Cite this article
Nguyen D-B, Liao W-C and Hoang V-L (2026)
Evaluation of prediction model and machine learning methods on the shrinkage of self-compacting concrete.
Magazine of Concrete Research **78(1–2)**: 1–17,
https://doi.org/10.1680/jmacr.25.00254

Research Article
Paper 2500254
Received 25/06/2025; Accepted 03/10/2025

Magazine of Concrete Research

emerald PUBLISHING ice Publishing

Evaluation of prediction model and machine learning methods on the shrinkage of self-compacting concrete

Doan-Binh Nguyen
Department of Building Materials Technology, Faculty of Building Materials, Hanoi University of Civil Engineering, Hanoi, Vietnam

Wen-Cheng Liao
Department of Civil Engineering, College of Engineering, National Taiwan University, Taipei, Taiwan

Vinh-Long Hoang
Department of Building Materials Technology, Faculty of Building Materials, Hanoi University of Civil Engineering, Hanoi, Vietnam (corresponding author: longhv@huce.edu.vn)

Shrinkage strain directly reduces the concrete volume and significantly affects the long-term durability of structures, particularly in self-compacting concrete (SCC). However, most existing models are not sufficiently accurate for predicting SCC shrinkage. To address this, a new SCC prediction model, called *fib*-SCC, was developed by calibrating the *fib* Model Code 2010 (MC2010) using an expanded laboratory database, comprising 1328 datasets for total shrinkage and 282 datasets for autogenous shrinkage, collected from a wide range of published studies. The *fib*-SCC model was evaluated through targeted SCC experiments based on varying fly ash contents (20–55% by weight of cement). Additionally, new prediction methods for SCC shrinkage behaviour are proposed using machine learning (ML) algorithms, including decision trees, the random forest, extra trees, gradient boosting (GB) and extreme gradient boosting (XGBoost). The results showed that the *fib*-SCC model had the highest accuracy among traditional models and other SCC prediction models. Nevertheless, the *fib*-SCC model was outperformed by the ML approaches, with GB and XGBoost demonstrating the highest accuracy in predicting total shrinkage.

Keywords: machine learning (ML)/non-destructive testing/plain concrete/prediction models/self-compacting concrete (SCC)/shrinkage prediction

Notation

A_c	cross-sectional area of concrete specimen
C	number of regression trees
cm	cementitious material in the prediction formulae
F	space of regression tree
$F(x)$	gradient boosting (GB) model
f_{cm28}	28-day compressive strength
$f_{cm28,T}$	temperature-adjusted 28-day compressive strength
f_j	fraction of value xj
f_k	independent tree structure and leaf weights
$\hat{f}^K_{RF}(x)$	linear random forest (RF) model
h	notional size
h_T	temperature-adjusted notional size
$h(x)$	base learner or weak learner
K	number of classes
k	number of trees
k_h	humidity factor for shrinkage
$k_{h,T}$	temperature-adjusted humidity factor for shrinkage
$L(y, F(x))$	loss function iteration
l	differentiable convex loss function
p_i	fraction or probability of samples belonging to class i
q	tree structure
S	entropy
T	temperature (°C)
T_{cur}	curing temperature
T_l	number of leaves
t	current age of concrete
t_c	age at exposure
U	cross-section perimeter of concrete specimen
w	leaf weight
x	sample
y_i	target value
$\tilde{y}_i$	pseudo-response
$\hat{y}_i$	predicted value
$\beta_{as}(t)$	time function of autogenous shrinkage
$\beta_s(t-t_c)$	time function for drying shrinkage
γ, λ	constants to control the degree of regularisation
$\varepsilon_{au}(t,t_c)$	autogenous shrinkage strain at t days
$\varepsilon_{au}(t_c)$	autogenous shrinkage strain at t_c days
$\varepsilon_{au0}(f_{cm28})$	notional autogenous shrinkage coefficient
$\varepsilon_{sh}(t,t_c)$	drying shrinkage strain
ε_{shu}	ultimate drying shrinkage strain
ρ	learning rate
$\Omega(A)$	set of categorical values of attribute A

Introduction

Shrinkage is the reduction in concrete volume owing to moisture loss in the absence of external stress. When concrete is exposed to ambient temperature and humidity, shrinkage strain can lead to internal cracking, which negatively affects dimensional stability and long-term durability (Abdalhmid *et al*., 2019; Mehta and Monteiro, 2006). This effect is particularly pronounced in mixtures with a high content of cementitious material and a low content of coarse aggregate – characteristics typical of self-compacting concrete (SCC) (Güneyisi *et al*., 2010; Leemann and Hoffmann, 2005; Leemann *et al*., 2011; Rozière *et al*., 2007). To address these issues, several models have been developed (Aashto, 2017; ACI, 1992, 2008; Baweja and Bažant, 1995; Bažant and Wan-Wendner, 2015; CEB, 1991; *fib*, 2013; Gardner and Lockman, 2001; JSCE, 2007). Additionally, several models have been proposed to predict the shrinkage of SCC, such as the modified JSCE model (Aslani and Maia, 2013), the modified CEB90 model (Poppe and Schutter, 2005) and the modified Aashto model for the shrinkage of steam-cured SCC (Khayat and Long, 2010). However, these SCC-specific models still exhibit limited accuracy in predicting shrinkage, largely owing to the constraints and variability of the experimental data on which they are based.

Estimating shrinkage behaviour using machine learning (ML) offers significant advantages for long-term predictions. As a result, several ML methods have been applied in studies of concrete structure deformation. ML involves enabling computers to modify or adapt their actions to become more accurate over time (Kristiawan and Aditya, 2015). In other words, ML allows software applications to improve predictive accuracy without being explicitly programmed, relying instead on the design of efficient and accurate prediction algorithms. These learning techniques are data-driven, combining fundamental concepts from computer science with ideas from statistics, probability and optimisation (Mohri *et al*., 2018). ML methods can be broadly categorised into supervised learning, unsupervised learning and reinforcement learning (Marsland, 2015; Russell, 2018). For regression tasks – a supervised learning technique – various models have been applied to predict the shrinkage of SCC, including decision trees (DTs), the random forest (RF), extra trees (ET), gradient boosting (GB) and extreme gradient boosting (XGBoost).

- DT algorithms include ID3 and C4.5, introduced by Quinlan (1986, 1993), and Cart (classification and regression tree) (Marsland, 2015). ID3 (Iterative Dichotomiser 3), developed from Hunt's concept learning system framework, was the predecessor to C4.5, which generates a classifier in the form of a DT algorithm.
- The RF, originally developed by Ho (1995) and significantly extended by Breiman (2001), is an ensemble tree-based method developed to address the limitations of Cart (Ahmad *et al*., 2018).
- The ET algorithm, reported by Geurts *et al*. (2006), is a relatively recent ML technique designed as an extension of the RF, offering a lower risk of overfitting (Ahmad *et al*., 2018).
- GB, developed by Friedman (2001), is another ensemble ML technique employed for regression and classification problems based on a boosting system.
- The XGBoost algorithm, proposed by Chen and Guestrin (2016), describes a scalable, end-to-end tree boosting system.

In recent years, a few studies have applied ML algorithms to model the time-dependent deformation behaviour of concrete. Bal and Buyle-Bodin (2013) used artificial neural networks to predict shrinkage, while Karthikeyan *et al*. (2008) used the same algorithm for modelling the time-dependent deformation of high-performance concrete.

In this work, a new SCC prediction model – named *fib*-SCC – was developed by calibrating the *fib* Model Code for Concrete Structures (MC2010) model (*fib*, 2013). The objective was to create a model that accurately fitted an expanded SCC laboratory database with minimum error. The performance of the *fib*-SCC model was compared against established models, including the *fib* MC2010 model (*fib*, 2013), the B4 model (Bažant and Wan-Wendner, 2015), the ACI model (ACI, 1992) and several other SCC-specific models. The F-SCC model was further validated through experimental data involving mixtures with varying fly ash (FA) contents (20–55% by weight of cement). Furthermore, the efficiency of the *fib*-SCC model was also compared with the previously listed ML techniques.

ML algorithms

DTs

DTs are one of the most important methods of supervised learning and are used widely for classification, which uses attribute values to partition the decision space into smaller subspaces in an iterative manner. A DT has a flowchart-like tree structure, starting with a root node and gradually growing to a final classification (or leaf) (Maimon and Rokach, 2005; Yang, 2019). The DT can be drawn in different ways, but they are almost all variants of the same principle – algorithms build the tree in a greedy manner, starting at the root and choosing the most informative feature at each step.

The split information of an attribute is measured by the entropy (S) as follows:

$$S = -\sum_{i=1}^{K} p_i \log_2 p_i \quad (1)$$

The Gini index (GI) can be expressed as:

$$GI = \sum_{i=1}^{K} p_i(1 - p_i) \quad (2)$$

The information gain is defined as:

$$\mathrm{IG}(A) = S - \sum\nolimits_{x_j \in \Omega(A)} f_j S_{x_j} \quad (3)$$

In Equations 1–3, p_i is the fraction or probability of samples belonging to class i, K is the number of classes, $\Omega(A)$ is the set of categorical values of attribute A, f_j is the fraction of value $x_j \in A$ with $j = 1, 2, \ldots, K$ and S_{xj} is entropy, which has a similar to formula to the entropy S (Yang, 2019).

RF

The RF adds a randomness algorithm to the model while growing the trees. The 'random' part of the name RF refers to the fact that each tree uses a random portion of the original data as training. Instead of searching for the most important feature based on node splitting, the method creates the best feature among a random subset of features for viable split points at each node in the trees produced (Watt *et al.*, 2020).

The RF regression can be presented as (Ahmad *et al.*, 2018; Ao *et al.*, 2019):

$$\hat{f}^{C}_{\mathrm{RF}}(x) = \frac{1}{C}\sum_{i=1}^{C} T_i(x) \tag{4}$$

where $\hat{f}^{C}_{\mathrm{RF}}(x)$ is the linear RF model, x is the sample and C is the number of regression trees.

ET

The ET algorithm builds totally randomised trees whose structures are independent of the output values of the learning sample (Geurts *et al.*, 2006) and chooses optimal cut-points, which are responsible for a large proportion of the variance of the induced tree (Berrouachedi *et al.*, 2019). The ET method is quite similar to a RF; it only differs in constructing the DT.

GB

GB was derived by Friedman (2001) and is an ensemble ML technique employed for regression and classification problems based on a boosting system. The principal boosting algorithm is AdaBoost (Marsland, 2015), which was first proposed by Freund and Schapire (1996), who reported that boosting can significantly reduce the error of any 'weak learner' algorithms that consistently generate classifiers that are only slightly better than random predictions. When a weak learner can be implemented efficiently, boosting provides a tool for aggregating weak hypotheses to approximate a gradually better prediction for larger information (Shalev-Shwartz and Ben-David, 2014). Therefore, GB can produce a competitive, highly robust and interpretable statistical process in many areas of ML. For a regression problem, the GB algorithm can be improved as follows (Friedman, 2001).

$$F_m(x) = F_{m-1}(x) + \rho_m h_m(x; a_m) \tag{5}$$

$$\rho_m = \mathrm{argmin}_{\rho} \sum_{i=1}^{N} L\left[y_i, F_{m-1}(x_i)\right] + \rho_h(x_i; a_m) \tag{6}$$

$$a_m = \mathrm{argmin}_{a,\beta} \sum_{i=1}^{N} \left[\tilde{y}_i - \beta h(x_i; a)\right]^2 \tag{7}$$

$$\tilde{y}_i = -\left(\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right)_{F(x)=F_{m-1}(x)} \tag{8}$$

$$F_0(x) = \mathrm{argmin}_{\rho} \sum_{i=1}^{N} L(y_i, \rho) \tag{9}$$

In these equations, $F(x)$ is the GB model, m is an iteration, ρ is the learning rate, $h(x)$ is a base learner or weak learner, $L(y, F(x))$ is the loss function and $\tilde{y}_i$ is the pseudo-response ($i = N$).

XGBoost

The XGBoost algorithm, proposed by Chen and Guestrin (2016), describes a scalable end-to-end tree boosting system. The method is similar to GB, as it builds an additive expansion of the objective function by minimising the loss function. However, XGBoost, which is known as one of the best-performing algorithms for supervised learning, focuses only on DTs as base classifiers, and a variation of the loss function is used to control the complexity of the trees (Bentéjac *et al.*, 2021).

The output function of the XGBoost algorithm can be expressed as:

$$\hat{y}_i = \varphi(x_i) = \sum_{k=1}^{K} f_k(x_i) \quad (f_k \in F) \tag{10}$$

$$F = \left\{f(x) = w_{q(x)}\right\} \quad \left(q : R^m \to T, w \in R^T\right) \tag{11}$$

The loss function is calculated as:

$$L(\varphi) = \sum\nolimits_i l(\hat{y}_i, y_i) + \sum\nolimits_k \Omega(f_k) \tag{12}$$

Regularisation of the model to avoid overfitting is expressed as:

$$\Omega(f) = \gamma T_1 + \frac{1}{2}\lambda \|w\|^2 \tag{13}$$

where $\hat{y}_i$ is the predicted value; y_i is the target value; k is the number of trees; f_k corresponds to an independent tree structure and leaf weights; q is the tree structure; w is the leaf weight; T_l is the number of leaves; F is the space of regression tree; l is a differentiable convex loss function, and γ and λ are constants to control the degree of regularisation (Chen and Guestrin, 2016).

Experimental study

Materials and mix proportions

The materials used included Portland cement (PC), FA, fine aggregate (river sand), coarse aggregate (limestone) and superplasticiser (SP), all supplied by local companies in Vietnam. The PC was blended with FA to form a cementitious material. The cement (PC40) conformed to type I (ASTM, 2019a), while the FA (obtained from a coal-fired thermal plant) satisfied the requirements for type F (ASTM, 2019b). The river sand and limestone (used as fine and coarse aggregate) met the requirements of ASTM C1611 (ASTM, 2018a). The main physical properties of the aggregates are summarised in Table 1. The liquid SP was Sika Viscocrete SKY 8713, which was used as a high-range water-reducing agent with a specific gravity of 1.15 g/cm^3.

In *fib* MC2010 (*fib*, 2013), the 28-day compressive strength is considered a key parameter for predicting concrete shrinkage. To investigate the influence of compressive strength on shrinkage behaviour, the mix proportions of the SCC were designed to achieve various strength grades based on the method proposed by Okamura and Ouchi (2003). The target strength grade was varied from 30 MPa to 70 MPa, encompassing both normal- and high-strength concretes (ACI, 1998). All the mixtures incorporated FA as a mineral admixture, ranging from 20% to 55% by weight of the cementitious material. Five mixes were developed, with the mix proportions presented in Table 2.

Experimental methods

The properties of the fresh SCC mixes were evaluated based on ASTM C33 (ASTM, 2018b) using the slump flow and the flow time at 500 mm spread (T_{50}). Concrete specimens were cast without rodding or tapping the sides of the moulds (ASTM, 2015). A plastic sheet was placed over the top surface of all specimens to minimise moisture loss owing to evaporation.

The compressive strengths were evaluated by testing 150 mm × 150 mm × 150 mm cube specimens according to TCVN 3118-2022 (VSQI, 2022). The specimens were demoulded after 24 h and then cured in a moisture room at a temperature of 23 ± 2°C and relative humidity (RH) above 95% until the designated test age.

Both autogenous shrinkage and total shrinkage strains were manually measured using a length comparator, satisfying ASTM C157 (ASTM, 2017). Shrinkage tests were conducted on prism specimens measuring 75 × 75 × 285 mm. After demoulding at 24 h, the specimens for autogenous shrinkage tests were sealed with plastic sheets to prevent water loss. The initial length of the specimens was measured at this point using a comparator. All specimens were then placed in a drying chamber maintained at a temperature of 23 ± 2°C and RH of 60 ± 4%. Shrinkage measurements were recorded at predefined intervals according to the designated schedule.

Experimental results

Workability and compressive strength

The slump flow values of all the mixes varied significantly, from 650 to 760 mm, while the flow time (T_{50}) was 4–8 s, as shown in Table 2. Based on visual inspection, there was no segregation of the SCCs in the fresh state. The increased flowability can be attributed to the spherical and glassy structure of the FA particles (El-Chabib and Syed, 2013; Khatib, 2008). The SP dosage also played a critical role in reducing the water-to-cementitious

Table 1. Physical properties of aggregate

Aggregate	Maximum size: mm	Specific density: g/cm^3	Bulk density: g/cm^3	Fineness modulus	Water absorption: %	Clay and dust content: %
Sand	5	2.65	1.57	3.01	0.65	2.09
Coarse aggregate	20	2.74	1.38	—	0.50	0.14

Table 2. Mix proportions and properties of SCCs

Mix	Mix proportion: kg/m^3						Slump flow: mm	T_{50}: s	Compressive strength: MPa	
	PC	FA	Sand	Limestone	Water	SP			28 days	180 days
M30	270	330	1020	415	184	1.1	720	4	31.5	43.9
M40	300	300	994	493	165	1.6	760	5	43.6	57.7
M50	360	240	882	644	196	2.3	730	6	53.4	64.8
M60	420	180	842	741	183	7.0	650	6	65.1	76.4
M70	480	120	805	839	167	7.8	730	8	75.9	85.6

materials (w/cm) ratio and ensuring the slump flow met the requirements for SCC.

The compressive strengths of the mixes are also provided in Table 2. The results indicate that FA enhanced the strength of the concrete, particularly at later ages, owing to its pozzolanic reactivity. Compared with the 28-day compressive strength, the 180-day compressive strength increased by 13% to 39%, depending on the FA content. These results demonstrate that a higher FA content contributes to greater long-term strength development in SCC. Similar findings were reported by Jawahar *et al*. (2013).

Shrinkage behaviour

As shown in Figure 1(a), the autogenous shrinkage of the specimens varied significantly depending on the FA content and the w/cm ratio. M30 mix, which had the highest FA content (55% by weight of cement) and a w/cm ratio of 0.31, exhibited the lowest autogenous shrinkage. In contrast, M70 mix (with the lowest FA content (20%) and a w/cm ratio of 0.28) exhibited the highest autogenous shrinkage. These results indicate that increasing the replacement level of FA reduces autogenous shrinkage, consistent with the findings of Chan *et al*. (1998). The observed increase in autogenous shrinkage may be attributed to the pozzolanic reactivity of both FA and cement at later ages. Consequently, the consumption of water in the micropores of the concrete continued through the hydration reaction of the cement (Tazawa, 1999). Additionally, the autogenous shrinkage values were reduced with a decrease in cement content, which was partially replaced by FA (Lee *et al*., 2003). Autogenous shrinkage also may be influenced by the presence of SP, as noted by Esping and Löfgren (2006).

Figure 1(b) shows the development of total shrinkage over the test period of up to 150 days. According to the experimental results, mixes M30 and M40 exhibited the lowest total shrinkage during the early stage (up to 28 days). M70 demonstrated the highest total shrinkage of all the mixes. These outcomes indicate that the combined effects of FA content, w/cm ratio and SP dosage significantly influence the total shrinkage of SCC. Şahmaran *et al*. (2007) also reported a decrease in the total shrinkage of SCC with an increase in FA.

Prediction model *fib*-SCC for the shrinkage of SCC

Shrinkage database of SCC

Total shrinkage and autogenous shrinkage are recorded in unsealed and sealed specimens, respectively. For sealed specimens, moisture evaporation is prevented, allowing shrinkage to occur primarily owing to the self-desiccation process within the pores of the hydrated cement. In contrast, the shrinkage in the unsealed specimens is mainly owing to water extraction from the hardened concrete, especially from voids smaller than 50 nm in diameter (Minless *et al*., 2002). The complied SCC shrinkage database included 1328 datasets for total shrinkage and 282 datasets for autogenous shrinkage. This comprehensive database was compiled from experimental results reported in numerous publications spanning from 1995 to early 2024. Each database was organised into two sheets: 'Info' and 'Data'. The former included information on each shrinkage dataset, structured in a format consistent with the Northwestern University database (Hubler *et al*., 2015a).

The parameter ranges of the expanded SCC shrinkage database were selected for normal-weight concrete without the use of expansive additives or shrinkage-reducing agents. These ranges of the parameters (28-day compressive strength (f_{cm28}), the cementitious material in the prediction formulae *cm*, w/cm ratio, aggregate-to-cementitious material ratio (a/cm), sand-to-aggregate ratio (s/a), test duration ($t - t_c$), volume-to-surface area ratio (V/S), RH and the age at the start of testing or exposure (t_c)) are listed in Table 3. Additionally, the table provides the median and mean values for each parameter. The distribution of these parameters is shown in Figure 2. The 28-day compressive strengths encompassed both normal- and high-strength concretes spanning a wide range (17–110 MPa) with a peak frequency around 50 MPa. The mean values of *cm*, w/cm, a/cm and s/a ratios were approximately 500 kg/m^3, 0.36, 3.4 and 0.5, respectively.

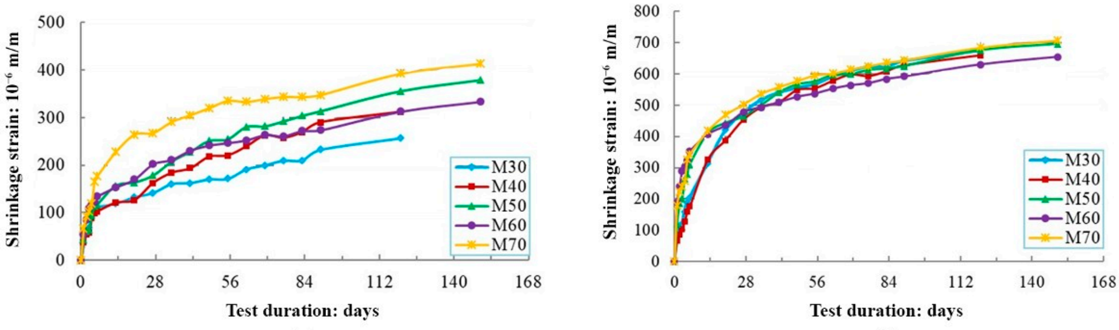

Figure 1. Shrinkage of SCC specimens: (a) autogenous shrinkage; (b) total shrinkage

Table 3. Parameter ranges of the SCC shrinkage database

	f_{cm28}: MPa	*cm*: kg/m^3	w/cm	a/cm	s/a	$t - t_c$: days	*V/S*: mm	RH	t_c: days
Minimum	17.0	275	0.18	1.60	0.28	14	11.4	0.27	0.14
Maximum	109.5	840	0.68	6.74	0.72	1,435	70	1.00	90
Median	55.6	500	0.35	3.34	0.50	112	22	0.50	1
Average	50.8	507	0.36	3.39	0.49	176	21	0.62	7

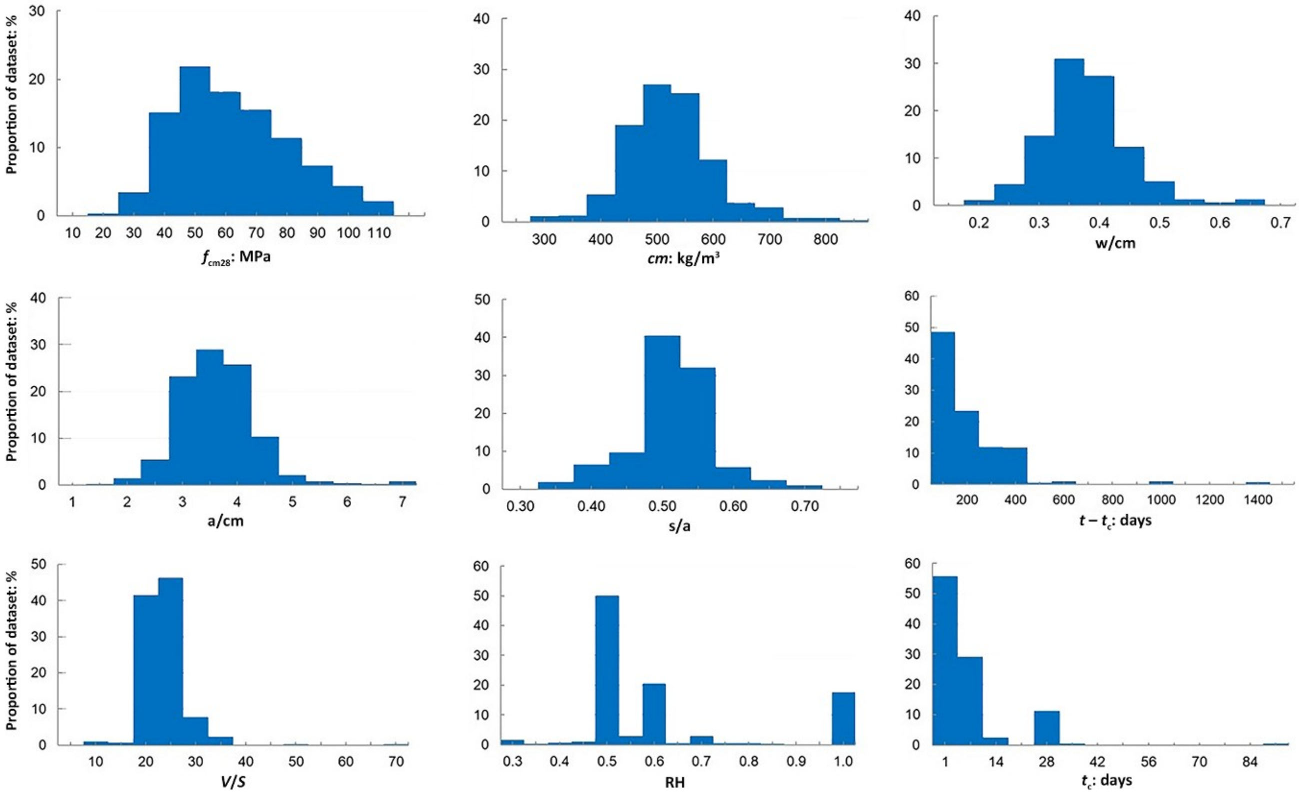

Figure 2. Distribution of shrinkage parameters

Supplementary cementitious materials are commonly incorporated into SCC. According to the collected data, approximately 79% of the SCC mixtures included fillers and mineral admixtures; pure cement was used in only 21% of the mixes. Among the fillers, limestone powder was the most frequently used, whereas slag and FA were the predominant additives, known for enhancing the properties of SCC. Another key parameter was the aggregate type, which varied in mineralogy and could be classified into diabase, granite, limestone, quartzite and sandstone.

The shrinkage strain was primarily measured in specimens produced with normal cement (type R), which accounted for approximately 76% of the SCC shrinkage database. In contrast, slow-hardening cement (type SL) and rapid-hardening cement (type RS) were used in only about, respectively, 15% and 9% of the samples in the database. Owing to the predominance of type R cement in the database, this study focused exclusively on optimising shrinkage prediction models using type R cement. On the other hand, the limited availability of data on cement types SL and RS suggests that further experimental studies are needed before these can be included in analytical or predictive modelling.

Optimum strategy

Statistical coefficients are commonly used to evaluate the performance of shrinkage prediction models. In this study, two primary metrics were employed – the determination coefficient (R^2) (Hami and Pougnet, 2020; Montgomery *et al*., 2012) and the normalised root mean square error (NRMSE) (Dietterich *et al*., 2017; Rizi and Granitzer, 2017). The optimisation process was based on minimising the prediction error according to these coefficients. However, statistical results can be significantly affected by the distribution of datasets. Typically, laboratory databases are dominated by short-term test results, small specimen sizes and early-age measurements. The percentage weights of these datasets was confirmed by the SCC database, as illustrated in Figure 2. Relying solely on unweighted error statistics may lead to biased model evaluation. To achieve unbiased data and error statistics, a weighting scheme was applied during model optimisation (Hubler *et al*., 2015b; Wendner *et al*., 2015). The weighing factors were based on key independent variables, including the test duration, environmental humidity, specimen size and age at the start of drying (Bažant and Li, 2007).

According to the SCC shrinkage database, data subsets were created to apply weightings, ensuring unbiased evaluation of statistical coefficients for a reliable model based on the independent variables of shrinkage type. Optimisation was first performed for dependent parameters on cement type. Then, an additional coefficient was introduced to improve the predictive accuracy of the original *fib* MC2010 model. In this study, the w/cm ratio – a key factor influencing shrinkage development – was selected to adjust the original model (Hubler *et al*., 2015a). The entire optimisation process was conducted sequentially for both autogenous shrinkage and drying shrinkage formulas.

Shrinkage formulas of the *fib*-SCC model

The *fib*-SCC model was developed within the parameter ranges derived from the SCC database, as presented in Table 4, including curing temperature (T_{cur}), f_{cm28}, cm and w/cm ratio.

The calibration was carried out on the *fib* MC2010 model, originally designed to predict the shrinkage of normal- and high-strength SCC. In the modified model, the w/cm ratio was incorporated in the formulas for both autogenous and total shrinkage. Additionally, the humidity factor of shrinkage in the *fib*-SCC model covered a broader range than that in the original *fib* MC2010 model, which had a minimum applicable RH of 40%. The updated model was adjusted to fit the SCC database with a minimum RH of 27%. Furthermore, shrinkage parameters dependent on cement type were optimised to achieve the lowest NRMSE, based on the appropriately filtered datasets. The F-SCC model for predicting shrinkage in SCC is defined by the following formulas.

Table 4. Ranges of *fib*-SCC model

	T_{cur}: °C	f_{cm28}: MPa	cm: kg/m^3	w/cm ratio
Minimum	20	17	275	0.18
Maximum	30	110	840	0.68

Total shrinkage strain is calculated as:

$$\varepsilon_{cs}(t, t_c) = \varepsilon_{sh}(t, t_c) + \varepsilon_{au}(t, t_c) - \varepsilon_{au}(t_c) \tag{14}$$

where t is the current age of concrete (days), t_c is the age at exposure (days), $\varepsilon_{sh}(t,t_c)$ is the drying shrinkage strain, and $\varepsilon_{au}(t,t_c)$ and $\varepsilon_{au}(t_c)$ are the autogenous shrinkage strains at t days and t_c days, respectively.

Autogenous shrinkage is expressed as:

$$\varepsilon_{au}(t) = \varepsilon_{au0}(f_{cm28}) \cdot \beta_{as}(t) \tag{15}$$

$$\varepsilon_{au0}(f_{cm28}) = -\alpha_{as}\left(\frac{0.1 f_{cm28,T}}{6 + 0.1 f_{cm28,T}}\right)^{2.5} \times 10^{-6} \tag{16}$$

$$\beta_{as}(t) = \frac{1 - \exp\left(-0.2\sqrt{t}\right)}{0.36(w/cm)^{\alpha_{aw}}} \tag{17}$$

where $\varepsilon_{au0}(f_{cm28})$ is the notional autogenous shrinkage coefficient, $\beta_{as}(t)$ is the time function of autogenous shrinkage, and α_{as} and α_{aw} are parameters dependent on the cement type.

The independent variables of the autogenous shrinkage database, including the test duration ($t - t_c$) and the start of the test (t_c), were divided into subsets, as summarised in Table 5. The new parameter, w/cm ratio, with an average value of 0.36 based on the SCC database, was also added to the formulas of the *fib*-SCC model. Based on the subsets of the database, the parameter α_{aw} in Equation 17 was optimised with values of 0.7 and 0.55 corresponding to, respectively, normal-strength concrete ($f_{cm28} \leq 55$ MPa) and high-strength concrete ($f_{cm28} > 55$ MPa), as can be seen from the results shown in Figures 3 and 4, respectively. Consequently, the prediction capacity of the modified model based on the w/cm ratio was improved compared with the original model, as shown in Figures 5(a) and 5(b).

The optimisation of drying shrinkage was carried out in the same way as autogenous shrinkage. However, compared with autogenous shrinkage, drying shrinkage involves the consideration of a greater number of independent variables, including test duration,

Table 5. Independent variables of autogenous shrinkage database

Test duration, $t - t_c$: days		Start of test, t_c: days	
Range	Number of data points	Range	Number of data points
≤90	212	<1	214
90–210	98	≥1	141
210–1500	45	—	—

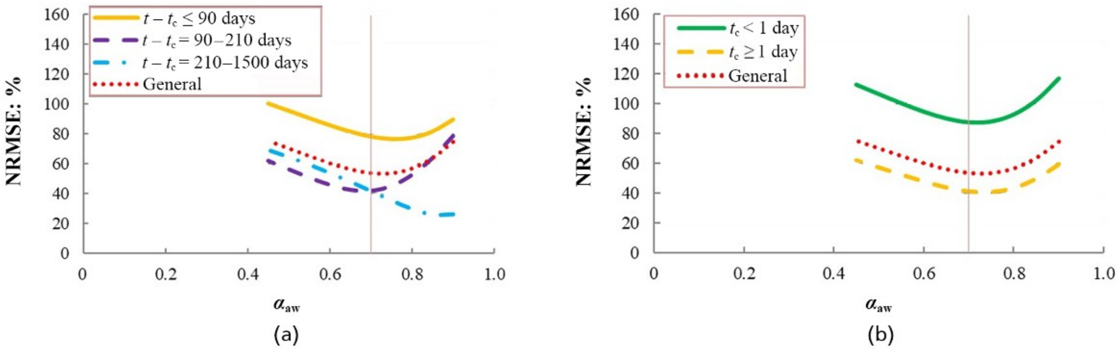

Figure 3. Optimum parameter (α_{aw}) of autogenous shrinkage of normal-strength concrete as a function of: (a) test duration; (b) start of test

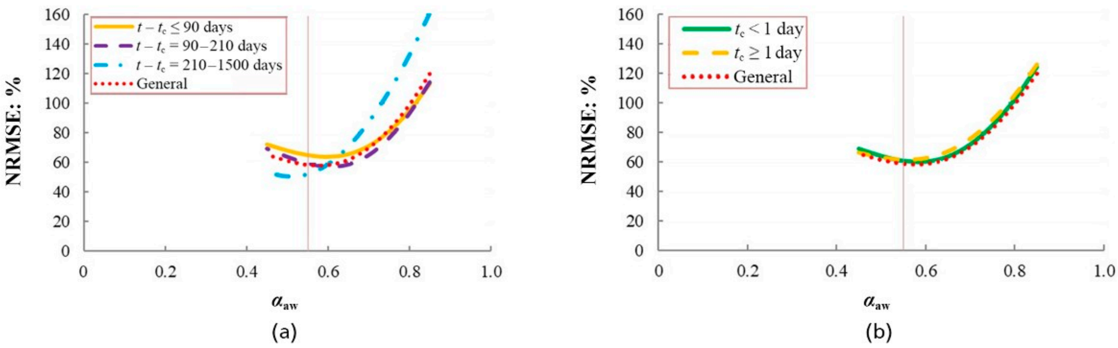

Figure 4. Optimum parameter (α_{aw}) of autogenous shrinkage of high-strength concrete as a function of: (a) test duration; (b) start of test

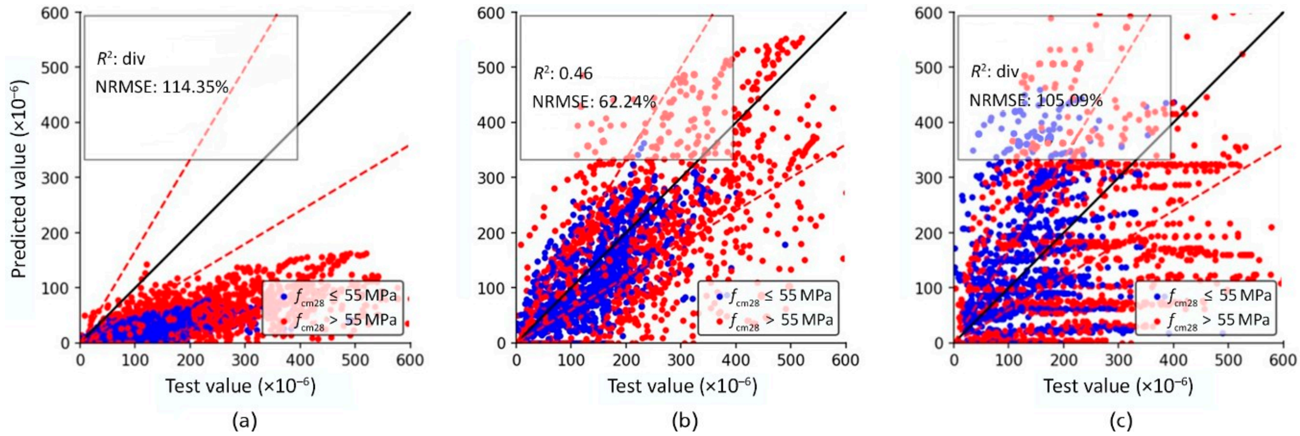

Figure 5. Evaluation of models for predicting autogenous shrinkage based on 245 datasets (2199 data points): (a) *fib* MC2010 model (*fib*, 2013); (b) *fib*-SCC model; (c) B4 model (Bažant and Wan-Wendner, 2015). A full-colour version of this figure can be found on Emerald Insight (www.emerald.com)

specimen size (V/S), RH and the start of the test. These variables were divided into subsets, as summarised in Table 6.

Drying shrinkage is expressed as:

18. $$\varepsilon_{\text{sh}}(t, t_c) = \varepsilon_{\text{shu}} k_{\text{h},T} \beta_{\text{s}}(t - t_{\text{c}})$$

19. $$\varepsilon_{\text{shu}} = \left(\frac{\text{w/cm}}{0.36}\right)^{\alpha_{\text{dw}}} \left[(220 + 110\alpha_{\text{ds1}})\exp\left(-\alpha_{\text{ds2}} f_{\text{cm28},T}\right)\right] \times 10^{-6}$$

20. $$k_{\text{h},T} = k_{\text{h}}\left[1 + \left(\frac{4}{103 - \text{RH}}\right)\left(\frac{T - 20}{40}\right)\right]$$

21. $$k_{\text{h}} = \begin{cases} -1.55\left[1 - \left(\frac{\text{RH}}{100}\right)^3\right] \text{for} \quad \text{RH} < \text{RH}_T \\ 0.25 \quad \text{for} \quad \text{RH} \geq \text{RH}_T \end{cases}$$

22. $$\text{RH}_T = 99\%\beta_{\text{s1}} + \left(\frac{T - 20}{25}\right)^3$$

23. $$\beta_{\text{s1}} = \left(\frac{35}{f_{\text{cm28},T}}\right)^{0.1} \leq 1.0$$

24. $$f_{\text{cm28},T} = f_{\text{cm28}}(1.06 - 0.003T)$$

25. $$\beta_s(t - t_c) = \left(\frac{t - t_c}{(t - t_c) + 0.035h_T^2}\right)^{0.5}$$

26. $$h_T = 0.035h^2 \exp\left[-0.06(T - 20)\right]$$

27. $$h = 2A_c/U$$

In Equations 18–27, ε_{shu} is the ultimate drying shrinkage strain (10^{-6} m/m), $k_{\text{h},T}$ is the temperature-adjusted humidity factor for shrinkage, k_{h} is the humidity factor for shrinkage, RH_T is the temperature-adjusted RH (%), T is the temperature (°C), $f_{\text{cm28},T}$ is the temperature-adjusted 28-day compressive strength (MPa), f_{cm28} is the 28-day compressive strength (MPa), $\beta_{\text{s}}(t - t_{\text{c}})$ is the time function for drying shrinkage, h_T is the temperature-adjusted notional size (mm), h is the notional size (mm), A_{c} is the cross-section area of the concrete specimen (mm^2), U is the cross-section perimeter of the concrete specimen (mm) and α_{ds1}, α_{ds2}, α_{dw}, α_{as} and α_{aw} are parameters dependent on the cement type, as provided in Table 7.

Comparison of *fib*-SCC with other models for shrinkage prediction

In the case of autogenous shrinkage, as illustrated in Figure 5, the NRMSE of the *fib*-SCC was 62.24%, which was the lowest error of the evaluated models. The *fib* MC2010 and B4 models yielded significantly higher NMRSE values of 114.35% and 105.09%, respectively, based on 245 datasets of SCC with cement type R. In other words, the errors of the *fib* MC2010 and B4 models were approximately 84% and 69% higher than that of the *fib*-SCC model, respectively. Notably, the *fib* MC2010 model tended to underestimate autogenous shrinkage, with many data points clustering along the x-axis. In contrast, the *fib*-SCC model displayed a more uniform scatter of data around the standard line and most of the data points were distributed well in an acceptable error range of ±40% (indicated by the two red dashed lines), indicating a remarkable improvement for the *fib*-SCC model. Although the B4 model showed a reasonable level of prediction, its data points were more widely spread, reducing overall accuracy.

Figure 6 shows the performance of the models in predicting total shrinkage. The *fib*-SCC model not only accurately captured autogenous shrinkage but also provided enhanced predictions for total shrinkage as compared with the original *fib* MC2010 model. The data points of the *fib*-SCC model were well-distributed around the reference line, while the *fib* MC2010 model strongly

Table 7. Coefficients of shrinkage formulas for type R cement

				α_{aw}	
α_{ds1}	α_{ds2}	α_{dw}	α_{as}	$f_{\text{cm28}} \leq 55$ MPa	$f_{\text{cm28}} > 55$ MPa
4	0.012	0.8	700	0.55	0.70

Table 6. Independent variables of total shrinkage database

Test duration, $t - t_c$: days		Specimen size, V/S: mm		RH: %		Start of test, t_c: days	
Range	Number of data points	Range	Number of data points	Range	Number of data points	Range	Number of data points
≤90	507	11.4–18	586	27–50	900	≤	615
90–210	494	18–23	597	50–60	400	1–7	512
210–1500	390	23–70	208	60–90	89	7–90	264

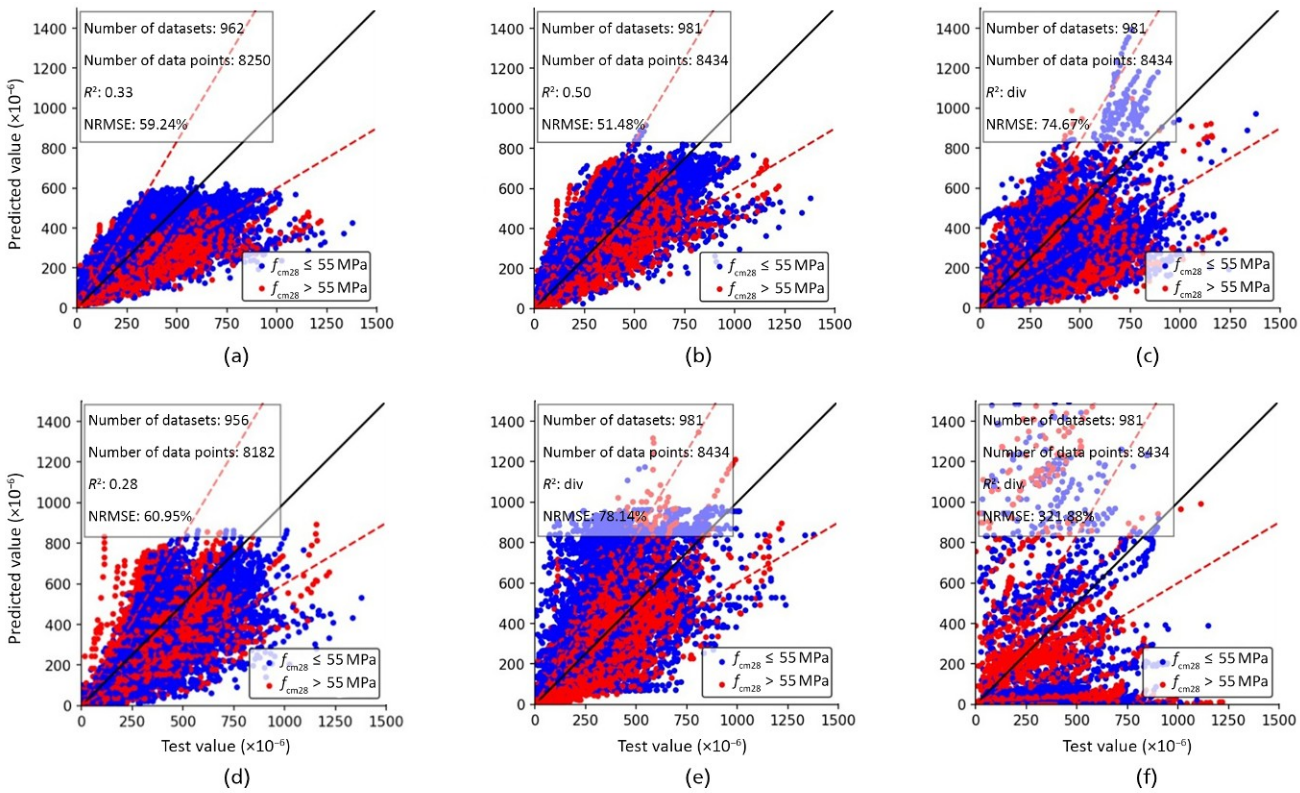

Figure 6. Evaluation of prediction models for total shrinkage: (a) *fib* MC2010 model (*fib*, 2013); (b) *fib*-SCC model; (c) B4 model (Bažant and Wan-Wendner, 2015); (d) ACI model (ACI, 1992); (e) modified JSCE model (Aslani and Maia, 2013); (f) modified CEB90 model (Poppe and Schutter, 2005). A full-colour version of this figure can be found on Emerald Insight (www.emerald.com)

underestimated the total shrinkage. The *fib*-SCC model achieved $R^2 = 0.50$ and NRMSE = 51.48%, the lowest error among the evaluated models. For comparison, the NRMSE values of the *fib* MC2010, B4, ACI model and modified JSCE models were approximately 59%, 75%, 61% and 78%, respectively. The B4 and ACI models generally underestimated total shrinkage, while the modified JSCE model showed a tendency to overestimate it. The modified CEB90 model demonstrated the lowest accuracy, with an NMRSE approaching 322%.

Evaluation of *fib*-SCC model based on experimental values

Based on the experimental results obtained in this study, the *fib*-SCC model demonstrated reliable performance in predicting both autogenous and total shrinkage. As shown in Figure 7, although the model slightly underestimated the measured shrinkage, the majority of the data points remained within the acceptable deviation range of ±40%. Statistical evaluation showed that the R^2 and NRMSE of the autogenous shrinkage were 0.77 and 30%, respectively. For total shrinkage, the corresponding values were 0.84 and 24.7%. These results confirm that the *fib*-SCC model effectively predicted not only across the shrinkage of the SCC database with a test duration up to 1435 days but also under the experimental conditions in the short term (150 days in this study).

ML methods for total shrinkage prediction

ML algorithms (DT (single-tree method), RF, ET, GB and XGBoost) were also used to predict the total shrinkage of SCC. These supervised learning algorithms were selected owing to their high predictive accuracy when handling multiple input parameters. The models were implemented using regression techniques based on bagging (bootstrap aggregating) and boosting algorithms. Bagging is an ensemble meta-algorithm designed to reduce model overfitting, while boosting incrementally improves model performance by training sequentially on different subsets of the database to minimise training errors (Aggarwal, 2015).

The typical stages of the learning process are illustrated in Figure 8 (Mohri *et al*., 2018). Each algorithm follows an iterative, step-by-step computational procedure that may involve descriptions, equations or a combination of both (Yang, 2019). The key motivation for applying ML in shrinkage prediction lies in the ability of algorithms to continuously learn and detect patterns by processing large volumes of

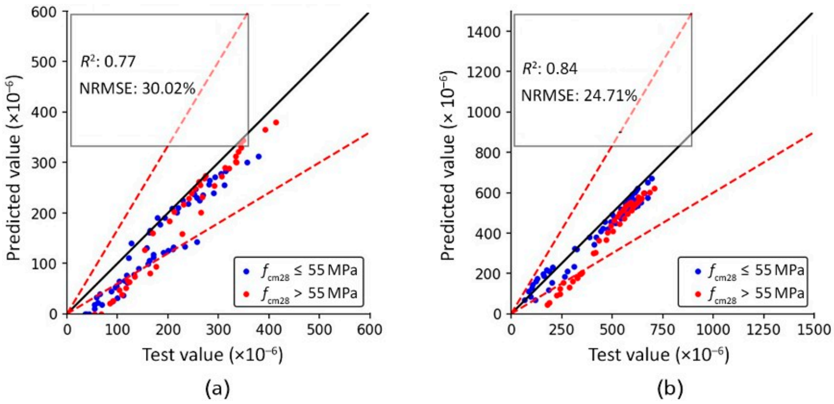

Figure 7. Evaluation of *fib*-SCC model on experimental results (five datasets, 98 data points): (a) autogenous shrinkage; (b) total shrinkage. A full-colour version of this figure can be found on Emerald Insight (www.emerald.com)

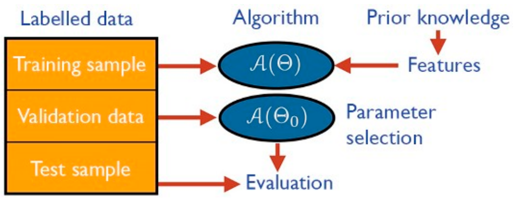

Figure 8. Typical stage of the learning process

experimental data, unconstrained by any mathematical formulas or framework. This capability allows for more accurate shrinkage predictions, making these methods promising for practical applications.

Optimum parameters of ML methods for total shrinkage

Eight input features (w/cm ratio, *cm*, f_{cm28}, RH, *V/S*, t_c, *T* and T_{cur}) were considered in the optimisation process of the ML methods. These features also correspond to the key parameters in the formulae of the *fib*-SCC model to compare the influence of features on different types of prediction models. The role of features in the SCC database for total shrinkage prediction was also dissimilar, corresponding to the types of methods carried out by their algorithm as presented. According to the results shown in Figure 9, all the ML methods considered the RH as the most important variable influencing total shrinkage behaviour, while temperature features (*T* and T_{cur}) were the least important variables.

In the SCC shrinkage database, the total shrinkage data (8250 data points) were divided into training and validation sets using an 80:20 ratio. As a result, the training and validation sets included 6600 and 1650 data points, respectively. As already mentioned, during optimisation of the ML methods, eight input features were selected, corresponding to the key influencing parameters in the *fib*-SCC model.

The main parameters of the ML methods that were investigated were

- the number of estimators (n_estimators)
- the maximum number of features (max_features)
- the maximum tree depth (max_depth)
- the maximum number of leaf nodes (max_leaf_nodes)
- the column subsampling rate (colsample_bytree).

For consistency, the number of boosting stages (n_estimators) was fixed at 100 as a standard default and max_features was set to 8, corresponding to the key parameters' influence on the *fib*-SCC model. The other parameters were optimised through a step-by-step tuning process aimed at minimising errors based on values of statistical coefficients. The parameters were also chosen to avoid overfitting of algorithms, ensuring robust prediction not only on the training data but also on new data.

In the optimisation process, the max_depth parameter limits the number of nodes in each tree, with the goal of achieving optimum accuracy for both the training and validation datasets, as shown in Figure 10. The optimal results showed that the max_depth values for the DT, RF and ET were significantly higher than those for GB and XGBoost. A higher max_depth value generally indicates increased complexity, which can lead to a greater risk of overfitting when applied to the SCC database. In addition, the max_leaf_nodes parameter was effective in reducing overfitting for algorithms such as RF, ET and GB, while the colsample_bytree parameter improved the predictive performance of the XGBoost method. The optimum values of these parameters are shown in Table 8.

Evaluation of ML methods on total shrinkage

Figure 11 shows the predicted results of the ML methods, evaluated on the validation dataset of 1650 points using the optimum parameter values. The prediction capacities of the DT, RF and ET were less efficient than those of GB and XGBoost. The data points predicted by the DT were not well scattered in the diagram,

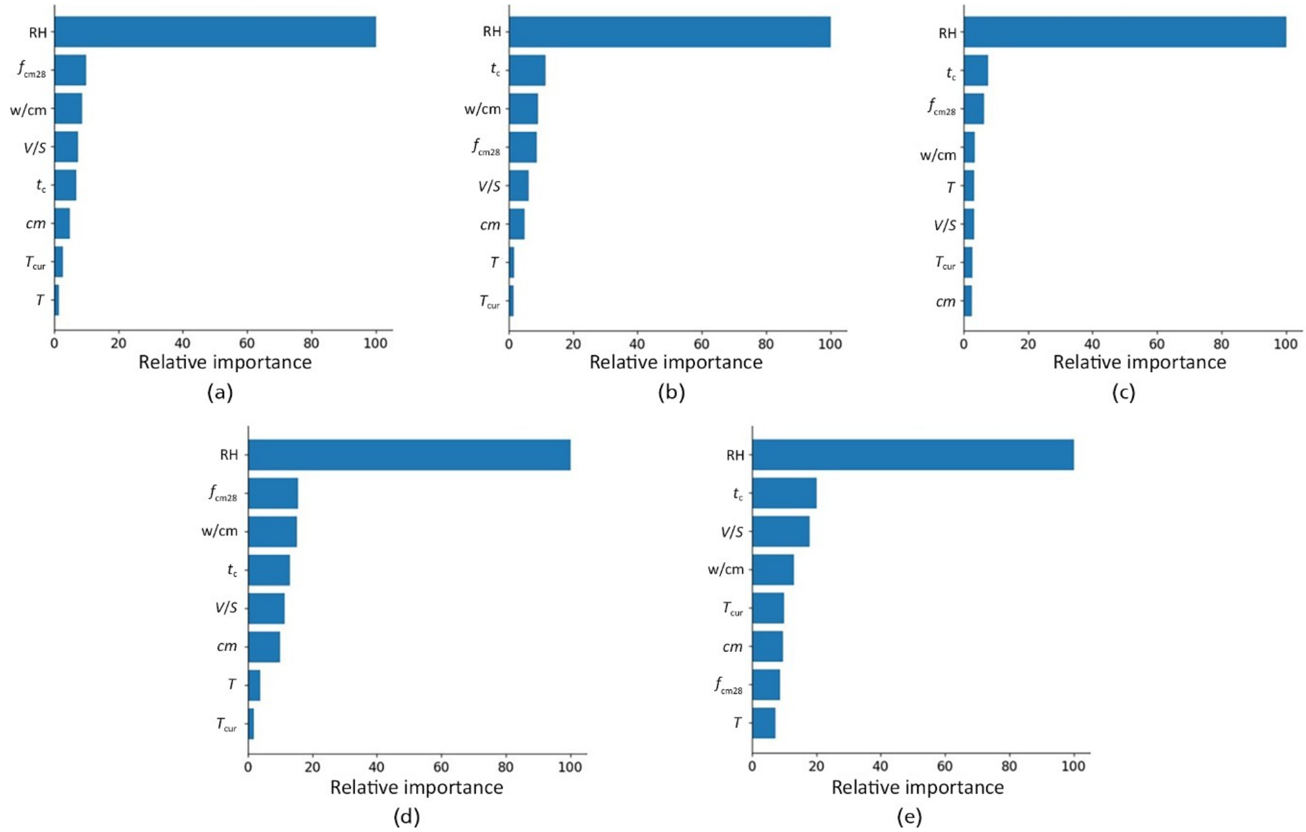

Figure 9. Importance of features in total shrinkage prediction by ML methods: (a) DT; (b) RF; (c) ET; (d) GB; (e) XGBoost

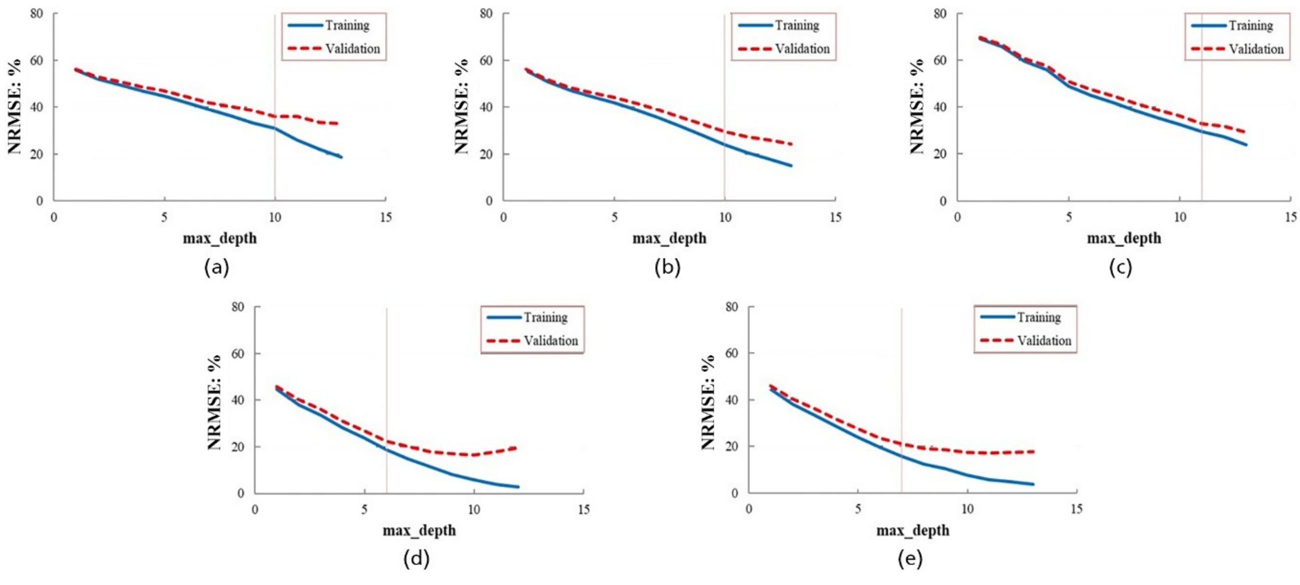

Figure 10. Optimisation of max_depth for total shrinkage prediction: (a) DT; (b) RF; (c) ET; (d) GB; (e) XGBoost

Table 8. Optimum parameters of ML methods for total shrinkage prediction

Parameter	DT	RF	ET	GB	XGBoost
n_estimators	—	100	100	100	100
max_features	8	8	8	8	8
max_depth	10	10	11	6	7
max_leaf_nodes	—	70	70	40	—
colsample_bytree	—	—	—	—	0.4

yielding $R^2 = 0.73$ and an NRMSE = 37.4%. The RF and ET, which use bagging techniques, tended to overestimate total shrinkage at later ages. The boosting algorithms (GB and XGBoost) predicted total shrinkage more effectively, with NRMSE values of approximately 23.8% and 23.7%, respectively. However, the data points predicted by XGBoost were more closely distributed around the standard line than those of the other methods.

When the SCC database with the ML methods was evaluated on k-fold cross-validation with n_splits of 5, 10 and 15 (Figure 12), the XGBoost algorithm exhibited the highest accuracy. When k-fold cross-validation was divided into ten folds (n_splits = 10), the accuracy of the algorithms in descending order was 0.52, 0.50, 0.49, 0.43, and 0.21, corresponding to XGBoost, GB, ET, RF and DT, respectively. In other words, XGBoost showed the best generalised performance among the ML methods. The DT algorithm again displayed the lowest accuracy based on k-fold validation.

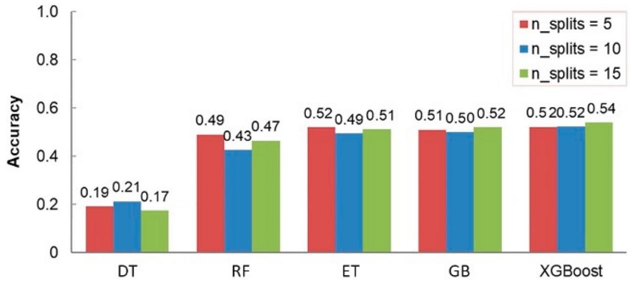

Figure 12. Evaluation of ML methods on k-fold validation.
A full-colour version of this figure can be found on Emerald Insight (www.emerald.com)

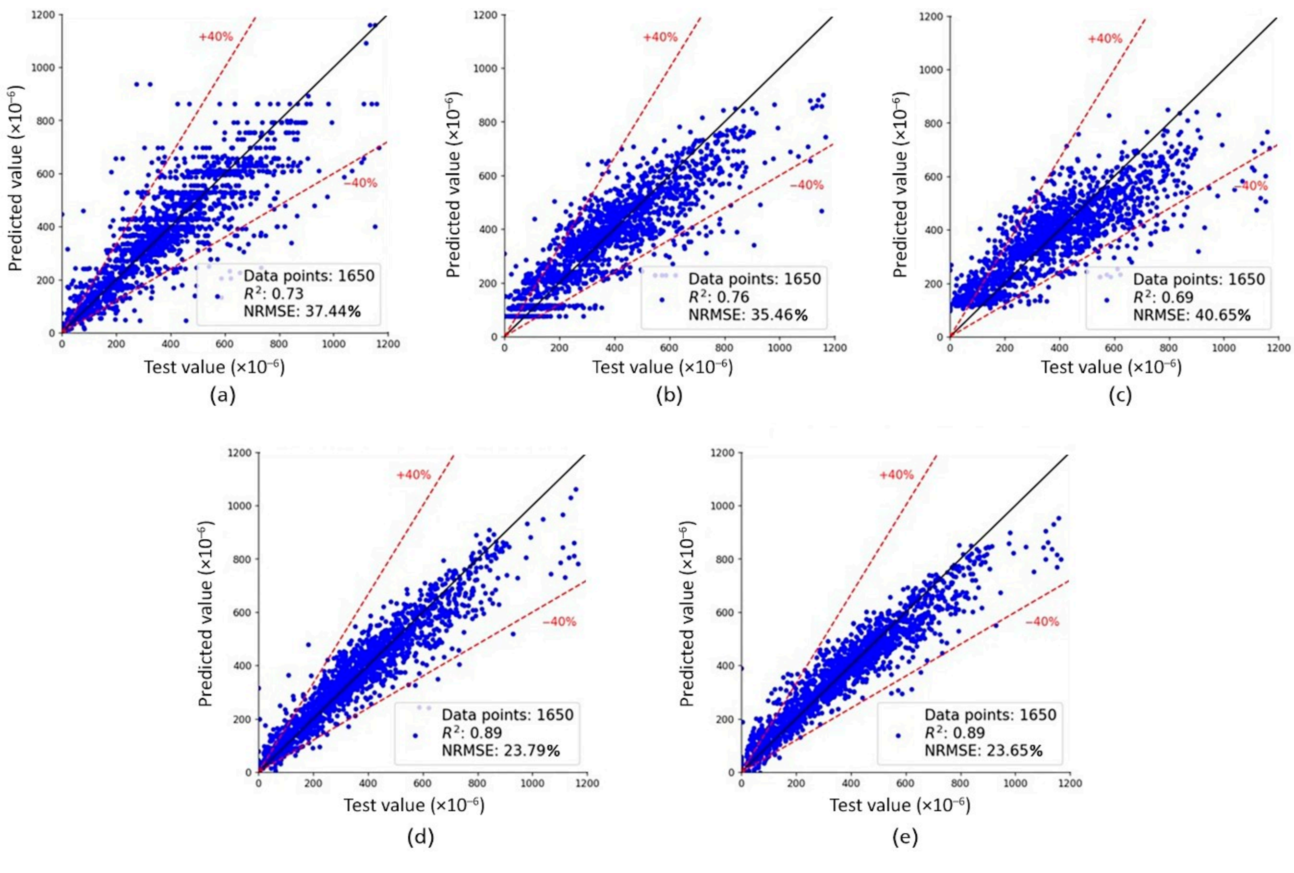

Figure 11. Evaluation of ML methods on validation data of total shrinkage: (a) DT; (b) RF; (c) ET; (d) GB; (e) XGBoost

Comparison of ML methods and the *fib*-SCC model for shrinkage prediction

Figure 13 shows the prediction results of the ML methods and the *fib*-SCC model applied to the entire dataset of total shrinkage (8250 data points). The NRMSEs of the DT, RF, ET, GB and XGBoost models were approximately 31.2%, 33.3%, 37.9%, 20.8% and 20%, respectively – significantly lower than the NRMSE of the *fib*-SCC model (51.5%), as illustrated in Figure 6(b). Moreover, the data points of the *fib*-SCC model showed greater dispersion around the standard line ($x = y$) compared with the ML methods, although the majority still fell within the boundary lines defined by the ±40% error margin (represented by the red dashed lines). This indicates that the *fib*-SCC model was less efficient than the supervised learning methods in predicting total shrinkage. This discrepancy can be attributed to the training methods of the algorithms. Bagging methods improve predictive performance by combining the predictions of multiple models based on the average values of multiple subsets of the training data, thereby improving the final result, while boosting methods enhance the prediction performance of weak learners by iteratively correcting the errors (Unpingco, 2016; Witten *et al*., 2017). In comparison, the *fib*-SCC model calculates shrinkage strain based solely on the relationships among the influential parameters in the explicit formulas.

In aspect of residual analysis, as shown in Figure 14, differences between the test data and the predicted values of ML algorithms are clear. The results indicated that the residual points of XGBoost and GB were distributed closely around the standard axis (horizontal red line). Consequently, these algorithms exhibited better performance than the other algorithms. The DT algorithm showed the least fit for the SCC database, with the points spread widely. The RF and ET tended to overestimate the total shrinkage in the early stage, so that the predicted values were higher than the test data, and the residual points were scattered around the standard axis. However, these algorithms underestimated the shrinkage of SCC in later stages of testing.

Conclusion

A new model (*fib*-SCC model) and ML methods for predicting the shrinkage of SCC were evaluated. The following main conclusions were drawn from this work.

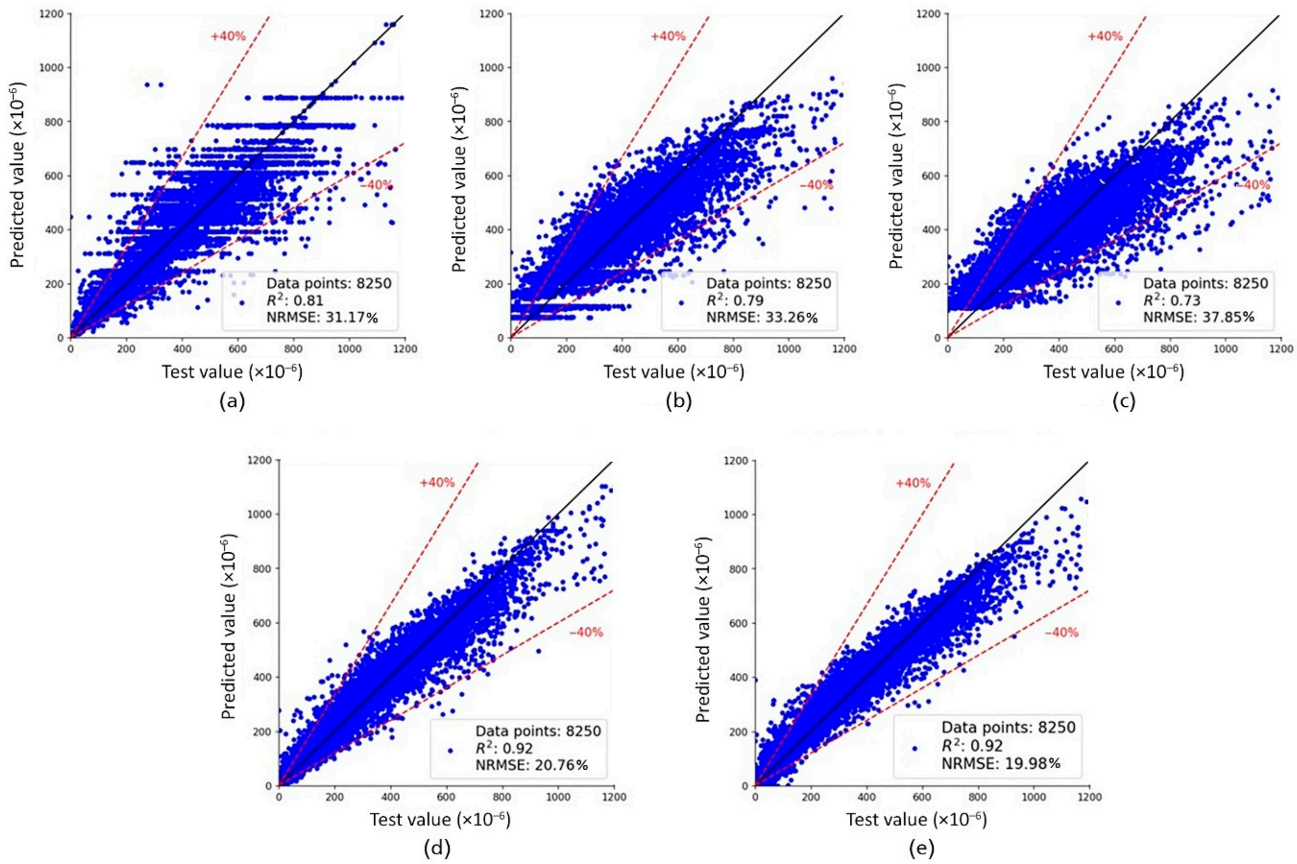

Figure 13. Evaluation of ML methods on total shrinkage: (a) DT; (b) RF; (c) ET; (d) GB; (e) XGBoost

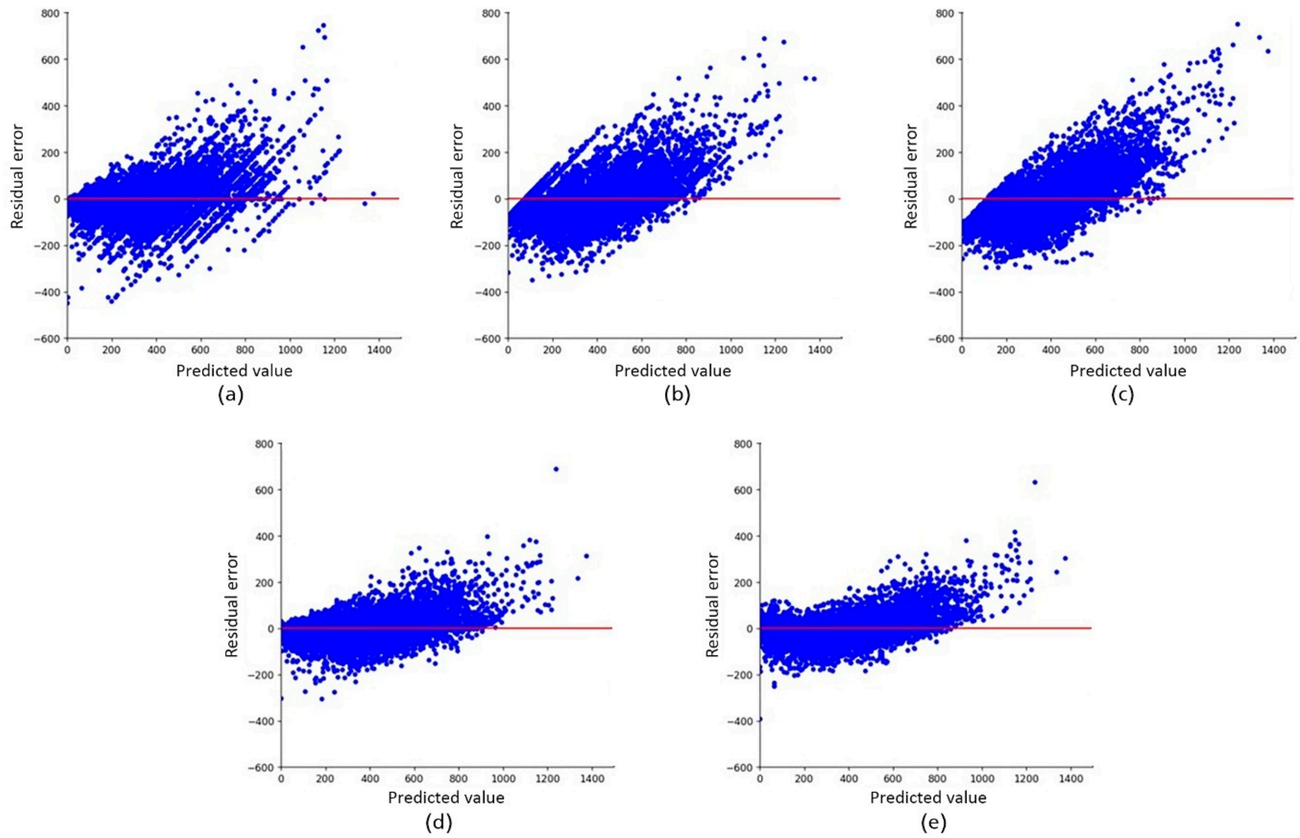

Figure 14. Evaluation of ML methods on residual analysis (8250 data points): (a) DT; (b) RF; (c) ET; (d) GB; (e) XGBoost

(*a*) The experimental results showed that the incorporation of FA improved the workability of concrete used in proportions of 20–55% by weight of cement. FA also enhanced the compressive strength at later ages. In addition to FA, the w/cm ratio and SP dosage significantly impacted both autogenous and total shrinkage.

(*b*) Numerous data were collected in an expanded laboratory database of SCC, comprising 1328 datasets for total shrinkage and 282 datasets for autogenous shrinkage.

(*c*) The *fib*-SCC model, developed through calibration of the *fib* MC2010 model, was adapted to fit over 1200 datasets using normal cement (type R). The modified model accurately predicted both the shrinkage of the SCC database and the experimental results. However, the *fib*-SCC model was less efficient than the ML methods.

(*d*) Of the ML methods, the DT, RF and ET were less accurate than GB and XGBoost in predicting the total shrinkage of SCC.

Data availability

All data, prediction models and algorithms generated or used during the study appear in the published article.

Funding

This research was sponsored by the financial support of the Ministry of Education and Training (MOET) of Vietnam under grant number B2024-XDA-07.

REFERENCES

Aashto (American Association of State Highway and Transportation Officials) (2017) *AASHTO LRFD Bridge Design Specifications*. Aashto, Washington, DC, USA.

Abdalhmid JM, Ashour A and Sheehan T (2019) Long-term drying shrinkage of self-compacting concrete: experimental and analytical investigations. *Construction and Building Materials* **202**: 825–837.

ACI (American Concrete Institute) (1992) ACI 209R-92: prediction of creep, shrinkage, and temperature effects in concrete structures. ACI, Farmington Hills, MI, USA.

ACI (1998) ACI 363.2R-98: Guide to quality control and testing of high-strength concrete. ACI, Farmington Hills, MI, USA.

ACI (2008) ACI 209.2R-08: Guide for modeling and calculating shrinkage and creep in hardened concrete. ACI, Farmington Hills, MI, USA.

Aggarwal CC (2015) *Data Classification: Algorithms and Applications*. CRC Press, Boca Raton, FL, USA.

Ahmad MW, Reynolds J and Rezgui Y (2018) Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production* **203**: 810–821.

Ao Y, Li H, Zhu L, Ali S and Yang Z (2019) The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. *Journal of Petroleum Science and Engineering* **174**: 776–789.

Aslani F and Maia L (2013) Creep and shrinkage of high-strength self-compacting concrete: experimental and analytical analysis. *Magazine of Concrete Research* **65(17)**: 1044–1058.

ASTM (2015) C1758/C1758M: Standard practice for fabricating test specimens with self-consolidating concrete. ASTM International, West Conshohocken, PA, USA.

ASTM (2017) C157/C157M: Standard test method for length change of hardened hydraulic-cement mortar and concrete. ASTM International, West Conshohocken, PA, USA.

ASTM (2018a) C1611/C1611M: Standard test method for slump flow of self-consolidating concrete. ASTM International, West Conshohocken, PA, USA.

ASTM (2018b) C33/C33M: Standard specification for concrete aggregates. ASTM International, West Conshohocken, PA, USA.

ASTM (2019a) C150/C150M: Standard specification for Portland cement. ASTM International, West Conshohocken, PA, USA.

ASTM (2019b) C618: Standard specification for coal fly ash and raw or calcined natural pozzolan for use in concrete. ASTM International, West Conshohocken, PA, USA.

Bal L and Buyle-Bodin F (2013) Artificial neural network for predicting drying shrinkage of concrete. *Construction and Building Materials* **38**: 248–254.

Baweja S and Bažant ZP (1995) Creep and shrinkage prediction model for analysis and design of concrete structures – model B3. *Materials and Structures* **28**: 357–365.

Bažant ZP and Li GH (2007) *Unbiased Statistical Comparison of Creep and Shrinkage Prediction Models*. Northwestern University, Evanston, IL, USA.

Bažant ZP and Wan-Wendner R (2015) Model B4 for creep, drying shrinkage and autogenous shrinkage of normal and high-strength concretes with multi-decade applicability. *Materials and Structures* **48**: 753–770.

Bentéjac C, Csörgő A and Martínez-Muñoz G (2021) A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review* **54(3)**: 1937–1967.

Berrouachedi A, Jaziri R and Bernard G (2019) Deep cascade of extra trees. In *Trends and Application in Knowledge Discovery and Data Mining* (Leong Hou U and Lauw HW (eds)). Springer Nature, Cham, Switzerland, pp. 117–129.

Breiman L (2001) Random forests. *Machine Learning* **45(1)**: 5–32.

CEB (Comité Euro-International du Béton) (1991) *CEB-FIP Model Code 1990*. CEB, Lausanne, Switzerland.

Chan Y, Liu C and Lu Y (1998) Effects of slag and fly ash on the autogenous shrinkage of high performance concrete. In *Proceedings of International Workshop Organized by JCI* (Tazawa E (ed.)). E & FN Spon, London, UK, pp. 221–228.

Chen T and Guestrin C (2016) XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Krishnapuram B, Shah M, Smola AJ *et al.* (eds)). Association for Computing Machinery, New York, NY, USA, pp. 785–794.

Dietterich HR, Lev E, Chen J, Richardson JA and Cashman KV (2017) Benchmarking computational fluid dynamics models of lava flow simulation for hazard assessment, forecasting, and risk management. *Journal of Applied Volcanology* **6(9)**: 1–14.

El-Chabib H and Syed A (2013) Properties of self-consolidating concrete made with high volumes of supplementary cementitious materials. *Journal of Materials in Civil Engineering* **25(11)**: 1579–1586.

Esping O and Löfgren I (2006) Investigation of early age deformation in self-compacting concrete. In *Proceedings of the 2nd International Symposium on Advances in Concrete through Science and Engineering* (Marchand J, Bissonnette B, Gagné R, Jolin M and Paradis F (eds)). RILEM Publications SARL, Champs-sur-Marne, France, pp. 207–224.

fib (fedération internationale du béton) (2013) *fib Model Code for Concrete Structures 2010*. fib, Lausanne, Switzerland.

Freund Y and Schapire RE (1996) Experiments with a new boosting algorithm. In *Proceedings of International Conference on Machine Learning* (Saitta L (ed.)). Morgan Kaufmann, San Francisco, CA, USA, pp. 148–156.

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *The Annals of Statistics* **29(5)**: 1189–1232.

Gardner NJ and Lockman MJ (2001) Design provisions for drying shrinkage and creep of normal-strength concrete. *ACI Materials Journal* **98(2)**: 159–167.

Geurts P, Ernst D and Wehenkel L (2006) Extremely randomized trees. *Machine Learning* **63(1)**: 3–42.

Güneyisi E, Gesoğlu M and Özbay E (2010) Strength and drying shrinkage properties of self-compacting concretes incorporating multi-system blended mineral admixtures. *Construction and Building Materials* **24(10)**: 1878–1887.

Hami AE and Pougnet P (2020) *Embedded Mechatronic Systems 2. Analyses of Failures, Modeling, Simulation and Optimization*. ISTE Press and Elsevier, London, UK.

Ho TK (1995) Random decision forests. In *Proceedings of the 3rd International Conference on Document Analysis and Recognition* (Kavanaugh M and Storms P (eds)). IEEE Computer Society Press, Los Alamitos, CA, USA, pp. 278–282.

Hubler MH, Wendner R and Bažant ZP (2015a) Comprehensive database for concrete creep and shrinkage: analysis and recommendations for testing and recording. *ACI Materials Journal* **112(4)**: 547–558.

Hubler MH, Wendner R and Bažant ZP (2015b) Statistical justification of model B4 for drying and autogenous shrinkage of concrete and comparisons to other models. *Materials and Structures* **48(4)**: 797–814.

Jawahar JG, Sashidhar C, Reddy IR and Peter JA (2013) Micro and macrolevel properties of fly ash blended self compacting concrete. *Materials & Design* **46**: 696–705.

JSCE (Japan Society of Civil Engineers) (2007) *Standard Specifications for Concrete Structures Design*. JSCE, Japan, Tokyo.

Karthikeyan J, Upadhyay A and Bhandari NM (2008) Artificial neural network for predicting creep and shrinkage of high performance concrete. *Journal of Advanced Concrete Technology* **6(1)**: 135–142.

Khatib J (2008) Performance of self-compacting concrete containing fly ash. *Construction and Building Materials* **22(9)**: 1963–1971.

Khayat KH and Long WJ (2010) Shrinkage of precast, prestressed self-consolidating concrete. *ACI Materials Journal* **107(3)**: 231–238.

Kristiawan SA and Aditya MT (2015) Effect of high volume fly ash on shrinkage of self-compacting concrete. *Procedia Engineering* **125**: 705–712.

Lee HK, Lee KM and Kim BG (2003) Autogenous shrinkage of high-performance concrete containing fly ash. *Magazine of Concrete Research* **55(6)**: 507–515.

Leemann A and Hoffmann C (2005) Properties of self-compacting and conventional concrete differences and similarities. *Magazine of Concrete Research* **57(6)**: 315–319.

Leemann A, Lura P and Loser R (2011) Shrinkage and creep of SCC the influence of paste volume and binder composition. *Construction and Building Materials* **25**: 2283–2289.

Maimon O and Rokach L (2005) *Data Mining and Knowledge Discovery Handbook*. Springer, New York, NY, USA.

Marsland S (2015) *Machine Learning an Algorithmic Perspective*. Taylor & Francis, New York, NY, USA.

Mehta PK and Monteiro PJ (2006) *Concrete Microstructure, Properties, and Materials*, 3rd edn. McGraw Hill, New York, NY, USA.

Minless S, Young J and Darwin D (2002) *Concrete*, 2nd edn. Pearson Education, London, UK.

Mohri M, Rostamizadeh A, Talwalkar A and Bach F (2018) *Foundations of Machine Learning*. MIT Press, Cambridge, MA, USA.

Montgomery DC, Peck EA and Vining GG (2012) *Introduction to Linear Regression Analysis*, 5th edn. John Wiley, Hoboken, NJ, USA.

Okamura H and Ouchi M (2003) Self-compacting concrete. *Journal of Advanced Concrete Technology* **1(1)**: 5–15.

Poppe AM and Schutter GD (2005) Creep and shrinkage of self-compacting concrete. In *Proceedings of the 1st International Symposium on Design, Performance and Use of Self-Consolidating Concrete* (Yu Z, Shi C, Khayat KH and Xie Y (eds)). RILEM Publications SARL, Champs-sur-Marne, France, pp. 329–336.

Quinlan J (1986) Induction of decision trees. *Machine Learning* **1(1)**: 81–106.

Quinlan JR (1993) *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Francisco, CA, USA.

Rizi FS and Granitzer M (2017) Properties of vector embeddings in social networks. *Algorithms* **10(4)**: 1–15.

Rozière E, Granger S, Turcry P and Loukili A (2007) Influence of paste volume on shrinkage cracking and fracture properties of self-compacting concrete. *Cement & Concrete Composites* **29(8)**: 626–636.

Russell R (2018) *Machine Learning Step-by-Step Guide to Implement Machine Learning Algorithms with Python*. CreateSpace, Scotts Valley, CA, USA.

Şahmaran M, Yaman Ö and Tokyay M (2007) Development of high-volume low-lime and high-lime fly-ash-incorporated self-consolidating concrete. *Magazine of Concrete Research* **59(2)**: 97–106.

Shalev-Shwartz S and Ben-David S (2014) *Understanding Machine Learning from Theory to Algorithms*. Cambridge University Press, New York, NY, USA.

Tazawa EI (1999) *Autogenous Shrinkage of Concrete*. E & FN Spon, London, UK.

Unpingco J (2016) *Python for Probability, Statistics, and Machine Learning*. Springer, Cham, Switzerland.

VSQI (Vietnam Standards and Quality Institute) (2022) TCVN 3118-2022: Hardened concrete. Test method for compressive strength. VSQI, Hanoi, Vietnam.

Watt J, Borhani R and Katsaggelos AK (2020) *Machine Learning Refined Foundations, Algorithms, and Applications*, 2nd edn. Cambridge University Press, Cambridge, UK.

Wendner R, Hubler MH and Bažant ZP (2015) Optimization method, choice of form and uncertainty quantification of model B4 using laboratory and multi-decade bridge databases. *Materials and Structures* **48(4)**: 771–796.

Witten IH, Frank E, Hall MA and Pal CJ (2017) *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, CA, USA.

Yang XS (2019) *Introduction to Algorithms for Data Mining and Machine Learning*. Academic Press, London, UK.

How can you contribute?

To discuss this paper, please submit up to 500 words to the editor at support@emerald.com. Your contribution will be forwarded to the author(s) for a reply and, if considered appropriate by the editorial board, it will be published as a discussion in a future issue of the journal.