

Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete

Liang, Minfei; Chang, Ze; Wan, Zhi; Gan, Yidong; Schlangen, Erik; Šavija, Branko

DOI

[10.1016/j.cemconcomp.2021.104295](https://doi.org/10.1016/j.cemconcomp.2021.104295)

Publication date

2022

Document Version

Final published version

Published in

Cement and Concrete Composites

Citation (APA)

Liang, M., Chang, Z., Wan, Z., Gan, Y., Schlangen, E., & Šavija, B. (2022). Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete. *Cement and Concrete Composites*, 125, Article 104295. <https://doi.org/10.1016/j.cemconcomp.2021.104295>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Interpretable Ensemble-Machine-Learning models for predicting creep behavior of concrete

Minfei Liang, Ze Chang, Zhi Wan, Yidong Gan^{*}, Erik Schlangen, Branko Šavija

MicroLab, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, 2628, CN, the Netherlands

ARTICLE INFO

Keywords:

Concrete
Creep
Prediction
Machine learning
Bayesian optimization
Interpretation

ABSTRACT

This study aims to provide an efficient and accurate machine learning (ML) approach for predicting the creep behavior of concrete. Three ensemble machine learning (EML) models are selected in this study: Random Forest (RF), Extreme Gradient Boosting Machine (XGBoost) and Light Gradient Boosting Machine (LGBM). Firstly, the creep data in Northwestern University (NU) database is preprocessed by a prebuilt XGBoost model and then split into a training set and a testing set. Then, by Bayesian Optimization and 5-fold cross validation, the 3 EML models are tuned to achieve high accuracy ($R^2 = 0.953, 0.947$ and 0.946 for LGBM, XGBoost and RF, respectively). In the testing set, the EML models show significantly higher accuracy than the equation proposed by the *fib* Model Code 2010 ($R^2 = 0.377$). Finally, the SHapley Additive exPlanations (SHAP), based on the cooperative game theories, are calculated to interpretate the predictions of the EML model. Five most influential parameters for concrete creep compliance are identified by the SHAP values of EML models as follows: time since loading, compressive strength, age when loads are applied, relative humidity during the test and temperature during the test. The patterns captured by the three EML models are consistent with theoretical understanding of factors that influence concrete creep, which proves that the proposed EML models show reasonable predictions.

1. Introduction

Creep is one of the most significant behaviors of concrete, which forms necessary basis for precise prediction and evaluation of both short- and long-term mechanical response of concrete structures under sustained loads. On one hand, for early-age concrete members that bear shrinkage-induced tension stress, creep essentially reduces the magnitude of tensile stress evolution [1,2]. On the other hand, for long-term mechanical behavior of concrete structures, creep is a major reason accounting for serious time-dependent engineering problems such as substantial deflection and prestress loss in bridges [3,4]. Therefore, an accurate estimation of creep behavior of concrete is important for long-term serviceability of concrete structures. For this, a number of experimental and numerical approaches have been developed, as described below.

1.1. Experimental methods

Creep and relaxation have been experimentally studied for decades.

A wide range of variables have been investigated, including different concrete mixtures (i.e., w/c ratio, supplementary cementitious materials (SCM), fiber, admixture.) [5–8], environmental conditions (i.e., relative humidity, temperature.) [9–11] and loading schemes (i.e., compressive or tensile loading, age or duration of loading.) [12,13]. These tests provided valuable data for understanding the creep behavior of concrete. However, accurate prediction remains a challenge because of high sensitivity of creep to wide ranges of parameters and interplay between these parameters [14].

In recent years, researchers have tried to decouple the influence of multiple variables to more accurately measure concrete creep under certain conditions. Hydration degree and its interaction with loading conditions can significantly influence the experimental results of creep tests [15,16]. To overcome this, Wyrzykowski et al. [17] partly replaced unhydrated cement with an inert quartz powder to form a non-aging system, on which the mechanical properties of pseudo cement paste were calibrated to mimic the real one and then uniaxial compressive creep tests were conducted. With the concern that most researchers assumed an axial load in their creep/relaxation tests, Liang et al. [18]

^{*} Corresponding author.;

E-mail addresses: M.Liang-1@tudelft.nl (M. Liang), Z.Chang-1@tudelft.nl (Z. Chang), Z.Wan-1@tudelft.nl (Z. Wan), Y.Gan@tudelft.nl (Y. Gan), Erik.Schlangen@tudelft.nl (E. Schlangen), B.Savija@tudelft.nl (B. Šavija).

<https://doi.org/10.1016/j.cemconcomp.2021.104295>

Received 15 June 2021; Received in revised form 1 September 2021; Accepted 5 October 2021

Available online 13 October 2021

0958-9465/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

performed biaxial creep tests of high strength concrete using ring test method and concluded that creep properties under biaxial stress condition were significantly lower than that under uniaxial stress condition. Besides, in order to decouple the influence of shrinkage in creep tests, Liang et al. [19] also devised the methodology of flexural deflection test to exclude the influence of shrinkage on measurement of creep properties under the assumption of a symmetric distribution of relative humidity about the neutral plane. As an important factor influencing both the hydration reaction and the time-dependent change of microstructure, temperature can significantly influence creep behavior of concrete. Ladaoui et al. [20] tested the basic creep of 4 types of high-performance concrete under different temperatures, and observed a factor of 2–3 in basic creep when rising the temperature 20 °C and 50 °C. Similarly, the studies [21,22] also observed significant increase of concrete creep under elevated temperatures. These observations are in line with [23] and can be explained by distortion of water from the CSH gel. Meanwhile, To exclude the influence of temperature and achieve the aim of in-situ and early-age measurement of creep, a series of temperature-regulated stress testing methods were employed to investigate the influencing mechanisms of loading age and loading scenarios on creep [24–26].

Except for macro-scale tests mentioned above, micro-scale indentation tests have been conducted to investigate the intrinsic creep properties of C–S–H in recent years [27,28]. Recently, Gan et al. [29–32] have proposed the methodology of micro-cantilever bending test to characterize the creep properties of cement paste, by which the effects of w/b ratio, binder type and stress level on creep measurement were investigated. Their results showed that microstructural features largely influence the global creep behavior.

To sum up, the state-of-the-art testing methods for creep of concrete can effectively improve the measurement accuracy by using more sophisticated loading scenarios, material processing methods, and advanced testing machines. However, an unavoidable challenge is that only a limited number of parameters can be tested and decoupled in laborious experimental works, which largely restrains the estimation accuracy of creep behavior of concrete under practical loading conditions.

1.2. Numerical methods

Based on experimental tests, a series of numerical, analytical and semi-empirical models were proposed to predict creep behavior of concrete. Bazant et al. [33] proposed a rate-type creep law to characterize the linear creep behavior of concrete, which can be interpreted by a Maxwell or a Kelvin Chain model. To avoid the need for storing all the stress or strain history of all FE elements and improve the computational efficiency, the exponential algorithm proposed by Carol et al. [34] was incorporated in the rate-type law. Meanwhile, in order to tackle the computational instability and non-uniqueness of a discrete Kelvin spectrum, Bazant et al. [35] proposed a continuous retardation spectrum to unambiguously fit the experimental creep curve, which can be discretized and then implemented in a complete creep analysis. Di Luzio et al. [36] systematically performed the rate-type creep analysis methods formed by Refs. [33–35] on 3 typical creep constitutive models (i.e., Eurocode 2, ACI Model, and Model B3), which further validated the effectiveness of rate-type creep model in characterizing the creep behavior of concrete under various loading scenarios. Similarly, Hede-gaard et al. [37] assumed the concrete and reinforcement as a composite material and implemented the rate-type creep law to approximate the creep properties of reinforced concrete. Besides the rate-type method, many models were also proposed to simulate creep/relaxation behavior with emphasis on various influencing mechanisms (i.e., degree of hydration [38], relative humidity [39], loading eccentricity [40] and water-aggregate interaction [41]) and different numerical expressions of the creep constitutive model (i.e., time-varying generalized Maxwell model [42], age adjusted effective modulus [43] and parallel creep

curve [44]).

Except for the modelling approaches that mainly consider a single influencing mechanism, more sophisticated models were developed to partly or fully characterize the interplay between creep, shrinkage/swelling, damage, hydration and heat transfer. Based on experimental observations, time-dependent softening models were developed to simulate the creep failure of concrete under different levels of sustained loads [45–47]. Moreover, multi-physics models [48–53] were developed to incorporate almost all significant aspects of concrete behaviors including creep, shrinkage, hydration and crack. These models coupled the hydro-thermo-chemo-mechanical mechanisms of concrete and therefore could characterize both the short- and long-term behavior of concrete. Specifically, for a practical creep test, which is enviably influenced by complex interplay of shrinkage, damage, and aging, such models will be useful in separating the influence of each individual factor on the creep behavior.

Subsequently, with the advancement of experimental micro-mechanical facilities, abundant creep data gathered from micro-scale tests motivated more micro-scale creep models to spring up, which consider the influence of aggregate [54,55], C–S–H gel [56] and hydration [57]. Recently, with the aim to reproduce the experimentally observed creep behavior of cement paste, a local force -based lattice fracture model was developed to model the creep behavior [58]. Based on the results of X-ray computed microtomography analysis and nano-indentation tests, the experimentally informed model was capable of explaining the experimental results.

To sum up, the sophisticated numerical models described above have not only provided an effective alternative to experimental methods with good precision under certain conditions or assumptions, but also delivered in-depth explanations for various creep mechanisms and interactions between multiple influencing parameters. Nevertheless, given the objectives of predicting the creep behavior of concrete structures, most numerical methods turn out to be inappropriate because of the gap between the ideal conditions presumed by numerical models and practical engineering conditions which also constantly change during the creep process.

1.3. Significance of this study

With the aim to supplement the insufficiency of experimental and numerical methods mentioned above, this study provides a Machine-Learning (ML) solution to predict the creep behavior of concrete materials. As a tributary of artificial intelligence, supervised ML has been considered as a promising data-driven approach to build robust models for predicting various properties and behaviors of heterogeneous materials.

In recent years, many researchers in the field of construction materials have been implementing ML models in prediction of drying shrinkage [59], cracking propagation [60], compressive strength [61–63], elastic modulus [64], breakout capacity [65], shear capacity [66], slump flow [67], flexural strength [68] and interfacial bond strength between FRPs and concrete [69]. In these studies, conventional machine learning (CML) models including artificial neural network (ANN), Gaussian process regression (GPR), support vector machine (SVM), multilayer perceptron artificial neural network (MLP-ANN) and random forests (RF) proved to be effective in making precise prediction based on an existing database. Furthermore, as a well-known approach to reduce the risk of overfitting that brought by CML model, ensemble machine learning (EML) models were adopted in predicting surface chloride concentration [70] and compressive strength [71,72]. The results showed that compared with CML models, EML models could better consider multiple influencing factors and thus had superior prediction performance.

In this study, based on the Northwestern University (NU) creep database [73] that constituted by parameters characterizing different loading conditions, mixture proportions and environmental conditions,

three decision tree-based EML models (i.e., Random Forest (RF), Extreme Gradient Boosting Machine (XGBM) and Light Gradient Boosting Machine (LGBM)) are established to learn the underlying patterns behind 26,458 creep test results. Then, by tuning the hyperparameters with Bayesian Optimization, the performance of 3 EML models is efficiently improved within 60 iterations of a cross validation process. Finally, based on the well-tuned EML models, a game-based theory SHapley Additive exPlanation (SHAP) is implemented to unravel the underlying pattern that EML models reveal from the database, which can provide an in-depth understanding for the prediction of creep behavior.

2. Data preprocessing

This study aims to build EML models by exhaustively learning the creep data from NU database [73], which contains 29,196 observations of creep compliance at different time points. Fourteen parameters are selected as model input: water-cement ratio, aggregate-cement ratio, cement content (kg/m^3), cement type, amount of superplasticizer (kg/m^3), compressive strength at the age of 28 days (MPa), volume-surface ratio, environmental humidity of specimen pre-conditioning (%), concrete age at loading (day), temperature during loading ($^{\circ}\text{C}$), humidity during test (%), creep test type, loading stress (MPa) and time since loading (day).

2.1. Categorical parameters

Compared with other numeric parameters, both cement type and creep test type are both treated as categorical parameters here, which enables this study to cover a wider range of parameters and promote robustness for predictions in practical conditions.

In the category of cement type, the number 1, 2, 3, 4 will be used to classify slowly hardening cement (SL), normal cement (N), rapid hardening cement (R) and rapid hardening high strength cement (RS). While in the category of creep test type, according to the original records of NU database, most creep tests are lying in the categories of “basic”, “basic?”, “total?” and “total”. Due to ambiguous experimental process described by the original papers, the NU database does not provide clear information regarding the type of every creep test. As basic creep refers to time-dependent deformation without the influence of drying creep, fully-sealed conditions are clearly required in basic creep tests. While on the other hand, total creep tests aim to measure all the deformation that induced by both basic and drying creep, which corresponds to no-sealed conditions. Therefore, for the creep test type “basic”, “basic?”, “total?” and “total”, the sealing conditions of fully-sealed, inadequately-sealed, partly-sealed and no-sealed can be reasonably determined, with the two terms “inadequately-sealed” and “partly-sealed” are tentatively assumed to tackle the uncertain annotation of creep test type in NU database. Subsequently, this study labels each creep test type according to their sealing condition, using the number 1, 2, 3, 4 to represent the creep test type of “total”, “total?”, “basic?”, “basic”.

2.2. Missing data

In the selected 14 input parameters mentioned above, much data is missing. Therefore, an appropriate data imputation procedure should be taken to supplement missing data. First of all, to ensure the reliability of data, any observations that have more than 3 missing data has been discarded in this study.

There are mainly two kinds of available strategies to tackle the problem of missing data, namely univariate imputation methods and multivariate imputation methods. The univariate imputation method is a simple approach to impute the missing data with the statistics (i.e., mean, median or most frequent) of each column where the missing data are located. While the multivariate imputation is a more sophisticated approach to impute the missing data by implementing a model to fit the

correlation of all the available data, which can then be used to estimate the missing data point so as to make the process of data imputation unbiased and statistically valid [74–76]. For parameters other than compressive strength, there are little correlations between the missing data and existing data, so a multivariate imputation would be inappropriate. Therefore, these missing values are imputed following the method of univariate imputation method with its median values.

Compressive strength is one of the most important parameters of concrete that can be correlated to multiple concrete properties [23], therefore an accurate estimation for missing compressive strength data will be carried out. The XGBoost model, which is widely reported to have superior performance in predicting compressive strength [77–81], will be established here to conduct imputation of missing compressive strength data. Based on, in total, 3051 sets of compressive strength data (1206 sets from NU database, 1030 sets from Ref. [82], and 815 sets from Refs. [83–91]), the XGBoost model for compressive strength is trained on 2440 samples and tested on the remaining 611 samples. Besides, Bayesian Optimization and 10-fold cross validation [92] will be implemented to tune the hyper parameters of the XGBoost model. The input for prediction of compressive strength are as follows: amount of cement, fly ash, slag, superplasticizer, w/c ratio and curing age. The hyperparameters are shown in Table 1. The training and testing performance evaluated by coefficient of determination (R^2) are shown in Fig. 1. Compared with the other ML-Based compressive strength prediction models [71,72,77–81], the XGBoost model is built on the largest database of concrete strength, and therefore its prediction should be robust and accurate. After the preprocessing, 26,458 observations of creep tests results from NU database are selected as final input data and will be used in prediction of creep, as is shown in Fig. 2.

3. Methodology

For prediction problems involving unstructured data (e.g., image, text), deep artificial neural network tends to outperform all other ML algorithms. However, when it comes to medium tabular data, tree-based ML algorithms remain to be best-in-class. In this study, three kinds of the state-of-the-art tree-based EML models, namely Random Forest (RF), Extreme Gradient Boosting Machine (XGBoost) and Light Gradient Boosting Machine (LGBM) will be built to predict creep compliance of concrete under various experimental conditions. During the process of cross validation, Bayesian Optimization will be implemented to efficiently tune the hyperparameters. Afterwards, with the well-tuned models, Shapley Additive Values will be calculated to characterize the influence of each parameter on the final inference of creep compliance.

3.1. Decision tree-based ensemble machine learning (EML) models

The basic strategy of a decision tree-based EML is to build multiple decision trees based on randomly selected subsamples of the whole data set, and then combine the prediction of each pre-built decision tree estimator to give a new inference corresponding to the input data. Specifically, there are mainly two training methods for EML models, namely parallel training and sequential training. Parallel learning will be used in training of RF, while sequential training will be used for XGBoost and LGBM.

3.1.1. Classification and regression tree (CART)

The classification and regression tree (CART) is implemented here as the basic estimators that will be further ensembled to build the EML model. For regression aims, CART consistently conducts binary partition

Table 1
Hyperparameters of XGBoost for estimating missing compressive strength data.

Parameter type	Max depth	Gamma	Number of trees	Learning rate
Value	154	0.43015	796	0.21746

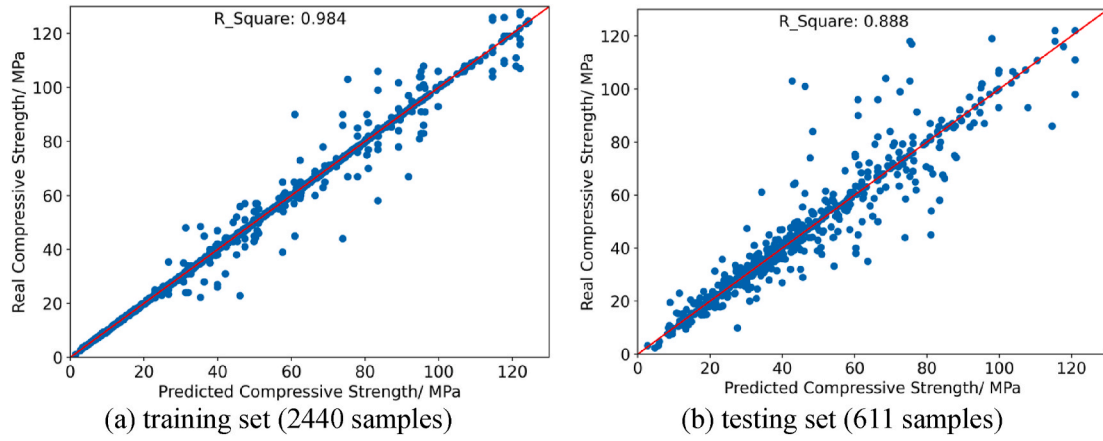


Fig. 1. Performance of XGBoost for imputing missing compressive strength data.

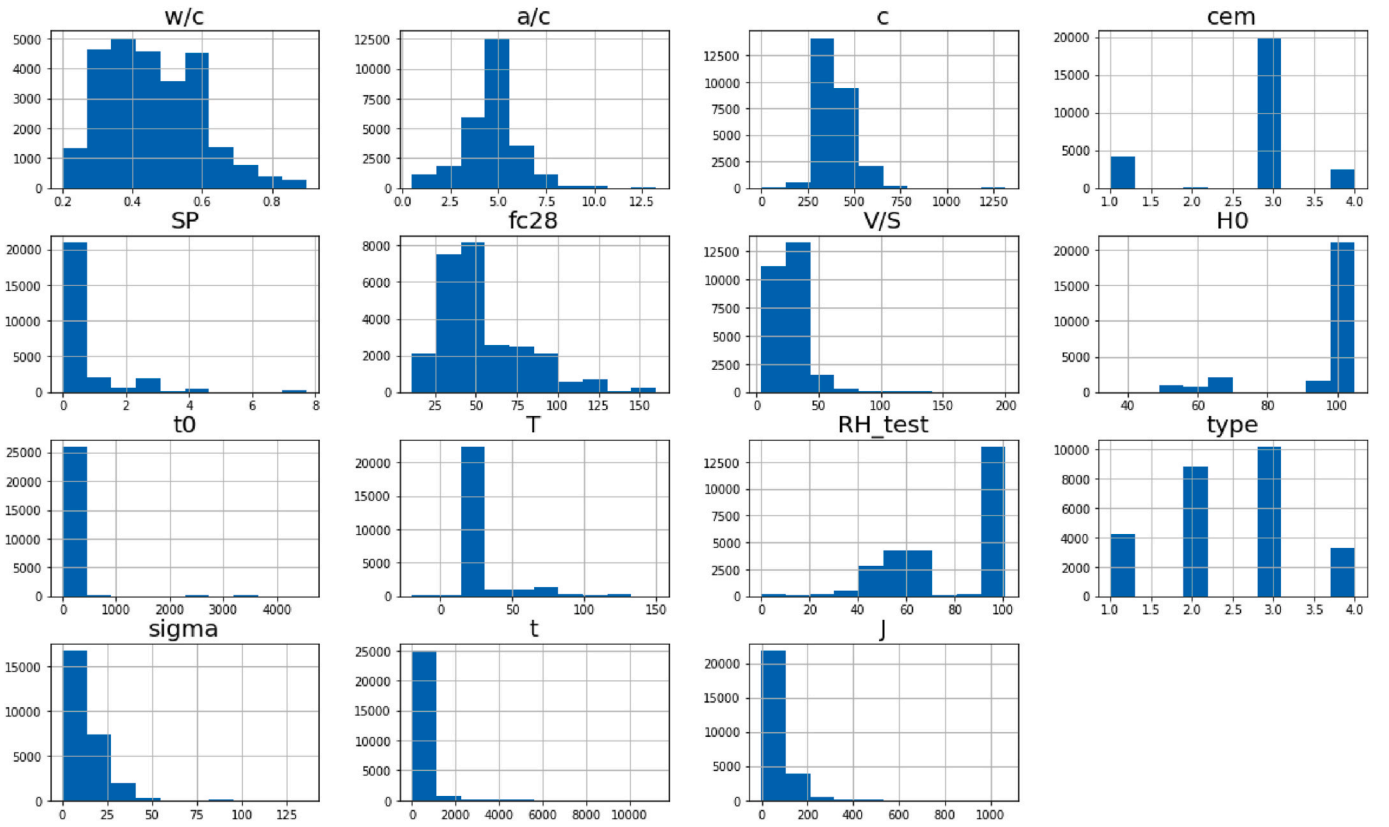


Fig. 2. Histogram of creep data selected from NU database

(The annotations of each parameter are as follows: 1). w

/c: water-cement ratio; 2). a

/c: aggregate-cement ratio; 3). c: cement amount (kg/m³); 4). cem: cement type; 5). SP: amount of superplasticizer (kg/m³); 6). fc28: compressive strength at 28 days (MPa); 7). V/S: volume-surface ratio; 8). H0: environmental humidity of specimen preconditioning; 9). t0: concrete age at loading (day); 10). T: temperature during loading (°C); 11). RH_test: humidity during test; 12). type: creep test type; 13). sigma: loading stress (MPa); 14). t: time since loading (day); 15). J: creep compliance (um/m/MPa)).

for each parameter of the dataset at each level until the prescribed maximum depth is reached [93]. During this process, the input data X will be partitioned at the split point s for splitting variable j :

$$R_1(j, s) = \{X|X_j \leq s\} \text{ and } R_2(j, s) = \{X|X_j > s\} \quad (1)$$

By repeating the binary partition process, the whole dataset will be divided into M regions (i.e., terminal nodes) R_1, R_2, \dots, R_m with each region R_i corresponding to a prediction response c_i . In order to find best

splitting point s for each splitting variable j , the mean squared error is used to build up the cost function:

$$J(j, s) = \min_{c_1} \sum_{x_i \in R_1(j, s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j, s)} (y_i - c_2)^2 \quad (2)$$

For any choice of j and s , the minimization in Eq. (2) is given by:

$$\hat{c}_1 = \frac{1}{N_1} \sum_{x_i \in R_1(j, s)} (y_i)^2 \text{ and } \hat{c}_2 = \frac{1}{N_2} \sum_{x_i \in R_2(j, s)} (y_i)^2 \quad (3)$$

where y_i is the i -th ground-truth value, and N_1, N_2 is the number of data points in region R_1, R_2 . Therefore, the best partition pair of (j, s) can be found by minimizing the cost function of Eq. (2).

An example of CART with a maximum depth of 3 and number of nodes of 4 is shown in Fig. 3. With a proper value of maximum depth, number of nodes and pruning strategy, a CART can serve as a qualified standalone estimator. However, with the intrinsic hierarchical nature of training process, the effect of an error in the top split is propagated down to all of the splits below, which means CART has very high variance and a slight change of input data can greatly change the output [93,94]. As a result, overfitting often happens and thus reduces the generalization capabilities of the CART model.

3.1.2. Random forest (RF)

RF is proposed as a straightforward EML model to tackle the instability of a standalone CART as described in section 3.1.1. The workflow of RF is shown in Fig. 4. Following the parallel learning method, RF uses a bagging strategy (i.e., bootstrap and aggregation) to build a number of CART models based on subsamples randomly selected from the training dataset [95]. Besides, the number of features that used as input for each CART model is also randomized from the original number of features. Finally, by taking averages, summation or following the majority vote, responses of all CART models are combined to provide an inference according to a certain input. Although the two injected randomness can yield the CART with certain decoupled errors, the aggregation procedure of combining all CART can be effective to cancel out most of these errors [96]. Thus, RF can effectively reduce variance at a very small cost of bias.

In this study, based on the Scikit-Learn API [97], RF will be implemented to build a baseline prediction model for concrete creep behavior. The selected hyperparameters for tuning process are `n_estimators` (the number of CART), `max_features` (the maximum size of every random subset of features for training each CART), `max_depth` (the maximum depth for each CART), `min_samples_split` (the minimum number of samples required to split an internal node of each CART) and `min_samples_leaf` (the minimum number of samples required in an internal node).

3.1.3. Extreme Gradient Boosting Machine (XGBoost)

As a counterpart to RF, gradient boosting machine adopts a sequential training strategy to create a strong learner by sequentially training a number of weak learners (i.e., CART in this study). In every step, a weak learner is trained to minimize the loss function by using gradient descent optimization algorithm [93]. XGBoost is recognized as an advanced implementation of gradient boosting machine, which can more effectively control over-fitting by using a more regularized model formation [98]. The objective function during the training process of

XGBoost is expressed as:

$$obj(t) = \sum_{i=1}^n L(y_i, \hat{y}_i^t) + \Omega(f_t) \quad (4)$$

in which L is the loss function; Ω is the regularization term for measuring the model complexity; t is the t -th iteration for generating t -th CART; n is the total number of CART; f_t is the mapping function of t -th CART. The model complexity of each CART can be expressed as:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (5)$$

in which T is the total number of nodes, w is the score of each node, γ and λ are the penalty terms for regularization to avoid over-fitting. When t -th CART is created to fit the residual error of the former CARTs, the prediction of the new tree can be expressed as:

$$\hat{y}_i^t = \hat{y}_i^{t-1} + f_t(x_i) \quad (6)$$

Then the objective function can be written as:

$$obj(t) = \sum_{i=1}^n L(y_i, \hat{y}_i^{t-1} + f_t(x_i)) + \Omega(f_t) \quad (7)$$

Expanding f_t with Taylor polynomial and assuming the loss function L uses the form of mean square error, the approximation of objective function with second-order accuracy can be expressed as:

$$obj(t) = \sum_{i=1}^n \left(L(y_i, \hat{y}_i^{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right) + \Omega(f_t) \quad (8)$$

in which g_i and h_i denote the first- and second-order derivative:

$$g_i = \partial_{\hat{y}_i^{t-1}} L(y_i, \hat{y}_i^{t-1}) \text{ and } h_i = \partial_{\hat{y}_i^{t-1}}^2 L(y_i, \hat{y}_i^{t-1}) \quad (9)$$

Because the value of objective function depends only on g_i and h_i , a customized loss functions can be supported by XGBoost. In this study, the mean square error is used as the loss function. Using the XGBoost API [98], the selected hyperparameters for tuning process are `gamma` (minimum loss reduction required to make a further partition on a leaf node of a CART), `n_estimators` (the number of CART), `learning_rate` (boosting learning rate), `max_depth` (the maximum depth for each CART) and `min_child_weight` (minimum sum of instance weight needed in a node).

3.1.4. Light Gradient Boosting Machine (LGBM)

As a major challenger to XGBoost, LGBM was developed to achieve faster training speed and lower memory usage, while maintain high accuracy. The main difference between LGBM and XGBoost lies in their ways of growing trees, as is shown in Fig. 5. XGBoost, as well as most

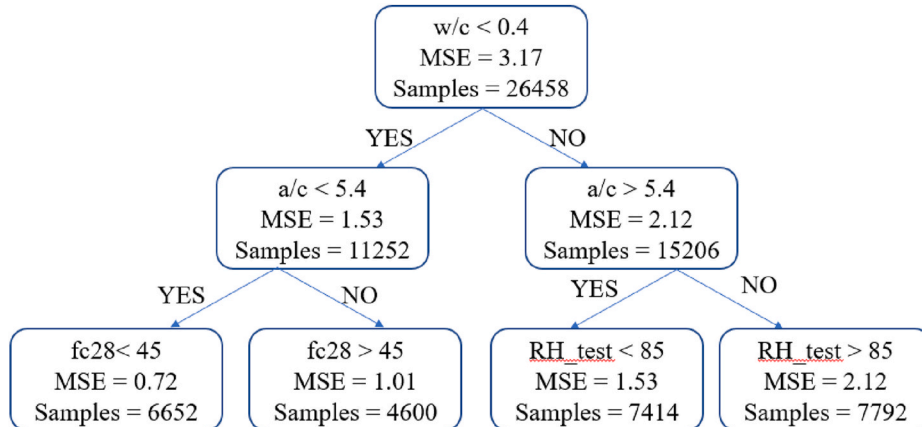


Fig. 3. An example of CART.

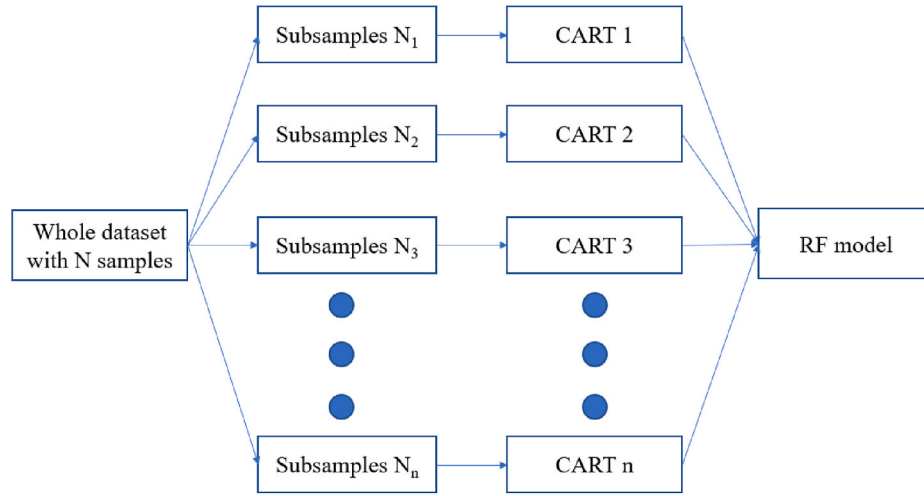


Fig. 4. Workflow of random forest.

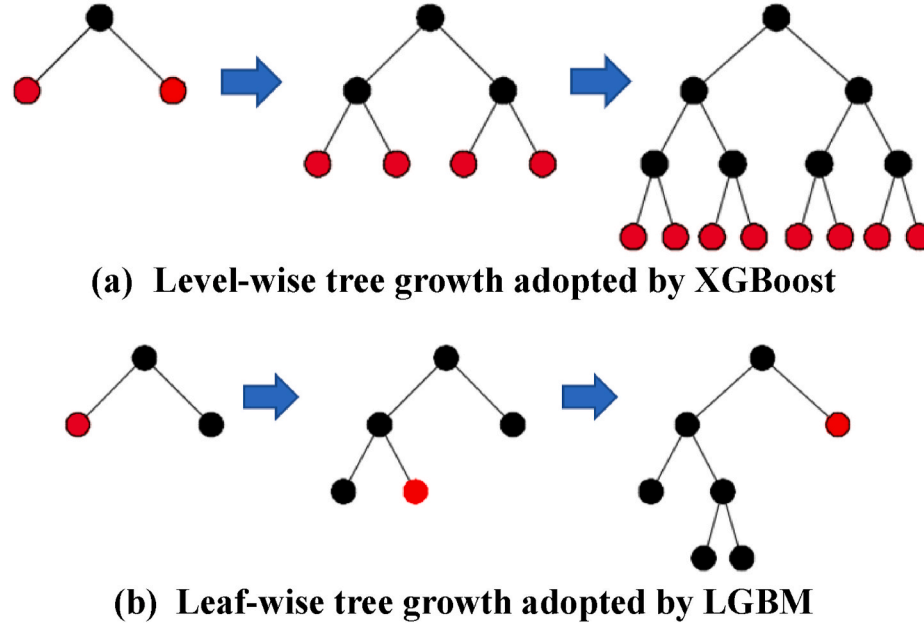


Fig. 5. Different ways for growing trees between XGBoost and LGBM.

gradient boosting algorithms, grows trees level-wise following the splitting procedure of CART, as described in section 3.1.1. While LGBM grows tree leaf-wise by splitting the leaf node using a histogram-based algorithm, which yields great advantages on both efficiency and memory consumption. As leaf-wise tree growth leads to more model complexity, LGBM can result in more accuracy gain in every iteration. However, it also means high risks of over-fitting, which is tackled by regularization terms. Two techniques, namely gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB), are also implemented to make the LGBM a fast, efficient and stable EML algorithm [99]. During the training process, because data samples with high gradient will contribute more to the information gain, GOSS selects the samples with high gradient and discards the ones with low gradient. On the other hand, EFB bundles exclusive features in sparse feature space, thus reduces the dimensionality and improve efficiency.

In this study, the API developed in Ref. [99] is used to build a LGBM model based on the NU database. The selected hyperparameters are num_leaves (the maximum tree leaves of each CART), max_depth (the maximum depth of each CART), n_estimators (number of boosted CART

to fit), learning_rate (boosting learning rate) and min_child_samples (minimum number of samples needed in a leaf node).

3.2. Bayesian Optimization

In most supervised learning tasks, a bias-variance tradeoff has always been one of the most tricky and essential objectives, since it represents the generalizing capabilities of the model to predict unseen new input data. To achieve the best bias-variance tradeoff, hyperparameters that represent the complexity and the training step of EML models should be well tuned to get highest prediction accuracy in the process of cross validation. For such optimization aim, methods like grid-search and random-search are broadly implemented by enumerating all combinations of presumed values for each hyperparameter [59–72]. For example, for the 5 hyperparameters selected in LGBM mentioned in section 3.1.4, if 3 values are assigned to each hyperparameter, then there is $3^5 = 243$ combinations of hyperparameters being required to be implemented in 243 times of training process. Furthermore, if a 10-fold validation is required, then 2430 times of training will be required to be

implemented. Therefore, grid-search and random-search can be highly inefficient and time-consuming when it comes to complex models with parameters of high dimensionality and large amount of data. Besides, such methods strongly depend on experience of the programmer, since only limited number of presumed values can be pre-assigned to each hyperparameter. In view of the limitations of grid-search and random-search, this study proposes to adopt Bayesian Optimization to find the optimal hyperparameters corresponding to the highest prediction accuracy in the process of cross-validation. The process of Bayesian Optimization mainly includes two components: a statistical model for modelling the objective function and an acquisition function for the next sampling point. The workflow of Bayesian Optimization used in this study is given in Fig. 6.

In Bayesian Optimization, a fundamental assumption is that the optimization objective follows the multivariate Gaussian distribution (i. e., $Y(X) \sim N(0, K)$, where Y is the prediction performance in cross-validation, while X is the hyperparameter vectors). Then the Gaussian Process (GP) regression can be implemented to provide a prior for the objective Y , which forms the GP model mentioned in Fig. 6. In GP, the kernel for the covariance term that is corrupted by noise with zero mean and standard deviation σ_{noise} can be expressed as:

$$K = \begin{bmatrix} k(X_1, X_1) & \cdots & k(X_1, X_n) \\ \vdots & \ddots & \vdots \\ k(X_n, X_1) & \cdots & k(X_n, X_n) \end{bmatrix} + \sigma_{\text{noise}}^2 \mathbf{I} \quad (10)$$

in which k is the covariance function that is calculated by an exponential function of the second norm of the difference value between two samples; n is the number of samples that are incorporated in the GP model. Under the assumption of Gaussian distribution, the distribution of the observation of next sample Y_{n+1} can be expressed as a joint Gaussian distribution with the observations in the first n iteration denoted as $D_{1:n} = (X_{1:n}, Y_{1:n})$ [100]:

$$Y_{n+1}|D_{1:n} \sim N(\mu(X_{n+1}), \sigma^2(X_{n+1}) + \sigma_{\text{noise}}^2) \quad (11)$$

in which

$$\mu(X_{n+1}) = k^T (K + \sigma_{\text{noise}}^2 \mathbf{I})^{-1} Y_{1:n} \quad (12)$$

$$\sigma^2(X_{n+1}) = k(X_{n+1}, X_{n+1}) - k^T (K + \sigma_{\text{noise}}^2 \mathbf{I})^{-1} k \quad (13)$$

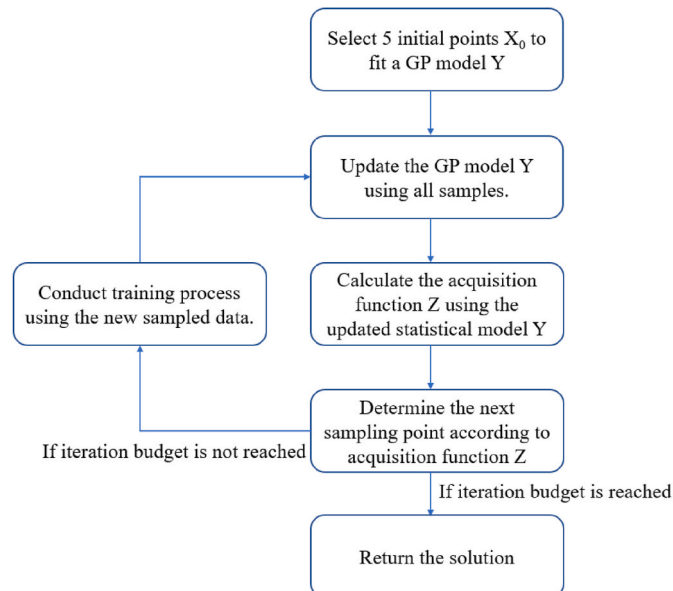


Fig. 6. Workflow of bayesian optimization.

$$k = [k(X_{n+1}, X_1) \ k(X_{n+1}, X_2) \ \cdots \ k(X_{n+1}, X_n)] \quad (14)$$

Thereby, with the GP model, a mapping from the selected hyperparameters to the prediction performance in cross-validation can be drawn, with a credible interval for every inference. Then, an acquisition function will be implemented to calculate the improving potential of every sample point. In this study, the Expect Improvement (EI) is adopted as acquisition function to achieve the balance between exploitation and exploration, which can be expressed as [101]:

$$EI(X_n) = (y_{\text{best}} - \mu(X_n)) \Phi\left(\frac{y_{\text{best}} - \mu(X_n)}{\sigma(X_n)}\right) + \sigma(X_n) \varphi\left(\frac{y_{\text{best}} - \mu(X_n)}{\sigma(X_n)}\right) \quad (15)$$

where $\Phi(\cdot)$ and $\varphi(\cdot)$ are the cumulative distribution function and probability distribution function of the standard Gaussian distribution; y_{best} is the tentative optimal value in current sample space. Note that the first term in Eq. (15) aims to exploit every sample point by evaluating the difference between the potential sample point and the best point in current sample space. The second term in Eq. (15) aims to quantify the uncertainty of each sample point. Therefore, the EI function tends to get higher values at sample points with larger posterior mean and credible intervals, and thus the balance of exploitation and exploration can be achieved.

4. Results and discussion

To evaluate the generalization capabilities of the 3 EML models, based on a pseudo random-number generator, the database is shuffled first and then divided into the training and testing set following the ratio of 8:2. For each EML model, 5 hyperparameters will be selected and tuned by Bayesian Optimization in the process of 5-fold cross validation [92,102]. And then, each tuned EML model will be examined by the 5292 unseen samples within the testing set. Furthermore, the performance of EML models on testing set will be compared with the broadly-used empirical model in the *fib* Model Code 2010 [103]. Finally, the game-based theory Shapley Additive Explanation will be implemented on each EML model to provide explanations for its prediction.

4.1. Model performance

The performance of the EML models will be evaluated using the following 3 metric indexes: coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE), whose formulations are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (16)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (17)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (18)$$

where x_i is the creep compliance of i -th sample point in the database; \hat{x}_i is the prediction value made by EML models for i -th sample point; \bar{x} is the averaged value of the creep compliance. Both MAE and RMSE explicitly characterize the residual error at each sample point and can give exact evaluation of model performance. In comparison, R^2 normalizes the squared residual error with the variance of the database and produces dimensionless score ranging from 0 to 1. By comparison, R^2 is more intuitive and convenient for comparing the performance of different models. Therefore, it's adopted as the main metric index in the following analysis.

Denote the black-box function represented by each EML model as f ,

which uses a hyperparameter vector θ_i as the input and then gives an output of the model performance indexed by the averaged R^2 of 5-fold cross-validation. Thereby, the form of the objective function of the Bayesian Optimization is given as follows:

$$\theta_{best} = \operatorname{argmax}_{\theta} f(\theta) \quad (19)$$

4.1.1. Optimization result and EML model comparison

Following Eq. (19), the performance of 3 EML models in the process of 5-fold cross-validation is optimized. For each EML model, 5 iterations are firstly conducted to form an initial sample space, which establishes the initial GP prior. Then in the following iterations, the acquisition function EI is calculated to select new sample point and then update the GP prior at the end of each iteration. According to Ref. [104], at least 30 iterations is recommended for most optimization aims. In this study, 60 iterations will be implemented to find the best combination of hyperparameters for each EML model. The optimizing history of the 3 EML models is shown in Fig. 7. With Bayesian Optimization, the 3 EML models successfully obtain high prediction accuracy measured by R^2 in the process of 5-fold cross-validation. However, there are distinct differences lying between them:

- 1) Compared with both RF and XGBoost, LGBM obtains highest prediction accuracy within significantly shortest calculation time, which is about 2.7% of the XGBoost and 3.4% of the RF. The significant improvement in computational efficiency as well as accuracy is attributed to the leaf-wise tree growing strategy, gradient-based subsampling method GOSS and feature bundling method EFB, according to Section 3.1.4.
- 2) Compared with RF, the two gradient boosting models XGBoost and LGBM shows much higher accuracy at the beginning of Bayesian Optimization process and ends with limited improvement of R^2 ,

which means that both XGBoost and LGBM are much more robust than RF and therefore require less effort in hyperparameter tuning.

The optimized hyperparameters for each EML model are listed in Table 2. The range of each hyperparameter is selected according to references [97–99], which is listed beneath the name of each hyperparameter in the form: (lower boundary, upper boundary). Note that although with different names, the parameters `min_samples_leaf` of RF, `min_child_weight` of XGBoost and `min_child_samples` of LGBM all refer to the minimum number of samples required in the leaf node of CART, which controls the complexity of the basic estimators and thus reduces the overfitting risk for each EML model. As shown in Table 2, all 3 well-tuned EML models show preference to larger amounts of basic estimators, which is logical since ensemble strategies (bagging for RF and gradient boosting for XGBoost and LGBM) play as inherent advantage to gain high accuracy for EML models. However, a significant difference between the ensemble strategies lies in their preference for complexity of basic estimators. For RF, which adopts bagging as the ensemble strategy, more complex basic estimators with more depth and smaller leaf nodes are preferred. As a standalone estimator, such complexity is highly prone to overfitting due to its high variance and low bias, especially for CART which has a hierarchical decision path and therefore is sensitive and unstable to fluctuation of input parameters. However, through bagging, the variance of all these complex estimators can be effectively counteracted and therefore produce a strong ensemble estimator. On the other hand, for gradient boosting models XGBoost and LGBM, a preference for less complex basic estimators (i.e., lower depth and large nodes) is observed. Less complex basic estimator, which has high bias and low variance, is prone to underfitting and therefore is called a weak learner. However, by gradient descent methods and strong regularization procedures, gradient boosting method can train the weak learners sequentially to compensate the residual errors of the former

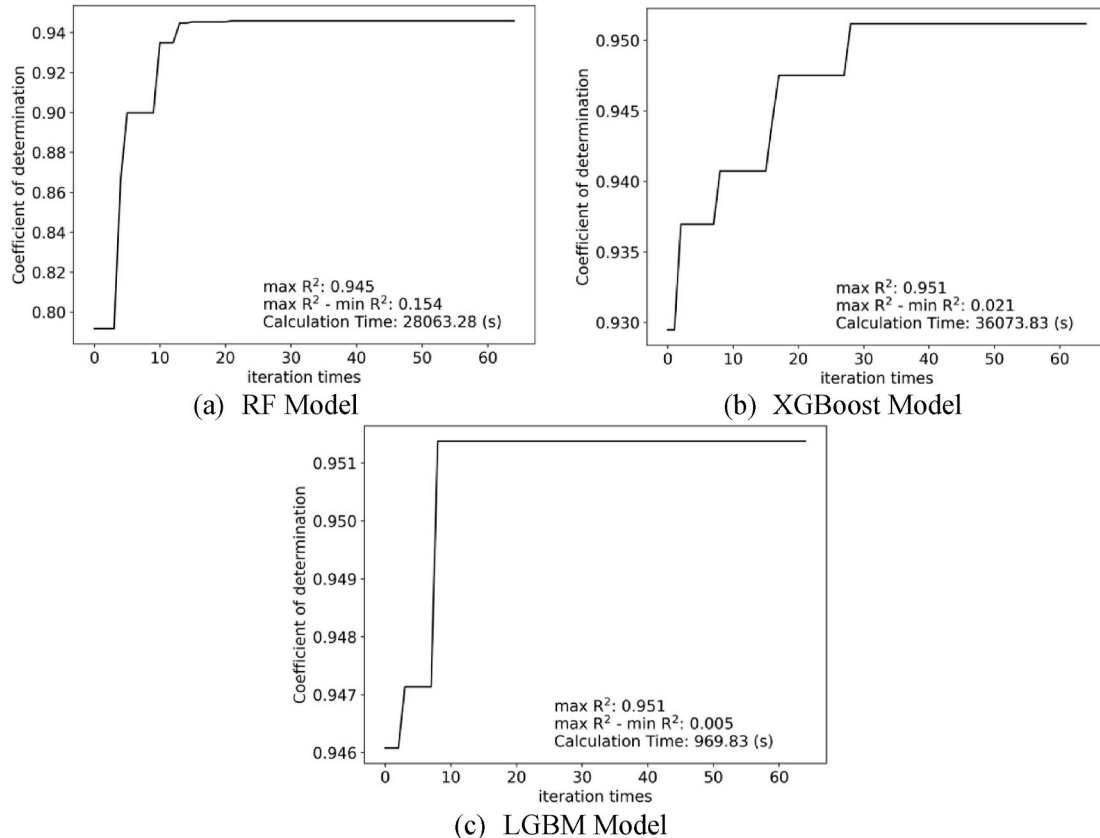


Fig. 7. Optimizing history for 3 EML models.

Table 2
Optimized hyperparameters of 3 EML models.

RF	max_depth (1,100)	n_estimators (1,3000)	max_features (7,14)	min_samples_leaf (2,35)	min_samples_split (2,60)
	73	2640	11	2	2
XGBoost	max_depth (1,100)	n_estimators (1,3000)	learning_rate (0.001,0.3)	min_child_weight (2,35)	Gamma (0.1,0.9)
	20	2799	0.0076	32	0.7516
LGBM	max_depth (1,100)	n_estimators (1,3000)	learning_rate (0.001,0.3)	min_child_samples (2,35)	num_leaves (5,50)
	22	2578	0.0333	24	40

ones, and therefore produce a strong learner at last.

4.1.2. Testing results and comparison with model code 2010

Using the 5292 data that was separated beforehand from the original database and therefore have never been exposed to the EML model, the generalization capabilities of the EML models can be examined. The validation results are shown in Table 3. As can be expected from the results of 5-fold cross validation, 3 EML models all obtain superior accuracy in both the training and the testing set. Because only minor discrepancies are observed between the accuracy of training set and testing set, the extent of overfitting is marginal and therefore and be reasonably neglected. Comparing the performance of the 3 EML models, RF and XGBoost obtain almost the same accuracy, while LGBM outperforms the others in both the training and the testing set in most accuracy indexes.

To further validate the performance of the EML models, the broadly-used empirical prediction model of creep in the Model Code 2010 (MC2010) [103] is also implemented on the testing set. There are 7 input parameters considered by MC2010, namely cement type, compressive strength at 28 days, concrete age at loading, volume surface ratio, relative humidity of the ambient environment, loading time and temperature. The comparison between MC2010 and the proposed EML models are shown in Fig. 8. The results show that EML models outperform MC2010 in the testing set. An intrinsic reason accounting for such gap lies in the sophisticated statistical advantages of EML models, which guarantees a balance between bias and variance during training process. Besides, another reason is the number of parameters that are taken into consideration. In EML models, there are 14 parameters being used for training, which is twice as many as MC 2010. Despite the significant gap when compared to EML models, MC2010 still shows certain precision in most data points, which can be attributed to the correct selection of the 7 input parameters. In the following section, 5 of the 7 input parameters of MC 2010 will be proved to be the most influential ones in the prediction of EML models.

4.2. Shapley Additive Explanations (SHAP)

Although many ML studies in concrete materials have achieved high accuracy in predictions of their objectives, inadequate attentions are paid to the interpretability of the ML models. For tree-based models, many studies calculate the feature importance based on the decision path, heuristic approaches, or model-agnostic approaches [105]. However, these methods are often infeasible and biased for EML models, especially for those ones with high bias [106]. In this study, the SHapley Additive exPlanation (SHAP) is adopted to illustrate both local and

global interpretation of each input parameter.

4.2.1. Overview of SHAP

Based on the cooperative game theory, SHAP is represented by the average marginal contribution of a feature value across all possible coalitions. Specifically, the SHAP value of a feature is the averaged prediction value of the samples with this feature minus the averaged prediction value of the other samples without this feature. To provide interpretability of an ML model, the output of the ML model is expressed by a linear addition of its input features multiplied by corresponding SHAP values:

$$f(x) = \varphi_0 + \sum_{i=1}^N \varphi_i X_i' \quad (20)$$

where f is the mapping function represented by the ML model; N is the number of input features; φ_0 is the average of all predictions; φ_i is the SHAP value for i -th feature; X_i' is the coalition vector of the i -th feature, which can be calculated from the original input X_i by a mapping function expressed as $X_i = h_x(X_i')$. Lunderberg et al. [107] proposed a unique solution for Eq. (20) based on the 3 desirable preconditions: local accuracy, missingness and consistency. Local accuracy ensures that the additive relation expressed by Eq. (20) is achieved. Missingness ensures that all missing values get a SHAP value of 0. And consistency ensures that changing a ML model does not change the relative SHAP value of that feature, which means that assuming $z' \setminus j = 0$ and $f_x(z') - f_x(z' \setminus j) \geq f_x(z') - f_x(z' \setminus j)$, then $\varphi_i(f', x) \geq \varphi_i(f, x)$. Thereby the unique solution to the SHAP value of a feature value is weighted and summed over all possible feature value combinations:

$$\varphi_i(f, x) = \sum_{z \subseteq z'} \frac{|z'|!(N - |z'| - 1)!}{N!} [f_x(z') - f_x(z' \setminus i)] \quad (21)$$

where $|z'|$ is the number of non-zero entries in z' . For computational efficiency, Lundberg et al. [94] estimated that $f_x(z') = f_x(h_x(z')) = E[f(z) | z_s]$, with S being the non-zero indexes in z' . In this study, the SHAP value of each input features are calculated on the testing set for each tree-based EML models based on the API [108].

4.2.2. Local interpretation

Local interpretation aims to give explanation for prediction of each individual sample. As is given by Eq. (20), the linear addition of SHAP values forms the output of each sample data, and therefore the SHAP value of each input parameter can characterize the contribution of each feature. As is shown in Fig. 9, based on LGBM, 3 samples that represent 3 typical scenarios are selected to be explained by their SHAP values. With a fixed base value of 68.59, the final prediction for creep compliance is the counteracting result of SHAP values of different features. Force plot is used here to visualize the counteracting process, in which red and blue bars present positive and negative contribution to the magnitude of prediction value, respectively.

In scenario 1 as shown by Fig. 9 (a), although beginning from the base value of 68.59, the final creep compliance predicted by LGBM is 99.06. As shown by the force plot of scenario 1, the most significant reason accounting for this prediction result is the high temperature ($T = 80^\circ\text{C}$), which is believed to induce desorption of water from the gel and

Table 3
Comparison of EML model performance.

RF	Training Set			Testing Set		
	RMSE	MAE	R ²	RMSE	MAE	R ²
	9.780	3.009	0.977	14.836	5.207	0.946
XGBoost	RMSE	MAE	R ²	RMSE	MAE	R ²
	10.751	3.789	0.972	14.816	5.409	0.947
LGBM	RMSE	MAE	R ²	RMSE	MAE	R ²
	9.761	3.469	0.977	13.901	4.830	0.953

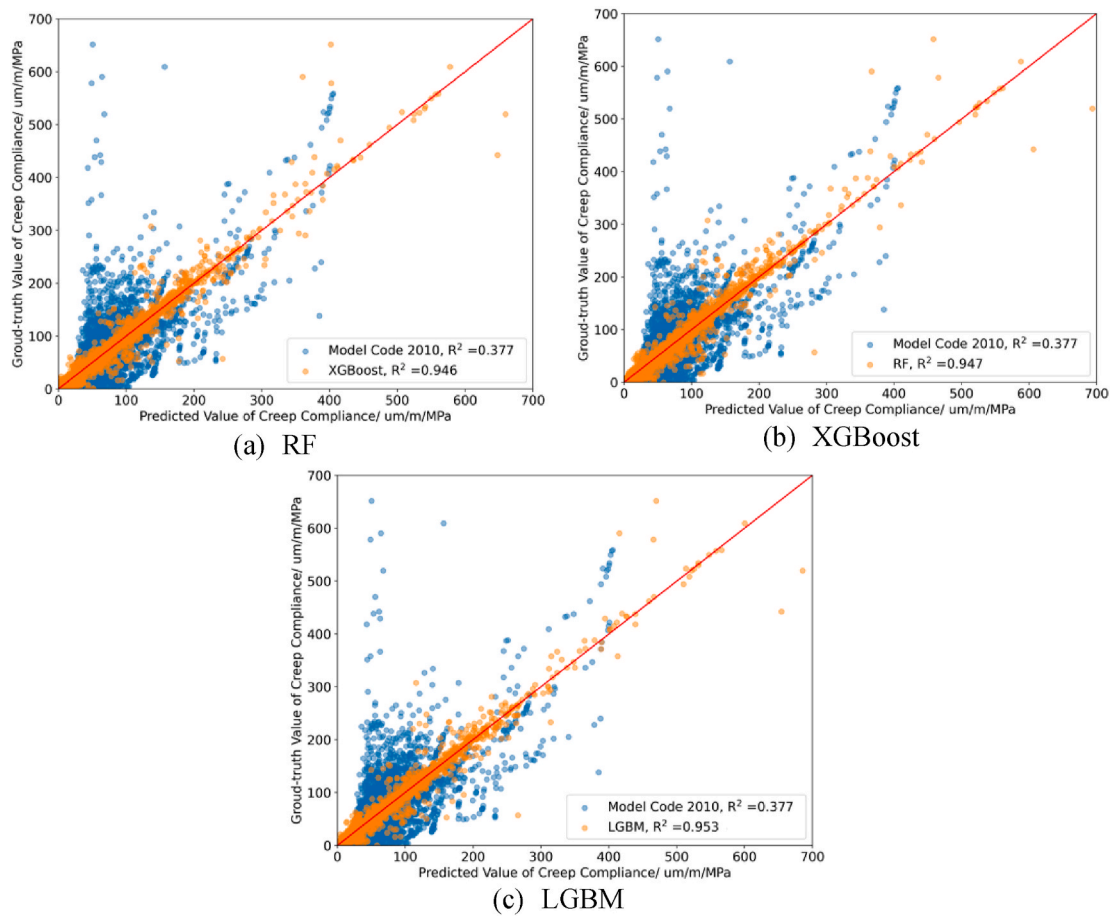


Fig. 8. Prediction of EML models and Fib Model Code 2010.

therefore the gel gradually becomes the sole phase subject to molecular diffusion and shear flow [23]. Another feature that is of positive contribution for scenario 1 is that drying-related parameters volume-surface ratio ($V/S = 30$) and relative humidity ($RH_{test} = 99\%$). High V/S means that less surface area of the specimen is exposed to the environment, which indicates less drying creep and therefore the magnitude of predicted creep compliance decreases. On the other hand, high RH_{test} can reduce the drying creep and hence is identified as a negative contribution in scenario 1. Besides, loading age (t_0) is considered as a negative contribution since scenario 1 has a late loading age ($t_0 = 90$ days).

However, in scenario 2 as shown by Fig. 9 (b), t_0 becomes the most significant positive contribution since its value 0.65 indicates loading at very early age, when the degree of hydration is low and the gel micro-structure is weak, indicating more deformation under the action of sustained loading. On the other hand, the leading negative contribution of scenario 2 is high strength at 28 days ($fc_{28} = 122$ MPa), which means that in the scenario of early age loading, high strength concrete has a lower creep compliance than normal concrete.

In scenario 3 as shown by Fig. 9 (c), when the loading stress is applied at the age of 28 days ($t_0 = 28$ days), high compressive strength ($fc_{28} = 97.3$ MPa) is identified as the most significant contribution to reduce the magnitude of predicted creep compliance. Besides, as mentioned in scenarios 1 and 2, normal temperature ($T = 20$ °C) and high relative humidity ($RH_{test} = 101\%$) are both considered as negative contribution in scenario 3. Note that the value 101% of RH_{test} means a steam experimental condition.

4.2.3. Global interpretation

Global interpretation aims to provide an overview of the SHAP

values for input parameters of all required samples. In this section, SHAP values of 5292 samples in the testing set are calculated based on the 3 EML models. The SHAP values of each feature are shown in Fig. 10, in which the features are sorted in the order of their averaged SHAP value. Therefore, the features listed on the top contribute more to the final prediction results of their corresponding samples. Comparing the results of the 3 EML models, 5 features are identified as the most significant input parameters influencing the prediction results of creep compliance: time since loading (t), compressive strength (fc_{28}), age when loads are applied (t_0), relative humidity during the test (RH_{test}) and temperature during the test (T). Moreover, for the 5 influential features, boundaries of feature values that can separate negative and positive contribution can be explicitly observed. The influence of the top 5 features on final prediction of creep compliance illustrated by global SHAP values here is consistent with what has already been discussed in local interpretation (see Fig. 9). Besides, for the influence of cement type (cem), the boundaries between the negative and positive contribution are not clear, which mainly due to: 1) limited influence of cement type on creep and 2) imbalanced distribution of the input parameter (see Fig. 2). While for the influence of creep test type ($type$), which corresponds to the sealing degree (described in detail in section 2.1), a very clear boundary can be observed. The distribution of SHAP values calculated by the 3 EML models for the input parameter creep test type ($type$) points out that high sealing degree can reduce the magnitude of creep compliance, which is exactly consistent with most observations in basic creep tests and total creep tests and can be explained by the influence of drying creep [9].

In order to look deeper into how a precise prediction is given by the EML model, the SHAP values of the most influential parameters need to be further analyzed. Although there is strong consensus between the 3

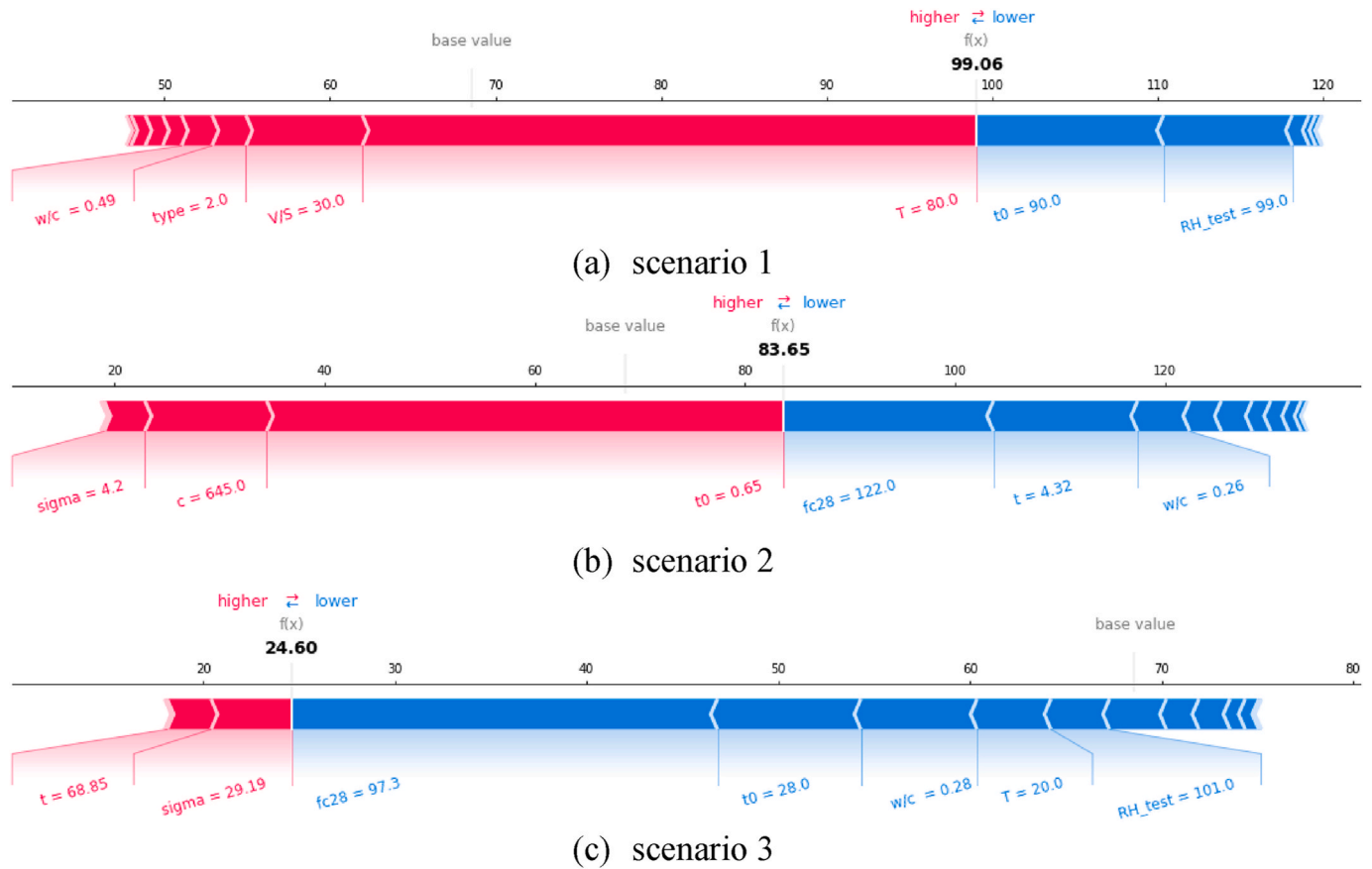


Fig. 9. Force plot for local interpretation of 3 typical scenarios based on LGBM.

EML models on the importance ranking and distribution of the SHAP values of input parameters, a slight difference between the SHAP results of the EML models indicates the importance of base model selection. Besides, according to Eq. (21), the calculation of SHAP values requires to enumerate all the possible coalition of the input features, which can be very time-consuming. Therefore, the SHAP results of LGBM, which shows the best accuracy and efficiency as mentioned in section 4.1, is adopted to show how the 6 most influential factors affects the final prediction of creep compliance (Fig. 11). The overall patterns of how the 6 parameters influence the creep compliance captured by the LGBM model are: 1) The increase of compressive strength, concrete age, and relative humidity will decrease the (predicted) creep compliance; 2) The increase of loading time, temperature and water-cement ratio will increase the predicted creep compliance. Moreover, some more specific patterns can also be captured: 1) The influence of loading time and concrete age will decrease as their values increase, which is in accordance with power function shape of creep compliance curve and hydration degree evolution in early-age stages [38]; 2) the influence of temperature becomes significant when the temperature reaches 70 °C [23]. Thereby, based on common knowledge of how concrete creep is influenced by various factors [23], the reasonability of the LGBM can be examined.

4.3. Exemplification of typical scenarios

To further exemplify the applicability of the proposed EML models for predicting the creep behavior of concrete, this section selects 6 typical scenarios from the testing set and makes predictions of the corresponding creep compliance ranging from 1 to 12,000 days (i.e., 32.8 years). The details of 6 selected scenarios (S1–S6) are shown in Table 4. The selected scenarios cover creep test results of the frequently-used w/c

ratio of 0.4–0.6 under a temperature of about 20 °C. S1 and S2 represent the total creep tests starting from the concrete age (t_0) of 28 days with loading durations of 1944 and 1278 days, respectively. S3 and S4 represent the basic creep tests starting from the concrete age of 7 days with the same loading duration of 28 days. S5 and S6 represent the basic creep tests starting from the concrete age of 14 days with loading durations of 11,076 and 4992 days, respectively.

4.3.1. Exemplification results

Based on the LGBM model, the predictions of the 6 selected scenarios are shown in Fig. 12. As comparison, the predictions made by MC2010 are also plotted. As reported by Refs. [109–111], the creep compliance curve can be effectively characterized by power law function and logarithmic function. Thereby, the predictions made by LGBM will also be fitted by the power law function and logarithmic function, as shown in Eq. (22) and Eq. (23):

$$J_{\log}(t, t_0) = a * \ln[(t - t_0) + c] + b \quad (22)$$

$$J_{\text{power}}(t, t_0) = a * (t - t_0)^c + b \quad (23)$$

where a , b , c are fitting parameters. The fitting procedure is based on Least Square Method and is solved by Levenberg-Marquardt algorithm, which is robust and efficient for solving unconstrained problems [112]. As shown in Fig. 12, for each scenario, the prediction results of LGBM are in conformity with the experimental results. In S1, S2, S5, and S6, where the experimental data cover wide time ranges, such conformity is much more obvious. Meanwhile, as already validated in Section 4.1.2, the prediction accuracy of LGBM significantly outperforms that of MC2010. Obvious gaps between MC2010 and experimental results can be seen in all selected scenarios. In S2, although the MC2010 shows certain accuracy in first days, it still cannot capture the long-term development of

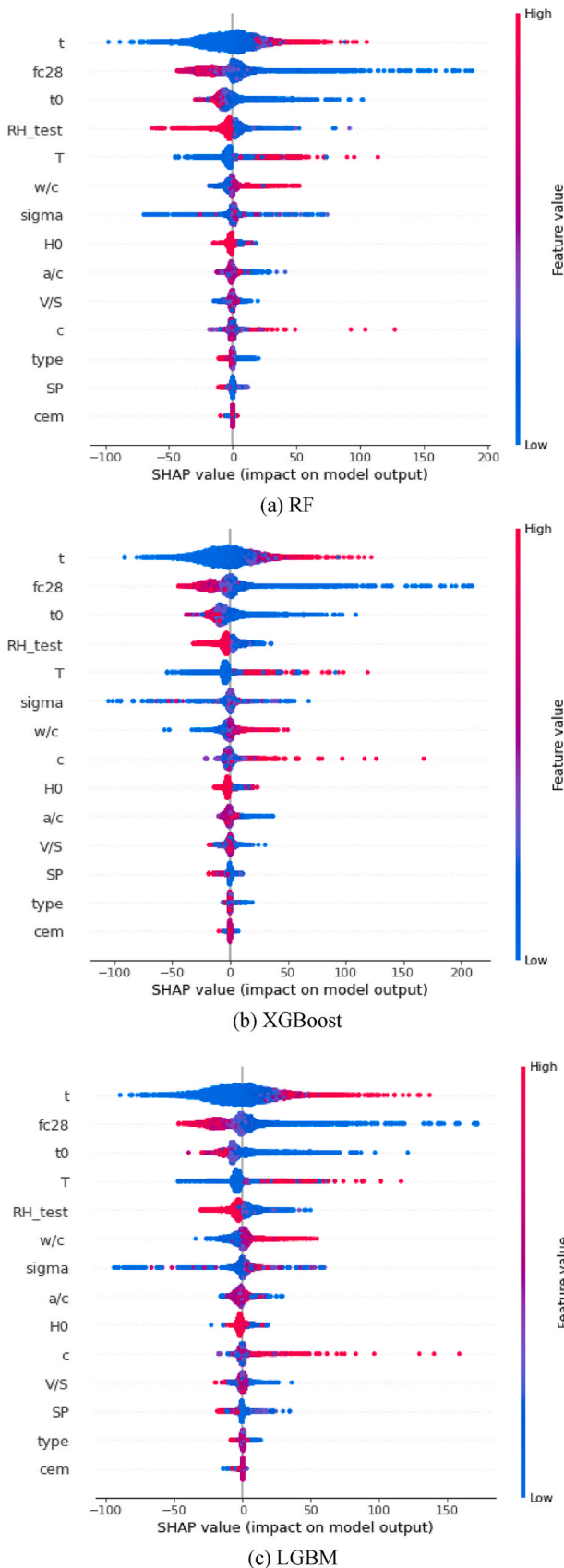


Fig. 10. Global SHAP values based on the 3 EML models.

creep compliance in later stages. By comparing the fitting results, both logarithmic and power functions can well characterize the prediction results of LGBM, with the R^2 all above 0.965. In S1–S4, the logarithmic function shows slight advantages compared to power function, but in S5–S6 the power function prevails with higher R^2 .

4.3.2. Limitations and discussion

As shown in Fig. 12, non-smooth creep compliance curves predicted by LGBM are observed. This is inherently due to the nature of the input database in the following 3 aspects: 1) The data of creep tests in the NU database are formed by scattered points distributed in different time ranges and characterized by multiple mixture parameters and experimental conditions, as shown in Fig. 2; 2) Derived from creep tests that last days, months or years, errors inevitably happen to the experimental results and therefore the scattered points often do not follow a presumed smooth curve, as shown in Fig. 12; 3) Some information is missing in the NU database, as processed in Section 2.1 and 2.2, which can also influence the prediction results. Because the performance of the EML models highly depends on the input database, the scattered, incomplete and possibly abnormal data can result in local numerical instability of EML models in prediction of the creep compliance at a long-term continuous time range. In view of the limitations mentioned above, 3 possible measures can be taken to promote the applicability of ML approaches in prediction of creep compliance:

- 1) Improve the quality of the database before training. This can be achieved by implementing a thorough data cleaning procedure to identify anomalies and interpolate the missing values. However, large-scale data cleaning procedure requires a highly-accurate algorithm, or the quality of the database may be significantly reduced. A potential method is to fit all the creep test data with a predefined function (e.g., logarithmic or power functions), and then using the fitting parameters as output of ML models. However, this can be risky because the fitting results are already biased (by the predefined function) and may often be very different from the original testing results. In this study, only slight and careful preprocessing measures have been implemented to tackle the missing information while maintaining the originality of the NU database (as discussed in Section 2.1 and 2.2).
- 2) Increase the robustness of ML models to decrease the influence of unstructured and possibly abnormal data points. This can be achieved by the “bagging and bootstrap” procedure typically used in EML models, which are reported to be highly effective in prevent overfitting (as discussed in Section 3.1). Besides, instead of training the ML models on scattered points as this study, another potential method can be to treat the experimental results as time-series, although the unstructured database may pose even more difficulties in this case.
- 3) Keep the originality of the database and filter the prediction results after training. This can be achieved by simply fitting the prediction results of EML models with a predefined function. In this study, the logarithmic and power functions are fitted by the prediction of LGBM. As shown in Fig. 12, these fitted results still preserve high accuracy when compared with original experimental results.

Overall, this study mainly adopts measure 2 and 3: the advantages of EML models are taken to prevent overfitting and obtain high accuracy through Bayesian Optimization. This is validated by both cross-validation and SHAP. For further application, the final results on a long-term continuous time range should be processed by a proper fitting function (e.g., logarithmic or power functions). As for measure 1, this study does not implement large-scale data cleaning procedures, while only conduct slight and careful measures to complement missing information. This keeps the originality of the database but also allows the anomalies to influence the training process. Further studies are recommended to focus on the data preprocessing techniques to effectively

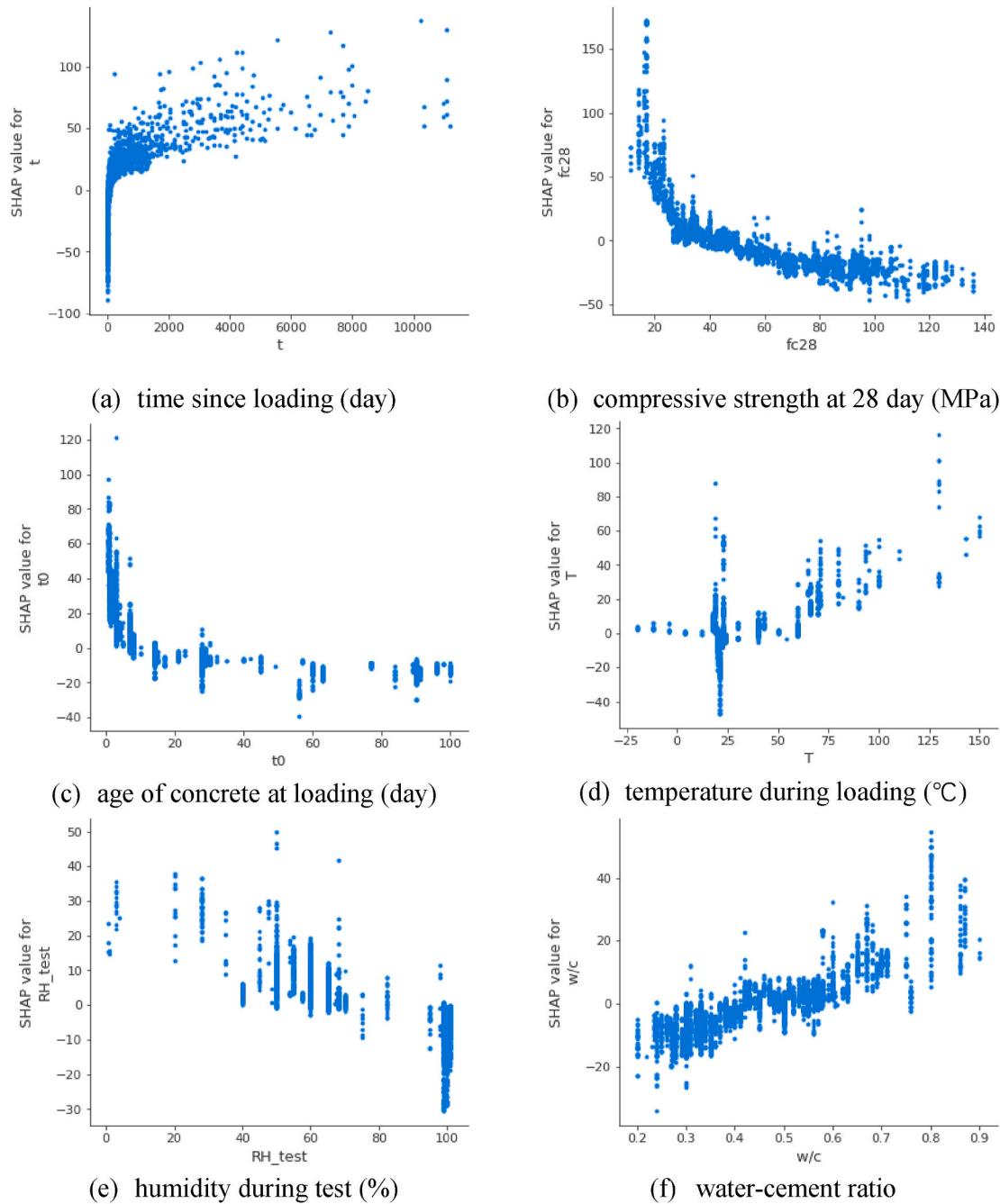


Fig. 11. SHAP values of the most influential features based on LGBM.

Table 4

Details of selected scenarios for exemplification.

	w/c	a/c	c	cem	SP	fc28	V/S	H0	t0	T	RH_Test	type	sigma
S1	0.47	1.37	390	3	0	50	75.0	99	28	20	60	1	7.00
S2	0.56	5.40	334	3	0	42	44.4	65	28	20	65	1	16.68
S3	0.40	3.92	438	3	0.6	42	30.0	98	7	20	99	4	9.14
S4	0.50	5.09	350	3	0.3	35	30.0	98	7	20	99	4	7.61
S5	0.58	4.75	337	3	0	35	16.5	100	14	21	100	4	7.89
S6	0.45	4.75	291	3	0	27	16.5	100	14	21	100	4	7.27

filter the anomalies, interpolate missing information while maintaining the originality of the database.

5. Conclusions

This study builds three Ensemble Machine Learning (EML) models (i. e., Random Forest (RF), Extreme Gradient Boosting Machine (XGBoost)

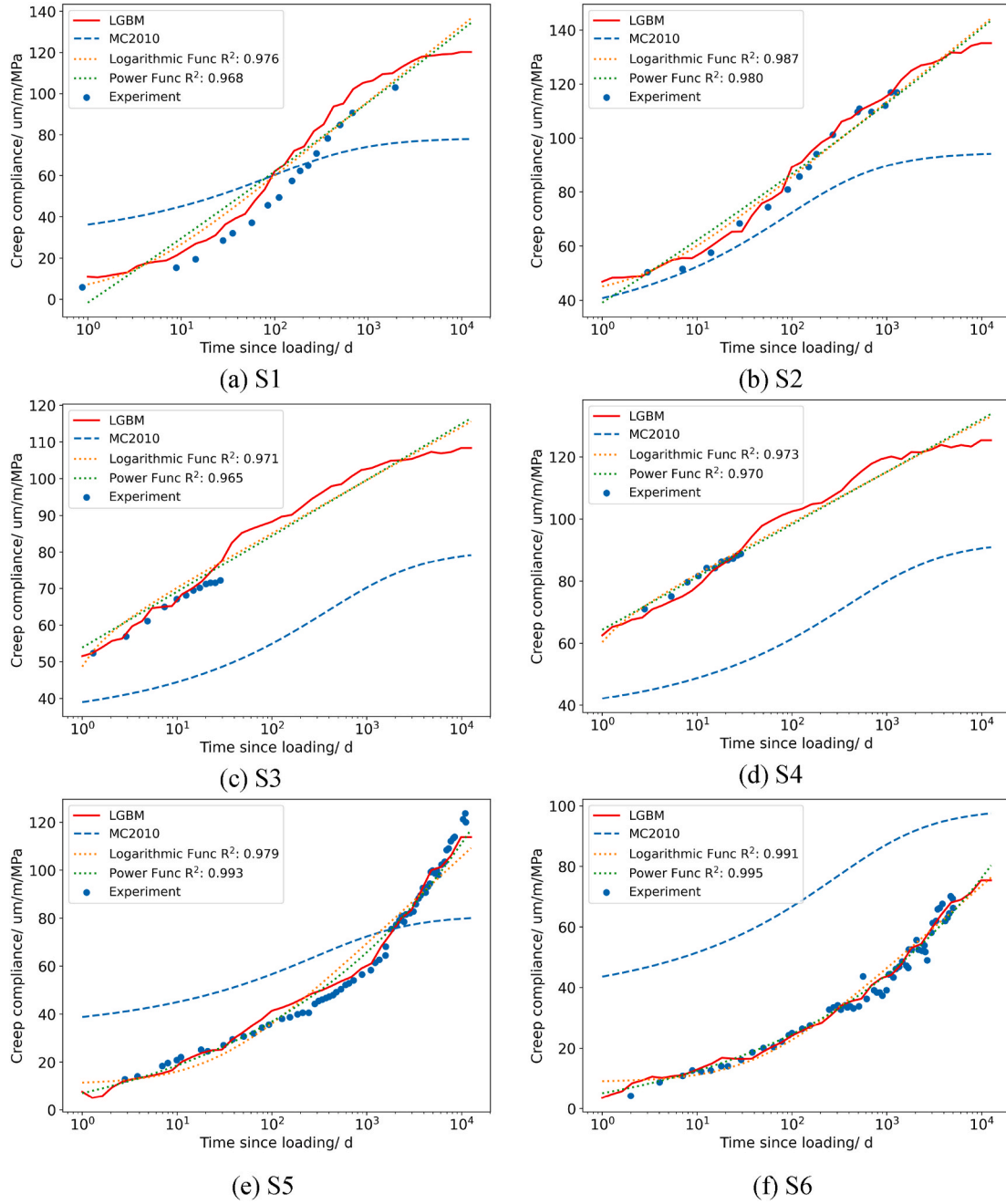


Fig. 12. Prediction of typical scenarios.

and Light Gradient Boosting Machine (LGBM)) for prediction of creep of concrete, based on an input vector constituted by parameters of concrete mix proportions, environmental conditions and loading conditions. Firstly, the NU database is preprocessed by a prebuilt XGBoost model and split into a training and a testing set. Then, through Bayesian Optimization and 5-fold cross validation, the EML models are tuned to achieve a high accuracy. Based on the cooperative game theories, the Shapley Additive Explanations (SHAP) is adopted to interpretate the predictions of the EML model and validate its reasonability by comparing its interpretations with common knowledges in the field of creep of concrete. Finally, the applicability of the EML models is exemplified by 6 typical scenarios and the limitations and potential improvement measures of the ML approaches are discussed. Through this study, the following conclusions can be obtained:

- (1) Through Bayesian Optimization and cross validation, three EML models all can achieve high accuracy in the testing set ($R^2 > 0.946$), which largely outperforms the broadly used empirical model code MC2010 ($R^2 = 0.378$).
- (2) Compared with XGB and RF, LGBM can achieve higher accuracy in significantly shorter calculation time on the NU database: the coefficient of determination of LGBM, RF and XGBoost are 0.953, 0.947, 0.946, while their calculation time is 969, 28,063, 36,073 s, respectively.
- (3) The well-tuned EML models can be effectively interpreted by its SHAP values both locally and globally. Although different EML models produces different SHAP results, all three EML models arrive at the same conclusion about the 5 most influential parameters for concrete creep compliance: time since loading (t), compressive strength (fc28), concrete age when loads are applied

(t_0), relative humidity during the test (RH_{test}) and temperature during the test (T).

- (4) The patterns that EML models captured are consistent with common knowledge about factors that influence concrete creep, which validates the reasonability of proposed EML models.
- (5) The prediction of EML models on a long-term continuous time range are in conformity with experimental observations. The creep compliance curves are not smooth due to the scattered, incomplete and possibly abnormal data in the NU database. Further improvements are required on the data preprocessing techniques for anomaly identification and missing data imputation.

Declaration of competing interest

The authors 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

Minfei Liang, Yidong Gan, Ze Chang and Zhi Wan would like to acknowledge the funding supported by China Scholarship Council under grant number 202007000027, 201706130140, 201806060129 and 201906220205.

References

- [1] S. Asamoto, K. Kato, T. Maki, Effect of creep induction at an early age on subsequent prestress loss and structural response of prestressed concrete beam, *Construct. Build. Mater.* 70 (2014) 158–164.
- [2] F. Benboudjema, F. Meftah, J.-M. Torrenti, A viscoelastic approach for the assessment of the drying shrinkage behaviour of cementitious materials, *Mater. Struct.* 40 (2007) 163–174.
- [3] Z.P. Bazant, M.H. Hubler, Q. Yu, Pervasiveness of excessive segmental bridge deflections: wake-up call for creep, *ACI Struct. J.* 108 (2011) 766–774.
- [4] Z.P. Bazant, L. Panula, Creep and shrinkage characterization for analyzing prestressed concrete structures, *Prestress. Concr. Inst.* 25 (1980) 86–122.
- [5] Dejian Shen, Jiao Yang, Yan Gao, Shuaishuai Zhu, Guoqing Jiang, Tensile creep and cracking resistance of concrete with different water-to-cement ratios at early age, *Construct. Build. Mater.* 146 (2017) 410–418.
- [6] A. Kamen, E. Denarié, H. Sadouki, E. Bruhwiler, UHPFRC tensile creep at early age, *Mater. Struct.* 42 (2009) 113–122.
- [7] Anja Estensen Klausen, Terje Kanstad, Øyvind Bjøntegaard, Erik Sellevold, Comparison of tensile and compressive creep of fly ash concretes in the hardening phase, *Cement Concr. Res.* 95 (2017) 188–194.
- [8] Chunping Gu, Yicong Wang, Gao Fan, Yang Yang, Tongyuan Ni, Jintao Liu, Xiaotian Lou, Chen Jin, Early age tensile creep of high performance concrete containing mineral admixtures: experiments and modeling, *Construct. Build. Mater.* 197 (2019) 766–777.
- [9] Y. Theiner, M. Drexel, M. Neuner, G. Hofstetter, Comprehensive study of concrete creep, shrinkage, and water content evolution under sealed and drying conditions, *Strain* (53) (2017), e12223.
- [10] Zhihui Zheng, Di Hu, Peng Liu, Fei Sha, Lei Liu, Zhiwu Yu, Considering the effect of the randomness of concrete strength and relative humidity on concrete creep, *Struct. Concr.* (2020) 1–15.
- [11] Matthieu Briffaut, Farid Benboudjema, Jean-Michel Torrenti, Georges Nahas, Concrete early age basic creep: experiments and test of rheological modelling approaches, *Construct. Build. Mater.* 36 (2012) 373–380.
- [12] Dejian Shen, Chengcai Li, Ci Liu, Ming Li, Jiacheng Kang, Experimental investigations on early-age tensile creep of internally cured high strength concrete under different initial stress/strength ratios, *Construct. Build. Mater.* 265 (2020) 120313.
- [13] Li Su, Yuan-feng Wang, Sheng-qi Mei, Peng-fei Li, Experimental investigation on the fundamental behavior of concrete creep, *Construct. Build. Mater.* 152 (2017) 250–258.
- [14] L. Granger, Comportement différé du béton dans les enceintes de centrales nucléaires: analyse et modélisation, Ph.D. thesis) Ecole Nationale des Ponts et Chaussées, 1995.
- [15] G. De Schutter, L. Taerwe, Fictitious degree of hydration method for the basic creep of early age concrete, *Mater. Struct.* 33 (2000) 370–380.
- [16] Basile T. Tamsia, James J. Beaudoin, Jacques Marchand, The early age short-term creep of hardening cement paste: load-induced hydration effects, *Cement Concr. Compos.* 26 (2004) 481–489.
- [17] Mateusz Wyrzykowski, Karen Scrivener, Pietro Lura, Basic creep of cement pastes at early age - the role of cement hydration, *Cement Concr. Res.* 116 (2019) 191–201.
- [18] Siming Liang, Ya Wei, Biaxial creep of high-strength concrete at early ages assessed from restrained ring test, *Cement Concr. Compos.* 104 (2019) 103421.
- [19] Yuan Gao, Jun Zhang, Po Han, Determination of stress relaxation parameters of concrete in tension at early-age by ring test, *Construct. Build. Mater.* 41 (2013) 152–164.
- [20] Wahid Ladaoui, Thierry Vidal, Alain Sellier, Xavier Bourbon, Effect of a temperature change from 20 to 50°C on the basic creep of HPC and HPFRC, *Mater. Struct.* 44 (2011) 1629–1639.
- [21] Minho Yoon, Gyu Yong Kim, Youngsun Kim, Taegyu Lee, Gyeongcheol Choe, Euichul Hwang, Jeongsoo Nam, Creep behavior of high-strength concrete subjected to elevated temperatures, *Materials* 10 (2017) 781.
- [22] Wook Lee Young, Gyu Yong Kim, Nenad Gucunski, Gyeong Choe Choe, Min Ho Yoon, Thermal strain behavior and strength degradation of ultra-high-strength-concrete, *Mater. Struct.* 49 (2016) 3411–3421.
- [23] A.M. Neville, *Properties of Concrete*, fifth ed., Trans-Atlantic Publications, Inc, 2012.
- [24] Zhu He, Qingbin Li, Yu Hu, Rui Ma, Double feedback control method for determining early-age restrained creep of concrete using a temperature stress testing machine, *Materials* 11 (2018) 1079.
- [25] M. Irfan-ul-Hassan, B. Pichler, R. Reihnsner, Ch Hellmich, Elastic and creep properties of young cement paste, as determined from hourly repeated minute-long quasi-static tests, *Cement Concr. Res.* 82 (2016) 36–49.
- [26] Brice Delsaute, Boulay Claude, Stephanie Staquet, Creep Testing of concrete since setting time by means of permanent and repeated minute-long loadings, *Cement Concr. Compos.* 73 (2016) 75–88.
- [27] Ya Wei, Siming Liang, Xiang Gao, Indentation creep of cementitious materials: experimental investigation from nano to micro length scales, *Construct. Build. Mater.* 143 (2017) 222–233.
- [28] Zhangli Hu, Mateusz Wyrzykowski, Michele Griffo, Karen Scrivener, Pietro Lura, Young's modulus and creep of calcium-silicate-hydrate compacts measured by microindentation, *Cement Concr. Res.* 134 (2020) 106104.
- [29] Yidong Gan, Matthieu Vandamme, Hongzhi Zhang, Klaas van Breugel, Branko Šavija, Micro-cantilever testing on the short-term creep behaviour of cement paste at micro-scale, *Cement Concr. Res.* 134 (2020) 106105.
- [30] Yidong Gan, Claudia Romero Rodriguez, Erik Schlangen, Klaas van Breugel, Branko Šavija, Assessing strain rate sensitivity of cement paste at the micro-scale through micro-cantilever testing, *Cement Concr. Compos.* 121 (2021) 104084.
- [31] Y Gan, H Zhang, B Šavija, E Schlangen, K Van Breugel. Static and fatigue tests on cementitious cantilever beams using nanoindenter. *Micromachines* 9 (12), 630.
- [32] Y Gan, H Zhang, B Šavija, E Schlangen, K Van Breugel. Micro-cantilever testing of cementitious materials under various loading conditions. *Proceedings of the 10th International Conference on Fracture Mechanics of Concrete and Concrete Structures*.
- [33] Z.P. Bazant, S.T. Wu, Rate-type creep law of aging concrete based on Maxwell chain Materials and Structures 7 (1974) 45–60.
- [34] I. Carol, J. Murcia, A model for the non-linear time-dependent behavior of concrete in compression based on a Maxwell chain with exponential algorithm, *Mater. Struct.* 22 (1989) 176–184.
- [35] Z.P. Bazant, Yunping Xi, Continuous retardation spectrum for solidification theory of concrete creep, *J. Eng. Mech.* 121 (2) (1995) 281–288.
- [36] Giovanni Di Luzio, Luigi Cedolin, Carlo Beltrami, Tridimensional long-term finite element analysis of reinforced concrete structures with rate-type creep approach, *Appl. Sci.* 10 (2020) 4772.
- [37] Brock D. Hedegaard, Carol K. Shield, Catherine E.W. French, Smeared-bar model for viscoelastic analysis of uncracked reinforced concrete structures, *ASCE Journal of Structural Engineering* (2014), 04014167.
- [38] Wei Jiang, Geert De Schutter, Yong Yuan, Degree of hydration based prediction of early age basic creep and creep recovery of blended concrete, *Cement Concr. Compos.* 48 (2014) 83–90.
- [39] Ya Wei, Jingsong Huang, Siming Liang, Measurement and modeling concrete creep considering relative humidity effect, *Mech. Time-Dependent Mater.* 24 (2020) 161–177.
- [40] Haidong Huang, Reyes Garcia, Shan-Shan Huang, Maurizio Guadagnini, Kypros Pilakoutas, A practical creep model for concrete elements under eccentric compression, *Mater. Struct.* 52 (2019) 119.
- [41] Muhammad Irfan-ul-Hassan, Markus Königsberger, Roland Reihnsner, Christian Hellmich, Bernhard Pichler, How water-aggregate interactions affect concrete creep: multiscale Analysis, *J. Nanomech. Micromech.* 7 (4) (2017), 04017019.
- [42] H. Do, S. Bishnoi, K.L. Scrivener, Microstructural modeling of early-age creep in hydrating cement paste, *ASCE Journal of Engineering Mechanics* 142 (11) (2016), 04016086.
- [43] Weinan Wang, Jinxin Gong, New relaxation function and age-adjusted effective modulus expressions for creep analysis of concrete structures, *Eng. Struct.* 188 (2019) 1–10.
- [44] Y. Park, Y. Lee, Description of concrete creep under time-varying stress using parallel creep curve, *Advances in Materials Science and Engineering* (2016) 9370514.
- [45] A. Loukili, M. Omar, G. Pijaudier-Cabot, in: Z.P. Bazant, F.-J. Ulm, F.H. Wittmann (Eds.), *Basic Creep of Ultra-high Strength Concrete Experiments and Modeling*, Proc., Concreep 6, Elsevier, Amsterdam, The Netherlands, 2001, pp. 545–550.
- [46] N. Challamel, C. Lanos, C. Casandjian, Creep damage modelling for quasi-brittle materials, *Eur. J. Mech. Solid.* 24 (2005) 593–613.
- [47] G. van Zijl, R. de Borst, J. Rots, A numerical model for the time-dependent cracking of cementitious materials, *Int. J. Numer. Methods Eng.* 38 (2001) 5063–5092.

- [48] M. Cervera, J. Oliver, T. Prato, Thermo-chemo-mechanical model for concrete. I: hydration and aging, *Journal of Engineering Mechanics*, ASCE 125 (9) (1999) 1018–1027.
- [49] M. Cervera, J. Oliver, T. Prato, Thermo-chemo-mechanical model for concrete. II: damage and creep, *Journal of Engineering Mechanics*, ASCE 125 (9) (1999) 1028–1039.
- [50] R. Lackner, H.A. Mang, Chemo-plastic material model for the simulation of early-age cracking: from the constitutive law to numerical analyses of massive concrete structures, *Cement Concr. Compos.* 26 (5) (2004) 551–562.
- [51] G. Di Luzio, G. Cusatis, Solidification-Microprestress-Microplane (SMM) theory for concrete at early age: theory, validation and application, *Int. J. Solid Struct.* 50 (6) (2013) 957–975.
- [52] D. Gawin, F. Pesavento, B.A. Schrefler, Hygro-thermo-chemo-mechanical modelling of concrete at early ages and beyond. Part I: hydration and hygro-thermal phenomena, *Int. J. Numer. Methods Eng.* 67 (3) (2006) 299–331.
- [53] D. Gawin, F. Pesavento, B.A. Schrefler, Hygro-thermo-chemo-mechanical modelling of concrete at early ages and beyond. Part II: shrinkage and creep of concrete, *Int. J. Numer. Methods Eng.* 67 (3) (2006) 332–363.
- [54] F. Laverne, K. Sab, J. Sanahuja, M. Bornert, C. Toulemonde, Investigation of the effect of aggregates' morphology on concrete creep properties by numerical simulations, *Cement Concr. Res.* 71 (2015) 14–28.
- [55] Alain B. Giorla, Cyrille F. Dunant, Microstructural effects in the simulation of creep of concrete, *Cement Concr. Res.* 105 (2018) 44–53.
- [56] Petr Havlásek, Milan Jirásek, Multiscale modeling of drying shrinkage and creep of concrete, *Cement Concr. Res.* 85 (2016) 55–74.
- [57] Wei Jiang, Geert De Schutter, Yong Yuan, Degree of hydration based prediction of early age basic creep and creep recovery of blended concrete, *Cement Concr. Compos.* 48 (2014) 83–90.
- [58] Yidong Gan, Claudia Romero Rodriguez, Hongzhi Zhang, Erik Schlangen, Klaas van Breugel, Branko Šavija, Modelling of microstructural effects on the creep of hardened cement paste using an experimentally-informed lattice model, *Comput Aided Civ Inf* 35 (2020) 1085–1100.
- [59] Lyes Bal, François Buyle-Bodin, Artificial neural network for predicting drying shrinkage of concrete, *Construct. Build. Mater.* 38 (2013) 248–254.
- [60] Gökhan Bayar, Turhan Bilir, A novel study for the estimation of crack propagation in concrete using machine learning algorithms, *Construct. Build. Mater.* 215 (2019) 670–685.
- [61] Mostafa Jalal, Zachary Grasley, Charles Gurganus, Jeffrey W. Bullard, Experimental investigation and comparative machine-learning prediction of strength behavior of optimized recycled rubber concrete, *Construct. Build. Mater.* 256 (2020) 119478.
- [62] Afshin Marani, Armin Jamali, L. Moncef, Nehdi, Predicting ultra-high-performance concrete compressive strength using tabular generative adversarial networks, *Materials* 13 (2020) 4757.
- [63] Khoa Tan Nguyen, Quang Dang Nguyen, Tuan Anh Le, Jiuk Shin, Kihak Lee, Analyzing the compressive strength of green fly ash based geopolymer concrete using experiment and machine learning approaches, *Construct. Build. Mater.* 247 (2020) 118581.
- [64] Taihao Han, Ashfia Siddique, Kamal Khayat, Jie Huang, Aditya Kumar, An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete, *Construct. Build. Mater.* 244 (2020) 118271.
- [65] Oladimeji B. Olalusi, Panagiotis Spyridis, Machine learning-based models for the concrete breakout capacity prediction of single anchors in shear, *Adv. Eng. Software* 147 (2020) 102832.
- [66] Roya Solhmirzaei, Hadi Salehi, Venkatesh Kodur, M.Z. Naser, Machine learning framework for predicting failure mode and shear capacity of ultra high performance concrete beams, *Eng. Struct.* 224 (2020) 111221.
- [67] Eslam Gomaa, Taihao Han, Mohamed ElGawady, Jie Huang, Aditya Kumar, Machine learning to predict properties of fresh and hardened alkali-activated concrete, *Cement Concr. Compos.* 115 (2021) 103863.
- [68] Min-Chang Kang, , Doo-Yeol Yoo, Rishi Gupta, Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete, *Construct. Build. Mater.* 266 (2021) 121117.
- [69] Miao Su, Qingyu Zhong, Hui Peng, Shaoan Li, Selected machine learning approaches for predicting the interfacial bond strength between FRPs and concrete, *Construct. Build. Mater.* 270 (2021) 121456.
- [70] Caia Rong, Taihao Han, Wenyu Liao, Jie Huang, Dawang Li, Aditya Kumar, Hongyan Ma, Prediction of surface chloride concentration of marine concrete using ensemble machine learning, *Cement Concr. Res.* 136 (2020) 106164.
- [71] Panagiotis G. Asteris, Athanasia D. Skentou, Abidhan Bardhan, Pijush Samui, Kypros Pilakoutas, Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models, *Cement Concr. Res.* 145 (2021) 106449.
- [72] Hoang Nguyen, Thanh Vu, P. Thuc, Vo, Thai Huu-Tai, Efficient machine learning models for prediction of concrete strengths, *Construct. Build. Mater.* 266 (2021) 120950.
- [73] Z.P. Bazant, G.-H. Li, Comprehensive database on concrete creep and shrinkage, *ACI Mater. J.* 105 (2008) 635–637.
- [74] Roderick J.A. Little, Donald B. Rubin, Statistical Analysis with Missing Data, John Wiley & Sons, Inc., New York, NY, USA, 1986.
- [75] J. Carpenter, M. Kenward, Multiple Imputation and its Application, John Wiley & Sons, 2012.
- [76] Stef van Buuren, Karin Groothuis-Oudshoorn, Mice: multivariate imputation by chained equations in R, *J. Stat. Software* 45 (2011) 1–67.
- [77] Zhi Wan, Yading Xu, Branko Šavija, On the use of machine learning models for prediction of compressive strength of concrete: influence of dimensionality reduction on the model performance, *Materials* 14 (2021) 713.
- [78] Wassim Ben Chaabene, Majdi Flah, L. Moncef, Nehdi, Machine learning prediction of mechanical properties of concrete: critical review, *Construct. Build. Mater.* 260 (2020) 119889.
- [79] Nguyen-Sy Tuan, Jad Wakim, Quy-Dong To, Minh-Ngoc Vu, The-Duong Nguyen, Thoi-Trung Nguyen, Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method, *Construct. Build. Mater.* 260 (2020) 119757.
- [80] Jin Duan, Panagiotis G. Asteris, Hoang Nguyen, Xuan-Nam Bui, Hossein Moayed, A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model, *Eng. Comput.* (2020), <https://doi.org/10.1007/s00366-020-01003-0>.
- [81] R. Mosbeh, Kaloop, deepak kumar, pijush samui, jong wan hu, dongwook kim, Compressive strength prediction of high-performance concrete using gradient tree boosting machine, *Construct. Build. Mater.* 264 (2020) 120198.
- [82] I.-C. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, *Cement Concr. Res.* 28 (1998) 1797–1808.
- [83] Panagiotis G. Asteris, Konstantinos G. Kolovos, Self-compacting concrete strength prediction using surrogate models, *Neural Comput. Appl.* 31 (Suppl 1) (2019) S409–S424.
- [84] Mengxiao Shi, Qiang Wang, Zhikai Zhou, Comparison of the properties between high-volume fly ash concrete and high-volume steel slag concrete under temperature matching curing condition, *Construct. Build. Mater.* 98 (2015) 649–655.
- [85] Okan Karahan, Transport properties of high volume fly ash or slag concrete exposed to high temperature, *Construct. Build. Mater.* 152 (2017) 898–906.
- [86] M. Alaa, Rashad, An investigation of high-volume fly ash concrete blended with slag subjected to elevated temperatures, *J. Clean. Prod.* 93 (2015) 47–55.
- [87] S. Mukherjee, S. Mandal, U.B. Adhikari, Comparative study on physical and mechanical properties of high slump and zero slump high volume fly ash concrete (HVFA), *Global NEST Journal* 15 (4) (2013) 578–584.
- [88] Jelena Dragas, Nikola Tosic, Ivan Ignjatovic, Snezana Marinkovic, Mechanical and time-dependent properties of high-volume fly ash concrete for structural use, *Mag. Concr. Res.* 68 (12) (2016) 632–645.
- [89] Yu Jing, Cong Lu, K. Christopher, Y. Leung, Gengying Li, Mechanical properties of green structural concrete with ultrahigh- volume fly ash, *Construct. Build. Mater.* 147 (2017) 510–518.
- [90] Yasmin Hefni, Yehia Abd El Zaher, Mona Abdel Wahab, Influence of activation of fly ash on the mechanical properties of concrete, *Construct. Build. Mater.* 172 (2018) 728–734.
- [91] P. Nath, P. Sarker, Effect of fly ash on the durability properties of high strength concrete, *Procedia Engineering* 14 (2011) 1149–1156.
- [92] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, *Int. Joint Conf. Artif. Intell.* 14 (2) (1995) 1137–1145.
- [93] Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction, second ed., Springer, 2008.
- [94] M.A. DeRousseau, E. Laftchiev, J.R. Kasprzyk, B. Rajagopalan, W.V. Srubar III, A comparison of machine learning methods for predicting the compressive strength of field-placed concrete, *Construct. Build. Mater.* 228 (2019) 116661.
- [95] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (2) (1996) 123–140.
- [96] L. Breiman, Random forests, *Mach. Learn.* 45 (1) (2001) 5–32.
- [97] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Thirion Bertrand, Olivier Grisel, Mathieu Blondel, Prettenhofer Peter, Ron Weiss, Dubourg Vincent, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay, Scikit-learn: machine learning in Python, *J. Mach. Learn. Res.* 12 (85) (2011) 2825–2830.
- [98] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, 2016, pp. 785–794.
- [99] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, T.Y. Liu, LightGBM: a highly efficient gradient boosting decision tree, *Adv. Neural Inf. Process. Syst.* (2017) 3146–3154.
- [100] C.E. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, MIT Press, Cambridge, MA, 2006.
- [101] Donald R. Jones, Matthias Schonlau, J. William, Welch, Efficient global optimization of expensive black-box functions, *J. Global Optim.* 13 (1998) 455–492.
- [102] Markus Ojala, C. Gemma, Garriga, Permutation tests for studying classifier performance, *J. Mach. Learn. Res.* 11 (2010) 1833–1863.
- [103] International Federation for Structural Concrete (fib), Fib Model Code for Concrete Structures 2010, Ernst & Sohn, Wiley, Berlin, Germany, 2013.
- [104] E. Brochu, V.M. Cora, N. De Freitas, A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning, *arXiv Preprint arXiv*, 2010, 1012.2599.
- [105] M.T. Ribeiro, S. Singh, C. Guestrin, Why should I trust you?: explaining the predictions of any classifier, in: *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2016, pp. 1135–1144.
- [106] S. Haufe, et al., On the interpretation of weight vectors of linear models in multivariate neuroimaging, *Neuroimage* 87 (2014) 96–110.
- [107] S.M. Lundberg, S.-I. Lee, A Unified Approach to Interpreting Model Predictions. 31st Conference on Neural Information Processing Systems (NIPS 2017), 2017. Long Beach, CA.
- [108] S.M. Lundberg, G.G. Erion, S.-I. Lee, Consistent Individualized Feature Attribution for Tree Ensembles, *arXiv preprint arXiv:1802.03888*, 2018.

- [109] Z.P. Bazant, M. Jirásek, *Creep and Hygrothermal Effects in Concrete Structures*, 225, Springer, Basingstoke, UK, 2018.
- [110] Z.P. Bazant, J.C. Chern, Log double power law for concrete creep, *ACI Journal* (82-59) (1985) 665–675.
- [111] Z.P. Bazant, E. Osman, Double power law for basic creep of concrete, *Mater. Struct.* 9 (1976) 3–11.
- [112] J.J. More, The levenberg-marquardt algorithm: implementation and theory, in: G. A. Watson (Ed.), *Numerical Analysis. Lecture Notes in Mathematics*, 630, Springer Verlag, 1977, pp. 105–116.