Data-Centric Engineering (2025), 6: e37
doi:10.1017/dce.2025.10018

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Predicting concrete strength using machine learning and fresh-state measurements

Bahdan Zviazhynski[1], Callum White[2], Janet M. Lees[2] and Gareth J. Conduit[1,3]

[1]Cavendish Laboratory, University of Cambridge, Cambridge, UK
[2]Department of Engineering, University of Cambridge, Cambridge, UK
[3]Intellegens, Cambridge, UK
Corresponding author: Bahdan Zviazhynski; Email: bogdanzwier@gmail.com

Received: 31 October 2024; **Revised:** 23 March 2025; **Accepted:** 25 June 2025

Keywords: concrete; machine learning; slump; strength prediction

Abstract

Understanding the properties of lower-carbon concrete products is essential for their effective utilization. Insufficient empirical test data hinders practical adoption of these emerging products, and a lack of training data limits the effectiveness of current machine learning approaches for property prediction. This work employs a random forest machine learning model combined with a just-in-time approach, utilizing newly available data throughout the concrete lifecycle to enhance predictions of 28 and 56 day concrete strength. The machine learning hyperparameters and inputs are optimized through a novel unified metric that combines prediction accuracy and uncertainty estimates through the coefficient of determination and the distribution of uncertainty quality. This study concludes that optimizing solely for accuracy selects a different model than optimizing with the proposed unified accuracy and uncertainty metric. Experimental validation compares the 56-day strength of two previously unseen concrete mixes to the machine learning predictions. Even with the sparse dataset, predictions of 56-day strength for the two mixes were experimentally validated to within 90% confidence interval when using slump as an input and further improved by using 28-day strength.

Impact Statement

This paper demonstrates that machine learning can predict concrete strength, even for small datasets, by utilizing new test results as they become available over the concrete lifetime. This has important implications for accelerating the decision-making process currently adopted in concrete practice, accelerating the adoption of novel low-carbon materials with limited test data. Furthermore, the machine learning methodology developed in this paper, including a newly proposed accuracy-uncertainty metric to compare various machine learning models, has application to the broader field of emerging materials design, allowing bespoke materials to be designed rapidly for particular applications.

1. Introduction

Concrete usage causes around 8% of global greenhouse gas emissions, primarily due to the production of ordinary Portland cement (OPC) (Turner and Collins, 2013; Miller et al., 2021). Efforts are underway to

B.Z. and C.W. contributed equally to this work.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

reduce and replace OPC with alternative materials. Lower carbon OPC replacement materials include calcined clay, ground granulated blast furnace slag (GGBS), and other supplementary cementitious materials. Extensive historical testing of OPC-based mixes enabled researchers to establish general relationships between the constituent proportions and the properties of the resulting concrete. From this, prescriptive design principles are used in many aspects of concrete mix formulation, ensuring appropriate short-term and longer-term properties, including strength and durability (Teychenne et al., 1997).

Current guidelines for OPC mixes may not reflect the performance of newly developed concrete mixes that incorporate novel materials to reduce OPC content and embodied carbon. Experimental testing of lower carbon mixes, including such materials, can reduce uncertainty, but testing all possible mixes is prohibitively time-consuming and expensive. Therefore, accurate predictive methods are required to reduce experimental testing requirements. Machine learning is a powerful tool that can capture complex non-linear relationships between inputs (such as mix proportions) and outputs (desired physical properties) from historical data, even if the available data is sparse and noisy (Conduit et al., 2017, 2018, 2019). The approach has been successfully used to predict the physical properties of many different materials (Bhadeshia et al., 1995; Sourmail et al., 2002; Agrawal et al., 2014; Ward et al., 2016, 2017; Kim et al., 2018), including concrete (Taffese et al., 2015; Chaabene et al., 2020; Prayogo et al., 2020; Liu et al., 2021; Tran et al., 2022). Furthermore, machine learning has been exploited to design concrete with high confidence in meeting requirements (Forsdyke et al., 2023).

One of the most important properties of concrete is the compressive strength, which increases over time. In practical settings, strength is conventionally assessed at 28 days via compressive testing, but a target strength at 56 days can also be specified, particularly for materials with lower early-age strength development. Machine learning algorithms are increasingly being used to predict concrete compressive strength for mixes with a range of OPC replacement materials, including fly ash (Chou et al., 2014; Deng et al., 2018; Young et al., 2019; Feng et al., 2020; Khursheed et al., 2021; Salami et al., 2021; Wan et al., 2021; Nguyen et al., 2021a; Liu, 2022; Lee et al., 2023; Li et al., 2023; Pakzad et al., 2023; Hariri-Ardebili et al., 2024), blast furnace slag (Chou et al., 2014; Feng et al., 2020; Salami et al., 2021; Wan et al., 2021; Nguyen et al., 2021a; Lee et al., 2023; Li et al., 2023; Hariri-Ardebili et al., 2024) and other materials including silica fume, metakaolin and rubber (Chou et al., 2014; Hadzima-Nyarko et al., 2020; Mansouri et al., 2022). Some attention has also been given to predicting the properties of high-strength mixes (Nguyen et al., 2021b), binary and ternary blended concretes (Murthy et al., 2024) and self-compacting mixes (Chakravarthy et al., 2023) using machine learning algorithms.

Despite a wealth of academic literature in this area, the adoption of machine learning models in practical settings remains limited. Most existing predictive tools have been developed using large test banks (greater than 1000 data points) offering high prediction accuracy. Due to the rapid introduction of new materials and the evolving nature of concrete mixes, there will be limited data for each OPC replacement material. As such, prediction models must be robust for smaller datasets, with highly accurate strength predictions, but also realistic uncertainty estimates.

From a review of the 20 articles highlighted in the area of machine learning strength prediction, it emerges that existing models commonly rely on the mix aggregates, binder (ordinary Portland cement and other cementitious materials), and water contents to estimate strength at a specific age (Figure 1). Due to the increasing complexity of concrete mix behaviour and the limitations of traditional mix design principles, it may be necessary to include more than three primary input variables to obtain accurate predictions for a variety of concrete mixes with varying materials, especially when test data is sparse. Therefore, this work proposes utilizing information obtained along the concrete life-cycle, such as fresh state measurements, to improve strength predictions without having to wait extended periods. This potentially reduces the risk of adopting mixes with sparse historical data, given strength estimations can be updated in real time.

Practitioners need good estimates of model uncertainty to be confident in the prediction outputs and to assess the risk profile of different concrete products. Most existing prediction models assess performance using a single accuracy metric, i.e., how close on average all of the predictions are to the actual strength. A measure of how trustworthy any particular prediction is, i.e. its uncertainty, is currently limited.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

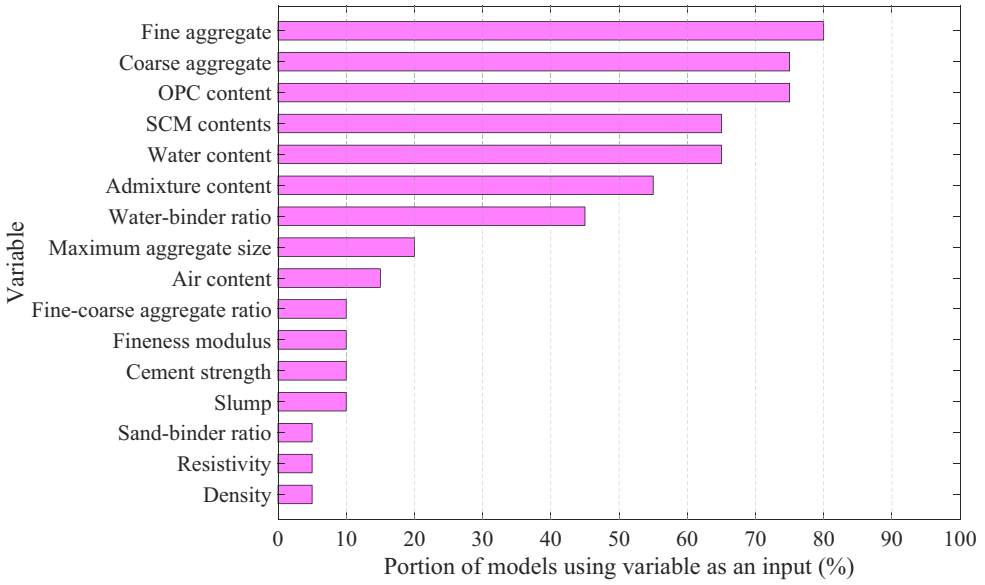


Figure 1. *Prevalence of model input variables used to predict compressive strength in the literature (Chou et al., 2014; Deng et al., 2018; Young et al., 2019; Feng et al., 2020; Hadzima-Nyarko et al., 2020; Pham et al., 2020; Khursheed et al., 2021; Nunez et al., 2021; Salami et al., 2021; Wan et al., 2021; Nguyen et al., 2021a; Han et al., 2022; Khan et al., 2022; Liu, 2022; Mansouri et al., 2022; Chi et al., 2023; Lee et al., 2023; Li et al., 2023; Pakzad et al., 2023; Hariri-Ardebili et al., 2024).*

Reviewing existing machine learning models revealed that most work considers a very limited range of accuracy metrics (Alkayem et al., 2024). Almost all of the papers relied solely on accuracy metrics such as coefficient of determination R^2, mean absolute error, or root mean square error, to compare model performance (Chou et al., 2014; Chopra et al., 2018; Deng et al., 2018; Young et al., 2019; Chaabene et al., 2020; Feng et al., 2020; Hadzima-Nyarko et al., 2020; Pham et al., 2020; Nunez et al., 2021; Salami et al., 2021; Wan et al., 2021; Nguyen et al., 2021a; Han et al., 2022; Khan et al., 2022; Liu, 2022; Mansouri et al., 2022; Chi et al., 2023; Lee et al., 2023). A very small number of papers directly considered the uncertainty in model predictions (Pakzad et al., 2023; Hariri-Ardebili et al., 2024). Where uncertainty was considered, it was to minimize prediction uncertainty, rather than assess the quality of uncertainty estimates and metrics such as the scatter index were adopted. From this review, it is clear that the practical considerations of applying machine learning for concrete strength prediction are often overlooked, resulting in predictive tools that are insufficiently robust for real-world use.

In this work, a sparse dataset (fewer than 50 data points) is adopted containing both OPC and GGBS materials to train a machine learning model to predict concrete compressive strength. The dataset used in this study is comparable to what a practitioner would have access to for an emerging lower-carbon material. This work aims to determine how machine learning models can be informed in real time over the life cycle of the concrete product as newly obtained information, such as fresh state measurements or earlier age strength data, becomes available. To do this, the following objectives are addressed:

- Which properties are most helpful to predict longer-term (56-day) strength?
- When predicting the strength of concrete mixes with OPC with or without GGBS binders, are better predictions achieved when using datasets separated by binder type or a single unified dataset?
- What is the role of uncertainty in the performance assessment of machine learning models and for concrete strength prediction? Can a broader range of predictive performance measures, including accuracy and uncertainty, be effectively captured in a single metric?

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

This paper is organized into several sections to achieve its objectives. Section 2 highlights the key properties of concrete throughout its life cycle and reviews the correlations between various concrete properties within the adopted dataset. Initially, the correlations for fresh state properties are examined separately for Ordinary Portland Cement (OPC) only mixes and mixes containing Ground Granulated Blast Furnace Slag (GGBS). This is followed by an assessment of the correlations within a unified dataset that includes both types of binder materials. Guidance is presented for processing datasets for property prediction in practical settings. Section 3 discusses the random forest machine learning methodology employed in this study. It begins with an overview of the two-layer model and the decision tree process. Next, the accuracy and uncertainty metrics are evaluated, and a unified accuracy-uncertainty metric is proposed. Finally, leave-one-out cross-validation is performed to optimize model performance. In Section 4, two previously unseen concrete mixes are used for experimental validation of the proposed approach, demonstrating the advantages of the just-in-time prediction method. Section 5 addresses the limitations of the presented approach and discusses potential future applications of this work within the concrete industry and beyond.

2. Key concrete properties

This section discusses the key properties of concrete, which change throughout its lifecycle from the fresh to hardened state. Correlations between the properties are then reviewed. After the design and mixing of the constituent materials, concrete is initially in a fluid state for the first few hours (fresh state) and then solidifies into a hardened state (Figure 2). Because of the different states passed through, several performance measures are used throughout the life cycle of concrete. Research has shown that shorter-term fresh-state performance impacts longer-term performance due to compaction effects and

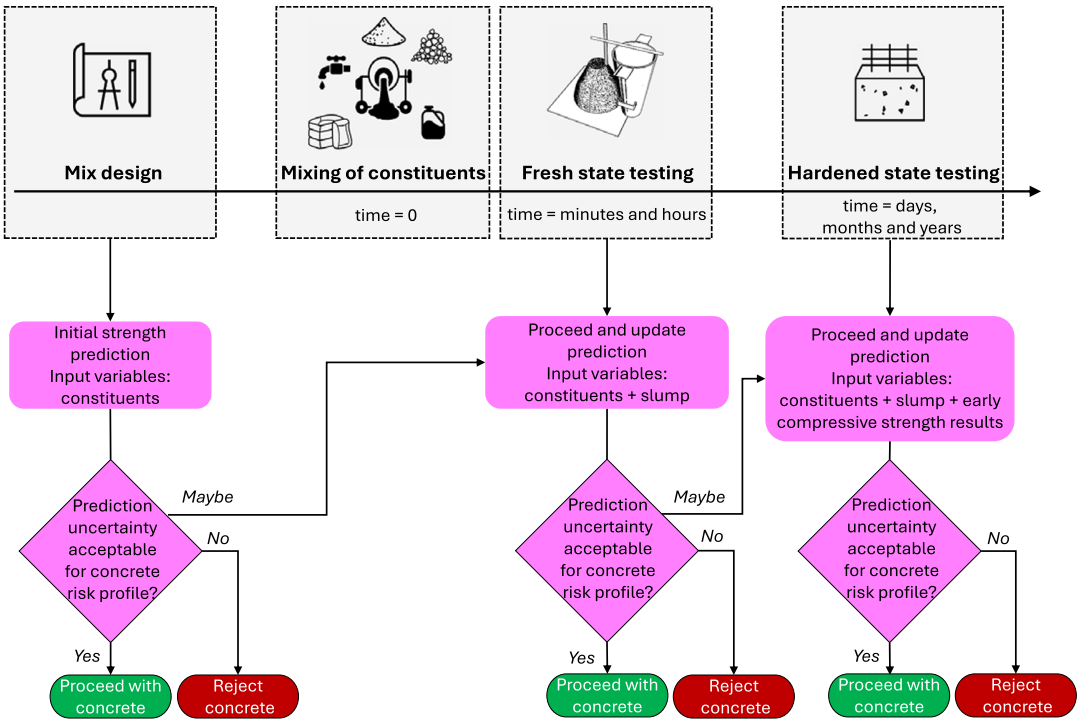


Figure 2. *Life cycle of concrete, from mixing to the hardened state, with the step-wise prediction approach and decision tree highlighted at each concrete age.*

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

microstructure development (Lydon, 1972; Neville, 1988). Fresh state testing, therefore, may aid understanding of the development of hardened state properties, leading to the early identification of issues in the later stages of the life cycle of a concrete product (Tattersall, 1991).

Figure 2 outlines a flow chart of how longer-term compressive strength predictions could be updated when new information becomes available over the life cycle of the concrete product. This offers a mechanism to reduce risk via several decision stages and, therefore, increase the adoption of mixes whose behavior is relatively unknown. The most basic input for the strength predictions is the mix constituents. Inputting solely this information to the machine learning model is advantageous as predictions can be performed before mixing. However, for novel mixes, these inputs alone will likely cause the prediction output to have large uncertainty as the training data will be primarily formed of traditional concrete mixes whose behaviour may not be representative of emerging materials. If large prediction uncertainty is observed, the next step is to update the strength predictions with fresh state results (such as slump testing) measured at the placement stage. Updating the strength predictions at this stage is advantageous as a decision can be made whether to proceed with the pour based on the improved estimates of strength and uncertainty. If the pour proceeds based on the output of this prediction step, then the strength predictions can be updated with the hardened state strength results taken at early ages. Progressively, the strength predictions will be closer to the actual compressive strength, with reduced uncertainty in the estimates. This step-wise prediction approach is adopted in this paper, with the experimental validation demonstrating the improved prediction performance as the model is updated with new information in real time.

First, in Section 2.1, we describe the preparation of concrete mixes used in the dataset and experimental validation performed in this study. Second, in Section 2.2, we describe fresh state performance testing of the mixes. Third, in Section 2.3, we set out hardened testing following the setting process of the mixes. Finally, in Section 2.4, we describe the correlations between the concrete properties observed in the sparse dataset, including mix proportions, fresh state properties, and hardened state properties, to form a sound basis for developing the machine learning model.

2.1. Preparation of concrete mix dataset

A total of 29 concrete mixes, 20 of which had 56-day strength data, were produced to form the training dataset (see the Supplementary Material) adopted in this study. These concrete mixes were produced in the University of Cambridge Civil Engineering Building laboratories. The experimental series adopted CEMI 52.5 ordinary Portland cement and sand fine aggregates with a maximum aggregate size of 2mm. A coarse crushed aggregate with a maximum aggregate size of 10mm was also adopted. A 100-litre planetary concrete mixer with variable rotor and drum speed was utilized throughout the test series. To produce a suitable range of fresh and hardened state properties, the water–binder ratio (0.50–0.91), total binder content (260–418 kg/m^3), sand–aggregate ratio (0.48–0.62), aggregate–binder ratio (3.9–7.0), and the GGBS cement replacement content of the mix (0–70%) were manipulated. These values are typical of those used in the wider large-scale concrete activity. The full table of data is included in the Supplementary Material.

The mix proportions, including the total binder content (ordinary Portland cement plus other cementitious materials), water–binder ratio, sand–aggregate ratio, aggregate–binder ratio, and supplementary cementitious material (SCM) percentage, were designed to achieve adequate variation in the fresh and hardened properties and represent the diversity in mix design seen in practice, whilst maintaining the scale of a small dataset. Several mixes exhibit comparable mix proportions but were batched at different total volumes, replicating the real-life production processes. This also replicates the nature of data within a practical setting, where the potential of novel mixes is scoped before scaling up production. Where SCMs were adopted, ground granulated blast furnace slag (GGBS) was utilized, given it is one of the primary materials used in practice to replace OPC and reduce embodied carbon as it is a waste material from the steel industry. GGBS is used as an example, but the process is developed so that it is valid for other SCMs.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

2.2. Fresh state properties

Immediately, and up to around two hours after water is added to the mix, concrete is in its fresh state so that it can flow with a relatively small applied stress (Tattersall, 1991). This allows concrete to be poured and shaped within molds to form the desired geometry as the mix hardens. Tests are completed in the fresh state to ensure that the concrete is suitable for placement and that a stable mix has been produced. Fresh state testing often takes the form of the Abrams slump test, conducted in this study and in practice, according to BS EN 12350-2 (BSI, 2019a). In this test, concrete is added to a conical mold in three layers, with each layer tamped sequentially. The slump cone is lifted steadily to allow the material to flow due to its weight, and the concrete is classified according to the difference between the height of the deformed material and the height of the slump cone. This classification is termed the total slump height. Other properties, including the rheological measures of yield stress and viscosity, can further describe the fresh state characteristics of concrete as a non-Newtonian fluid. A concrete rheometer was used in this study to measure yield stress and viscosity directly, although alternative easier-to-access mechanisms to derive the rheological properties have been developed (White and Lees, 2023; White and Lees, 2025). The fresh state properties of each mix within the adopted dataset are provided as the Supplementary Material.

2.3. Hardened state properties

Following the addition of water, a chemical hydration process begins, and over time, structural build-up results in a hardened state as the concrete develops strength. After the design concrete strength is specified, the actual concrete mix placed during construction must achieve at least this value. The strength development of the concrete is tracked by conducting hardened state testing at various ages on 100 mm × 100 mm × 100 mm cube specimens cast at the same time as the structure. In this study, compressive strength testing is conducted according to BS EN 12390-3 at 28 and 56 days via the application of axial force onto a concrete cube through parallel platens (BSI, 2019b). For each concrete mix, three samples are taken, and the maximum load reached by each sample is recorded. The average of these maximum load values is considered as the compressive strength of the specimen. The 28-day strength is taken as an intermediate value to be used as an input for machine learning; the 56-day strength is taken as representative of the longer-term strength performance. The actual compressive strength values attained via experimental testing, and the predicted strength for each mix are provided as the Supplementary Material.

2.4. Correlations between properties

The previous sections demonstrated how the properties of concrete change throughout its lifecycle. Since these properties pertain to the same concrete mix, they can be related to one another, which is discussed in this section.

The correlations between properties within the dataset, calculated using the Pearson correlation coefficient, are shown in Figure 3. Color intensity indicates the correlation coefficient, with the lightest shade indicating the strongest positive correlation and the darkest shade indicating the strongest negative correlation. Three matrices are presented for the OPC mixes (Figure 3a), the GGBS mixes (Figure 3b), and the unified dataset comprising OPC and GGBS mixes (Figure 3c). The first seven rows of the matrix (eight rows for GGBS mixes and the unified dataset) display the proportions of the constituent materials present in the concrete. The following four rows display the fresh state testing completed, such as the slump test. The final three rows of the correlation matrix display the hardened state testing completed, including density and strength measurements.

In all three datasets, the fresh state properties of the mix are only weakly correlated with mix proportions. There is, as expected, a moderate positive correlation between water content and total slump height during fresh state testing, which is increased in the presence of GGBS (0.51 for OPC, 0.72 for GGBS and 0.52 for unified). There are similar physically justified negative correlations between water content and static yield stress (−0.55 for OPC, −0.65 for GGBS, and −0.52 for unified). Viscosity is more

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

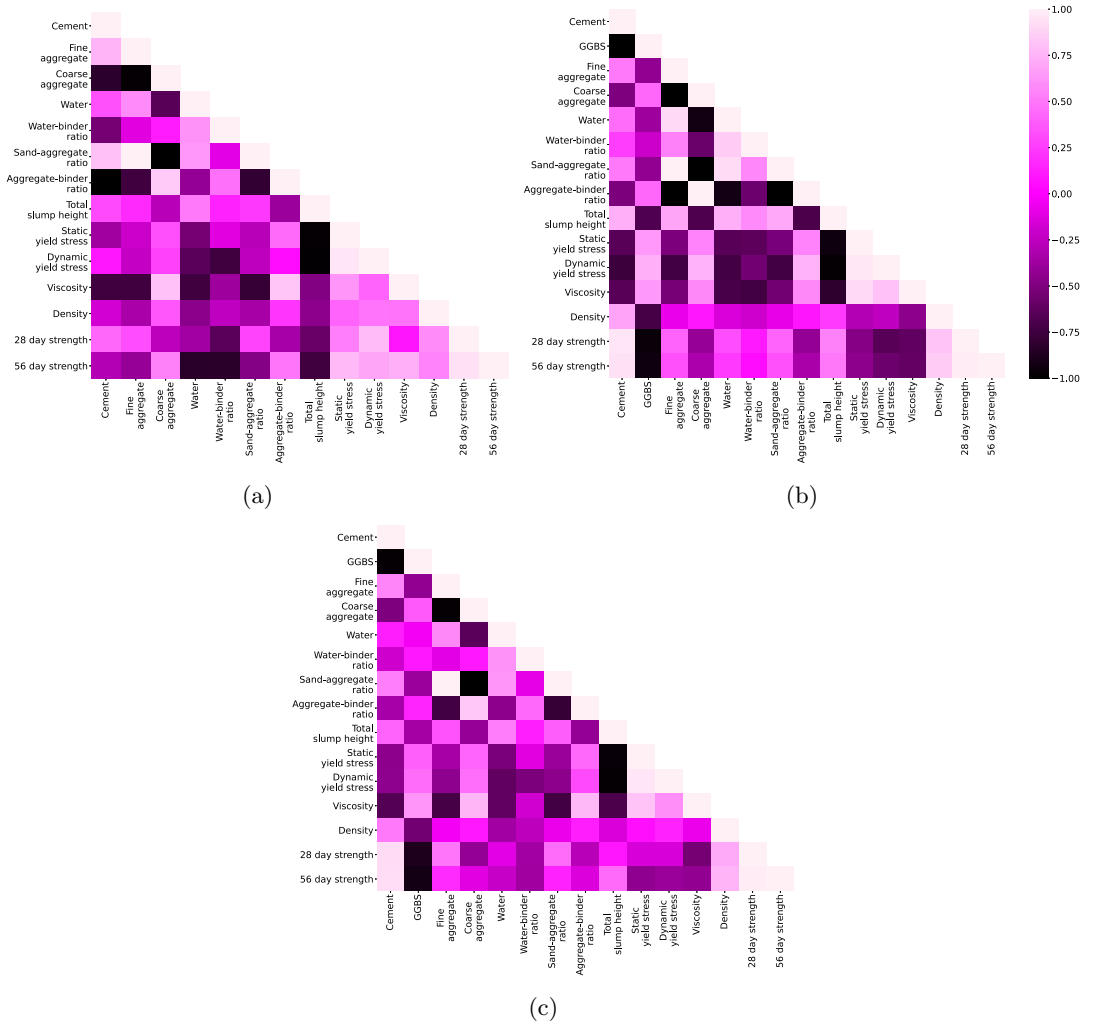


Figure 3. *(a) Correlation map for the OPC mixes within the dataset, (b) correlation map for the GGBS mixes within the dataset, and (c) correlation map for the unified dataset. The right-hand scale bar calibrates the correlation values.*

strongly correlated with the mix proportions, particularly with coarse aggregate content (0.80 for OPC, 0.56 for GGBS, and 0.75 for unified) and with aggregate–binder ratio (0.82 for OPC, 0.56 for GGBS, and 0.76 for unified). Furthermore, a strong negative correlation of −0.98 for OPC, −0.94 for GGBS, and −0.96 for the unified dataset was observed between total slump height and static yield stress. This effect has been observed elsewhere (Ferraris and Larrard, 1998; Laskar, 2009; White and Lees, 2023).

In the GGBS dataset, a much stronger correlation between cement content and compressive strength at both 28 days (0.43 for OPC, 0.95 for GGBS) and 56 days (−0.30 for OPC, 0.93 for GGBS) is observed. Combined with the fact that strong negative correlations between GGBS content and compressive strength at both 28 days (−0.95) and 56 days (−0.94) mean that including GGBS within the mix has a significant negative effect on strength. Therefore, GGBS mixes should be included in the training data if this material is to be adopted in practice to enable accurate predictions of compressive strength with the presence of a novel material.

A negative correlation between the water–binder ratio and compressive strength is expected in both OPC and GGBS mixes (Neville, 1988; Newman and Choo, 2003; Fan et al., 2021). Increasing the water–

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

binder ratio means more water is available for the chemical reaction between binder and other materials, but too much water results in weak concrete with pores and large matrix spacing. In the OPC dataset, a much stronger and negative correlation between the water-binder ratio and compressive strength at both 28 days (−0.65 for OPC, 0.11 for GGBS) and 56 days (−0.83 for OPC, 0.05 for GGBS) is observed. The lack of this correlation in the GGBS dataset could spuriously arise due to the small size of the dataset. Therefore, including OPC mixes and forming a unified dataset would allow machine learning to capture the expected fundamental relationships based on non-material specific elements while still capturing any differences that appear due to the inclusion of alternative binder materials besides ordinary Portland cement.

As expected, in all three datasets, the 28 day strength is highly correlated to the 56 day strength, meaning this earlier strength measure could be used to accurately predict the 56 day strength, following the workflow in Figure 2. Overall, Figure 3 displays strong correlations between many of the variables within the dataset, which can be utilized by machine learning to make predictions. These correlations are overall most accurately reflected when both GGBS and OPC mixes are included in the dataset. Therefore, moving forward, we adopt the unified dataset to train machine learning models, and a unified dataset is suggested for practical adoption of machine learning for emerging materials with sparse empirical data. This section has provided an understanding of the key properties of concrete and their interrelationships. The following section will outline the machine learning methodology used to predict concrete strength. It will cover the training process, the performance metrics adopted for the model, and comparisons of the model's performance.

3. Machine learning methodology

Following the identification of the key relationships between concrete properties through its fresh and hardened states, the machine learning methodology used to predict concrete strength and uncertainty is discussed in this section. First, in Section 3.1, we describe the machine learning algorithm used in this work based on the property relationships discussed previously. Second, in Section 3.2, we describe the development of a single metric that incorporates both the accuracy of model predictions and the quality of uncertainty estimates. Finally, in Section 3.3, we demonstrate that optimization of the single metric leads to a model that gives overall better predictions of strength and its uncertainty. Based on the single metric, we demonstrate that using intermediate measures such as slump as an input improves predictions of longer-term strength (56 day strength) and its uncertainty. This informs the strategy of real-time strength prediction, as discussed in later sections.

3.1. Model training

This section first reviews the available tools for concrete strength prediction, a process which was guided by existing reviews of soft computing processes and the results of other prediction models for these purposes, e.g.Alkayem et al. (2024)). Supervised learning processes are most common and examples of widely used machine learning models are *k*-means clustering (Cohn and Holm, 2021), neural networks (Hastie et al., 2001), and Gaussian processes (Tancret, 2013). This work adopts a two-layer random forest model (Zviazhynski and Conduit, 2023). This model is computationally cheap, robust against outliers, and can uncover non-linear relationships from sparse data. This method has been used previously to successfully design concrete mixes (Forsdyke et al., 2023) and, when compared to k-nearest neighbor and other models, has produced the best accuracy and more stability with less susceptibility to preprocessing of data (Ghunimat et al., 2023; Verma et al., 2023). The Scikit-Learn Python package (Pedregosa et al., 2011) is adopted for the random forest model.

The two-layer random forest model can be seen in Figure 4a. The first layer random forest model is trained on the concrete mix proportions (Flow A2) to predict all of the intermediate and output variables, such as 28-day strength, alongside their uncertainties (Flow A3). The second layer random forest model is then used, which takes the concrete mix proportions (Flow A1) and other recently predicted variables and

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

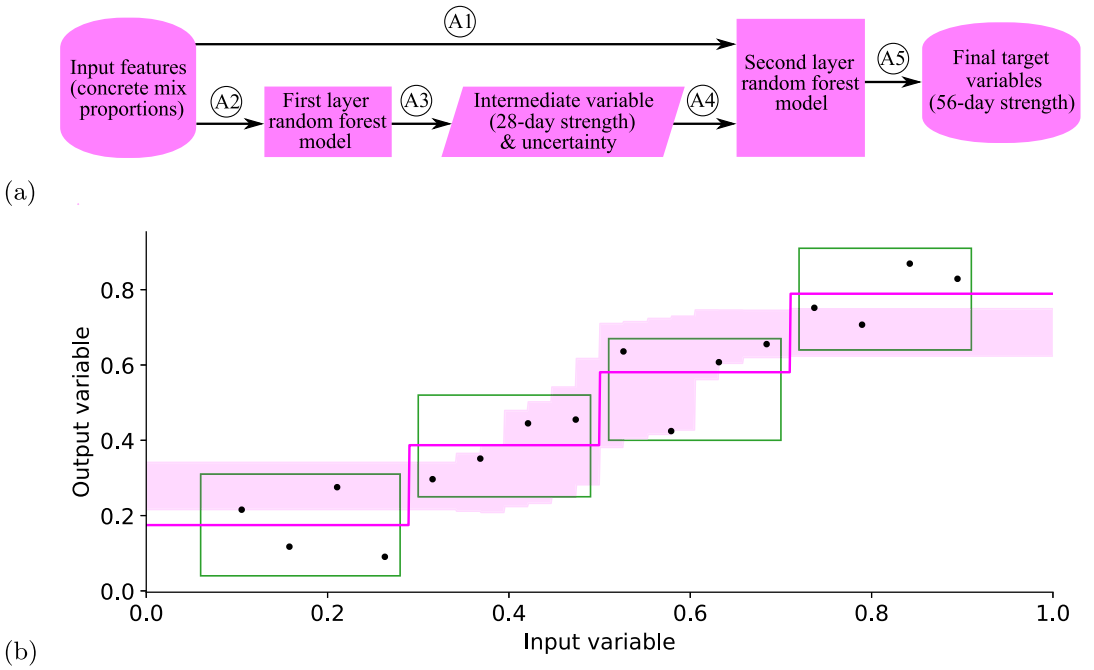


Figure 4. *(a) The two-layer random forest model. (b) A graphical description of a tree with min_samples_leaf = 4. Green boxes are tree leaves, black points are training data, the pink stepped line is the prediction of a single tree with all data present, and the pink shaded area is the range of the predictions from the ensemble of trees.*

uncertainties (Flow A4) to predict the final output, such as 56-day strength (Flow A5). Random forest model predictions are affected by their hyperparameters. In this work, to achieve the best overall predictions of the output variable and its uncertainty, we vary the *min_samples_leaf* hyperparameter, which is the minimum number of samples in a leaf of each constituent decision tree within a random forest, which, when increased, allows it to average out the noise.

An example of fitting a decision tree with *min_samples_leaf* = 4 is illustrated in Figure 4b. The tree leaves are represented by green boxes, each containing at least four training data points (black points). The resulting prediction is calculated as the average value of the output variable within each leaf and is represented by the pink stepped line. Our final model is an ensemble of such decision trees and is called random forest. Each decision tree within random forest is trained on a bootstrap sample from the training set, which is obtained by repeatedly drawing an entry from the training set randomly with replacement (Hastie et al., 2001). The differences among the bootstrap samples lead to differing decision trees that give a range of predictions represented by the pink shaded area in Figure 4b. The predictions are averaged to give the overall prediction along the center of the pink shaded area (without overfitting because they average across different samples); their standard deviation is the uncertainty in the overall prediction, which is represented by the width of the pink-shaded area. The performance of the adopted model is reviewed in the next section.

3.2. Model performance metrics

To evaluate the performance of the developed random forest model, three metrics are used. The first metric–R^2–captures how close the predictions are to the true values. The second metric–Γ–captures how accurately the model estimates uncertainty. The third metric is proposed in this work, α, which combines the first two metrics. Each of these metrics is discussed in turn.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

3.2.1. Coefficient of determination (R^2)

The coefficient of determination, R^2, assesses the quality of model predictions of strength. The formula for R^2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2} \tag{1}$$

where Y_i is the true value of strength for the i^{th} mix of N, $\hat{Y}_i$ is machine learning prediction of Y_i, and $\bar{Y}$ is the mean value of Y. R^2 ranges from 1 for perfect predictions to $-\infty$ for arbitrarily inaccurate predictions.

The coefficient of determination (R^2) is chosen over other metrics, in particular the widely used Pearson correlation coefficient (r^2), because the coefficient of determination not only accounts for the slope of the predictions versus experimental values, but also accounts for any possible shift. The Pearson correlation coefficient, on the other hand, only reports whether there is a linear relationship between predictions and experimental values, so it would report a perfect fit even if either the data were shifted or had different slopes (Neter et al., 1989; Whitehead et al., 2019).

3.2.2. Distribution of uncertainty quality (Γ)

The quality of uncertainty estimates from the model is now considered. We first calculate the error in machine learning prediction scaled by uncertainty–ε_i–for each mix:

$$\varepsilon_i = \frac{Y_i - \hat{Y}_i}{\sigma_{\hat{Y}_i}} \tag{2}$$

where Y_i is the true value of strength for the i^{th} mix, $\hat{Y}_i$ is the machine learning prediction, and $\sigma_{\hat{Y}_i} = \sqrt{\sigma_{i,ml}^2 + \sigma_{i,exp}^2}$ is the uncertainty, calculated as the quadrature sum of machine learning uncertainty $\sigma_{i,ml}$ and experimental uncertainty $\sigma_{i,exp}$.

The values of ε_i are accumulated across all the predictions and then binned into a histogram. This histogram is then compared to the histogram of a standard normal distribution, as shown schematically in Figure 5.

When uncertainty estimates are accurate, the weight in each bin of the ε histogram and the normal distribution histogram is equal. Therefore, the similarity between the two histograms is represented by the distribution of uncertainty quality Γ:

$$\Gamma = \frac{1}{2} \frac{M}{M-1} \sum_{m=1}^{M} \left| \frac{n_m}{N} - \frac{1}{M} \right| \tag{3}$$

where n_m is the number of entries in the m^{th} bin of M. Bins are chosen so that the normal distribution has equal weight of $1/M$ in each. The number of bins M is typically chosen to be close to $\sqrt{N}$ (Lohaka, 2007). The quantity Γ is known as the distribution of uncertainty quality (Taylor and Conduit, 2022). The quantity Γ ranges from 0 for perfect uncertainty estimates to 1 for poor uncertainty estimates.

3.2.3. Single metric for model performance

A single metric α that combines R^2 and Γ is proposed to identify the overall best model. The metric favors high R^2 value and low Γ value, and is given by:

$$\alpha = R^2 - \frac{1}{2}\Gamma \tag{4}$$

The factor $\frac{1}{2}$ is chosen so that variation of each term contributes approximately equally to the variation of α (Taylor and Conduit, 2022). The final metric achieves the peak value of 1. If all predictions are shifted by a constant, both R^2 and Γ get worse, so the metric will award uncertainty estimates becoming larger, which is the desired behavior in this scenario.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

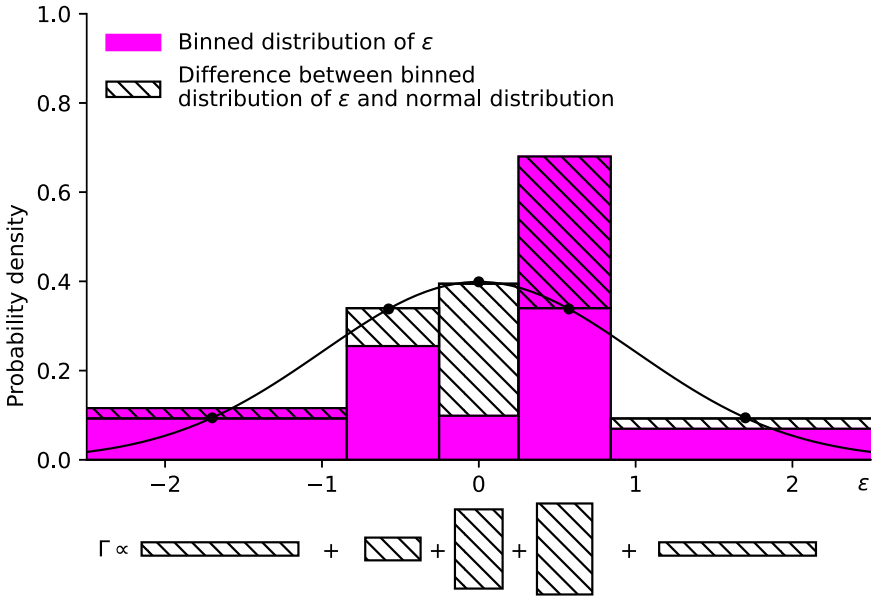
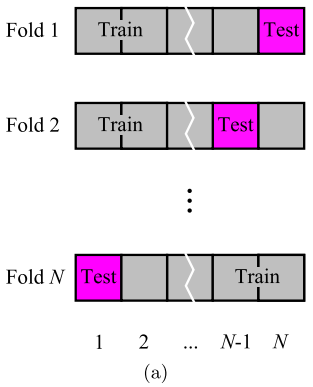


Figure 5. *Binned distribution of ε (magenta rectangles) and standard normal distribution (white transparent rectangles).*

3.3. Comparison of model performance

In the previous section, a combined performance metric integrating accuracy and uncertainty measurement was introduced. This section utilizes this metric to evaluate the performance of the proposed random forest model. The machine learning model has hyperparameters that must be selected to give optimal model performance. In order to select hyperparameters leave-one-out cross-validation is performed (Hastie et al., 2001): we start with a putative set of hyperparameters, train a model on all but one row of the dataset, and then blind test against the held-out row. As shown in Figure 6, the procedure is then repeated N times, with a different held-out row each time, until we obtain predictions for the entire dataset of N rows, and the accuracy of those predictions is found using the metrics R^2, Γ, and α. The blind testing in cross-validation is essential, as it ensures that we are not simultaneously training and testing the model against the same piece of data to avoid data leakage, and to not overfit the training data, and thereby



min_samples_leaf	R^2	Γ	α
1	0.844	0.267	0.711
2	0.834	0.200	0.734
3	0.790	0.267	0.657

(b)

Figure 6. *(a) Schematic of leave-one-out cross-validation. Gray squares are entries in the existing data and magenta squares are the test entries for each fold. (b) Model performance for different values of min_samples_leaf .*

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

replicating the real-life usage of the machine learning model to design a concrete formulation that has never before been produced. The suggested hyperparameters are then adjusted and the metrics R^2, Γ, α recomputed iteratively until the hyperparameters that give the best metric are found, here, we focus on the α metric. Below we illustrate the procedure by focusing on just one hyperparameter, which also serves to highlight the benefits of optimizing the α metric.

We vary the *min_samples_leaf* hyperparameter of the model, which is the minimum number of samples in a leaf of each tree within a random forest, and identify its values that lead to the optimum R^2, Γ, and α for the 56-day strength prediction. The metrics for each value of *min_samples_leaf* considered are shown in Figure 6.

The table in Figure 6 shows that *min_samples_leaf* $= 1$ it offers the highest quality of predictions R^2 but a relatively poor quality of uncertainty Γ. Increasing *min_samples_leaf* to 2 better averages the noise, so gives an improvement in Γ, but a relatively small decrease in R^2, resulting in the best value of α. Therefore, the model selected by optimizing R^2 is generally different from the model chosen by optimizing α. This conclusion indicates that the inclusion of uncertainty quality affects the appropriate selection of the machine learning model. This has significant implications for previous work that has omitted consideration of the quality of the uncertainty estimates. Therefore, going forward, we consider the performance metric α, as α incorporates both R^2 and quality of uncertainty as defined in Equation 4. This metric is calculated during the leave-one-out cross-validation procedure, and model selection determines the hyperparameters that maximize its value.

The input variables that yield the best overall results for strength prediction are now investigated, using the previously proposed just-in-time approach. Table 1 demonstrates model performance depending on whether slump and/or 28 day strength inputs are used. Mix proportions include ordinary Portland cement, GGBS, fine aggregate, coarse aggregate, and water content used alongside the water-binder, sand-aggregate, and aggregate-binder ratios. The last row of the table shows the results for 56-day predictions when 28-day strength is included as an additional input. For each set of inputs and the output, the best value of *min_samples_leaf* is selected by optimizing α on leave-one-out cross-validation. This value of *min_samples_leaf* was different from the value of *min_samples_leaf* that optimizes R^2 for 28-day predictions using slump and 56-day predictions without using slump.

Including slump improves R^2 from 0.820 to 0.848 and also improves Γ, which reduces from 0.207 to 0.172, demonstrating the utility of fresh-state concrete properties for 28-day strength predictions. The inclusion of slump as an input for 56 day strength predictions improves Γ (0.200 to 0.133) at a cost of a marginally smaller R^2 (0.834 to 0.821). The noise observed in the slump measurements reflects the underlying variability of concrete constituents mixed together and is captured by the model, resulting in better uncertainty estimates for strength. When 28-day strength is used as an input to predict 56-day strength, we see an improvement in R^2 from 0.821 to 0.897, driven by the strong correlation between 28-day strength and 56-day strength. We show the performance of this final model in Figure 7, where we plot predicted versus experimentally measured 56-day strength. We observe that predictions agree with experimental results within model uncertainty, confirming the strong correlations between properties and the utility of fresh-state measurements to predict strength using the random forest model.

Table 1. *Model performance for different inputs and outputs*

Day of strength prediction	Uses slump as input?	Uses 28-day strength as input?	R^2	Γ	α
28	No	N/A	0.820	0.207	0.717
	Yes	N/A	0.848	0.172	0.762
56	No	No	0.834	0.200	0.734
	Yes	No	0.821	0.133	0.754
56	Yes	Yes	0.897	0.200	0.797

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

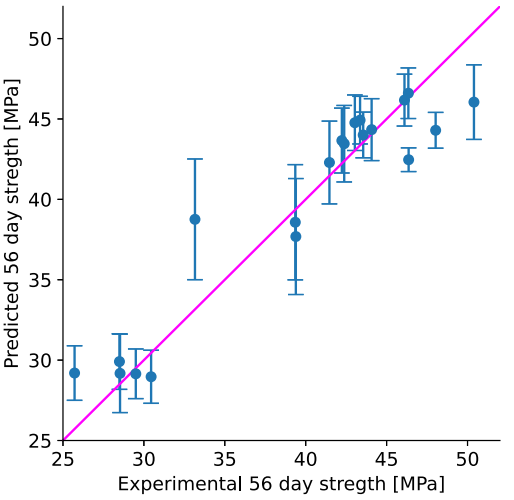


Figure 7. *The predicted versus experimental 56 day strength with model uncertainty (blue points). The magenta line shows the ideal trend.*

The effectiveness of this approach to predict the performance of unseen concrete mixes is evaluated in the following section.

4. Experimental validation

Given the optimization of the machine learning model in the previous section, this section considers the performance of the model to predict the strength of previously unseen concrete mixes. In particular, the benefit of incorporating intermediate data (slump and earlier age strength) to improve longer-term strength predictions and confidence intervals of the two concrete mixes shown in Table 2 is considered. The two mixes are expected to have significantly different compressive strengths and, hence, represent a good opportunity to test the predictions of the developed model across a strength range. Furthermore, these mixes are interesting from a concrete technology perspective as they both contain high levels of OPC replacement materials (50% GGBS content) and, therefore, have a lower carbon content compared to OPC-only mixes. Mix B contains a low aggregate-binder ratio of 3.0, so it would be of particular interest to practitioners as compressive strength is expected to increase with a decreased aggregate-binder ratio and hence offers an opportunity to somewhat offset the reduced strength offered by GGBS inclusion while maintaining sustainability credentials (Poon and Lam, 2008). Furthermore, this mix sits outside the binder content and aggregate-binder ratio within the training data, so it offers an opportunity to test the

Table 2. *Comparison of Mix A, Mix B, and training data parameters*

Property	Mix A	Mix B	Training data
Percent OPC replacement (%)	50	50	0–70
Total binder content (kg/m^3)	338	512	260–418
Aggregate–binder ratio	5.0	3.0	3.9–7.0
Water–binder ratio	0.67	0.49	0.50–0.91
Sand–aggregate ratio	0.50	0.50	0.48–0.62
Total slump height (mm)	150–230[a]	150–230[a]	40–240
28 day compressive strength (MPa)	30[a]	40[a]	24.5–51.4

[a]The property target values.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

Table 3. *Mix composition and fresh and hardened testing results for the experimental series*

Property	Mix A	Mix B
Mix proportions		
OPC (kg/m^3)	169	256
GGBS (kg/m^3)	169	256
Total binder content (kg/m^3)	339	512
Fine aggregate (kg/m^3)	847	768
Coarse aggregate (kg/m^3)	847	768
Water (kg/m^3)	227	249
Aggregate–binder ratio	5.0	3.0
Water–binder ratio	0.67	0.49
Sand–aggregate ratio	0.50	0.50
Fresh state measurements		
Total slump height (mm)	180	150
Hardened state measurements		
Mean hardened density (kg/m^3)	2270	2275
Day compressive strength (MPa)	28.0 ± 0.6	39.3 ± 1.0
Day compressive strength (MPa)	32.8 ± 0.9	43.7 ± 0.9

extension capabilities of the model in the presence of sparse training data, replicating practical conditions. The mix proportions and test results attained during the experimental series are shown in Table 3, and further details are provided in the Supplementary Material.

Machine learning predictions of 28-day strength and the experimental value of compressive strength obtained for Mix A are presented in Figure 8a. Without using slump as an input, machine learning predicts a 90% confidence interval that does not overlap with the 90% confidence interval of the experimental result. When adding slump as an input, machine learning takes advantage of the additional information to not only give a predicted strength that is closer to the experimental value but also a more realistic uncertainty estimate, which results in a 90% confidence interval that includes the experimental value. This is advantageous given the predicted strength is above the experimental results in both models. The machine learning predictions of 28 day strength and its experimental value for Mix B are presented in Figure 8b. Both models give 90% confidence intervals that include the experimental value. However,

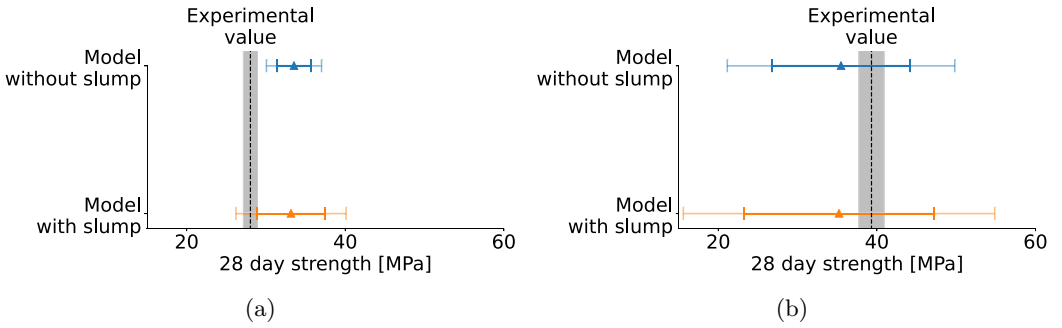


Figure 8. *Predictions of 28-day strength for (a) Mix A using the model without slump (blue) and model with slump (orange) and (b) Mix B using the model without slump (blue) and model with slump (orange). Small error bars represent standard deviation of predictions, large error bars correspond to 90% confidence intervals of predictions. Dashed vertical line is the experimental value, and gray shaded area is 90% confidence interval of the experimental value.*

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

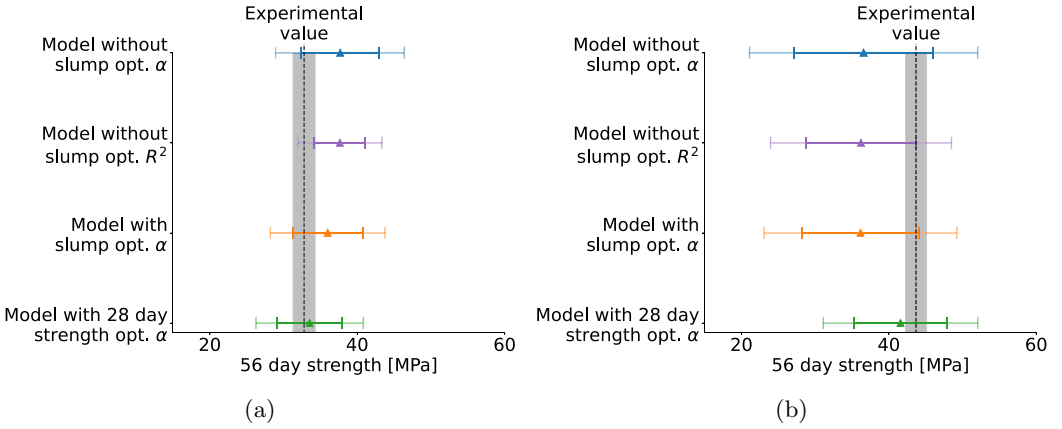


Figure 9. *Predictions of 56 day strength for (a) Mix A using the model without slump that optimizes α (blue), model without slump that optimizes R^2 (purple), model with slump that optimizes α (orange), and model with slump and 28 day strength that optimizes α (green); (b) Mix B using the model without slump that optimizes α (blue), model without slump that optimizes R^2 (purple), model with slump that optimizes α (orange), and model with slump and 28 day strength that optimizes α (green). Small error bars represent standard deviation of predictions, large error bars correspond to 90% confidence intervals of predictions. Dashed vertical line is the experimental value, and gray shaded area is 90% confidence interval of the experimental value.*

error bars are large for predictions from both models. This is expected since Mix B has a formulation far from those within the training dataset. The model with slump gives a slightly more conservative strength prediction with a higher uncertainty, which is better for practical applications.

The machine learning predictions of 56-day strength and the experimental value of compressive strength obtained for Mix A are presented in Figure 9a. First of all, prediction from the model that optimizes α agrees with the experimental value within one standard deviation, whereas prediction from the model that optimizes R^2 does not. This endorses the approach to optimize α rather than R^2 to select the model that gives both accurate predictions and good uncertainty estimates. Furthermore, prediction improves as fresh state data is added to the model input. When 28-day strength is included as input, the prediction becomes even closer to experimental value, again supporting the live updating of strength predictions as information becomes available. The machine learning predictions of 56-day strength and the experimental value of compressive strength obtained for Mix B are presented in Figure 9b. Again, with 28-day strength as an input, the predicted strength significantly improved, and moreover, the uncertainty is reduced owing to the strong correlation between 28 and 56-day strength, demonstrating the utility of the proposed machine learning approach to accurately predict strength and associated uncertainty. In the following section, the practical limitations and suggested uses for the developed model and approach are discussed.

5. Discussion

Following the demonstration of the utility of fresh state properties and a combined accuracy-uncertainty metric in a newly developed machine learning approach for concrete strength prediction, this section discusses the limitations and future improvements of the work. In Section 5.1, the limitations of the work are reviewed. In Section 5.2, the ramifications for the concrete industry are set out, and for machine learning practitioners in Section 5.3.

5.1. Limitations and future improvements

In this section, the key limitations of the modeling are reviewed, and suggestions for future improvements are presented. Data is key for machine learning to train a detailed, accurate model and to produce a model

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

with the breadth to apply to many concrete mixes. In this study, we adopted a dataset with only 29 mixes, which is comparable to the datasets that exist in practice for materials where extensive mix trials have not been conducted. Although the adopted dataset and training approach led to accurate strength predictions, the developed model may have limited extension capabilities to other lower-carbon materials. The full extension capabilities to alternative materials, e.g., calcined clays and fly ashes, have not been tested in the current work. Therefore, a key topic for future work is to gather more data to train the machine learning model on alternative emerging materials, and to conduct experimental validation on the proposed machine learning model for such materials.

The paper demonstrates the benefits of optimizing the model that gives excellent uncertainty estimates, and how uncertainty estimates can be effectively used to focus on robust predictions. However, the simultaneous optimization of quality and uncertainties in the metric α, with relative weighting of $1/2$ in Equation 4, should be explored further in a theoretical work. This future work should ascertain whether the optimum weighting factor should dynamically change. For example, when model quality is poor, greater weight should be assigned to the quality of uncertainty to deliver an improved understanding of robustness.

The paper focused solely on predicted strength at 28 and 56 days. However, many other properties of interest in the commercial use of concrete, including carbon footprint, density, shrinkage, durability, workability, permeability, and cost, are also of interest. Follow-up work should extend the proposed machine learning approach for these properties to develop models that can be used practically in the real world.

5.2. Uses for concrete industry digitization

This work demonstrates that using fresh state variables as inputs for machine learning can improve the accuracy of strength predictions and the quality of uncertainty estimates in these predictions. The uncertainty estimates provided by the tool can be compared with the risk profile of the concrete mix, for example, the cost and development time associated with the rejection of the mix if deemed unsuitable. For practical applications, this means that the inclusion of recently obtained slump test results as a model input can improve the accuracy of strength predictions and estimates of uncertainty. Therefore, by updating strength predictions with the slump attained at delivery, a real-time decision on whether to accept the concrete mix for use at the construction site, or to revisit the question after 28 days. This is particularly valuable in the context of growing uncertainty with smaller datasets containing concrete mixes with a wider range of constituent materials.

Just-in-time predictions support the growing interest in digitizing the concrete sector and the broader construction industry. Recently, methods have been developed to track concrete materials across the supply chain, linking constituents to test data, including slump and strength testing (White et al., 2024). This work's machine learning strength prediction methodology can be embedded in such modular digital frameworks. For the end user, automated strength predictions for each concrete element could be observed and updated as time progresses, and new test data added. Furthermore, in a similar fashion to strength prediction, the prediction of carbonation permeability or long-term durability may be aided by intermediate measures such as slump height or fresh density and is, therefore, the subject of further investigation.

5.3. Uses for machine learning applications

This work has demonstrated that it is crucial to have a metric that can assess both the quality of predictions and uncertainty estimates concurrently to determine the optimum machine learning model. The approach set out in this study allows for selecting the most suitable model for predicting concrete strength, and its inputs can be adjusted as new test data becomes available to improve strength predictions. This process can help decide whether the concrete should be approved at a construction site. This method represents a significant improvement over previous efforts, which solely focused on prediction accuracy and relied on

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

large amounts of test data that may not be available for emerging low-carbon materials. The approach to assess both predictions and uncertainty estimates has potential applications in areas beyond concrete mix design, including financial markets (Sharpe, 1966), cancer diagnosis (Hunter et al., 2022), additive manufacturing (Rasiya et al., 2021), information engineering (Płaczek and Bernaś, 2014), and language processing (Zerva et al., 2017).

6. Conclusions

Limited empirical data exist for emerging lower-carbon materials, which presents challenges in adopting machine learning approaches to predict key properties, such as concrete strength. This work used a random forest machine learning model to predict concrete strength when trained on a sparse dataset, reflecting the datasets observed for emerging materials. A novel metric to assess prediction performance that combines accuracy and uncertainty estimates is proposed to reflect the pertinent concerns of sparse datasets. Model selection using this metric differs from the traditional method, which relies solely on accuracy metrics.

In addition to a model assessment method that combined accuracy and uncertainty, fundamental understanding of the concrete lifecycle has led to developing a just-in-time machine learning prediction technique. This approach uses intermediate test results, such as slump and early age strength, as inputs for a machine learning model, meaning long-term strength prediction accuracy and uncertainty are improved through time. Cross validation and experimentation of unseen concrete mixes outside the sparse training data demonstrates that the just-in-time approach results in progressively more accurate confidence intervals of strength predictions. Predictions of 56-day strength for the two mixes were experimentally validated to within a 90% confidence interval when using slump as an input and further improved by using 28-day strength. Overall, these results are promising for the practical implementation of machine learning in the concrete industry, using real-time measurements throughout the life cycle.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/dce.2025.10018.

Acknowledgements. The authors acknowledge the support of Heidelberg Materials UK for the provision of cementitious materials and Wisdom Asotah for useful discussion.

Author contribution. Conceptualization: B.Z., C.W., J.M.L., G.J.C.; Data curation: C.W.; Formal analysis: B.Z., C.W.; Funding acquisition: J.M.L., G.J.C.; Investigation: C.W.; Methodology: B.Z., C.W.; Project administration: J.M.L., G.J.C.; Software: B.Z.; Supervision: J.M.L., G.J.C.; Validation: B.Z., C.W.; Visualization: B.Z., C.W.; Writing—original draft: B.Z., C.W.; Writing—review and editing: J.M.L., G.J.C.

Competing interests. B.Z., C.W., and J.M.L. confirm there are no competing interests to declare. G.C. is a Director of the machine learning company Intellegens.

Data availability statement. Supplementary material containing the data utilized in this work has been provided.

Funding statement. This work was supported by the Engineering and Physical Sciences Research Council and Costain PLC [grant number EP/S02302X/1], the Engineering and Physical Sciences Research Council [grant number EP/N017668/1], and the Harding Distinguished Postgraduate Scholars Programme Leverage Scheme.

References

Agrawal A, **Deshpande PD**, **Cecen A**, **Basavarsu GP**, **Choudhary AN and Kalidindi SR** (2014) Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. *Integrating Materials and Manufacturing Innovation 3*, 1–19. https://doi.org/10.1186/2193-9772-3-8.

Alkayem NF, **Shen L**, **Mayya A**, **Asteris PG**, **Fu R**, **Di Luzio G**, **Strauss A and Cao M** (2024) Prediction of concrete and frc properties at high temperature using machine and deep learning: A review of recent advances and future perspectives. *Journal of Building Engineering 83*, 108369. https://doi.org/10.1016/j.jobe.2023.108369.

Bhadeshia H, **MacKay D and Svensson L** (1995) Impact toughness of C-Mn steel arc welds – Bayesian neural network analysis. *Materials Science and Technology 11*, 1046–1051. https://doi.org/10.1179/mst.1995.11.10.1046.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

BSI (2019a) *BS EN 12350-2:2019. Testing Fresh Concrete –Slump Test*. London: The British Standards Institution. Available at https://knowledge.bsigroup.com/products/testing-fresh-concrete-slump-test-2/tracked-changes (Accessed 10.08.2024).

BSI (2019b) *BS EN 12390-3:2019. Testing Hardened Concrete –Compressive Strength of Test Specimens*. London: The British Standards Institution. Available at https://knowledge.bsigroup.com/products/testing-hardenedconcrete-compressive-strength-of-test-specimens-1?version=tracked (Accessed 10.08.2024).

Chaabene WB, **Flah M and Nehdi ML** (2020) Machine learning prediction of mechanical properties of concrete: Critical review. *Construction and Building Materials 260*, 119889. https://doi.org/10.1016/j.conbuildmat.2020.119889.

Chakravarthy HG, **Seenappa KM**, **Naganna SR and Pruthviraja D** (2023) Machine learning models for the prediction of the compressive strength of self-compacting concrete incorporating incinerated bio-medical waste ash. *Sustainability 15*(18). https://doi.org/10.3390/su151813621.

Chi L, **Wang M**, **Liu K**, **Lu S**, **Kan L**, **Xia X and Huang C** (2023) Machine learning prediction of compressive strength of concrete with resistivity modification. *Materials Today Communications 36*, 106470. https://doi.org/10.1016/j.mtcomm.2023.106470.

Chopra P, **Sharma RK**, **Kumar M and Chopra T** (2018) Comparison of machine learning techniques for the prediction of compressive strength of concrete. *Advances in Civil Engineering 2018*, 1–9. https://doi.org/10.1155/2018/5481705.

Chou J-S, **Tsai C-F**, **Pham A-D and Lu Y-H** (2014) Machine learning in concrete strength simulations: Multi-nation data analytics. *Construction and Building Materials 73*, 771–780. https://doi.org/10.1016/j.conbuildmat.2014.09.054.

Cohn R and Holm E (2021) Unsupervised machine learning via transfer learning and k-means clustering to classify materials image data. *Integrating Materials and Manufacturing Innovation 10*, 231–244. https://doi.org/10.1007/s40192-021-00205-8.

Conduit BD, **Jones NG**, **Stone HJ and Conduit GJ** (2017) Design of a nickel-base superalloy using a neural network. *Materials & Design 131*, 358. https://doi.org/10.1016/j.matdes.2017.06.007.

Conduit BD, **Jones NG**, **Stone HJ and Conduit GJ** (2018) Probabilistic design of a molybdenum-base alloy using a neural network. *Scripta Materialia 146*, 82. https://doi.org/10.1016/j.scriptamat.2017.11.008.

Conduit BD, **Illston T**, **Baker S**, **Duggappa DV**, **Harding S**, **Stone HJ and Conduit GJ** (2019) Probabilistic neural network identification of an alloy for direct laser deposition. *Materials & Design 168*, 107644. https://doi.org/10.1016/j.matdes.2019.107644.

Deng F, **He Y**, **Zhou S**, **Yu Y**, **Cheng H and Wu X** (2018) Compressive strength prediction of recycled concrete based on deep learning. *Construction and Building Materials 175*, 562–569. https://doi.org/10.1016/j.conbuildmat.2018.04.169.

Fan H, **Dhir R and Hewlett P** (2021) Ggbs use in concrete as cement constituent: Strength development and sustainability effects – part 1. *Magazine of Concrete Research 24*, 1–41. https://doi.org/10.1680/jmacr.21.00009.

Feng D-C, **Liu Z-T**, **Wang X-D**, **Chen Y**, **Chang J-Q**, **Wei D-F and Jiang Z-M** (2020) Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials 230*, 117000. https://doi.org/10.1016/j.conbuildmat.2019.117000.

Ferraris C and Larrard F (1998) *Testing and Modelling of Fresh Concrete Rheology (Tech. Rep.)*. Gaithersburg: National Institute of Standards and Technology. Available at https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir6094.pdf (Accessed 5.08.2024).

Forsdyke JC, **Zviazhynski B**, **Lees JM and Conduit GJ** (2023) Probabilistic selection and design of concrete using machine learning. *Data Centric Engineering 4*, 1–18. https://doi.org/10.1017/dce.2023.5.

Ghunimat D, **Alzoubi A and Alzboon A** (2023) Prediction of concrete compressive strength with GGBFS and fly ash using multilayer perceptron algorithm, random forest regression and *k*-nearest neighbor regression. *Asian Journal of Civil Engineering 24*, 169. https://doi.org/10.1007/s42107-022-00495-z.

Hadzima-Nyarko M, **Nyarko EK**, **Lu H and Zhu S** (2020) Machine learning approaches for estimation of compressive strength of concrete. *The European Physical Journal Plus 135*, 682. https://doi.org/10.1140/epjp/s13360-020-00703-2.

Han B, **Wu Y and Liu L** (2022) Prediction and uncertainty quantification of compressive strength of highstrength concrete using optimized machine learning algorithms. *Structural Concrete 23*, 3772–3785. https://doi.org/10.1002/suco.202100732.

Hariri-Ardebili MA, **Mahdavi P and Pourkamali-Anaraki F** (2024) Benchmarking automl solutions for concrete strength prediction: Reliability, uncertainty, and dilemma. *Construction and Building Materials 423*, 135782. https://doi.org/10.1016/j.conbuildmat.2024.135782.

Hastie T, **Tibshirani R and Friedman J** (2001) *The Elements of Statistical Learning* (2). New York: Springer.

Hunter B, **Hindocha S and RW L** (2022) The role of artificial intelligence in early cancer diagnosis. *Cancers (Basel) 14*(6), 1524. https://doi.org/10.3390/cancers14061524.

Khan K, **Salami BA**, **Jamal A**, **Amin MN**, **Usman M**, **Al-Faiad MA**, **Abu-Arab AM and Iqbal M** (2022) Prediction models for estimating compressive strength of concrete made of manufactured sand using gene expression programming model. *Materials 15*, 5823. https://doi.org/10.3390/ma15175823.

Khursheed S, **Jagan J**, **Samui P and Kumar S** (2021) Compressive strength prediction of fly ash concrete by using machine learning techniques. *Innovative Infrastructure Solutions 6*, 149. https://doi.org/10.1007/s41062-021-00506-z.

Kim K, **Ward L**, **He J**, **Krishna A**, **Agrawal A and Wolverton C** (2018) Machine-learning-accelerated high-throughput materials screening: Discovery of novel quaternary heusler compounds. *Physical Review Materials 2*, 123801. https://doi.org/10.1103/PhysRevMaterials.2.123801.

Laskar AI (2009) Correlating slump, slump flow, vebe and flow tests to rheological parameters of high-performance concrete. *Materials Research 12*, 75–81. https://doi.org/10.1590/S1516-14392009000100009.

Lee S, **Nguyen N**, **Karamanli A**, **Lee J and Vo TP** (2023) Super learner machine-learning algorithms for compressive strength prediction of high performance concrete. *Structural Concrete 24*, 2208–2228. https://doi.org/10.1002/suco.202200424.

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

Li D, **Tang Z**, **Kang Q**, **Zhang X and Li Y** (2023) Machine learning-based method for predicting compressive strength of concrete. *Processes 11*, 390. https://doi.org/10.3390/pr11020390.

Liu Y (2022) High-performance concrete strength prediction based on machine learning. *Computational Intelligence and Neuroscience 2022*, 1–7. https://doi.org/10.1155/2022/5802217.

Liu K, **Alam MS**, **Zhu J**, **Zheng J and Chi L** (2021) Prediction of carbonation depth for recycled aggregate concrete using ANN hybridized with swarm intelligence algorithms. *Construction and Building Materials 301*(124382). https://doi.org/10.1016/j.conbuildmat.2021.124382.

Lohaka H (2007) *Making a Grouped-Data Frequency Table: Development and Examination of the Iteration Algorithm*. Doctoral dissertation, Ohio University.

Lydon FD (1972) *Concrete Mix Design*. London: Applied Science Publishers.

Mansouri E, **Manfredi M and Hu J-W** (2022) Environmentally friendly concrete compressive strength prediction using hybrid machine learning. *Sustainability 14*, 12990. https://doi.org/10.3390/su142012990.

Miller SA, **Habert G**, **Myers RJ and Harvey JT** (2021) Achieving net zero greenhouse gas emissions in the cement industry via value chain mitigation strategies. *One Earth 4*(10), 1398–1411. https://doi.org/10.1016/J.ONEEAR.2021.09.011.

Murthy MN, **Amruth SK**, **Marulasiddappa SB and Naganna SR** (2024) Modeling the compressive strength of binary and ternary blended high-performance concrete mixtures using ensemble machine learning models. *Soft Computing 28*(9). https://doi.org/10.1007/s00500-023-09521-x.

Neter J, **Wasserman W and Kutner M** (1989) *Applied Linear Regression Models*. Burr Ridge, IL: Irwin Professional Publishing.

Neville AM (1988) *Properties of Concrete*, Essex: Longmann Scientific and Technical.

Newman J and Choo BS (2003) *Advanced Concrete Technology: Constituent Materials*, London: Elsevier.

Nguyen H, **Vu T**, **Vo TP and Thai H-T** (2021a) Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials 266*, 120950. https://doi.org/10.1016/j.conbuildmat.2020.120950.

Nguyen N-H, **Vo TP**, **Lee S and Asteris PG** (2021b) Heuristic algorithm-based semi-empirical formulas for estimating the compressive strength of the normal and high performance concrete. *Construction and Building Materials 304*, 124467. https://doi.org/10.1016/j.conbuildmat.2021.124467.

Nunez I, **Marani A**, **Flah M and Nehdi ML** (2021) Estimating compressive strength of modern concrete mixtures using computational intelligence: A systematic review. *Construction and Building Materials 310*, 125279. https://doi.org/10.1016/j.conbuildmat.2021.125279.

Pakzad SS, **Roshan N and Ghalehnovi M** (2023) Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. *Scientific Reports 13*, 3646. https://doi.org/10.1038/s41598-023-30606-y.

Pedregosa F, **Varoquaux G**, **Gramfort A**, **Michel V**, **Thirion B**, **Grisel O**, **Blondel M**, **Müller A**, **Nothman J**, **Louppe G**, **Prettenhofer P**, **Weiss R**, **Dubourg V**, **Vanderplas J**, **Passos A**, **Cournapeau D**, **Brucher M**, **Perrot M**, **Duchesnay E** (2011) Scikit-learn: Machine learning in python. *Journal of Machine Learning Research 12*, 2825–2830.

Pham A-D, **Ngo N-T**, **Nguyen Q-T and Truong N-S** (2020) Hybrid machine learning for predicting strength of sustainable concrete. *Soft Computing 24*, 14965–14980. https://doi.org/10.1007/s00500-020-04848-1.

Płaczek B and Bernaś M (2014) Uncertainty-based information extraction in wireless sensor networks for control applications. *Ad Hoc Networks 14*, 106–117. https://doi.org/10.1016/j.adhoc.2013.11.009.

Poon CS and Lam CS (2008) The effect of aggregate-to-cement ratio and types of aggregates on the properties of pre-cast concrete blocks. *Cement and Concrete Composites 30*, 283–289. https://doi.org/10.1016/j.cemconcomp.2007.10.005.

Prayogo D, **Santoso D**, **Wijaya D**, **Gunawan T and Widjaja J** (2020) Prediction of concrete properties using ensemble machine learning methods. *Journal of Physics: Conference Series 1625*(012024). https://doi.org/10.1088/1742-6596/1625/1/012024.

Rasiya G, **Shukla A and Saran K** (2021) Additive manufacturing – a review [International Conference on Advances in Design, Materials and Manufacturing]. *Materials Today: Proceedings 47*, 6896–6901. https://doi.org/10.1016/j.matpr.2021.05.181.

Salami BA, **Olayiwola T**, **Oyehan TA and Raji IA** (2021) Data-driven model for ternary-blend concrete compressive strength prediction using machine learning approach. *Construction and Building Materials 301*, 124152. https://doi.org/10.1016/j.conbuildmat.2021.124152.

Sharpe W (1966) Mutual fund performance. *The Journal of Business 39*, 119–138. https://doi.org/https://doi.org/10.1086/294846.

Sourmail T, **Bhadeshia H and MacKay D** (2002) Neural network model of creep strength of austenitic stainless steels. *Materials Science and Technology 18*, 655–663. https://doi.org/10.1179/026708302225002065.

Taffese W, **Sistonen E and Puttonen J** (2015) CaPrM: Carbonation prediction model for reinforced concrete using machine learning methods. *Construction and Building Materials 100*, 70–82. https://doi.org/10.1016/j.conbuildmat.2015.09.058.

Tancret F (2013) Computational thermodynamics, gaussian processes and genetic algorithms: Combined tools to design new alloys. *Modelling and Simulation in Materials Science and Engineering 21*, 045013. https://doi.org/10.1088/0965-0393/21/4/045013.

Tattersall G (1991) *Workability and Quality Control of Concrete*. CRC Press, London.

Taylor PL and Conduit G (2022) Machine learning predictions of superalloy microstructure. *Computational Materials Science 201*, 110916. https://doi.org/10.1016/j.commatsci.2021.110916.

Teychenne DC, **Franklin RE and Emtroy HC** (1997) *Design of Normal Concrete Mixes*, Watford: BRE Press.

Tran VQ, **Mai H-VT**, **To QT and Nguyen MH** (2022) Machine learning approach in investigating carbonation depth of concrete containing fly ash. *Structural Concrete*. https://doi.org/10.1002/SUCO.202200269

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press

Turner LK and Collins FG (2013) Carbon dioxide equivalent (CO2-e) emissions: A comparison between geopolymer and OPC cement concrete. *Construction and Building Materials 43*, 125–130. https://doi.org/10.1016/J.CONBUILDMAT.2013.01.023.

Verma M, **Upreti K**, **Khan MR**, **Alam MS**, **Ghosh S and Singh P** (2023) Prediction of compressive strength of geopolymer concrete by using random forest algorithm. In Shaw RN, Paprzycki M and Ghosh A (eds.), *Advanced Communication and Intelligent Systems*. Cham: Springer Nature Switzerland, pp. 170–179.

Wan Z, **Xu Y and Šavija B** (2021) On the use of machine learning models for prediction of compressive strength of concrete: Influence of dimensionality reduction on the model performance. *Materials 14*, 713. https://doi.org/10.3390/ma14040713.

Ward L, **Agrawal A**, **Choudhary A and Wolverton C** (2016) A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials 2*, 1–7. https://doi.org/10.1038/npjcompumats.2016.28.

Ward L, **Liu R**, **Krishna A**, **Hedge V**, **Agrawal A**, **Choudhary A and Wolverton C** (2017) Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Physical Review B 19*. https://doi.org/10.1103/PhysRevB.96.024104.

White C and Lees JM (2023) Yield stress prediction from 3d reconstruction of fresh concrete slump. *Cement and Concrete Research 174*, 107331. https://doi.org/10.1016/J.CEMCONRES.2023.107331.

White C and Lees JM (2025) The concrete slump test – a rheometer by definition? *Construction and Building Materials 462*, 139839. https://doi.org/10.1016/j.conbuildmat.2024.139839.

White C, **Forsdyke J**, **Harrison M**, **Desnerck P and Lees JM** (2024) Digital threads for sustainable concrete: A prototype traceability system. *Structural Concrete 20*. https://doi.org/10.1002/suco.202400398

Whitehead T, **Irwin B**, **Hunt P**, **Segall M and Conduit G** (2019) Imputation of assay bioactivity data using deep learning. *Journal of Chemical Information and Modeling 59*, 1197.

Young BA, **Hall A**, **Pilon L**, **Gupta P and Sant G** (2019) Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cement and Concrete Research 115*, 379–388. https://doi.org/10.1016/j.cemconres.2018.09.006.

Zerva C, **Batista-Navarro R**, **Day P and Ananiadou S** (2017) Using uncertainty to link and rank evidence from biomedical literature for model curation. *Bioinformatics 33*(23), 3784–3792. https://doi.org/10.1093/bioinformatics/btx466.

Zviazhynski B and Conduit G (2023) Unveil the unseen: Exploit information hidden in noise. *Applied Intelligence 53*, 11966–11978. https://doi.org/10.1007/s10489-022-04102-1.

Cite this article: Zviazhynski B, White C, Lees JM and Conduit GJ (2025). Predicting concrete strength using machine learning and fresh-state measurements. *Data-Centric Engineering*, 6, e37. doi:10.1017/dce.2025.10018

https://doi.org/10.1017/dce.2025.10018 Published online by Cambridge University Press