testing, effectively representing an out-of-bag (OOB) hold-one-out validation approach. OOB sampling ensures the measured errors are representative of the true model performance, eliminating the risk of data snooping [160]. Ten folds were applied to all

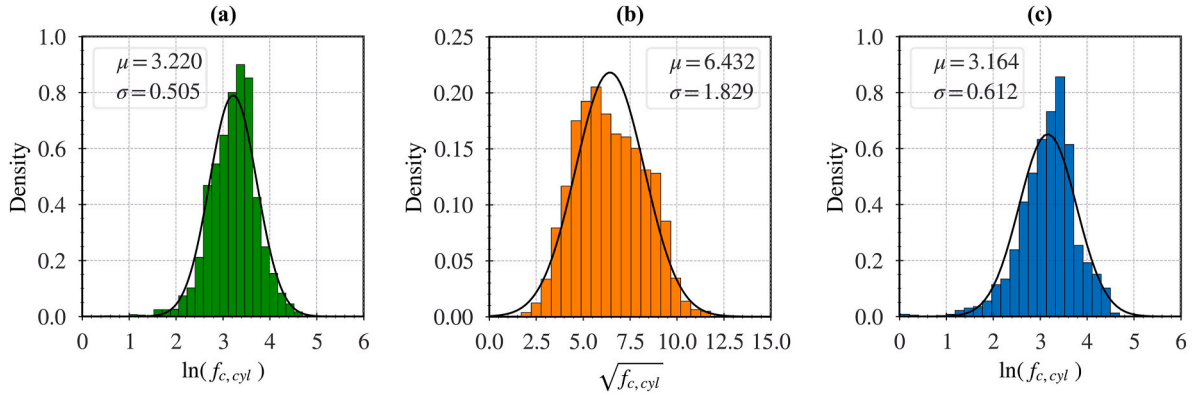


Fig. 9. Transformed concrete compressive strength distributions for UPV (a); RH (b); and SonReb (c) alongside idealised normal distribution curves.

Table 3
Machine learning models in Phases 1 and 2.

Typology	Model	Phase 1			Phase 2		
		UPV	RH	SonReb	UPV	RH	SonReb
Network-based	ANN	X	X	X			
	MANN				X	X	X
Ensemble Tree-based	GBRT	X	X	X			
	HGBRT				X	X	X
	XGBoost	X	X	X	X	X	X
	LightGBM	X	X	X	X	X	X
	CatBoost	X	X	X	X	X	X
Kernal-based	SVR	X	X	X			
	SVR + MICE				X	X	X
	k-NN	X	X	X	X	X	X

models, such that 90 % of the data was allocated for training while 10 % was reserved for testing in each iteration. For ANNs, the training sets were further partitioned into training and validation subsets at an 80:20 ratio. Hyperparameter tuning was conducted within the training loop using a standard grid-search optimisation approach, wherein appropriate search ranges were defined for each hyperparameter. Thus, ten models were trained and evaluated for each model.

Five performance metrics were used to evaluate the predictive capabilities of each model, namely the coefficient of determination (R^2), mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and the mean absolute percentage error (MAPE), given in Eqs. (20)–(24).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (20)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (21)$$

$$RMSE = \sqrt{MSE} \quad (22)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (23)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100 \% \quad (24)$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted response, and $\bar{y}$ is the mean true response. A higher R^2 indicates stronger predictive

performance, while lower MSE, RMSE, MAE, and MAPE signify more accurate predictions. The primary difference between RMSE and MAE is that MAE treats all errors equally, whereas RMSE assigns greater weight to larger errors due to its quadratic derivation, making it more sensitive to outliers.

4.1.4. Model performance and interpretability

Table 4 presents the average performance metrics of the Phase 1 models across the ten testing folds for each NDT. In general, the ensemble tree-based models demonstrated the most consistent predictive strength. CatBoost regression achieved the best performance for the UPV and RH databases, ranking second for SonReb based on the average R^2 – although it produced the best MAE and MAPE outcomes. CatBoost is a boosting aggregation approach that integrates categorical string data in its raw state, eliminating the need for manual encoding, which leverages the proportionately large number of categorical features (~50 %) included in each database. CatBoost sequentially develops many ‘weak’ learners, with each tree building upon the residual error of its predecessor, culminating into one final strong predictor with a globally minimised residual. The differentiating aspect from other boosting methods comes from growing ‘oblivious’ trees, which impose a rule across all nodes at a given level to test the same covariate under the same conditions, enabling more efficient binary indexing [161].

Fig. 10 presents the box and whisker distributions for R^2 from all ten

Table 4

Phase 1 machine learning results with models ordered from best to worst for each NDT.

NDT	Model	R^2	σ_e	MSE	RMSE (MPa)	MAE (MPa)	MAPE (%)
UPV	CatBoost	0.874	5.54	31.0	5.56	3.86	16.2
	XGBoost	0.872	5.59	31.4	5.60	3.86	16.1
	LightGBM	0.842	6.18	39.0	6.24	4.49	18.3
	ANN	0.833	6.36	40.9	6.39	4.55	19.0
	k-NN	0.825	6.57	43.5	6.55	4.54	18.7
	SVR	0.822	6.58	43.7	6.60	4.66	19.6
	GBRT	0.796	6.99	50.2	7.08	5.10	20.4
RH	CatBoost	0.891	7.74	60.0	7.74	5.43	14.6
	XGBoost	0.889	7.82	61.3	7.82	5.51	14.8
	ANN	0.885	7.92	63.1	7.94	5.73	15.6
	LightGBM	0.881	8.09	65.7	8.10	5.92	16.1
	SVR	0.874	8.33	69.5	8.33	5.89	15.9
	GBRT	0.872	8.38	70.6	8.40	6.20	16.9
	k-NN	0.866	8.58	73.9	8.59	6.32	17.5
SonReb	SVR	0.935	3.54	12.6	3.54	2.63	10.2
	CatBoost	0.934	3.57	12.8	3.58	2.40	9.38
	ANN	0.929	3.72	13.8	3.71	2.67	10.2
	XGBoost	0.926	3.78	14.3	3.77	2.54	10.1
	k-NN	0.920	3.91	15.4	3.91	2.66	10.3
	LightGBM	0.919	3.95	15.8	3.96	2.79	10.7
	GBRT	0.903	4.30	18.8	4.32	3.10	11.8

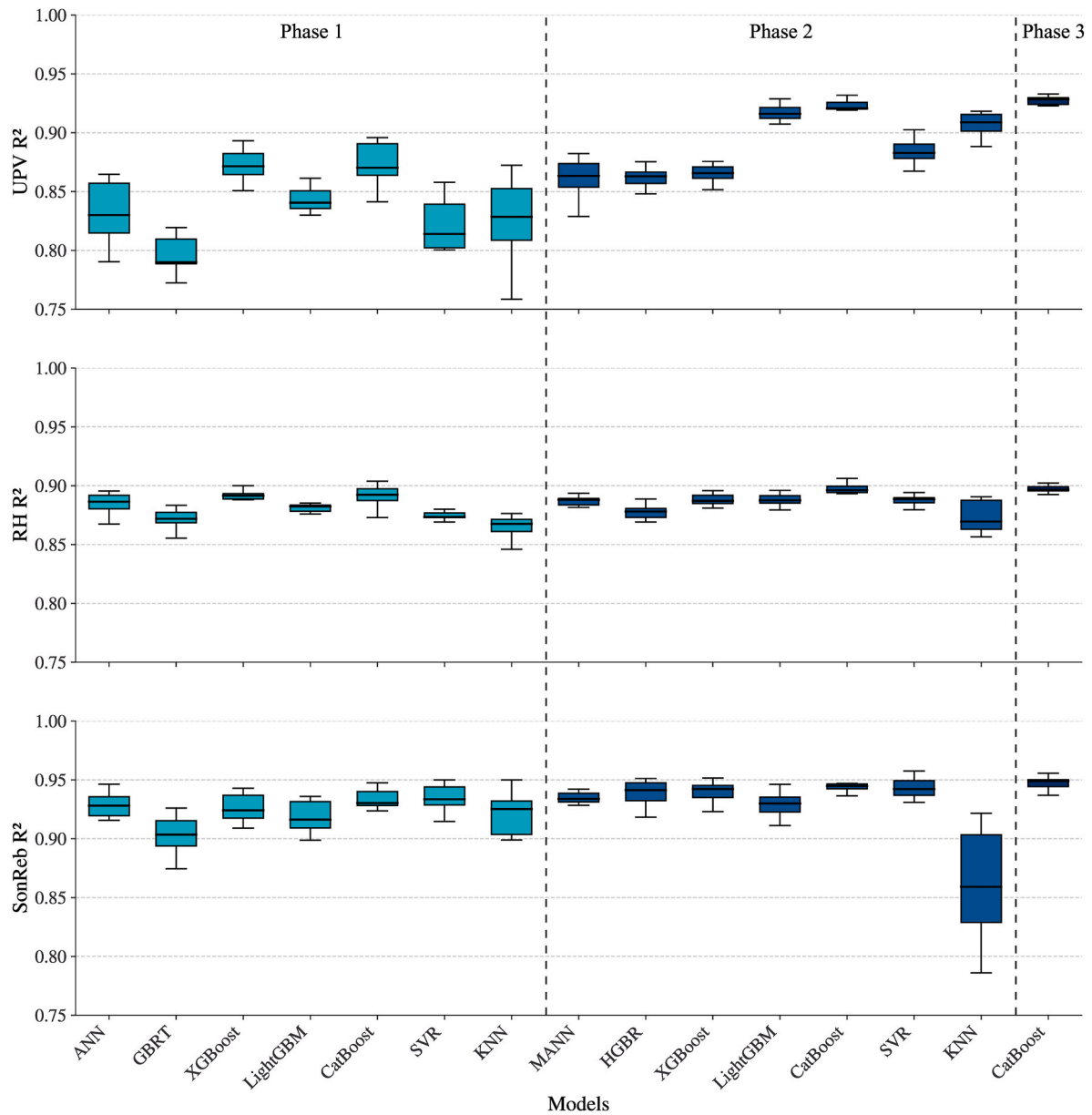


Fig. 10. Box and whisker distributions of the model correlation coefficients for each NDT across all phases.

folds and Phase 1 models. For SonReb, the SVR model outperformed all alternatives, with an R^2 of 0.934 and an MAE of 2.63 MPa. By comparison, the best empirical SonReb model (exponential function, Table 2) yielded an R^2 of 0.683 and an MAE of 6.05 MPa, representing a relative MAE reduction of 56.5 % for the SVR. UPV experienced the largest growth relative to the new empirical models proposed in Table 2, increasing R^2 from 0.427 (exponential model) to 0.874 (CatBoost model). The MAE subsequently reduced by 57.7 % from 9.13 MPa to 3.86 MPa.

During the EDA, variance inflation factors (VIFs) consistently remained below 10 for Phase 1 features, confirming acceptable levels of multicollinearity to evaluate feature influence. Consequently, SHapley Additive exPlanations (SHAP) were applied to each CatBoost regressor to quantify feature influence on the response. Fig. 11 displays the combined global feature importances (mean absolute SHAP values) in ascending order. The RN exerts substantially greater influence on compressive strength than V_p , corroborating the behaviour observed in Fig. 10 and Table 4. The RH testing standard also markedly affects strength predictions, indicating that procedural and statistical reporting

conventions have a nontrivial impact. UPV transmission type and the presence of rebar demonstrate the lowest global importance, despite the common consensus of their effect on test outcomes [16,31,71]. This behaviour is likely due to their under-representation within the training data (e.g., only 155 indirect tests among 5681), rather than a negligible mechanical effect, suggesting that more targeted experiments could clarify their true importance.

Fig. 12 presents the ridgeline kernel density distributions for the residual errors ($y_i - \hat{y}_i$) of each NDT method across all modelling phases. Empirical models exhibit non-normal residual error distributions for UPV and SonReb, displaying heavy-tailed left-skewing (indicating a tendency toward under-estimation). The mean of the residuals does not converge at zero either, confirming that the empirical models fail to fit the empirical datasets adequately – despite this being the best achievable outcome among existing empirical approaches (Appendix A). RH demonstrates a closer approximation to normality, likely due to the larger sample population, but still exhibits significant heavy tails in both directions. Given that the RH residual errors are normally distributed about a mean of 0 MPa (Fig. 12), approximately 95 % of all predictions

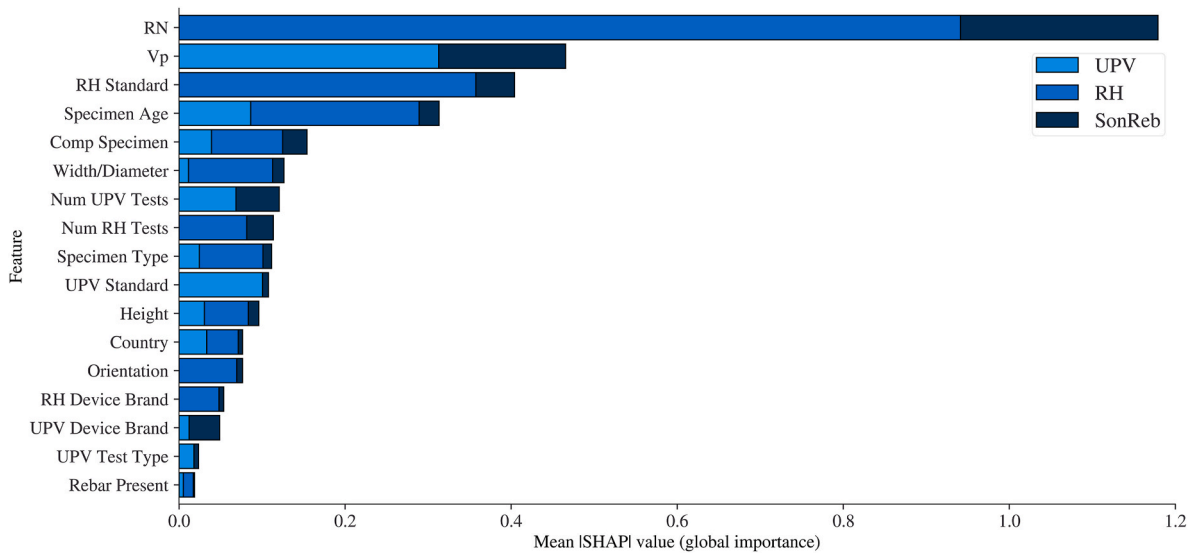


Fig. 11. Global ranked feature importance for each NDT.

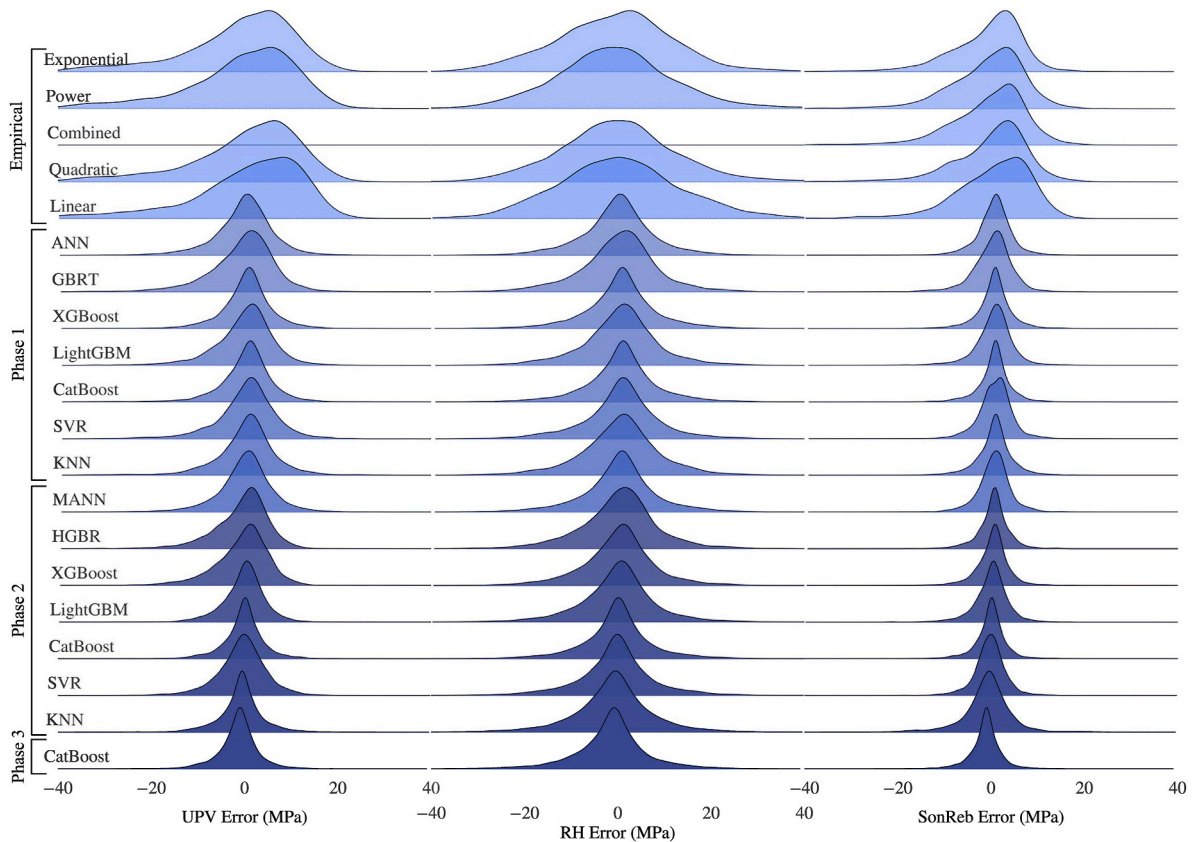


Fig. 12. Ridgeline kernel density plots for all models fitted or trained within this study for UPV (left); RH (middle); and SonReb (right).

fall within ± 15.48 MPa (two standard deviations). The SonReb method offers the most reliable consistency in predictive performance, despite being trained on approximately half the volume of data as the UPV models and one-third of the RH datasets.

4.2. Phase 2: all features

4.2.1. Feature selection

Phase 2 incorporates additional incomplete features into model

training, supplementing the feature set used in Phase 1. These newly introduced variables include UPV device diameter and frequency, maximum aggregate size, water/cement ratio, and the 28-day target strength. Meanwhile, features manually imputed in Phase 1 (e.g., specimen age and device brand/model) were left unmodified to assess their impact in an incomplete state.

The RH database contained the highest proportion of missing values, primarily due to Szilágyi [33], who contributed 4236 test results, many of which were collected from historical publications lacking detailed

reporting. For RH, the proportion of missing inputs was 62.4 % (specimen age), 18.8 % (device brand/model), 73.3 % (max aggregate size), 74.2 % (W/C ratio), and 87.4 % (target strength). Hence, the purpose of Phase 2 is to evaluate whether including incomplete features can still improve model performance, leveraging different intrinsic and extrinsic strategies for handling missing values. The final sets comprised 17 features for UPV, 15 for RH, and 22 for SonReb.

4.2.2. Model selection

The strategies for handling missing values can be categorised into two broad approaches: 1) native handling of missing data (built-in mechanisms within specific models) and 2) imputation-based methods (explicitly filling in missing data). Tree-based models typically provide robust intrinsic strategies for handling missing values, making them particularly suitable for datasets with incomplete information. For instance, CatBoost does not require explicit imputation for missing numerical values, instead substituting a default value learned during training [161]. Missing categorical values are processed through an ordered target encoding mechanism, allowing the model to infer patterns from missing entries as if they were distinct categories.

The Phase 2 model list incorporates both native and imputation-based techniques:

- The ANN model was adapted into a masked artificial neural network (MANN), which applies a masking layer directly after the input layer. Missing values are indicated using a unique placeholder ‘mask’, preventing the need for explicit imputation. This method is particularly effective when the missing values are systematic and not random.
- GBRT relies on exact feature splits when constructing decision trees and cannot inherently handle missing values. Therefore, the GBRT was replaced with histogram-based gradient boosting regression (HGBR), which groups continuous features into discrete bins, dynamically assigning splits for missing values to minimise the loss function [162].
- XGBoost and LightGBM handle missing data via strategies similar to HGBR and were retained from Phase 1.
- Kernel-based methods (e.g. SVR) lack in-built mechanisms for handling missing values and require explicit imputation before training. For the k-NN model, a simple imputation method was followed, taking the feature mean for missing numerical values and feature mode for missing categorical values.
- Multivariate imputation by chained equations (MICE) is a powerful method that preserves the uncertainty associated with missing data by generating multiple imputed datasets, rather than applying fixed values (e.g., the mean) [163]. MICE can be particularly beneficial when considering multiple features with missing values, when the missing entries are correlated or systematic, and works for both numerical and categorical variables. Given that SVR performed well for SonReb in Phase 1 (Table 4), an augmented model was developed by integrating MICE with SVR.

Thus, seven models were trained and optimised in Phase 2. The same k-fold cross-validation training scheme, grid-search optimisation algorithms, and performance metrics established in Phase 1 were maintained.

4.2.3. Model performance

Table 5 summarises the average performance metrics for each model trained and tested in Phase 2, ranked from best to worst predictor. CatBoost emerged as the top-performing model for all NDTs, surpassing SVR for SonReb, which exhibited only minor improvement after integrating the additional features and MICE framework. The tree-based models with native capabilities for handling missing data showed notable improvements relative to Phase 1. Similar to Phase 1, UPV demonstrated the most improvement after including incomplete

Table 5

Phases 2 and 3 ML results with models ordered from best to worst for each NDT.

NDT	Model	R ²	σ_e	MSE	RMSE (MPa)	MAE (MPa)	MAPE (%)
UPV	CatBoost + TPE	0.926	4.20	18.0	4.24	2.96	12.3
	CatBoost	0.923	4.34	18.9	4.34	2.99	12.6
	LightGBM	0.917	4.49	20.4	4.51	3.25	13.6
	k-NN	0.910	4.71	22.2	4.71	3.19	13.3
	SVR + MICE	0.880	5.42	29.5	5.42	3.88	16.4
	XGBoost	0.865	5.67	33.3	5.77	4.22	17.3
	MANN	0.862	5.80	34.1	5.83	4.17	17.2
	HGBR	0.862	5.68	34.0	5.83	4.22	17.1
	CatBoost + TPE	0.897	7.54	56.7	7.53	5.26	14.2
	CatBoost	0.897	7.54	56.9	7.54	5.21	14.0
RH	SVR + MICE	0.887	7.87	61.9	7.87	5.57	14.9
	XGBoost	0.887	7.87	62.2	7.88	5.70	15.6
	LightGBM	0.887	7.88	62.3	7.89	5.70	15.5
	MANN	0.887	7.88	62.4	7.89	5.62	15.3
	HGBR	0.876	8.24	68.2	8.25	6.08	16.7
	k-NN	0.874	8.33	69.6	8.33	5.96	16.2
	CatBoost + TPE	0.948	3.18	10.1	3.17	2.19	8.64
	CatBoost	0.945	3.26	10.7	3.25	2.19	8.55
	SVR + MICE	0.940	3.40	11.6	3.38	2.41	9.33
	XGBoost	0.939	3.42	11.8	3.42	2.38	9.22
SonReb	HGBR	0.938	3.45	11.9	3.44	2.34	9.11
	MANN	0.935	3.56	12.7	3.56	2.55	9.92
	LightGBM	0.929	3.69	13.7	3.70	2.58	9.99
	k-NN	0.862	5.16	26.6	5.10	3.48	13.1

features, reducing the MAE from 3.86 MPa to 2.99 MPa, representing a further 22.5 % improvement. Conversely, RH models exhibited only minor performance changes (4.12 % CatBoost improvement), likely due to the higher proportions of missing data.

From Tables 4 and 5, SonReb consistently exhibits the strongest correlation to the cylinder compressive strength, demonstrated by the significantly lower variance depicted in Fig. 12. Every metric presented is discernibly better than either UPV or RH individually (Fig. 10). Although the extensive review work performed by Breysse [18] identified inconsistent findings regarding SonReb’s efficacy, these findings appear to be limited to traditional univariate or bivariate empirical models. Machine learning with expanded feature sets demonstrates that SonReb provides a significantly stronger predictive capability when leveraging complex multivariate modalities.

Across all NDTs, Phase 2 consistently outperforms Phase 1, as demonstrated by Figs. 10 and 11, confirming that incorporating incomplete features positively influences model competency. Phase 1 comprises the core variables that are universally reported across the literature. Given the exhaustiveness of the 115-study databases, this core set is unlikely to change materially given further expansion. Phase 2 selectively augments the models with partially reported variables that demonstrate statistical association during EDA, and is therefore expected, and observed, to outperform Phase 1. Thus, the top-performing models from Phase 2 (CatBoost regression) will be included in Phase 3, integrating a more rigorous optimisation algorithm into the training process to further maximise predictive potential.

4.3. Phase 3: best model optimisation

To further augment the performance of the best models from Phases 1 and 2, a Tree-Structured Parzen Estimator (TPE) was integrated into the k-fold cross-validation training algorithm for each NDT, replacing the earlier grid-search algorithm. TPE is a practical multi-objective Bayesian optimiser widely employed for advanced hyperparameter

tuning in complex ML tasks [164]. Unlike grid-search, which exhaustively evaluates predefined hyperparameter sets, TPE constructs probabilistic models of the objective function, distinguishing between favourable and unfavourable hyperparameter configurations. These models are iteratively refined to identify an optimal set of hyperparameters. Given a set of hyperparameters, x , and an objective function, y , (e.g. validation loss or MAE), TPE models the following conditional probabilities:

$$p(y|x) = \begin{cases} l(x), & \text{if } y < y^* \\ g(x), & \text{if } y \geq y^* \end{cases} \quad (25)$$

$$l(x) = p(x|y < y^*) \quad (26)$$

$$g(x) = p(x|y \geq y^*) \quad (27)$$

Where $l(x)$ and $g(x)$ are the likelihoods of a given hyperparameter configuration resulting in a *good* (low) or *bad* (high) objective function value, with y^* denoting a threshold used to separate the best hyperparameter trials from the rest, typically expressed as a quantile value. The ratio $l(x)/g(x)$ guides the search toward hyperparameter spaces with higher potential, and the next hyperparameter configuration is selected by maximising the Expected Improvement (EI) [164].

$$x_{i+1} = \arg\max_x \frac{l(x)}{g(x)} \quad (28)$$

Although TPE is based on Bayesian Optimisation, it substitutes the traditional Gaussian Process with a more scalable Parzen Estimator, improving efficiency in high-dimensional (feature) spaces [165]. In exchange, TPE requires a minimum number of trials to build meaningful probability functions before it can effectively exploit promising regions, causing small datasets (<1000 observations) to converge prematurely and lead to suboptimal results. TPE is a particularly effective framework for gradient-boosting models, such as CatBoost, LightGBM, and XGBoost, due to their similar tree-based learning structures and hyperparameter characteristics. The CatBoost hyperparameters tuned using TPE within each training fold include the loss function, the number of

trees in the model, the learning rate, depth, and L2 regularisation (to prevent overfitting).

The performance results of the TPE-augmented CatBoost models are summarised in Table 5. The additional TPE optimisation yielded negligible changes to the RH and SonReb models, as illustrated in Figs. 10 and 11, although UPV benefited more from its integration. This outcome implies that the volume and reliability of the collected databases allow the models to perform near their theoretical limits, leaving only slight gains achievable through additional hyperparameter tuning. Fig. 13 presents the final regressions and residual error distributions of the CatBoost-TPE models compared to the actual compressive strengths.

The RH method visibly demonstrates a considerable residual scatter with heteroscedastic behaviour, indicating a systematic change in variance. Such heteroscedasticity could result from the optimised model not fully capturing the underlying variance within RH-based compressive strength predictions, limitations in the model structure, or inherent constraints of the NDT technique itself. For example, due to the non-linear relationship between the RN and compressive strength, there is paucity of data beyond an RN of 60, while an RN of 50 is associated with a scatter of approximately ± 30 MPa (Fig. 5). Without supplementing the training with synthetic data to augment high-strength domains, improving model training and reliability in these regions will require the collection of additional experimental data. Technological improvements and the emergence of rebound hammers specifically designed for high-strength concretes may also offer a pathway for improving predictive error.

Because of the clear exponential relationship between V_p and $f_{c,cyl}$ (Fig. 5), a similar trend is anticipated for the UPV at higher concrete strengths, which will limit the applicability of this testing technique. To enhance prediction reliability for higher-strength concretes, it may be beneficial to combine these techniques with an additional approach, such as penetration testing [13,14]. Alternatively, model performance in regions with significant scatter could be improved by employing repeated k-fold cross-validation to train and aggregate a larger ensemble of models.

A native CatBoost feature importance metric was computed for all

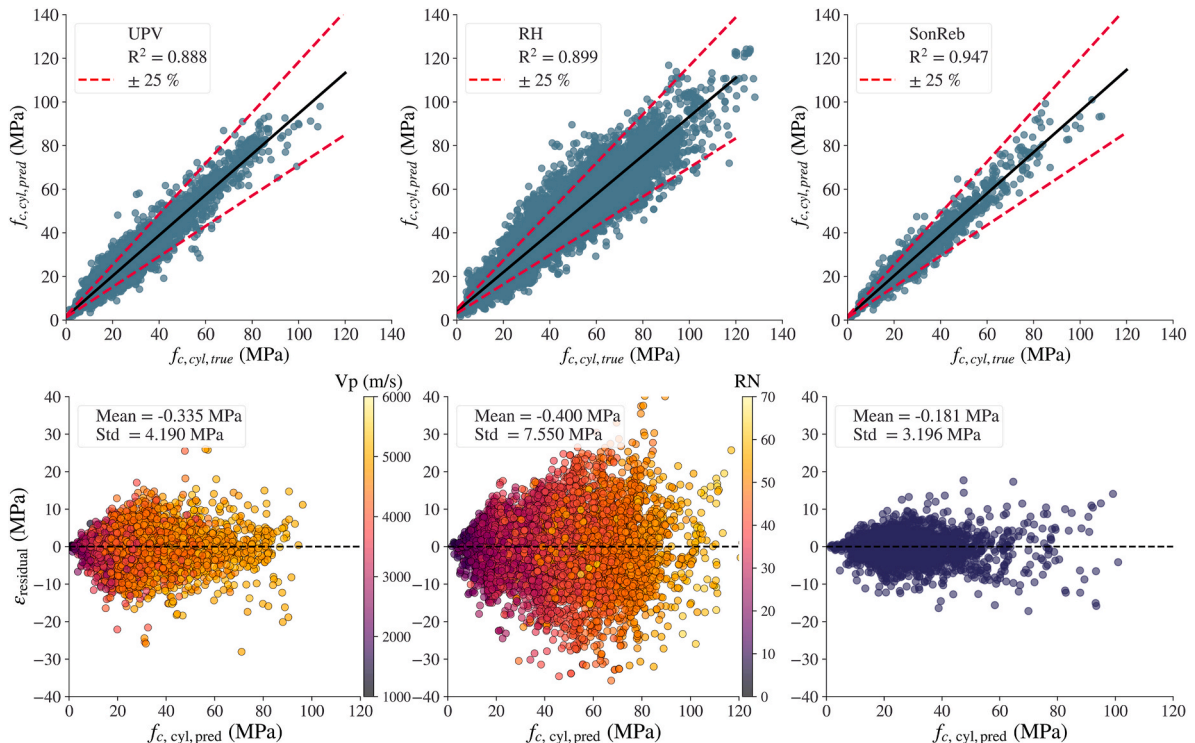


Fig. 13. CatBoost-TPE regression (top) and residual error distributions (bottom) for UPV (left); RH (middle); and SonReb (right).

ten folds and averaged to obtain a robust in-house ensemble feature ranking. This approach quantifies the average change in the model output attributable to each feature, directly reflecting the influence of features with missing data. The top 15 features by mean importance are shown in Table 6, where the value represents the percentage contribution to the prediction. The results corroborate the same behaviour observed in Fig. 11, with the additional concrete details – maximum aggregate size, w/c ratio, and target 28-day strength, all demonstrating measurable impact on the model performance, despite the varying levels of missing data.

5. Web application

The final step of this study involves providing researchers and practitioners with a practical tool for predicting the cylinder compressive strength, utilising the optimised TPE-CatBoost models given in Table 5 and Section 4.3. While ML models consistently outperform conventional empirical equations, their dissemination and practical implementation can be challenging, as they do not produce simple, closed-form expressions. To address this issue, a web-based GUI has been developed and deployed on a cloud server, hosted on Render, which can be accessed at the URL: <https://recreate-ndt.onrender.com/>

The frontend of the application is designed using HTML, CSS, and JavaScript modules to create a simple and intuitive user interface. A Django-based backend retrieves the user-supplied information from the frontend, processes the data, runs the TPE-CatBoost models and returns real-time predictions to the user interface. The user can select the NDT method, initiating the relevant input features. Once the necessary feature information is provided, the TPE-CatBoost models retrieve the relevant features from the backend, process the input, and generate a prediction of the cylinder compressive strength. The trained TPE-CatBoost models from all ten cross-validation folds are saved offline and stored in a secure cloud storage system. The final compressive strength prediction displayed to the user is the average output from all ten models, each contributing with equal weighting. The accuracy of the generated predictions is presented in Table 5 for each NDT. It should be stressed that these performance metrics apply only within the domains of the available training data (Table 2) for each NDT method; outside of which, the models must extrapolate, and accuracy will diminish. Refer to Fig. 13 for the relationship between $f_{c,cyl}$ and model error. For deployment, the backend and frontend code have been added to a public GitHub repository as per the open research requirements of the EU-funded ReCreate project. The repository can be accessed at: <https://github.com/bma114/recreate-ndt>.

The incomplete features incorporated in Phase 2 are left as optional fields, while all other features are treated as required fields before generating predictions. Concrete compression specimen details are also optional, enabling predictions based solely on NDT parameters and specimen metadata. All numerical features are normalised using the

Table 6

Top 15 features influencing compressive strength through CatBoost's native importance test.

Feature	SonReb	UPV	RH
RN	30.63	–	47.48
V_p	24.67	38.43	–
UPV Device & Brand	6.376	0.953	–
Specimen Age (days)	4.996	12.19	4.508
W/C Ratio	4.470	9.754	4.842
Max Agg. Size (mm)	3.375	3.084	3.990
Compression Specimen	3.288	5.046	2.733
No. UPV Tests	3.171	5.362	–
NDT Specimen Type	2.895	4.030	2.352
No. RH Tests	2.389	–	2.706
Height (mm)	2.266	1.940	3.532
Width/Diameter (mm)	2.184	1.676	1.785
Design Strength (MPa)	2.129	3.051	2.306
RH Standard	1.814	–	2.522
RH Device & Brand	1.665	–	15.82

same scaler functions applied during model training, ensuring consistency with the original preprocessing pipeline.

Since CatBoost regression models take categorical variables in their raw string formats, no encoding of the user-supplied inputs is necessary. However, to mitigate user-input inconsistencies, categorical features are normalised during testing (Phase 3) and within the Django backend. This step offsets the risk of user-induced errors such as inadvertent spaces, inconsistent capitalisation, or formatting discrepancies. For more information regarding the expected formatting of the numerical and categorical features, for optimal model performance, refer to the databases and supplementary documentation stored in the Zenodo repository at: <https://doi.org/10.5281/zenodo.14921019>. Since the app is deployed on Render's free tier, and due to the size of the ensemble models being loaded from AWS, a short cold start delay of approximately 2–3 min is expected each time the app is opened.

6. Model and app validation

To assess the predictive performance and underlying learned behaviour of the final models (Fig. 13), the web app is applied to a small set of randomly selected tests spanning a wide range of compressive strengths. Although these tests were included in the training data, they serve as a simple demonstration of the model and app performance. Table 7 presents the predictive results of the tests for each NDT, generated through the cloud app described in Section 5.

Prediction accuracy declines modestly as strength increases. This behaviour reflects the training domains summarised in Table 1 – model performance degrades near the empirical boundaries as data becomes more sparse, and deteriorates more substantially when extrapolating beyond them.

A complementary verification check, summarised in Table B.1 (Appendix B), represents a rudimentary feature influence test to ensure the models behave as expected. Since multicollinearity is not addressed in this paper, actual feature influences on the response may be suppressed, depending on their inter-feature collinearity. Nonetheless, this brief analysis provides an indication of the functional capacity of the models and learned understanding of patterns between the independent and dependent variables. The test from Fawzi et al. [71] in Table 7 is applied for this purpose (Specimen ID: B40_S3), where one feature is varied while holding all else constant, observing the ensuing influence on the predicted compressive strength.

7. Conclusions

This study evaluates the global predictive capability of various non-destructive test (NDT) methods for indirectly estimating concrete compressive strength. The investigation focused on ultrasonic pulse velocity (UPV), rebound hammer (RH), and SonReb (UPV and RH combined) due to their widespread adoption in both research and practice. Comprehensive databases were compiled from 90 studies (6103 test results) for UPV, 87 studies (10,428 tests) for RH, and 53 studies (3299 tests) for SonReb. In each case, the response variable was the concrete compressive strength, normalised to a reference cylinder of 150×300 mm dimensions. The databases were first used to assess the performance of existing empirical models proposed in the literature, which were initially calibrated to more localised datasets. New relationships were then derived to capture the 'global' behaviour of each NDT method across the typical typologies – exponential, power, quadratic, and linear. A three-phase machine learning (ML) framework was then employed to investigate the impact of incorporating incomplete features (variables with a high proportion of missing values) on predictive performance. Phase 1 examined seven ML models from three different typologies – network-based, tree-based, and kernel machines – trained on a core feature set with complete or near-complete reporting. Phase 2 involved seven models that implemented different imputation strategies to handle substantial volumes of missing data. A k-fold cross-

Table 7

Model performance obtained through the web application.

NDT	Study	V_p (m/s)	RN	$f_{c,cyl,act}$ (MPa)	$f_{c,cyl,pred}$ (MPa)	Error (MPa)	Error (%)
UPV	Poorarabai [23]	4390	–	16.3	16.4	0.13	0.79
	Nash't et al. [9]	4609	–	27.4	32.5	5.13	15.8
	Fawzi et al. [71]	4959	–	35.1	34.1	–0.96	–2.81
	Cianfrone & Facaoaru [38]	5190	–	58.2	50.2	–8.05	–16.0
	de Almeida [53]	5000	–	81.8	90.8	8.97	9.88
RH	Poorarabai [23]	–	20	16.3	15.3	–1.04	–6.81
	Nash't et al. [9]	–	37	27.4	32.1	4.70	14.7
	Fawzi et al. [71]	–	44	35.1	39.3	4.21	10.7
	Cianfrone & Facaoaru [38]	–	41	58.2	50.5	–7.75	–15.4
	de Almeida [53]	–	55	81.8	98.1	16.3	16.6
SonReb	Poorarabai [23]	4390	20	16.3	17.8	1.51	8.47
	Nash't et al. [9]	4609	37	27.4	30.6	3.27	10.7
	Fawzi et al. [71]	4959	44	35.1	35.0	–0.06	–0.17
	Cianfrone & Facaoaru [38]	5190	41	58.2	51.2	–7.05	–13.8
	de Almeida [53]	5000	55	81.8	88.8	6.99	7.87

validation training algorithm, paired with a simple grid-search optimisation, was implemented across Phases 1 and 2. Finally, the top model from Phases 1 and 2 was integrated with a Tree-Structured Parzen Estimator (TPE) optimisation algorithm further to advance its performance for deployment in a web-based application. The following conclusions can be drawn from the investigation:

- Literature-proposed static models reveal a significant diversity related to their ability to predict the normalised compressive strength. Although many models boast high performance metrics after fitting to their local datasets or a small collection of validation sets, few models produce metrics close to their original values when compared against the global behaviour of each NDT. Due to the substantial variability within each test method, empirical models cannot adequately capture the concrete compressive behaviour strength through univariate or bivariate relationships alone. The best-fitted models to the global datasets achieved maximum R^2 values of 0.385, 0.746, and 0.648 for UPV, RH, and SonReb, respectively.
- Incorporating additional predictive features in the machine learning models resulted in substantial performance improvements compared to the empirical models. The top-performing models in Phase 1 were the CatBoost regression model for UPV and RH, and the support vector regression (SVR) model for SonReb. The UPV method experienced the greatest growth, increasing the correlation coefficient by 220.5 % to 0.849 when using CatBoost regression. The mean absolute error (MAE) was reduced by 51.1 %, 36.4 %, and 56.0 % in UPV, RH, and SonReb models. The SonReb method consistently demonstrated the highest aptitude for predicting compressive strength, reaching an MAE of 2.44 MPa and an R^2 of 0.933 in Phase 1. These results are valid over the available training data domains, which are concentrated within [0.05, 0.95] fractiles for $f_{c,cyl}$ in MPa of [8.47, 59.6] for UPV, [11.7, 85.0] for RH, and [11.3, 58.7] for SonReb. A small SHAP and native CatBoost feature importance analysis highlights the hierarchical influence of features on the compressive strength predictions.
- After incorporating several heavily incomplete features into model training (Phase 2), tree-based models, which could natively handle missing values, consistently outperformed their counterparts trained on only complete features (Phase 1). CatBoost regression was the best performer for all NDTs and experienced a further improvement of 16.3 % in MAE for UPV compared to Phase 1. Some imputation approaches (e.g., multivariate imputation by chained equations, MICE), which were used to augment the SVR model, did not yield any notable improvements across phases, likely due to the severity of the missing data for some features.
- Overall, CatBoost regression emerged as the best-performing model across Phases 1 and 2. An advanced TPE optimisation algorithm was then applied during k-fold training, further refining the hyperparameters. After optimisation, only minor improvements were

observed compared to Phase 2, suggesting that the models had already reached their near-optimal performance, given the dataset volume and quality. Finally, a Django-based web application was developed to assimilate the optimised models, allowing users to generate real-time predictions based on new NDT inputs through a convenient and straightforward interface.

CRedit authorship contribution statement

Benjamin Matthews: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Diego Allaix:** Writing – review & editing, Supervision. **Simon Wijte:** Writing – review & editing, Supervision, Conceptualization. **Marcel Vullings:** Writing – review & editing, Supervision, Conceptualization.

Code availability

The Python code used in this study to build, train, and test the CatBoost-TPE model, taking UPV as an example, has been added to a public GitHub repository, which can be accessed at: <https://github.com/bma114/upv-compressive-strength-predictions>.

The full suite of code written in Python, HTML, JavaScript and CSS for the web application developed as part of this study is publicly available at the following GitHub repository: <https://github.com/bma114/recreate-ndt>.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Benjamin Matthews reports financial support was provided by Horizon 2020 European Innovation Council Fast Track to Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 958200.

Disclaimer: The content presented herein reflects the authors' views. The European Commission is not responsible for any use that may be made of the information this publication contains.

The authors gratefully acknowledge Ramboll Finland OY for providing additional industry test data, which helped improve the reliability and robustness of our models.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ndteint.2025.103549>.

APPENDIX A. EXISTING EMPIRICAL FORMULAE

Table A1

Published empirical models for the relationship between UPV and compressive strength.

Model Type	Equation	Eq.	Local R ²	Global R ²	Reference
Exponential	$f_c(V_p) = 1.146 \cdot e^{(0.77V_p)}$	(A1)	0.920	0.309	Turgut [146]
	$f_c(V_p) = 1.19 \cdot e^{(0.715V_p)}$	(A2)	0.590	0.417	Nash't et al. [9]
	$f_c(V_p) = 0.0854 \cdot e^{(1.288V_p)}$	(A3)	0.640	0.310	Trtnik et al. [61]
	$f_c(V_p) = 11.804 \cdot e^{(0.2601V_p)}$	(A4)	0.081	0.036	Al-Nu'man et al. [147]
	$f_c(V_p) = 1.2288 \cdot e^{(0.726V_p)}$	(A5)	0.700	0.421	Ali-Benyahia et al. [126]
Power	$f_c(V_p) = 1.2 \cdot 10^{-5} \cdot (1000V_p)^{1.7447}$	(A6)	0.409	0.308	Kheder [20]
	$f_c(V_p) = 5.4942 \cdot V_p^{0.9874}$	(A7)	0.780	0.101	Kumavat & Chandak [148]
	$f_c(V_p) = 0.6401 \cdot V_p^{2.5654}$	(A8)	0.720	0.392	Ali-Benyahia et al. [21]
Polynomial	$f_c(V_p) = 13.906 \cdot V_p^2 - 96.467 \cdot V_p + 176.9$	(A9)	0.820	0.194	Logothetis [149]
	$f_c(V_p) = 28.9 \cdot V_p^2 - 221.6 \cdot V_p + 440.1$	(A10)	0.882	-0.556	Trezos et al. [150]
Linear	$f_c(V_p) = 36.73 \cdot V_p - 129.077$	(A11)	0.956	-0.189	Qasrawi [151]
	$f_c(V_p) = 15.533 \cdot V_p - 34.58$	(A12)	0.919	0.315	Shariati et al. [31]

Note V_p is in km/s.

Table A2

Published empirical models for the relationship between RH and compressive strength.

Model Type	Equation	Eq.	Local R ²	Global R ²	Reference
Exponential	$f_c(RN) = 3.9622 \cdot e^{(0.0504RN)}$	(A13)	0.903	0.346	Pucinotti & Lorenzo [152]
	$f_c(RN) = 2.6113 \cdot e^{(0.0558RN)}$	(A14)	0.770	0.036	Ali-Benyahia et al. [21]
Power	$f_c(RN) = 0.4030 \cdot RN^{1.2083}$	(A15)	0.805	0.347	Kheder [20]
	$f_c(RN) = 0.788 \cdot RN^{1.03}$	(A16)	0.770	0.317	Nash't et al. [9]
	$f_c(RN) = 0.0307 \cdot RN^{1.952}$	(A17)	NR	0.681	Szilágyi [33]
	$f_c(RN) = 0.0238 \cdot RN^{1.8781}$	(A18)	0.770	-0.146	Ali-Benyahia et al. [126]
Polynomial	$f_c(RN) = 0.02 \cdot RN^2 + 0.52 \cdot RN - 9.40$	(A19)	0.952	0.700	Logothetis [149]
	$f_c(RN) = 0.032 \cdot RN^2 - 0.164 \cdot RN + 5.9$	(A20)	0.851	0.640	Trezos et al. [150]
	$f_c(RN) = -0.0177 \cdot RN^2 + 2.0481 \cdot RN - 19.303$	(A21)	0.810	0.136	Erdal [11]
Linear	$f_c(RN) = 1.353 \cdot RN - 17.393$	(A22)	0.880	0.469	Qasrawi [151]
	$f_c(RN) = 1.7206 \cdot RN - 26.595$	(A23)	0.936	0.641	Shariati et al. [31]
	$f_c(RN) = 2.832 \cdot RN - 35.472$	(A24)	0.878	-0.937	Kocáb et al. [105]

NR = Not reported.

Table A3

Published empirical models for the relationship between SonReb and compressive strength.

Model Type	Equation	Eq.	Local R ²	Global R ²	Reference
Exponential & Power-Exponential	$f_c(V_p, RN) = 0.0981 e^{(1.78 \ln(V_p) + 0.85 \ln(RN) - 0.02)}$	(A25)	0.934	0.589	Logothetis [149]
	$f_c(V_p, RN) = 0.0981 \cdot 18.6 e^{(0.515V_p + 0.019RN)}$	(A26)	0.954	0.528	Arioglu & Manzak [153]
	$f_c(V_p, RN) = 0.356 \cdot RN^{0.866} \cdot e^{(0.302V_p)}$	(A27)	0.800	0.620	Nash't et al. [9]
Double Power	$f_c(V_p, RN) = 0.0158 \cdot (1000V_p)^{0.4254} \cdot RN^{1.1171}$	(A28)	0.907	0.581	Kheder [20]
	$f_c(V_p, RN) = 2.6199 \cdot 10^{-8} \cdot (1000V_p)^{2.2878} \cdot RN^{0.5341}$	(A29)	0.339	0.380	Faella et al. [154]
	$f_c(V_p, RN) = 1.6411 \cdot 10^{-9} \cdot (1000V_p)^{2.29366} \cdot RN^{1.30768}$	(A30)	0.867	0.208	Nikhil et al. [84]
	$f_c(V_p, RN) = 0.0543 \cdot V_p^{1.286} \cdot RN^{1.171}$	(A31)	0.840	0.395	Ali-Benyahia et al. [21]
	$f_c(V_p, RN) = 9.043 \cdot 10^{-6} \cdot (1000V_p)^{1.381} \cdot RN^{1.036}$	(A32)	0.598	0.417	Breccolotti & Bonfigli [32]
Bilinear	$f_c(V_p, RN) = 0.951 \cdot V_p + 0.745 \cdot RN - 0.544$	(A33)	0.940	0.490	Tanigawa et al. [155]
	$f_c(V_p, RN) = 13.166 \cdot V_p + 0.42 \cdot RN - 40.255$	(A34)	0.857	0.540	Erdal [11]
	$f_c(V_p, RN) = 7.666 \cdot V_p + 1.017 \cdot RN - 38.653$	(A35)	0.974	0.631	Fawzi et al. [71]
	$f_c(V_p, RN) = 5.752 \cdot V_p + 0.653 \cdot RN - 24.674$	(A36)	0.840	0.320	Ali-Benyahia et al. [21]

Note V_p is in km/s.

APPENDIX B. PHASE 3 FEATURE INFLUENCE VERIFICATION

Table B.1

Simple verification of feature influence on predicted compressive strength, to determine expected relationships.

Feature Name	Model	Original Value	Prediction, $f_{c,cyl,pred1}$ (MPa)	New Value	Prediction, $f_{c,cyl,pred2}$ (MPa)	Observation
Specimen Type	SonReb	Cylinder - Laboratory	36.07	Cylinder - In-situ	34.88	A 3.3 % decrease is observed due to the effects of drilling, demonstrating that cored cylinders are correctly interpreted as having lower compressive strengths [50].
Rebar Present	UPV	FALSE	38.68	TRUE	36.96	Inclusion of rebar in the concrete reduces the predicted strength by 4.5 %, as expected from wave-path interference. The effect is smaller than the 0.8–0.9 corrections in Ref. [16], perhaps due to the unknown bar orientation.
UPV Transmission Type	UPV	Direct	38.68	In-direct	38.49	Negligible change observed, contrary to the expected lower wave velocities in indirect tests. The absence of this change indicates that the models require additional indirect test data for calibration and to improve understanding, aligning with Fig. 11.
RH Device Brand & Model	RH	Original Schmidt Type N	40.54	Silver Schmidt	36.00	Silver Schmidt hammers yield lower rebound numbers, resulting in smaller estimated compressive strengths that align with the expected device behaviour.
Maximum Aggregate Size (mm)	RH	10	38.02	30	41.10	An 8.1 % increase is observed when the maximum aggregate size increases for an RH test. The presence of larger aggregates near the surface can cause artificial inflation of the measured surface hardness.
W/C Ratio	SonReb	0.65	30.69	0.40	36.40	Reducing W/C ratios increases predicted strength, consistent with established mix-design principles.
28-Day Target (Design) Strength (MPa)	UPV	40	38.68	60	46.32	The model accurately predicts higher strength for increased 28-day design values, despite substantial rates of missing data.

Data availability

The complete collection of databases has been made publicly available and open source at: <https://doi.org/10.5281/zenodo.14921019>. A user's guide is provided alongside the database to provide full descriptions of all assumptions, nomenclature, equations, and abbreviations used throughout the databases. The authors hope that these databases may serve as a strong foundation for future research and the accumulation of experimental data, which is necessary to continually improve the reliability of non-destructive methods for estimating concrete compressive strength.

References

- [1] Küpfer C, Bertola N, Brütting J, Fivet C. Decision framework to balance environmental, technical, logistical, and economic criteria when designing structures with reused components. *Front Sustain* 2021;2. <https://doi.org/10.3389/frsus.2021.689877>.
- [2] Addis B. Building with reclaimed components and materials: a design handbook for reuse and recycling. Sterling, VA: Earthscan; 2006. p. 225.
- [3] United Nations Environment Program. 2023 global status report for buildings and construction: beyond foundations. UNEP; 2024. <https://doi.org/10.59117/20.500.11822/45095>.
- [4] Allam AS, Nik-Bakht M. From demolition to deconstruction of the built environment: a synthesis of the literature. *J Build Eng* 2023;64. <https://doi.org/10.1016/j.jobbe.2022.105679>.
- [5] Gordon M, Batallé A, de Wolf C, Sollazzo A, Dubor A, Wang T. Automating building element detection for deconstruction planning and material reuse: a case study. *Autom Constr* 2023;146. <https://doi.org/10.1016/j.autcon.2022.104697>.
- [6] Küpfer C, Bastien-Masse M, Fivet C. Reuse of concrete components in new construction projects: critical review of 77 circular precedents. *J Clean Prod* 2023;383. <https://doi.org/10.1016/j.jclepro.2022.135235>.
- [7] Tomsett HN. The practical use of ultrasonic pulse velocity measurements in the assessment of concrete quality. *Mag Concr Res* 1980;32(110):7–16. <https://doi.org/10.1680/mac.1980.32.110.7>.
- [8] Oktar ON, Moral H, Taqdemir MA. Sensitivity of concrete properties to the pore structure of hardened cement paste. *Cement Concr Res* 1996;26(11):1619–27.
- [9] Nash't H, Hameed A'bour, Anwar S, Sadoon A, Dawod YK, Mansor SL. Finding an unified relationship between crushing strength of concrete and non-destructive tests. 3rd Middle East nondestructive testing conference & exhibition (MENDT). Bahrain: Manama; 2005.
- [10] Masi A, Vona M. La stima della resistenza del calcestruzzo in-situ: impostazione delle indagini ed elaborazione dei risultati [Estimating the strength of in-situ concrete: setting up investigations and processing results]. *Progettazione Sismica* 2008;1:53–67 [in Italian].
- [11] Erdal M. Prediction of the compressive strength of vacuum processed concretes using artificial neural network and regression techniques. *Sci Res Essays* 2009;4 (10):1057–65. <http://www.academicjournals.org/sre>.
- [12] Garnier V, Piwakowski B, Abraham O, Villain G, Payan C, Chaix JF. Acoustic techniques for concrete evaluation: improvements, comparisons and consistency. *Constr Build Mater* 2013;43:598–613. <https://doi.org/10.1016/j.conbuildmat.2013.01.035>.
- [13] Craeye B, van de Laar H, van der Eijk J, Gijbels W, Lauriks L. On-site strength assessment of limestone based concrete slabs by combining non-destructive techniques. *J Build Eng* 2017;13:216–23. <https://doi.org/10.1016/j.jobbe.2017.08.006>.
- [14] Demir T, Ulucan M, Alyamaç KE. Development of combined methods using non-destructive test methods to determine the In-Place strength of high-strength concretes. *Processes* 2023;11(3). <https://doi.org/10.3390/pr11030673>.
- [15] Pucinotti R. Reinforced concrete structure: non destructive in situ strength assessment of concrete. *Constr Build Mater* 2015;75:331–41. <https://doi.org/10.1016/j.conbuildmat.2014.11.023>.
- [16] Hong S, Yoon S, Kim J, Lee C, Kim S, Lee Y. Evaluation of condition of concrete structures using ultrasonic pulse velocity method. *Appl Sci* 2020;10(2). <https://doi.org/10.3390/app10020706>.
- [17] Mikulic D, Pause Z, Ukraincik V. Determination of concrete quality in a structure by combination of destructive and non-destructive methods. *Mater Struct* 1992; 25:65–9.
- [18] Breyse D. Non-destructive evaluation of concrete strength: an historical review and a new perspective by combining NDT methods. *Constr Build Mater* 2012;33: 139–63. <https://doi.org/10.1016/j.conbuildmat.2011.12.103>.
- [19] Chitti A, Burra SG, Chu TP, Kolay P, Kumar S. Combined NDT correlation to estimate the compressive strength of concrete. *Am Soc Non-Destruct Testing ASNT* 2019;1–10.
- [20] Kheder GF. A two stage procedure for assessment of in situ concrete strength using combined non-destructive testing. *Mater Struct* 1999;32:410–7.
- [21] Ali-Benyahia K, Sbartaï ZM, Breyse D, Kenai S, Ghrici M. 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 Constr Mater* 2017;6:109–19. <https://doi.org/10.1016/j.cscm.2017.01.004>.
- [22] Asteris PG, Mokos VG. Concrete compressive strength using artificial neural networks. *Neural Comput Appl* 2020;32(15):11807–26. <https://doi.org/10.1007/s00521-019-04663-2>.
- [23] Poorarabi A, Ghasemi M, Azhdary Moghaddam M. Concrete compressive strength prediction using non-destructive tests through response surface methodology. *Ain Shams Eng J* 2020;11(4):939–49. <https://doi.org/10.1016/j.asej.2020.02.009>.

- [24] Long R. The improvement of ultrasonic apparatus for the routine inspection of concrete. London, United Kingdom: University of London; 2000. p. 256. PhD Thesis.
- [25] Karaiskos G, Deraemaeker A, Aggelis DG, van Hemelrijck D. Monitoring of concrete structures using the ultrasonic pulse velocity method. *Smart Mater Struct* 2015;24(11). <https://doi.org/10.1088/0964-1726/24/11/113001>.
- [26] Popovics S, Rose JL, Popovics JS. The behavior of ultrasonic pulses in concrete. *Cement Concr Res* 1990;20:259–70.
- [27] EN 12504-4. Testing concrete in structures, part 4: determination of ultrasonic pulse velocity. Brussels, Belgium: European Committee for Standardisation (CEN); 2021.
- [28] ASTM C597. Standard test method for pulse velocity through concrete. West Conshohocken, PA, USA: American Society for Testing and Materials; 2022.
- [29] BS 1881-203. Testing concrete Part 203. Recommendations for measurement of velocity of ultrasonic pulses in concrete. London, UK: British Standards Institute; 1986.
- [30] Ariöz Ö, Tuncan A, Tuncan M, Kavas T, Ramyar K, Kilinç K, Karasu B. Use of combined non-destructive methods to assess the strength of concrete in structures. *AKU - J Sci* 2009;147–54.
- [31] Shariati M, Ramli-Sulong NH, Arabnejad MM, Shafigh P, Sinaei H. Assessing the strength of reinforced concrete structures through ultrasonic pulse velocity and Schmidt Rebound Hammer tests. *Sci Res Essays* 2011;6(1):213–20. <https://doi.org/10.5897/SRE10.879>.
- [32] Breccolotti M, Bonfigli MF. I-SonReb: an improved NDT method to evaluate the in situ strength of carbonated concrete. *Nondestr Test Eval* 2015;30(4):327–46. <https://doi.org/10.1080/10589759.2015.1046872>.
- [33] Szilágyi K. Rebound surface hardness and related properties of concrete. Budapest, Hungary: Budapest University of Technology and Economics; 2013. PhD Thesis.
- [34] Kairu WM. *Non-destructive testing of concrete structures using schmidt hammer and profometer 5+.* Master's thesis. Nairobi, Kenya: University of Nairobi, Institute of Nuclear Science and Technology; 2016.
- [35] ASTM C 805/C805M-13a. Standard test method for rebound number of hardened concrete. West Conshohocken, PA: American Society for Testing and Materials; 2013. 2013.
- [36] EN 12504-2. Testing concrete in structures, part 2: non-destructive testing - determination of rebound number. European Committee Standardis (CEN) 2012. Brussels, Belgium. 1-18.
- [37] Facaoaru I. Non-destructive testing of concrete in Romania. In: Symp on NDT of concrete and timber. London. Institute of Civil Engineers; 1970. p. 39–49. 1970.
- [38] Cianfrone F, Facaoaru I. Study on the introduction into Italy on the combined non-destructive method, for the determination of in situ concrete strength. *Materiaux et Construct* 1979;12(71):413–24.
- [39] Amini K, Jalalpour M, Delatte N. Advancing concrete strength prediction using non-destructive testing: development and verification of a generalisable model. *Constr Build Mater* 2016;102:762–8. <https://doi.org/10.1016/j.conbuildmat.2015.10.131>.
- [40] Ivanchev I. Investigation with non-destructive and destructive methods for assessment of concrete compressive strength. *Appl Sci* 2022;12(23). <https://doi.org/10.3390/app122312172>.
- [41] Boussahoua Y, Kenai S, Sbartaï ZM, Breyse D, Ali-Benyahia K. Influence of the number of cores on concrete strength assessment by non-destructive tests in old existing structures. *Asian J Civil Eng* 2023;24(6):1731–45. <https://doi.org/10.1007/s42107-023-00599-0>.
- [42] ACI Committee 228. In-Place methods to estimate concrete strength. ACI 228.1R-03. Farmington Hills, MI, USA: American Concrete Institute; 2003.
- [43] Cristofaro MT, Viti S, Tanganelli M. New predictive models to evaluate concrete compressive strength using the SonReb method. *J Build Eng* 2020;27. <https://doi.org/10.1016/j.jobe.2019.100962>.
- [44] Haavisto J, Husso A, Laaksonen A. Compressive strength of core specimens drilled from concrete test cylinders. *Struct Concr* 2021;22(S1):E683–95. <https://doi.org/10.1002/suco.202000428>.
- [45] EN 1992-1-1. Eurocode 2: design of concrete structures - part 1-1: general rules and rules for buildings, bridges and civil engineering structures. European committee for standardisation (CEN). 2023. Brussels, Belgium.
- [46] Reineck KH, Kuchma DA, Fitik B. Research report NER050: extended databases with shear tests on structural concrete beams without and with stirrups for the assessment of shear design procedures. United States of America Nuclear Regulatory Commission; 2010.
- [47] DIN 1045-1. Concrete, reinforced and prestressed concrete structures – part 1: design. Normenausschuss Bauwesen (NABau) im DIN Deutsches Institut für Normung e.V. Beuth Verl. 2001. Berlin.
- [48] fib Bulletin 12. Punching of structural concrete slabs – technical report by the CEB/fib task group “Utilisation of concrete tension in design”. Lausanne: International Federation for Structural Concrete; 2001.
- [49] ASCE-ACI Committee 445. *Recent approaches to shear design of structural concrete.* state-of-the-art report by ASCE-ACI committee 445 on shear and torsion. ASCE J Struct Eng 1998;124(12):1375–417.
- [50] Chakiri H. *Appraisal of the compressive strength for concrete in situ.* Master's graduation thesis. Eindhoven, the Netherlands: Eindhoven University of Technology; 2016.
- [51] Malhotra VM, Carrette G. Comparison of pull-out strength of concrete with compressive strength of cylinders and cores, pulse velocity and rebound number. Canada Centre for Mineral and Energy Technology (CANMET). Research Report; 1975.
- [52] Knaze P, Beno P. The use of combined non-destructive testing methods to determine the compressive strength of concrete. *Materiaux et Construct* 1984;17(99):207–10.
- [53] De Almeida IR. Non-destructive testing of high-strength concretes: rebound (Schmidt hammer) and ultra-sonic pulse velocity. In: Lambotte H, Taerwe L, editors. Quality control of concrete structures, proceedings of the second international RILEM/CEB symposium; 1991. p. 387–97.
- [54] del Río LM, Jiménez A, López F, Rosa FJ, Rufo MM, Paniagua JM. Characterisation and hardening of concrete with ultrasonic testing. *Ultrasonics* 2004;42(1–9):527–30. <https://doi.org/10.1016/j.ultras.2004.01.053>.
- [55] Hobbs B, Tchoketch Kebir M. Non-destructive testing techniques for the forensic engineering investigation of reinforced concrete buildings. *Forensic Sci Int* 2007;167(2–3):167–72. <https://doi.org/10.1016/j.forsciint.2006.06.065>.
- [56] Baqer AHA. Assessment of concrete compressive strength by ultrasonic non-destructive test. 2008. Master's Thesis. University of Baghdad, Baghdad, Iraq.
- [57] Musmar MA, Abed Alhadi N. Relationship between ultrasonic pulse velocity and standard concrete cube crushing strength. *JES J Eng Sci* 2008;36(1):51–9. <https://doi.org/10.21608/jesaun.2008.115591>.
- [58] Domingo R, Hirose S. Correlation between concrete strength and combined non-destructive tests for concrete using high-early strength cement. <https://www.researchgate.net/publication/237412181>; 2009.
- [59] Monteiro AV, Gonçalves A. *Assessment of characteristic compressive strength in structures by the rebound hammer test according to EN 13791: 2007.* NDTCE'09, non-destructive testing in civil engineering. Nantes, France; 2009. <https://www.researchgate.net/publication/241751360>.
- [60] Na UJ, Park TW, Feng MQ, Chung L. Neuro-fuzzy application for concrete strength prediction using combined non-destructive tests. *Mag Concr Res* 2009;61(4):245–56. <https://doi.org/10.1680/jmagr.2007.00127>.
- [61] Trtnik G, Kavčić F, Turk G. Prediction of concrete strength using ultrasonic pulse velocity and artificial neural networks. *Ultrasonics* 2009;49(1):53–60. <https://doi.org/10.1016/j.ultras.2008.05.001>.
- [62] Bilgehan M, Turgut P. Artificial neural network approach to predict compressive strength of concrete through ultrasonic pulse velocity. *Res Nondestr Eval* 2010;21(1):1–17. <https://doi.org/10.1080/09349840903122042>.
- [63] Madandoust R, Ghavidel R, Nariman-Zadeh N. Evolutionary design of generalized GMDH-type neural network for prediction of concrete compressive strength using UPM. *Comput Mater Sci* 2010;49(3):556–67. <https://doi.org/10.1016/j.commatsci.2010.05.050>.
- [64] Kurtulus C, Bozkurt A. Determination of concrete compressive strength of the structures in Istanbul and Izmit Cities (Turkey) by combination of destructive and non-destructive methods. *Int J Phys Sci* 2011;6(16):4044–7.
- [65] Lawson I, Danso KA, Odoi HC, Adjei CA, Quashie FK, Mumuni II, Ibrahim IS. Non-destructive evaluation of concrete using ultrasonic pulse velocity. *Res J Appl Sci Eng Technol* 2011;3(6):499–504.
- [66] Agunwamba JC, Adagba T. A comparative analysis of the rebound hammer and ultrasonic pulse velocity in testing concrete. *Niger J Technol* 2012;31(1):31–9.
- [67] Gkotsamanis A, Zoidis N, Matikas TE, Tatsis E, Vlachopoulos C. *Concrete compressive strength estimation in-situ at precast beams using direct and indirect testing methods according to EN 13791.* Emerging technologies in non-destructive testing V. 2012. p. 137.
- [68] Hajjeh HR. Correlation between destructive and non-destructive strengths of concrete cubes using regression analysis. *Contemporary Eng Sci* 2012;5(10):493–509.
- [69] Hannachi S, Guetteche MN. Application of the combined method for evaluating the compressive strength of concrete on site. *Open J Civ Eng* 2012;2(1):16–21. <https://doi.org/10.4236/ojce.2012.21003>.
- [70] Yeşilmen S. Evaluation of rebound hammer test as a combined procedure used with drill core testing for evaluation of existing structures. In: Büyükoztürk, et al., editors. Nondestructive testing of materials and structures, 6. RILEM Bookseries; 2012. p. 341–6. https://doi.org/10.1007/978-94-007-0723-8_49.
- [71] Fawzi NM, Ali KJ, Said AI. Prediction of compressive strength of reinforced concrete structural elements by using combined non-destructive tests. *J Eng* 2013;19(10).
- [72] Jain A, Kathuria A, Kumar A, Verma Y, Murari K. Combined use of non-destructive tests for assessment of strength of concrete in structure. *Procedia Eng* 2013;54:241–51. <https://doi.org/10.1016/j.proeng.2013.03.022>. The 2nd International Conference on Rehabilitation and Maintenance in Civil Engineering.
- [73] Lal T, Sharma S, Naval S. Reliability of non-destructive tests for hardened concrete strength. *Int J Eng Res Technol* 2013;2(3):1–7.
- [74] Nguyen NT, Sbartaï ZM, Lataste JF, Breyse D, Bos F. Assessing the spatial variability of concrete structures using NDT techniques - laboratory tests and case study. *Constr Build Mater* 2013;49:240–50. <https://doi.org/10.1016/j.conbuildmat.2013.08.011>.
- [75] Belagraa L, Bouzid A, Logzit N, Badache N, Belguendouz A. Local concrete characterization assessment by the means of non-destructive tests (NDT) methods. Concrete solutions - proceedings of concrete solutions, 5th international conference on concrete repair. 2014. p. 761–9. <https://doi.org/10.1201/b17394-118>.
- [76] Brozovsky J. High-strength concrete-NDT with rebound hammer: influence of aggregate on test results. *Nondestr Test Eval* 2014;29(3):255–68. <https://doi.org/10.1080/10589759.2014.926897>.
- [77] Giannini R, Sguerri L, Paolacci F, Alessandri S. Assessment of concrete strength combining direct and NDT measures via Bayesian inference. *Eng Struct* 2014;64:68–77. <https://doi.org/10.1016/j.engstruct.2014.01.036>.

- [78] Holčapek O, Litoš J, Zatloukal J. Destructive and nondestructive characteristics of old concrete. *Adv Mater Res* 2014;1054:243–7. <https://doi.org/10.4028/www.scientific.net/AMR.1054.243>.
- [79] Jaggerwal H, Bajpai Y. Assessment of characteristic compressive strength in concrete Bridge girders using rebound hammer test. *Int J Comput Eng Res* 2014;4(4):70–5.
- [80] Osman MH, Alaraji WA, Saim AA, Nor W, Majid AWA. In-Situ strength of concrete using the correlation of different NDT test methods. *Int J Appl Eng Res* 2014;9(23):19767–80.
- [81] Pfister V, Tundo A, Luprano VAM. Evaluation of concrete strength by means of ultrasonic waves: a method for the selection of coring position. *Constr Build Mater* 2014;61:278–84. <https://doi.org/10.1016/j.conbuildmat.2014.03.017>.
- [82] Samson D, Moses T. Correlation between non-destructive testing (NDT) and Destructive Testing (DT) of compressive strength of concrete. *Int J Eng Sci Invent* 2014;3(9):12–7.
- [83] Sanchez K, Tarranza N. Reliability of rebound hammer test in concrete compressive strength estimation. *Int J Adv Agricult Environ Eng* 2015;1(2). <https://doi.org/10.15242/ijaaee.c1114040>.
- [84] Nikhil MV, Deep CS, Vijay GD, Vishal TS, Shweta P. The use of combined non-destructive testing in the concrete strength assessment from laboratory specimens and existing buildings. *Int J Curr Eng Sci Res* 2015;2(5):55–9.
- [85] Nobile L. Prediction of concrete compressive strength by combined non-destructive methods. *Meccanica* 2015;50(2):411–7. <https://doi.org/10.1007/s11012-014-9881-5>.
- [86] Patil H, Khairnar D, Thube R. Comparative study of effect of curing on compressive strength of concrete by using NDT & DT. *Int J Sci Adv Res Technol* 2015;1(6). <https://www.researchgate.net/publication/316855440>.
- [87] Abdullah BI, Abdulkadir MR. Correlation between destructive and non-destructive tests results for concrete compressive strength. *J Zankoi Sulaimani* 2016;18(4).
- [88] Bhosale N, Salunkhe PA. To establish relation between destructive and non-destructive tests on concrete. *Int J Eng Res Gen Sci* 2016;4(2).
- [89] El-Mir A, Nehme SG. A comparative study on ultrasonic pulse velocity for normally vibrated and self-compacting concretes. *Concrete Struct* 2016;8–12. <https://www.researchgate.net/publication/310460942>.
- [90] Gehlot T, Sankhla SS, Gupta A. Study of concrete quality assessment of structural elements using rebound hammer test. *Am J Eng Res* 2016;5:192–8.
- [91] Gehlot T, Sankhla Dr SS, Gehlot Dr SS, Gupta A. Study of concrete quality assessment of structural elements using ultrasonic pulse velocity test. *IOSR J Mech Civ Eng* 2016;13(5):15–22. <https://doi.org/10.9790/1684-1305071522>.
- [92] Lin YC, Lin Y, Chan CC. Use of ultrasonic pulse velocity to estimate strength of concrete at various ages. *Mag Concr Res* 2016;68(14):739–49. <https://doi.org/10.1680/jmacr.15.00025>.
- [93] Lopes YD, Vanalli L, Ferrari VJ. Concrete compressive strength estimation by means of non-destructive testing: a case study. *Open J Civ Eng* 2016;6(4):503–15. <https://doi.org/10.4236/ojce.2016.64043>.
- [94] Masi A, Chiauzzi L, Manfredi V. Criteria for identifying concrete homogeneous areas for the estimation of in-situ strength in RC buildings. *Constr Build Mater* 2016;121:576–87. <https://doi.org/10.1016/j.conbuildmat.2016.06.017>.
- [95] Saleem MA, Siddiqi ZA, Aziz M, Abbas S. Ultrasonic pulse velocity and rebound hammer testing for non-destructive evaluation of existing concrete structure. *Pak J Eng Appl Sci* 2016;18:89–97.
- [96] El Mir A, Nehme SG. Repeatability of the rebound surface hardness of concrete with alteration of concrete parameters. *Constr Build Mater* 2017;131:317–26. <https://doi.org/10.1016/j.conbuildmat.2016.11.085>.
- [97] Ju M, Park K, Oh H. Estimation of compressive strength of high strength concrete using non-destructive technique and concrete core strength. *Appl Sci* 2017;7(12). <https://doi.org/10.3390/app7121249>.
- [98] Najim KB. Strength evaluation of concrete structures using IsonReb linear regression models: laboratory and site (case studies) validation. *Constr Build Mater* 2017;149:639–47. <https://doi.org/10.1016/j.conbuildmat.2017.04.162>.
- [99] Rashid K, Waqas R. Compressive strength evaluation by non-destructive techniques: an automated approach in construction industry. *J Build Eng* 2017;12:147–54. <https://doi.org/10.1016/j.jobbe.2017.05.010>.
- [100] Sabbag N, Uyanik O. Prediction of reinforced concrete strength by ultrasonic velocities. *J Appl Geophys* 2017;141:13–23. <https://doi.org/10.1016/j.jappgeo.2017.04.005>.
- [101] Djamila B, Mohamed G. The use of non-destructive tests to estimate self-compacting concrete compressive strength. *MATEC Web Conf* 2018;149:01036. <https://doi.org/10.1051/mateconf/201814901036>.
- [102] Kurtuluş C, Çiçek S, İrmak TS. Experimental study on compressive strength, ultrasonic pulse velocity and water content of concrete at early ages after A 28 day curing period. *Eastern Anatolian J Sci* 2018;31–9. VI.
- [103] Ali MS, Kakkure GG. Comparison between destructive and non-destructive test on concrete. *Int Res J Eng Technol* 2019;6(10):264–80.
- [104] Atoyebi OD, Ayanrinde OP, Oluwafemi J. Reliability comparison of schmidt rebound hammer as a non-destructive test with compressive strength tests for different concrete mix. *International conference on engineering for sustainable world*. 2019. <https://doi.org/10.1088/1742-6596/1378/3/032096>. Ota, Nigeria. 1378(3).
- [105] Kocáb D, Misák P, Cikrle P. Characteristic curve and its use in determining the compressive strength of concrete by the rebound hammer test. *Materials* 2019;12(7). <https://doi.org/10.3390/ma12172705>.
- [106] Lasisi A, Sadiq O, Balogun I, Tunde-Lawal A, Attah-Okine N. A boosted tree machine learning alternative to predictive evaluation of nondestructive concrete compressive strength. *Proceedings - 18th IEEE international conference on machine learning and applications, ICMLA 2019*. 2019. p. 321–4. <https://doi.org/10.1109/ICMLA.2019.00060>.
- [107] Pal P. Dynamic poisson's ratio and modulus of elasticity of Pozzolana Portland cement concrete. *Int J Eng Technol Innov* 2019;9(2):131–44.
- [108] Bamogo S, Toguyeni DYK, Zoma F, Yerbanga M. In situ measurement of the compressive strength of local concrete: correlation between non-destructive and destructive tests. *Phys Sci Int J* 2020;1–10. <https://doi.org/10.9734/psij/2020/v24i830205>.
- [109] Chandak NR, Kumavat HR. *SonReb method for evaluation of compressive strength of concrete*. ICEMEM 2019 proceedings. IOP Conf Ser Mater Sci Eng 2020;810(1). <https://doi.org/10.1088/1757-899X/810/1/012071>.
- [110] Fodil N, Chemrouk M. Relevance of the ultrasonic pulse velocity test for strength assessment of high strength concretes. *IOP Conf Ser Mater Sci Eng* 2020;960(3). <https://doi.org/10.1088/1757-899X/960/3/032011>.
- [111] Jedidi M. Evaluation of the quality of concrete structures by the rebound hammer method. *Current Trends Civil Struct Eng* 2020;5(5). <https://doi.org/10.33552/ctcse.2020.05.000621>.
- [112] Karahan Ş, Büyüksaraç A, Işık E. The relationship between concrete strengths obtained by destructive and non-destructive methods. *Iranian J Sci Technol Transact Civil Eng* 2020;44(1):91–105. <https://doi.org/10.1007/s40996-019-00334-3>.
- [113] Mendes SES, Oliveira RLN, Cremonese C, Pereira E, Pereira E, Medeiros-Junior RA. Mixture design of concrete using ultrasonic pulse velocity. *Int J Civ Eng* 2020;18(1):113–22. <https://doi.org/10.1007/s40999-019-00464-9>.
- [114] Onyekwa FC. A comparative analysis of the rebound hammer and pullout as non-destructive method in testing concrete. *Eur J Eng Res Sci* 2020;5(5):554–8. <https://doi.org/10.24018/ejers.2020.5.5.1903>.
- [115] Rahim MA, Shahidan S, Onn LC, Saiful Bahari NAA, Rahman NA, Ayob A. The behavior of non-destructive test for different grade of concrete. *Int J Integrated Eng* 2020;12(9):1–8. <https://doi.org/10.30880/ijie.2020.12.09.001>.
- [116] Sethy SK, Kishore MV, Garg V, Raja, Kumar V. Comparison of compressive strength of hardened concrete using schmidt rebound hammer and conventional testing method. In: Khan, et al., editors. *Advances in industrial safety: proceedings of HSFEA 2018*. Springer; 2020. p. 203–11. https://doi.org/10.1007/978-981-15-6852-7_17.
- [117] Shaikh MF, Pathak R, Pandey A. Non-destructive assessment of structural health of university building using rebound hammer. *J Civil Construct Eng* 2020;6(1):11–8.
- [118] Wang YR, Lu YL, Chiang DL. Adapting artificial intelligence to improve in situ concrete compressive strength estimations in rebound hammer tests. *Front Mater* 2020;7. <https://doi.org/10.3389/fmats.2020.568870>.
- [119] Mohammed AA, Rafiq SK, Hamid NA. The assessment of concrete subjected to preloading using non destructive testing methods. *Case Stud Constr Mater* 2021;15. <https://doi.org/10.1016/j.cscm.2021.e00705>.
- [120] Biswas R, Rai B, Samui P. Compressive strength prediction model of high-strength concrete with silica fume by destructive and non-destructive technique. *Innov Infrastruct Sol* 2021;6(2). <https://doi.org/10.1007/s41062-020-00447-z>.
- [121] Fodil N, Chemrouk M, Ammar A. Influence of steel reinforcement on ultrasonic pulse velocity as a non-destructive evaluation of high-performance concrete strength. *Eur J Environ Civil Eng* 2021;25(2):281–301. <https://doi.org/10.1080/19648189.2018.1528890>.
- [122] Kumavat HR, Chandak NR, Patil IT. Factors influencing the performance of rebound hammer used for non-destructive testing of concrete members: a review. *Case Stud Constr Mater* 2021;14. <https://doi.org/10.1016/j.cscm.2021.e00491>.
- [123] Ofuyatan O, Olowofoyeku A, Oluwafemi J, Ighalo J. Predicting the compressive strength of concrete by ultrasonic pulse velocity. *ICSID proceedings, 2020. IOP Conf Ser Mater Sci Eng* 2021;1036(1). <https://doi.org/10.1088/1757-899X/1036/1/012053>.
- [124] Paktiawala A, Alam M. A study of non-destructive test (NDT) results on old concrete specimens under loaded and unloaded conditions and their crushing strength. In: Emamian, et al., editors. *Icscdems, 2019*. Singapore: Springer Nature; 2021. p. 67–82.
- [125] Zárate DM, Cárdenas F, Forero EF, Peña FO. Strength of concrete through ultrasonic pulse velocity and uniaxial compressive strength. *Int J Technol* 2022;13(1):103–14. <https://doi.org/10.14716/ijtech.v13i1.4819>.
- [126] Ali-Benyahia K, Kenai S, Ghrici M, Sbartaï ZM, Elachachi SM. Analysis of the accuracy of in-situ concrete characteristic compressive strength assessment in real structures using destructive and non-destructive testing methods. *Constr Build Mater* 2023;366. <https://doi.org/10.1016/j.conbuildmat.2022.130161>.
- [127] Bensaber A, Boudaoud Z, Seghir NT, Czarnecki S, Sadowski Ł. The assessment of concrete subjected to compressive and flexural preloading using nondestructive testing methods, correlation between concrete strength and combined method (SonReb). *Measurement: J Int Measure Confeder* 2023;222. <https://doi.org/10.1016/j.measurement.2023.113659>.
- [128] Chen J, Jin Q, Dong B, Dong C. Research on the rebound hammer testing of high-strength concrete's compressive strength in the Xinjiang region. *Buildings* 2023;13(2). <https://doi.org/10.3390/buildings13122905>.
- [129] Gunes B, Karatosun S, Gunes O. Drilling resistance testing combined with SonReb methods for nondestructive estimation of concrete strength. *Constr Build Mater* 2023;362. <https://doi.org/10.1016/j.conbuildmat.2022.129700>.
- [130] Hidayat I, Suangga M, Suwondo R, Hamiah A, Kosalim V. Correlation of normal concrete compressive strength with destructive and non-destructive test methods. *AIP Conf Proc* 2023;2846(1). <https://doi.org/10.1063/5.0154337>.
- [131] Mata R, Ruiz RO, Nuñez E. Correlation between compressive strength of concrete and ultrasonic pulse velocity: a case of study and a new correlation method.

- Constr Build Mater 2023;369. <https://doi.org/10.1016/j.conbuildmat.2023.130569>.
- [132] Shrestha B, Giri OP. Study on concrete compressive strength through destructive and non-destructive testing. *Int J Sci Math Technol Learn* 2023;31(2):336–56. <https://doi.org/10.5281/zenodo.8366195>.
- [133] Abazarsa M, Raisi K, Yu T. *Estimation of compressive strength of Portland cement concrete using synthetic aperture radar, ultrasonic testing, and rebound hammer*. In non-destructive characterization and monitoring of advanced materials, aerospace, civil infrastructure, and transportation XVIII (p. 8). *Proc SPIE* 2024. <https://doi.org/10.1117/12.3010516>.
- [134] Alavi SA, Noel M, Moradi F, Layssi H. Development of a machine learning model for on-site evaluation of concrete compressive strength by SonReb. *J Build Eng* 2024;82. <https://doi.org/10.1016/j.jobe.2023.108328>.
- [135] Chouhan A, Verma SK. Condition evaluation of concrete through ultrasonic pulse velocity. *Lect Notes Civil Eng* 2024;352:179–89. https://doi.org/10.1007/978-981-99-2676-3_15.
- [136] Khoeri H, Isvara W, Pradana R, Sofiana D. The influence of building age on the formulation of concrete compressive strength using non-destructive testing. *Clean Energy Smart Technol* 2024;2(2):71–80. <https://doi.org/10.58641/cleanenergy.2024.11.1.041>.
- [137] Kim W, Jeong K, Lee T. Statistical reliability analysis of ultrasonic velocity method for predicting residual strength of high-strength concrete under high-temperature conditions. *Materials* 2024;17(6). <https://doi.org/10.3390/ma17061406>.
- [138] Mangasi YJU, Kirana NK, Sjah J, Handika N, Vincens E. Correlation of rebound hammer and ultrasonic pulse velocity methods for instant and additive-enhanced concrete. *Struct Monitor Mainten* 2024;11(1):41–55. <https://doi.org/10.12989/smm.2024.11.1.041>.
- [139] Mulira AS, Kiguli AV, Rucukye A, Matovu M. Comparative analysis of concrete compressive strength using a destructive compression method and non-destructive rebound hammer. *J Civil Eng Mater App* 2024;8(2):71–81. <https://doi.org/10.22034/jcema.2024.466048.1137>.
- [140] Räsänen A, Lahdensivu J, Gudmundsson K, Dervishaj A, Westerlind H, Lambrechts T, Vullings M, Arnold V, Huuhka S. Properties and quality of precast concrete elements deconstructed in ReCreate's pilots. *The ReCreate project*; 2024.
- [141] Singh S, Alhussainy AK, Panduri B, Rajalakshmi B, Gupta M, Singh H, Reddy GC. Compressive strength prediction model of high strength concrete by destructive and nondestructive technique. *ICMPC - E3S Web Conf* 2024;552. <https://doi.org/10.1051/e3sconf/202455201111>.
- [142] Kairu WM, Mumenya SW, Njoroge KD, Kaniu IM. Application of schmidt hammer and ultrasonic pulse velocity for structural integrity assessment in water dams. *J Infrastruct Preserv Resil* 2025;6(1). <https://doi.org/10.1186/s43065-025-00123-5>.
- [143] Bellavista P, Berrocal J, Corradi A, Das SK, Foschini L, Zanni A. A survey on fog computing for the internet of things. *Pervasive Mob Comput* 2019;52:71–99. <https://doi.org/10.1016/j.pmcj.2018.12.007>. Elsevier B.V.
- [144] EN 13791. Assessment of in-situ compressive strength in structures and precast concrete components. *Eur Committee Standardis (CEN) Brussels Belgium* 2019. 21–24.
- [145] Debroy S, Sil A. Empirical models for non-destructive evaluation of in-situ compressive strength of structural concrete with the objective of encompassing global uncertainty factors. *Asian J Civil Eng* 2024. <https://doi.org/10.21203/rs.3.rs-4813626/v1> [Pre-Print].
- [146] Turgut P. Research into the correlation between concrete strength and UPV value. *NDT net* 2004;12(12).
- [147] Al-Nu'man BS, Aziz BR, Abdulla SA, Khaleel SE. Compressive strength formula for concrete using ultrasonic pulse velocity. *Article in Int J Eng Trends Technol* 2015; 26(1). <https://doi.org/10.14445/22315381/IJETT-V26P203>.
- [148] Kumavat HR, Chandak NR. Statistical analysis for evaluating concrete strength of existing structure using non-destructive and destructive test. *Innov Infrastruct Sol* 2024;9(5). <https://doi.org/10.1007/s41062-024-01490-w>.
- [149] Logothetis LA. *Combination of three non-destructive methods for the determination of the strength of concrete*. Athens, Greece: National Technical University of Athens; 1978. Ph.D. thesis.
- [150] Trezos C, Papakyriakopoulos P, Spanos C. Calibration of the rebound hammer and pulse velocity methods through in situ concrete cores and standard cube specimens. 1993. Technical Chamber of Greece. [in Greek].
- [151] Qasrawi HY. Concrete strength by combined non-destructive methods simply and reliably predicted. *Cement Concr Res* 2000;30:739–46.
- [152] Pucinotti R, de Lorenzo RA. In situ non-destructive testing: the steel and concrete resistance assessment of “ancient” r/c structures. *Ninth Int Conf Struct Studies Repair Mainten Heritage Architect* 2005;83:355–64.
- [153] Arioglu E, Manzak O. Application of ‘Sonreb’ method to concrete samples produced in Yedpa construction site. *Prefabrication Union*; 1991. p. 5–12 [in Turkish].
- [154] Faella G, Guadagnuolo M, Donadio A, Ferri L. Calibrazione sperimentale del metodo SonReb per costruzioni della Provincia di Caserta degli anni “60-80” [Experimental calibration of the SonReb method for buildings in the Province of Caserta from the 1960s to the 1980s]. *Proceedings of the 14th ANDIS conference*. 2011 [in Italian].
- [155] Tanigawa Y, Baba K, Mori H. Estimation of concrete strength by combined non-destructive testing method, in situ non-destructive testing of concrete. *Detroit: American Concrete Institute*; 1984. p. 195–9.
- [156] Kromrey JD, Hogarty KY. Analysis options for testing group differences on ordered categorical variables: an empirical investigation of type I error control and statistical power. *Multiple Linear Regression Viewpoints* 1998;25(1):70–82.
- [157] Judd CM, McClelland GH, Ryan CS. *Data analysis: a model comparison approach to regression, ANOVA, and beyond*. New York, NY: Routledge; 2017.
- [158] Kutner MH, Nachtsheim CJ, Neter J, Wasserman W. *Applied linear regression models*, 4. New York. Pg: McGraw-Hill/Irwin; 2004. p. 289.
- [159] Matthews B, Palermo A, Logan T, Scott A. Experimental testing and predictive machine learning to determine the mechanical characteristics of corroded reinforcing steel. *Constr Build Mater* 2024;438:137023. <https://doi.org/10.1016/j.conbuildmat.2024.137023>.
- [160] Géron A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. first ed. ed. Sebastopol, CA: O'Reilly Media, Inc.; 2022.
- [161] Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. *J Big Data* 2020;7(1). <https://doi.org/10.1186/s40537-020-00369-8>.
- [162] Guryanov A. *Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees*. Analysis of images, social networks and texts. 8th Int Conf 2019;11832:39–50. <https://doi.org/10.1007/978-3-030-37334-4>.
- [163] Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res* 2011;20(1):40–9. <https://doi.org/10.1002/mpr.329>.
- [164] Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for hyper-parameter optimization. *Advances in neural processing systems*, 24. Spain: (NIPS). Grenada; 2011.
- [165] Watanabe S. *Tree-structured parzen estimator: understanding its algorithm components and their roles for better empirical performance*. arXiv preprint: 2304.11127. 2023 <http://arxiv.org/abs/2304.11127>.