Cleaner Materials 4 (2022) 100072

Contents lists available at ScienceDirect

Cleaner Materials

journal homepage: www.elsevier.com/locate/clema

An engineered ML model for prediction of the compressive strength of Eco-SCC based on type and proportions of materials

Ehsan Sadrossadat [a,*], Hakan Basarir [b], Ali Karrech [a], Mohamed Elchalakani [a]

[a] Department of Civil, Environmental and Mining Engineering, University of Western Australia, WA 6009, Australia
[b] Department of Petroleum and Geosciences, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Keywords:
Self-compacting concrete
Supplementary cementitious material
Prediction
Uniaxial compressive strength
Machine learning

ABSTRACT

Recently, various waste materials and industrial by-products such as supplementary cementitious materials (SCMs) have been proposed to improve the properties of self-compacting concrete (SCC). This profitable waste management strategy results in lowering the costs and carbon emission, and a more sustainable, cleaner and eco-friendly production of SCC (Eco-SCC). The properties of such a complex material are commonly measured through costly experiments. Researchers also proposed experimental data analysis and predictive modeling methods such as machine learning (ML) algorithms for prediction of the properties of concrete. However, proposed models commonly relate the properties to the proportion of constituents only and ignore the effect of their type and properties, and other influential factors. This paper aims to engineer the concept and develop a more efficient ML model for prediction of the 28-day uniaxial compressive strength (UCS_{28d}) of SCC containing SCMs. A comprehensive dataset is collected through a precise literature survey. Some dimensionless ratios are proposed to reduce the dimensionality of variables and reflect the effects of considered influential factors in different ML models. Two separate datasets are considered to test the predictability of models where one has new proportions of materials only and the other contains new type of material with new properties. After validation and comparison between various ML models, Gaussian process regression (GPR) model proved to perform well on both considered Test datasets with R^2, RMSE and MAE of around 0.96, 3.66 and 2.49 respectively. Sensitivity analysis results confirm the contribution and importance of considering type and properties of materials as model variables. This paper demonstrates and highlights that all influential factors must be considered to develop engineered ML models to use as universal tools for indirect estimation of properties of composite materials such as Eco-SCC.

1. Introduction

Self-Compacting Concrete (SCC), also named as self-consolidating concrete, is a class of concrete that is designed to flow and pass through congested reinforcement under its own weight without the need for getting externally vibrated after casting (Okamura et al., 2000). It offers many advantages over conventional concrete such as ease of placement resulting in reducing labour and overall cost, better segregation resistance, producing a denser and more homogenous concrete (Okamura et al., 2000). SCC is composed of aggregates and a binder paste consisting of water and cementitious material, typically Portland cement. The mixture design and choosing the type of constituents are dependent upon the requirements. Traditional SCC often requires higher amount of cement which increases its production cost and risk of thermal cracks (Long et al., 2015). Cement production is very energy-intensive and leads to emitting a large volume of carbon dioxide (CO_2) that is not environmentally-friendly. CO_2 is a reason of global warming and the reduction of CO_2 emission has become increasingly crucial nowadays. Researchers have recently proposed partial replacement of cement with some waste material and by-products, known as supplementary cementitious materials (SCMs). This profitable waste management strategy results in lowering the production costs and carbon emission and a more sustainable and eco-friendly SCC (Eco-SCC). SCMs may enhance the properties of a cementitious mixture like SCC through their pozzolanic and or hydraulic activity, or they can be used as filler to provide an improved mixture (Pacewska and Wilińska, 2020). SCMs include a wide range of materials and by-products such as ground granulated blast-furnace slag

* Corresponding author.
E-mail addresses: ehsan.sadrossadat@research.uwa.edu.au (E. Sadrossadat), hakan.basarir@ntnu.no (H. Basarir), ali.karrech@uwa.edu.au (A. Karrech), mohamed.elchalakani@uwa.edu.au (M. Elchalakani).

https://doi.org/10.1016/j.clema.2022.100072
Received 15 November 2021; Revised 23 January 2022; Accepted 26 March 2022

(GGBS or GGBFS), fly ash (FA), meta-kaolin (MK), silica fume (SF) lime powder (LP), and waste glass powder (WGP). FA and GGBS are by-products of coal combustion and steel manufacturing respectively while MK is a product of thermal treatment and calcination of kaolin clay in temperatures between 500 °C and 900 °C.

Specifications classify cement differently. In terms of compressive strength, EN 197–1 (EN197) classifies cement into three classes, namely 32.5, 42.5 and 52.5. According to ASTM C150 / C150M-20, cement is classified into ten types including type I, IA, II, IIA, II (MH), II(MH)A, III, IIIA, IV and V with different properties (C150M-20, 2020). The chemical composition of cement in this specification must include: aluminum oxide or alumina (AL_2O_3), iron oxide (Fe_2O_3), magnesium oxide (MgO), sulfur trioxide (SO_3), dicalcium silicate (Ca_2SiO_4 or C_2S), tricalcium aluminate ($Ca_3Al_2O_6$ or C_3A), tricalcium silicate (Ca_3SiO_5 or C_3S), and tetracalcium aluminoferrite (4CaO.$Al_2O_3Fe_2O_3$ or C_4AF). During the hydration process and reactions between ions in cementitious material, various compounds such as calcium silicate hydrates (CSH), calcium aluminate hydrates (CAH) and calcium aluminium silicate hydrates (CASH) may form and strengthen the final product (Harrisson, 2019; L'Hôpital et al., 2015). Calcium silicate hydrate (CaO-SiO_2-H_2O or CSH) which is the product of the reaction of C_3S and C_2S in Portland cement with water is the main compound responsible for gaining strength, hydraulic and self-cementing properties in presence of water (Harrisson, 2019). Calcium hydroxide ($Ca(OH)_2$ or CH) is another product of the hydration process; it may crystallize with CSH in an appropriate pH and produce a form of crystals interlocked with other constituents such as aggregates. Although carbonation may heighten the strength of concrete, the pH decreases and causes corrosion in steel reinforcements such as rebars.

SCMs contain different amounts of calcium oxide (CaO), silicon oxide or silica (SiO_2), both amorphous and crystalline, AL_2O_3, Fe_2O_3, MgO and other oxides in form of different minerals or compounds which are responsible for the cementitious properties. SCMs are usually in form of powder and as fine as cement with a considerable specific surface area. Pozzolanic SCMs such as SF, volcanic FA, low-calcium FA and MK mainly contain silica or aluminosilicate (Al_2SiO_5) components. They do not have self-cementing properties per se but they are able to react with CH in presence of water (Pacewska and Wilińska, 2020). Therefore, they commonly need a cementing agent such as Portland cement to harden. SCMs such as GGBS or high-calcium fly ash can exhibit hydraulic and self-cementing properties that develop the strength of concrete (Pal et al., 2003; Saleh Ahari et al., 2015). The hydration of hydraulic SCMs is different from PC. For example, GGBS may contain minerals which can produce CSH in the presence of water, or consume CH to form additional CSH which is responsible for self-cementing (Lee and Lee, 2020; Pal et al., 2003). Pozzolanic SCMs may also react with CH to reduce carbonation and use it to produce other compounds which can improve the performance of concrete. Some SCMs such as SF may assist in gaining faster strength at early age, while some like GGBS may cause an increase in later ages (Saleh Ahari et al., 2015; Siddique and Bennacer, 2012). Based on the physical and chemical properties and proportion of SCMs in the mixture, strength may increase or decrease (Dadsetan and Bai, 2017; Vivek and Dhinakaran, 2017). They may also enhance the workability, pumpability, durability of concrete such as SCC (Dinakar et al., 2013; Siddique and Bennacer, 2012). Importantly, by partially replacing cement with one or more SCMs, a binder that is more sustainable and eco-friendlier but with comparable or better properties can be produced as opposed to the mixture containing only cement as the binder (Ahmed et al., 2021; Elchalakani et al., 2017; Rahla et al., 2019).

The chemical composition of cementitious materials has the main role in the formation of reaction and properties of concrete, assuming that physical properties such as specific surface area and particle size distribution are appropriate. Reviewing the literature indicates that five oxides of CaO, SiO_2, MgO, Fe_2O_3 and Al_2O_3 are commonly considered by researchers to evaluate the chemical composition of cementitious materials. Researchers and specifications proposed different ratios or conditions based on chemical composition to classify and evaluate cementitious materials used in concrete. Some of the proposed indices and terms for classification of ashes, raw or natural pozzolans and slags as material with pozzolanic and hydraulic properties are summarised in Table 1 (ASTM International, 2019a; EN197; Pal et al., 2003; Xie and Visintin, 2018).

The proportion of fine and coarse aggregates in concrete must be designed in a way to reduce voids and achieve the optimal particle packing. Regardless of the mineralogy or morphology of aggregates, SCC mixes containing aggregates with higher packing density show higher UCS, while there is no significant change when aggregates have the same packing density (Nanthagopalan and Santhanam, 2012). Several other raw or waste materials and by-products such as recycled glass, limestone powder, marble powder, pumice powder, quartz powder and so forth have been proposed to use as SCMs or as inert filler in Eco-SCC to improve properties such as density and strength (Bani Ardalan et al., 2017; Elemam et al., 2020; Long et al., 2015). LP can be used as an inert filler to increase the paste content or viscosity of the mixture or it can be used as a replacement of unreacted cement (Bonavetti et al., 2003; Elemam et al., 2020; Moosberg-Bustnes et al., 2004). In general, inert fillers are used to enhance particle packing and fresh state properties with no significant influence on the UCS and hydration process (Bonavetti et al., 2003; Moosberg-Bustnes et al., 2004). In this paper, LP is considered to be an inert filler used as an admixture in Eco-SCC mixture.

Moreover, superplasticiser (SP) is a necessary chemical admixture in Eco-SCC to get the required properties, particularly in fresh state such as flowability, viscosity and workability. Polycarboxylate SP is one of the widely used SPs consisting of water and polycarboxylate polymers, mainly polyethylene glycols (Ilg and Plank, 2019). They typically benefit the properties of concrete through reducing the water to cement/binder ratio, e.g. by 40%, without negatively affecting the workability but increasing the strength in the same proportion of constituents (Houst et al., 2008). Moreover, there are several other admixtures such as those provided in ASTM C494 / C494M-19 (ASTM International, 2019b) that can be used in concrete. According to (Xie et al., 2021), there are significant variations in the fresh properties of Eco-SCC such as V-funnel, slump flow, U-Box, L-Box and J-ring results (Meko et al., 2021). Such variations can be due to many factors such as mixture proportions, quality of materials but above all is the admixture content and their chemical compositions. Vital information such as chemical composition of the admixtures including SP, wide range of water reducing admixture, and viscosity modifying agents (VMA) are not commonly published in the literature (Xie et al., 2021; Xie and Visintin, 2018). Fresh state properties of Eco-SCC can be improved using some well-known admixtures such as SP with no significant change of the hardened state properties. But, it would be difficult to do the same for properties at hardened state such as compressive strength as it takes time to prepare specimens and test them.

Compressive strength is the major mechanical property of concrete which reflects its quality in different applications, particularly for structural concrete. Compressive strength of Eco-SCC and other types of concrete is typically measured through lab uniaxial compressive strength (UCS) tests on specimens, typically after curing for 28 days (UCS_{28d}). In order to reduce the testing costs and logistics, researchers proposed predictive modeling, data analysis methods that enable prediction of the properties of concrete such as UCS_{28d}. In this regard, artificial intelligence (AI) and machine learning (ML) approaches have been applied for predictive modeling of different properties of SCC such as (Siddique et al., 2008). These algorithms are able to find high-accuracy nonlinear models relating the output to input variables in a set of data. Several ML algorithms such as support vector machines (SVMs), artificial neural networks (ANNs), genetic programming (GP), and so forth have been successfully utilized for solving engineering problems (Abbassi et al., 2013; Ilyas et al., 2021; Javed et al., 2020;

Table 1
Classification of slags and fly ash and natural pozzolans based on their chemical composition for use in concrete.

Material	Ratio/Index	Requirement		
Coal fly ash and raw or calcined natural pozzolan	$SiO_2 + Al_2O_3 + Fe_2O_3$ (%)	Class N ≥ 70	Class F ≥ 70	Class C ≥ 50
	Max. Sulfur trioxide (SO_3) (%)	Class N ≤ 4	Class F ≤ 5	Class C ≤ 5
	Max. $Na_2O + 0.658\ K_2O$	Class N ≤ 1.5	Class F ≤ 1.5	Class C ≤ 1.5
	Max. loss on ignition	Class N ≤ 10	Class F ≤ 6	Class C ≤ 6
Hydraulicity of slags	$(CaO + MgO)/SiO_2$	>1.0		
	CaO/SiO_2	1.3–1.4		
	$(CaO + MgO + Al_2O_3)/SiO_2$	>1.0		
		In Japan, ≥ 1.4		
		In South Korea, ≥ 1.6		
	$(CaO + 0.56Al_2O_3 + 1.4MgO)/SiO_2$	≥1.65		

Onyelowe et al., 2021; Sadrossadat et al., 2019). Although several efforts have been made to propose efficient models, whether and how complicated properties of concrete, such as the compressive strength can be modelled to use for prediction aims are still open research questions (DeRousseau et al., 2018; Young et al., 2019).

Researchers have demonstrated the capability of different approaches for predictive modeling of properties of concrete (Aslam et al., 2020; Farooq et al., 2021a; Farooq et al., 2021b; Farooq et al., 2020; Nafees et al., 2021; Sadrossadat and Basarir, 2019). Proposed models often relate the properties of concrete such as UCS to only the amount of constituents of the mixture. This is merely correct in case the only influential parameter is the amount or proportion of materials. However, problem arises when it comes to use such models and make predictions. They cannot differentiate the type of material and the influence of their properties and output the same results e.g. the same UCS_{28d} for an Eco-SCC containing Cement type I and Type V with the same mixture proportions. There are several other uncertainties and influential factors which must be reduced and considered to develop reliable models. When dealing with different types of materials in collected datasets, particularly to develop predictive models, factors such as the type of constituents pertaining to their chemical or physical properties are influential. However, they have been overlooked in the existing literature of AI-based concrete modelling. This paper aims to better engineer the problem and propose a more efficient model which can consider the proportion and type of constituents of Eco-SCC to predict the UCS_{28d} using different ML and regression methods. For this aim, a database consisting of UCS_{28d} test results conducted by different researchers worldwide on cubic SCC specimens with various mixture proportions of different types of materials is collected. Some indices and dimensionless ratios are reasonably considered and proposed to reduce the dimensionality of variables and reflect the effects of considered factors in models. Different ML algorithms are used and evaluated for predictive modeling of the UCS_{28d} of Eco-SCC. In order to develop reliable ML models, several steps must be implemented and the accuracy model must be validated. The most important step is to verify whether the model can perform accurately when it comes to unseen proportion and type of materials which can be called predictability. After a process of validation, the predictability of the efficient model is confirmed.

2. Method

Predictive modeling of the properties of concrete is challenging due to the large number of influential factors and nonlinear relationships. In this paper, quadratic regression with interaction terms (QRI), and two variants of powerful ML algorithms namely, Gaussian process regression (GPR) and support vector regression (SVR) algorithms are used to develop UCS_{28d} prediction models.

2.1. Quadratic regression with interaction terms

The general form of the quadratic regression with interaction terms (QRI) model used here is as follows:

$$y = \sum a_i x_i^2 + \sum b_i x_i + \sum c_i x_i x_j + \varepsilon \quad (1)$$

where y is called dependent or model output variable, x_i is the independent or design variables, and a_i, b_i and c_i are coefficients obtained using statistical analysis of variances after fitting the model on data. ε is the constant or bias term of model which is the mean for the response when all of the explanatory variables are kept equal to 0. These coefficients or parameters can be found using analysis of variance (ANOVA) using the ordinary least squares (OLS) approach through minimising the sum of the squares of the error between the measured and estimated value of dependent variable for a given dataset.

2.2. Svr

SVR is a data-driven and supervised ML algorithms which aims to find a model, f(x), with at most ε deviation from the target (y) which is the output variable in the database (Cortes and Vapnik, 1995; Li et al., 2020). Note that support vector machines (SVMs) are typically used for classification aims with support vector methods, and SVR is commonly used for regression. Consider that a linear relationship exists between vectors of $X = \{x_1, x_2, x_3, \ldots, x_i, \ldots, x_n\}$ as input and corresponding $Y = \{y_1, y_2, y_3, \ldots, y_i, \ldots, y_n\}$ as output variables. The general form of the SVR model can be represented as follows:

$$Y = f(X) = W^T X + b \quad (2)$$

W is the vector of coefficients and b is the constant. As there might be a nonlinear relationship, the vector X can be transformed using kernel function φ(X).

Therefore, the above equation can be changed to the following expression:

$$y = f(X) = W^T \varphi(X) + b \quad (3)$$

In SVR, the input space is mapped to a higher infinite dimensional feature space through a nonlinear kernel function, namely φ(X), to find a higher accuracy model. There are various types of kernel functions such as polynomial, Gaussian, and radial basis functions (Chou et al., 2011). The general form of polynomial kernel function is as follows:

$$K(W, X) = (1 + W^T X)^d \quad (4)$$

where d is equal to 1, 2 and 3, the polynomial kernel functions is called linear, quadratic and cubic kernel function. In simple form, a least squares algorithm is used to minimize the existing error between f(X) and Y. In order to minimize deviations within the dimensional feature space and accordingly overfitting, Lagrange multipliers are used. Finally, the SVR problem is considered as a constrained mathematical optimization problem using the structural risk minimization principle.

Solving the following risk optimisation problem yield in finding the coefficients w and b (Heidaripanah et al., 2017; Vapnik, 1999):

$$Minimise \frac{1}{2}\|W\|^2 + C_p \sum_{i=1}^{n}(\xi_i + \xi_i^*) \quad (5)$$

subject to:

$$\begin{cases} y_i - Wx_i - b \le \varepsilon + \xi_i \\ Wx_i + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 i = 1, 2, 3, \cdots, n \end{cases} \quad (6)$$

where $\xi_i + \xi_i^*$ are positive slack variables; ε is the dimension of the insensitive zone; and C_P is a penalty parameter and determines the trade-off for ε.

2.3. Gpr

GPR is a nonparametric supervised machine learning algorithm for predictive modeling, regresion and data analysis (Hoang et al., 2016; Zhang et al., 2018). Given a linear function as follows:

$$y = wx + \varepsilon \quad (7)$$

this approach specifies a prior distribution, p(w), on the parameters or coefficient matrix, *w*, and apply probabilities depending on data using Bayes' Rule:

$$P(w|y, X) = \frac{p(y|x, w)p(w)}{p(y|x)} \quad (8)$$

p(w|y,X) is named posterior distribution which gets data from both prior distributions of w and dataset, i.e. p(y|x,w) and p(y|x). Posterior probability is the conditional probability based on the Bayes theorem (Sadrossadat et al., 2021). Therefore, GPR is considered as a Bayesian approach. In order to predict new x, i.e. x*, with f* values, the prediction distribution is updated through weighting all possible x values using their obtained posterior distribution as follows:

$$p(f^*|x^*, y, x) = \int_w p(f^*|x^*, w)p(w|y, x)dw \quad (9)$$

By considering the prior distribution as a Gaussian distribution, a prediction can be made using the mean and variance values. The Gaussian processes term is due to the use of the Gaussian distribution for transforming the multiple-dimensional generalization of multiple-variable normal distributions (Sadrossadat et al., 2021; Zhang et al., 2018). GPR calculates the probability distribution over all functions fitting the data used, i.e. training data, instead of merely the coefficients or parameters. This is why this approach is called non-parametric.

Similar to other regression methods, GPR can be used to find models existing between data. Given a dataset $D = \{(x_i, y_i)\} \ |i = 1,.., n\}$, where $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}^n$ are the input and output vectors, taken from an unknown distribution. GPR considers y as the following function (Hoang et al., 2016):

$$y = f(x) + \varepsilon \quad (10)$$

where $\varepsilon \sim N(0, \sigma_n^2)$. ε follows the Gaussian distribution with an average value of 0 and error variance of σ_n^2 considering the output values or y_i.

If $\{f(x), x \in \mathbb{R}^d\}$ is a Gaussian process, then for a set of i observations $x_1, x_2, \ldots, x_i$, Gaussian transformed function is a combined distribution of variables $f(x_1), f(x_2), \ldots, f(x_i)$. Given two x_i and x_j, a GP can be defined by m(x) and $k(x_i, x_j)$ which are the mean and covariance kernel functions respectively as represented as follows:

$$m(x) = E(f(x)) \quad (11)$$

$$k(x_i, x_j) = Cov(f(x_i), f(x_j)) = E[\{f(x_i) - m(x_i)\}\{f(x_j) - m(x_j)\} \quad (12)$$

where $k(x_i, x_j)$ determines the covariance between x_i and x_j. Therefore:

$$f(x) \sim GP(m(x), k(x_i, x_j)) \quad (13)$$

and accordingly:

$$y \sim GP\ (m(x),\ k(x_i, x_j) + \sigma_n^2\ (x_i - x_j)) \quad (14)$$

m(x) is the mean function is normally a constant and can be zero or the average of the training dataset. The general forms of some common kernel functions used in GPR is represented in Table 2.

In the equations listed in Table 2, the maximum allowable covariance is considered as σ_f^2, σ_l is the length scale of the kernel function, α is a non-negative parameter of covariance, and r is defined as follows:

$$r = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (15)$$

The parameters or hyper-parameters of the GPR model, such as σ_n^2, σ_l, α or σ_f^2 are updated during the modeling process by Bayesian inference until convergence is reached or a stop criterion is met. These parameters are calculated by optimizing the log likelihood function using a Bayesian optimization and gradient-based optimization algorithms (Bishop, 2006; Liu et al., 2017).

3. Dataset and variable selection

A comprehensive dataset is collected through a careful survey of the literature. The dataset contains test results conducted on cubic specimens of Eco-SCC with different mix designs to measure their UCS_{28d} and a wide range of different parameters of constituents. The database includes 456 data taken from 20 published papers in the literature and is attached as a supplementary document (Almuwbber et al., 2018; Bani Ardalan et al., 2017; Bingöl and Tohumcu, 2013; Elemam et al., 2020; Esen and Orhan, 2016; Gill and Siddique, 2017; Güneyisi and Gesoğlu, 2011; Güneyisi et al., 2010; Kannan, 2018; Kannan and Ganesan, 2014a, Kannan and Ganesan, 2014b; Le and Ludwig, 2016; Liu, 2010; Nikbin et al., 2014; Niknezhad et al., 2017; Turk et al., 2012; Uysal, 2012; Uysal and Sumer, 2011; Vivek and Dhinakaran, 2017; Zhao et al., 2015). The database includes reported proportions of Eco-SCC's constituents, chemical compositions of cementitious material such as main oxides (CaO, SiO_2, Al_2O_3, Fe_2O_3, MgO), some physical properties of materials such as maximum size of coarse aggregates (D_{max}), water absorption (WA) of aggregates as reported by researchers. The description of materials and specimens are also given such as different types of cement (C) such as Portland type I, II or IV, different types of ash (A) such as rice husk ash (RHA), fly ash class F (FAF), pulverised fly ash (PFA), GGBS, SF, MK and LP, water (W), fine aggregates (FAgg) and coarse aggregates (CAgg) with different types such as natural or crushed aggregates, various superplasticisers (SP) and so forth. Finally, there are 38 independent variables in the collected database where the goal is to develop a model for predicting the UCS_{28d}.

The more data and variables, the more comprehensive the model. It was strived to collect as large number of data and influential factors as possible with the least uncertainties to develop ML models. The UCS_{28d} of cylindrical specimens with different aspect ratio is different from the cubic ones (Li et al., 2018). Therefore, the UCS of cubic specimens are merely considered in this database rather than cylindrical or other

Table 2
Common kernel functions commonly used in GPR models.

Type	Equation
Squared Exponential (Gaussian)	$k(x_i, x_j) = \sigma_f^2 exp\left(-\frac{r^2}{2\sigma_l^2}\right)$
Rational Quadratic	$k(x_i, x_j) = \sigma_f^2 \left(1 + \frac{r^2}{2\alpha\sigma_l^2}\right)^{-\alpha}$
Matern 5/2	$k(x_i, x_j) = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\sigma_l} + \frac{5r^2}{3\sigma_l^2}\right) \exp\left(\frac{\sqrt{5}r}{\sigma_l}\right)$
Exponential	$k(x_i, x_j) = \sigma_f^2 exp\left(-\frac{r}{\sigma_l}\right)$

shapes. Other factors and experimental procedure issues are compulsorily considered here as systematic experimental error in data. Note that such factors were not reported in some papers and therefore were not possible to be considered.

Fig. 1 represents a ternary plot of chemical composition of available types of cement and SCMs in the collected database. It can be seen that SF and RHA have a high content of SiO_2 and MK and FA are low in CaO + MgO. GGBS and Cement are close but cement has certainly a higher content of CaO + MgO. It can be seen that cement types do not have the same chemical composition. Different chemical composition may produce different compounds. Obviously, each type of material with a different chemical composition have a different influence on properties of mixture such as UCS. Therefore, chemical composition of cementitious materials are important and must be considered as variables to develop predictive models which have been ignored.

In terms of modeling, each oxide such as CaO can be considered as an independent variable which leads to 25 variables for 5 types of cementitious materials, i.e. cement, ash, GGBS, SF and MK. It is possible to consider CaO + MgO, SiO_2 and Al_2O_3 + Fe_2O_3 yielding in 15 variables. In order to reflect the chemical compositions and reduce the dimensionality, this paper proposes the following pozzolanic and hydraulic reactivity indices, PRI and HRI respectively, can be proposed as were also considered in Table 1:

$$PRI = (SiO_2 + Al_2O_3 + Fe_2O_3)/100 \quad (16)$$

$$HRI = (CaO + Al_2O_3 + MgO)/SiO_2 \quad (17)$$

Fig. 2 can be used for classification of cementitious materials in terms of pozzolanic or hydraulic reactivity only. HRI and PRI are considered to reflect the type of cementitious material in the database. The reason why PRI is divided by 100 is to scale it within the range of other variables as it affects the regression model performance. Considering HRI and PRI indices instead of chemical composition such as SiO_2 and CaO results in a meaningful dimensionality reduction of model variables, from 25 to 10 variables.

Typically, 4.75 mm is considered to classify aggregates in terms of size. Below this threshold, particles are considered as fine aggregates (FAgg) and above it they are considered as coarse aggregates (CAgg). The maximum size index (MSI) is the ratio of the maximum size of coarse to fine aggregates:

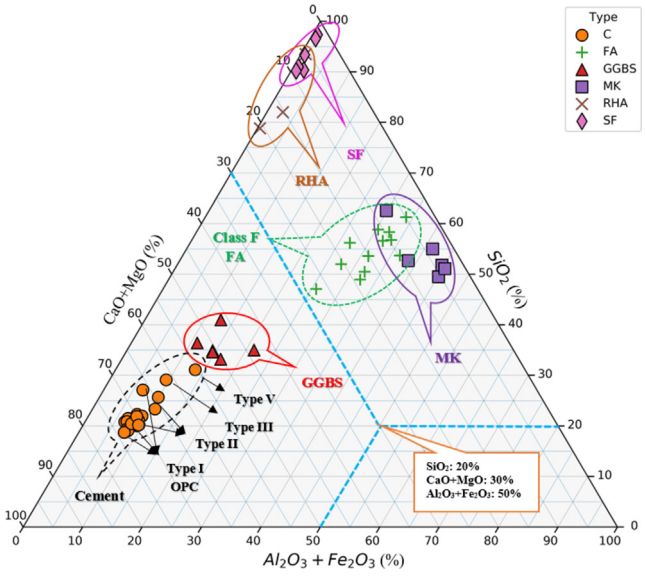

Fig. 1. A ternary plot of chemical composition of cementitious material in database.

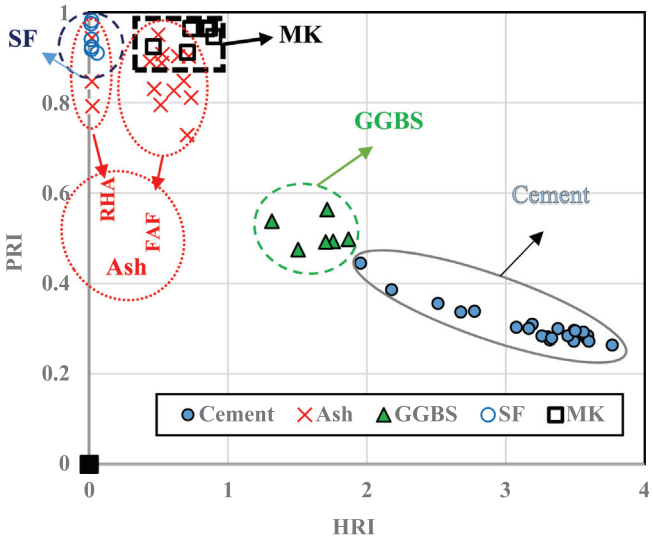

Fig. 2. A scatter plot to classify cementitious material in terms of hydraulic or pozzolanic reactivity.

$$MSI = D_{max}/4.75 \quad (18)$$

Additionally, water absorption of aggregates (WA) is considered as a factor reflecting the nature of aggregates in the model. Fig. 3 illustrates the PSD of some fine and coarse aggregates used by different researchers in the literature for producing SCC (Almuwbber et al., 2018; Anjos et al., 2020; Bani Ardalan et al., 2017; Güneyisi and Gesoğlu, 2011; Güneyisi et al., 2010; Kannan and Ganesan, 2014a, Kannan and Ganesan, 2014b; Nikbin et al., 2014; Zhao et al., 2015). When there is no idea about the optimal PSD of FAgg or CAgg, analysis of data can be helpful such as that provided in Fig. 3. Note that guidelines and specifications are also based on experimental data. In the present paper Fig. 3 and MSI are proposed to consider optimal particle size distribution of aggregates for mixture design. Packing density and specific gravity of constituents of the mixture are other important factors which can be considered in future studies where relevant data is avilable.

Furthermore, the weight (Wt) of water to weight of cement or binder ratio (Wt_W/Wt_C or Wt_W/Wt_B) is a dimensionless and a more interpretable variable compared to weight of water in the mixture, e.g. with the unit of kg or kg/m^3. Here, the weight of binder (B) is considered as the sum of weights of C, A, GGBS, SF and MK of SCC mixes reported. The ratio of the weight of constituent i to the binder (Wt_i/Wt_B) is a dimensionless which is considered instead of considering the amount of material (kg or kg/m^3) in the mix. This way, the control mix is the one where B equals C. Therefore, when cement is replaced with SCMs a binary, ternary or composite binder is produced.

Although the raw form of features can be considered as model input variables, some new indices and ratios are proposed in the present paper. The proposed variables are normalised and dimensionless. Additionally, the number of input variables is reduced from 38 to 23 input variables without ignoring any features. Obviously, the less the number of input variables, the less the chance of complexity and overfitting, particularly when there is a limited number of data. Although LP can be used as replacement of cement, it is considered as an admixture to indicate the feasibility of considering different variables. Due to the lack of reported characteristics, LP, water and SP are considered to be the same in terms of type. Therefore, only their proportions in the mixture were considered as variables. Finally, the UCS_{28d} can be represented as a function of the developed independent input variables as follows:

$$UCS_{28d} = f\left(\frac{Wt_C}{Wt_B}, \frac{Wt_A}{Wt_B}, \frac{Wt_{GGBS}}{Wt_B}, \frac{Wt_{SF}}{Wt_B}, \frac{Wt_{MK}}{Wt_B}, \frac{Wt_{LP}}{Wt_B}, \frac{Wt_{CAgg}}{Wt_B}, \frac{Wt_{FAgg}}{Wt_B}, \frac{Wt_W}{Wt_B}, \frac{Wt_{SP}}{Wt_B},\right.$$

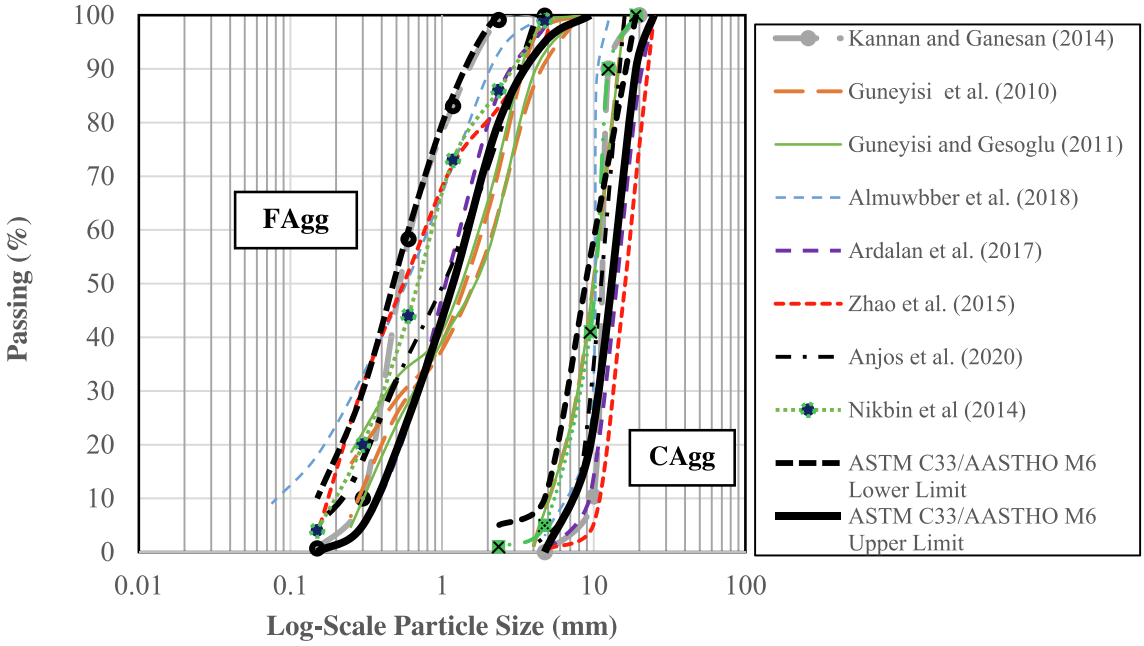

Fig. 3. Some PSD analysis results of aggregates in the database used to produce SCC.

$$MSI, WA_{CAgg}, WA_{FAgg}, HRI_C, PRI_C, HRI_A, PRI_A, HRI_S, PRI_S, HRI_{SF}, PRI_{SF}, HRI_{MK}, PRI_{MK}) \quad (19)$$

Note that Wt_i/Wt_B are proportional variables which reflect the amount and MSI, WA_i, HRI_i and PRI_i are representatives of characteristics and type of i^{th} material in the model.

Values of collected raw data and model variables are attached as supplementary document for more details and considerations. Descriptive statistics of variables considered here are given in Table 3-5.

4. Model development and results

In present study, the experimental data provided by (Kannan and Ganesan, 2014a), including 17 UCS_{28d} test results, is considered as Test data to examine the generalization performance of UCS_{28} models on mixtures with new material type and proportions. The remainder dataset is separated into two groups, Train and Validation data. 70% of the data published in each research paper including 328 data points are used to develop the models, Train data, and 30% including 111 experimental results are considered as Validation data. Models are trained using only Train Data. This scenario and data split form is considered to check the performance of the model on some unseen mixture proportions only while the material type is known in trained model, i.e. Validation Data. Test data is chosen to assess models on a dataset from consisting of new type of constituents. The coefficient of correlation (R), coefficient of determination (R^2), the root mean square error (RMSE) and the mean absolute error (MAE) between the predicted and experimental values of UCS_{28d} as the output are used for the initial accuracy evaluation of the models.

As already noted, the coefficients of variables and bias term in QRI model are obtained using OLS method and analysis of variances similar to linear regression but with quadratic and interaction terms. GPR and SVR are nonlinear models with a number of parameters and hyper-parameters which need to be adjusted based on data. ML techniques assumes that data are normally distributed. However, all dataset may not meet this. To transform the distribution of data to a normal and also scale their range, all data are standardised using z-score method. In order to develop GPR and SVR models, a 5-fold cross validation is implemented to train the models using Train data. Cross-validation aims to validate the performance of models on different subsets of data that was not used in the training process (Zhang et al., 2019; Zhang et al., 2018). In this process, the dataset is firstly divided to k equal-sized folds or subsets of data. Then, the algorithm trains the model using k-1 folds of data, namely train folds, and hold out one fold, i.e. validation fold. In other words, the parameters are tuned using k-1 subsets, i.e. training process, and the first fold is used for validation. Then, the second fold becomes the validation fold and the model is trained on the other k-1 folds. This process is summarised in Fig. 4. The parameters and hyper-parameters of the model are updated after each iteration. The number of iterations can be a user-defined parameter or a criterion such as a fitness function or condition, e.g. $R > 0.9$. The k-fold validation method enables updating the parameters of the model k times on different k folds of data which results in reducing overfitting.

After developing several GPR and SVR models with different kernel functions, those models with higher accuracy based on initial error evaluation using R, RMSE and MAE values are chosen for further evaluation. Amongst GPR models with different kernel functions, the model with rational quadratic kernel function, namely rational quadratic GPR (RQGPR) model, outperformed other models. Similarly, the SVR model with quadratic kernel function (QSVR) outperformed other SVR models with different kernel functions. These models are black-box due to the number of parameters and hyper-parameters, but they can be recalled by a computer for accuracy assessment, validation, prediction and further use.

Fig. 5 illustrates scatter plots of the predicted versus experimental values of UCS_{28d} obtained by models on Train and Validation data. The values of R, RMSE and MAE are also given in Table 6 to have an overall viewpoint on the accuracy and error of prediction. It is recommended by researchers that a regression model which gives $R > 0.8$ is acceptable and accordingly $R^2 > 0.64$ (Sadrossadat et al., 2020; Sadrossadat et al., 2013). According to Table 6 and Fig. 4 (a), all models perform well on Train data. Although QRI perform better than other models on Train data, it gives large error when it comes to Validation data as is represented in Fig. 4 (b) and Table 6. As was mentioned, the coefficients of QRI are obtained using the analysis of variances. Therefore, they are mainly useful on data they are calibrated and no more.

Table 3
Descriptive statistics of variables reflecting the proportions of materials in SCC.

Indicator	Wt_C/Wt_B	Wt_A/Wt_B	Wt_{GGBS}/Wt_B	Wt_{SF}/Wt_B	Wt_{MK}/Wt_B	Wt_{LP}/Wt_B	Wt_{CAgg}/Wt_B	$Wt_{FAgg}/Wt/_B$	Wt_W/Wt_B	Wt_{SP}/Wt_B
Mean	0.74	0.14	0.08	0.02	0.02	0.13	1.56	1.90	0.43	0.02
St. D.	0.18	0.14	0.18	0.04	0.05	0.17	0.47	0.38	0.07	0.01
Sample Var.	0.03	0.02	0.03	0.002	0.003	0.03	0.22	0.14	0.005	0.0001
Min.	0	0	0	0	0	0	0.893	0.84	0.26	0.0024
Max.	1	0.745	1	0.25	0.3	1	3.787	3.105	0.702	0.034

Table 4
Descriptive statistics of variables reflecting type and properties of materials in SCC.

Indicator	MSI	WA_{CAgg}	WA_{FAgg}	HRI_C	PRI_C	HRI_A	PRI_A	HRI_{GGBS}	PRI_{GGBS}	HRI_{SF}	PRI_{SF}	HRI_{MK}	PRI_{MK}
Mean	2.9	0.92	1.11	3.38	0.29	0.29	0.48	0.36	0.10	0.01	0.29	0.15	0.19
StD	0.53	0.44	0.48	0.25	0.02	0.29	0.42	0.69	0.20	0.02	0.43	0.31	0.37
Sample Var.	0.28	0.19	0.23	0.06	3e-4	0.09	0.17	0.48	0.04	5e-4	0.18	0.10	0.14
Min.	2.1	0.22	0.08	1.95	0.26	0	0	0	0	0	0	0	0
Max.	5.3	2.0	2.490	3.76	0.45	0.73	0.95	1.86	0.56	0.058	0.98	0.89	0.96

Table 5
Target or output variable.

Indicator	UCS_{28d} (MPa)
Mean	56.00
StD	16.92
Sample Var.	286.37
Min.	16
Max.	118.404

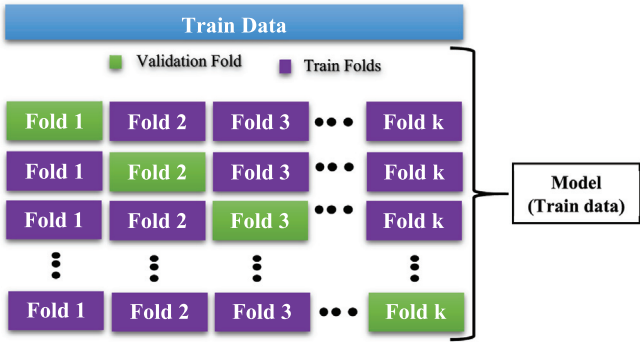

Fig. 4. K-fold cross validation used to develop SVR and GPR models using Train data.

RQGPR and QSVR still represent high accuracy on Validation data according to Fig. 5 (b) and Table 6. Hence, it can be concluded that these models can predict the UCS_{28d} for unseen mixture proportions when the type of constituents is known. This also confirms that these models neither underfit nor overfit for this scenario. The second scenario questions whether it is possible to use QSVR and RQGPR models for new materials with different properties or proportions. This can be called the predictability on new types of material. As already mentioned, some experimental data provided by (Kannan and Ganesan, 2014a) are reserved to address this question where type of FA and accordingly the mixtures are new to the model. The experimental and predicted values of UCS_{28d} and the residual error (RE) values for Test data is summarised in Table 7. The details of Test data is provided in attached supplementary document for more information.

According to Table 7, QSVR and QRI models produce large errors for mixtures containing new type of fly ash and only RQGPR model still performs with high accuracy. MAE values as overall error indicator for QRI, QSVR and RQGPR predictions are respectively 2.07, 12.05, 46.21, while RMSE values are 2.65, 15.24 and 60.01. These results

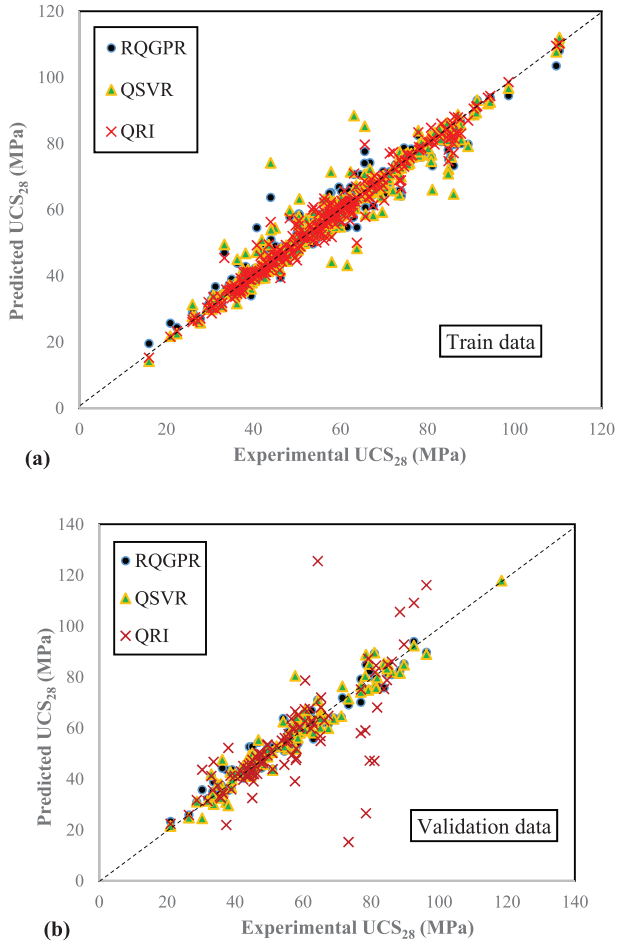

Fig. 5. Scatter plots of the predicted versus experimental values of UCS_{28d} obtained by GPR model on (a) Train and (b) Validation data.

impose the deficiency of QRI model for predictive modeling of the UCS. As already noted, QRI is the quadratic form of multiple variable polynomial regression methods, which also reflect the interaction effects of variables. The structure of the QRI model is known in advance. The coefficient of terms in QRI model are calculated using

Table 6
Performance of models on Train and Test data.

Parameter	Train Data			Validation data		
	QRI	QSVR	RQGPR	QRI	QSVR	RQGPR
R	0.99	0.96	0.98	−0.09	0.97	0.97
R^2	0.97	0.92	0.96	0.01	0.94	0.93
RMSE	2.90	4.70	3.66	175.16	4.41	4.55
MAE	1.70	2.86	2.49	166.50	3.02	3.26

analysis of variances of variables after fitting the model on data using OLS approach. Although QRI was able to successfully model the UCS on Train Data, its prediction performance considerably decreased when it came to Validation and Test datasets. This weakness arises due to the fact that parameters of QRI model are calibrated for only those data used to develop the model and are not validated for unseen data which is called overfitting.

ML models are commonly developed using a process of validation on different sets of data, which results in better prediction performance. GPR and SVR to apply kernel-based ML algorithms for predictive modeling of the UCS of SCC based on the properties and proportions of materials which have been less applied so far. Both QSVR and RQGPR models are non-parametric, i.e. they are not limited by a functional form, and are kernel-based models. Kernel function convert data from an original input space into a higher dimensional feature space in which a dominated hyper-plane can be found to consider a functional mapping between a set of input variables and the output. Therefore, kernel functions are useful to find the best model. ML methods also use regularisation techniques to add a penalty to model parameters, except for constants or intercepts, to make them robust against problems such as multicollinearity and overfitting. These are the reasons why they perform well on validation data. SVR models commonly map data into a higher dimensional space using kernel transformation functions which was a quadratic function in this paper. The quadratic kernel function was chosen in SVR and GPR models to allow for the quadratic response in the output. The reason why RQGPR outperforms the QSVR model can be realised form Fig. 6. The bar chart in Fig. 6 represents the experimental and predicted values of UCS_{28d} in Test data. The polynomial lines indicate the trends of experimental values UCS_{28d} when C is partially replaced with different percentages of FA, MK and FA + MK and those predicted by RQGPR and QSVR models. Both experimental and RQGPR increase and start decreasing after a specific replacement of C with FA while QSVR tends to merely decrease which does not conform to experimental results. It can be seen in Fig. 6 (a) and (c) that RQGPR model trends follow the experimental, while those of QSVR model are not correct. This incorrect sensitivity or parametric response of predictive models results in prediction errors and vice versa.

The better performance of RQGPR compared to QSVR is due to the fact that RQGPR uses probability distributions over all acceptable relationships fitting the data. This means that the Gaussian probability distribution and more accurate covariance kernels between variables helped better prediction for the considered new types of materials in this paper. Here, according to results, only RQGPR performed well on all subsets of data and met the conditions. Considering the results of model performance analyses in the present paper, RQGPR is proposed as a robust tool for predictive modeling of the UCS_{28d} of SCC based on type and proportions of mixture constituents.

5. Sensitivity analysis

In order to realise the contribution and importance of considered variables, a sensitivity analysis (SA) can be done on the RQGPR model as the best model found. Several SA approaches have been proposed for different purposes (Pianosi et al., 2016; Sadrossadat and Basarir, 2019; Sadrossadat et al., 2013) Typically, SA aims to investigate how and which input variables change the output. When it comes to nonlinear models such as RQGPR, a proper SA can be complicated as the output may change differently in local changes of an input variable or due to correlations of variables. According to (Pianosi et al., 2016), SA can be used to rank input factors with respect to their relative contribution to the output variability. The method proposed here is based on measuring the amount of output variation or prediction error when an input variable is removed from the model. RMSE between predicted and experimental values of output is chosen as the indictor of error. For this aim, RMSE is calculated for the output using the training data as they are assumed to be available, named Actual RMSE. Then, the values of one variable are changed to zero in the same dataset, while the amounts of all other variables are unchanged. After calculating the output using the RQGPR model, the new RMSE is calculated again.

Table 7
The performance of different models for predicting the UCS of SCC with unseen type of material.

No	Mix Label	Experimental UCS_{28} (MPa)	Predicted UCS_{28} (MPa)			Error (MPa)		
			RQGPR	QSVR	QRI	RQGPR	QSVR	QRI
1	SCC (100% OPC)	40.77	42.37	47.47	40.77	1.60	6.70	0
2	SCC-FA05	42.73	46.03	30.89	−37.93	3.30	−11.84	−80.66
3	SCC-FA10	44.31	45.12	28.81	−37.93	0.81	−15.51	−82.24
4	SCC-FA15	48.99	43.96	26.41	−37.93	−5.03	–22.58	−86.92
5	SCC-FA20	46.43	42.48	23.71	−37.93	−3.94	–22.72	−84.36
6	SCC-FA25	40.25	40.70	20.69	−37.93	0.46	−19.56	−78.18
7	SCC-FA30	39.43	38.70	17.36	−37.93	−0.73	–22.07	−77.36
8	SCC-MK05	48.28	47.64	49.65	46.44	−0.64	1.37	−1.84
9	SCC-MK10	51.91	51.42	51.48	51.89	−0.49	−0.43	−0.02
10	SCC-MK15	54.53	54.53	52.75	55.15	0.00	−1.78	0.62
11	SCC-MK20	57.17	54.72	53.46	56.21	−2.45	−3.71	−0.96
12	SCC-MK25	53.74	52.59	53.60	55.08	−1.15	−0.14	1.34
13	SCC-MK30	51.40	50.88	53.19	51.75	−0.52	1.79	0.35
14	SCC-MK05 + FA05	46.88	49.85	37.85	−24.61	2.97	−9.02	−71.49
15	SCC-MK10 + FA10	47.48	51.95	33.16	−24.56	4.46	−14.33	−72.04
16	SCC-MK15 + FA15	53.62	51.97	27.15	−24.51	−1.65	−26.47	−78.13
17	SCC-MK20 + FA20	44.68	49.61	19.83	−24.46	4.93	−24.85	−69.14

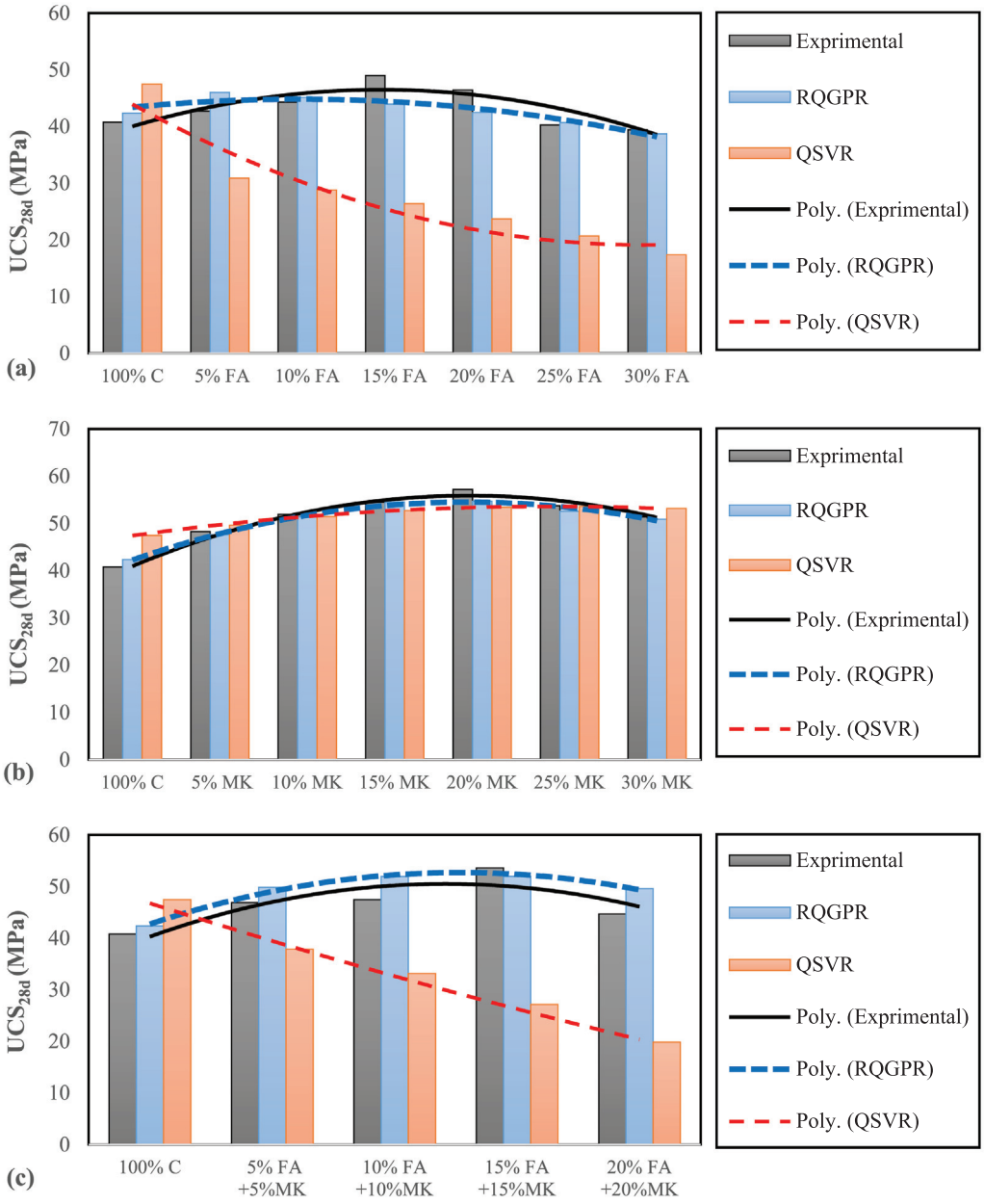

Fig. 6. Bar chart values and trend lines obtained by experimental and predicted UCS_{28d} on Test data where C is replaced with different percentages of (a) FA, (b) MK and (c) FA + MK.

Obviously, the larger variation of RMSE is caused by the most influential factor. This process is done for all variables in RQGPR model and the results are given in Fig. 7.

Fig. 7 shows that removing Wt_W/Wt_B from the RQGPR model changes the prediction error significantly from RMSE of 3.66 to 33.27. This increase in prediction error demonstrates the importance of Wt_W/Wt_B. Interestingly, the respective largest RMSE increase are caused by HRI_C, PRI_C and MSI which proves the prominence of considering properties and type of materials as influential factors. This indicates the contribution of the type of cement and particle size and grade of aggregates to the UCS_{28d} of Eco-SCC, which have been represented by several researchers using experimental studies such as those researched by (Niknezhad et al., 2017). However, such variables have commonly been ignored in previously proposed models. The main objective of this paper was to demonstrate that all influential factors such as type and properties of constituents of composite materials such as Eco-SCC must be considered as variables for predictive modeling of mixture factors such as UCS_{28d}. This issue is validated by the proposed SA and the results shown in Fig. 7. In order to have more comprehensive models, better understanding of effect of variables and decision-making, a more comprehensive dataset containing larger number of experimental data and variables is necessary.

6. Conclusion

This paper represented that each constituent of concrete mixtures has specific type, chemical and physical properties which influence its properties such as UCS_{28d} where previously proposed models merely relate such properties to the amount or proportions of con-

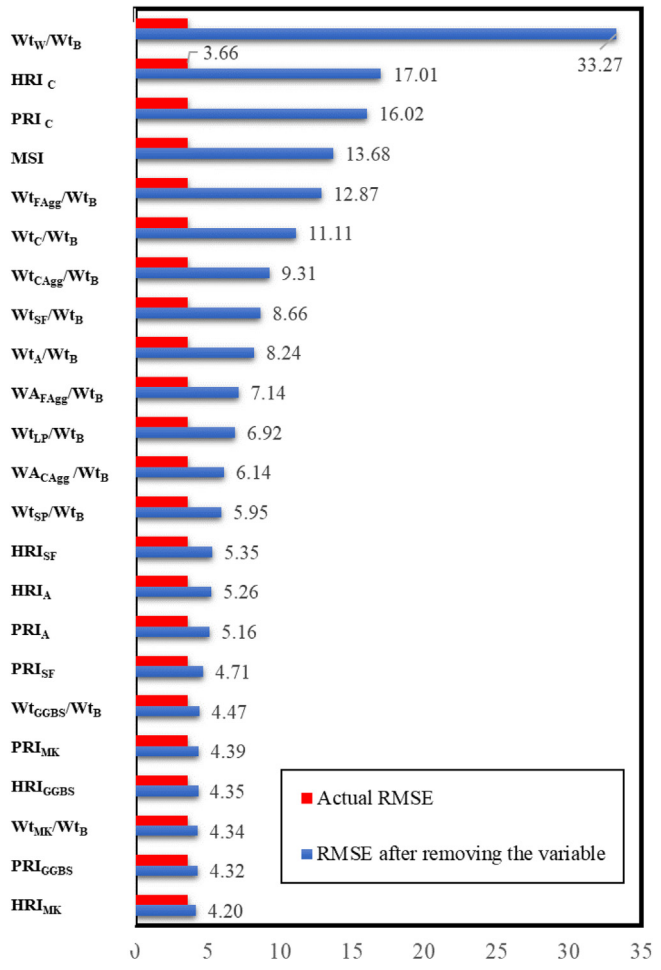

Fig. 7. SA for ranking and relative importance of considered factors for prediction of the UCS_{28d} using RQGPR model.

stituents only. Such models cannot differentiate the type of material and the influence of their properties and output the same results e.g. the same UCS_{28d} for an SCC containing Cement type I and Type V with the same mixture proportions. There are several other uncertainties and influential factors which must be reduced or considered as variables to develop reliable models. For example, this paper only considered UCS_{28d} of cubic specimens while there are experimental data conducted on cylindrical samples which have different UCS. Therefore, such results cannot be merged. This paper also provides new insights into data-pre-processing, feature transformation and selection and dimensionality reduction for development of ML models. HRI and PRI were proposed for classification of cementitious material in terms of hydraulic and pozzolanic reactivity and reducing the dimensionality of variables.

As a project, two separate datasets were considered to test the predictability of the models, validation data had new proportions of materials only and the other set, i.e. test data, contained a new type of material. As already noted, GPR was successful to find a model which can perform well on both considered Test datasets. Showing the capability of ML methods to find an efficient model working well when it comes to new proportion and new type of materials was the main objective of the paper, where only GPR is successful. SVR was used for comparisons and show that a model may perform well on data with new proportions, Validation data here, but it may fail when it comes to new type of material, i.e. Test data. QRI is used to represent the incompetence of classical ANOVA-based approaches for prediction aims. Such models are merely calibrated for the Train data and may not perform well on unseen data. Data pre-processing, regularisation and training process in ML methods enable them to perform accurately in prediction.

This paper aimed to demonstrate that all influential factors such as type and properties of constituents of composite materials such as Eco-SCC must also be considered as variables for an appropriate predictive modeling of mixture factors such as UCS_{28d}. However, such variables have commonly been ignored in previously proposed models. The SA confirmed the importance of considering properties of materials as influential factors as HRI_C, PRI_C and MSI had largest effects on prediction after the amount of water in the mixture compared to other variables. In this paper, RQGPR outperformed other models with regard to the considered variables, dataset and scenarios. Considering the results, RQGPR is proposed as a robust tool for predictive modeling of properties of concrete.

Although several other variables can be considered, this research confirmed that a comprehensive model is achievable that can predict the properties of concrete such as UCS_{28d} of Eco-SCC with high accuracy. Such models can be used worldwide to avoid conducting costly and time-consuming tests, or at least to find the first trial mix design which is difficult due to the large number of materials and factors. An efficient model requires sufficient number of data points, suitable variables and appropriate modeling tool which have been investigated in this paper. Note that the designer must be aware of the properties of materials such as cement and SCMs which are usually reported by their producers and retailers or can be acquired using experiments. Considering the large number of materials and variables, it is expected that a cloud repository is required to collect data and use for development of ML models, similar to image databases used for image processing and computer vision. Models such as RQGPR can also be used for optimisation and multi-objective mixture design purposes, which greatly helps sustainability of the construction and concrete industries.

Data availability statement

The data supporting this work has been reported in the manuscript and a dataset is attached as supplementary document.

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

This research was supported by an Australian Government Research Training Program (RTP) Scholarship and a University Postgraduate Award at University of Western Australia.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clema.2022.100072.

References

Abbassi, F., Belhadj, T., Mistou, S., Zghal, A., 2013. Parameter identification of a mechanical ductile damage using Artificial Neural Networks in sheet metal forming. Mater. Des. 45, 605–615.

Ahmed, T., Elchalakani, M., Basarir, H., Karrech, A., Sadrossadat, E., Yang, B.o., 2021. Development of ECO-UHPC utilizing gold mine tailings as quartz sand alternative. Cleaner. Eng. Technol. 4, 100176.

Almuwbber, O., Haldenwang, R., Mbasha, W., Masalova, I., 2018. The influence of variation in cement characteristics on workability and strength of SCC with fly ash and slag additions. Constr. Build. Mater. 160, 258–267.

Anjos, M.A.S., Camões, A., Campos, P., Azeredo, G.A., Ferreira, R.L.S., 2020. Effect of high volume fly ash and metakaolin with and without hydrated lime on the properties of self-compacting concrete. Journal of Building. Engineering 27, 100985.

Aslam, F., Farooq, F., Amin, M.N., Khan, K., Waheed, A., Akbar, A., Javed, M.F., Alyousef, R., Alabdulijabbar, H., 2020. Applications of gene expression programming for estimating compressive strength of high-strength concrete. Adv. Civil Eng. 2020, 1–23.

ASTM International, W.C., PA, 2019a. ASTM C618-19, Standard Specification for Coal Fly Ash and Raw or Calcined Natural Pozzolan for Use in Concrete,.

ASTM International, W.C., PA, 2019b. Standard Specification for Chemical Admixtures for Concrete, ASTM C494 / C494M-19.

Bani Ardalan, R., Joshaghani, A., Hooton, R.D., 2017. Workability retention and compressive strength of self-compacting concrete incorporating pumice powder and silica fume. Constr. Build. Mater. 134, 116–122.

Bingöl, A.F., Tohumcu, İ., 2013. Effects of different curing regimes on the compressive strength properties of self compacting concrete incorporating fly ash and silica fume. Mater. Des. 51, 12–18.

Bishop, C.M., 2006. Pattern Recognition And Machine Learning. Springer.

Bonavetti, V., Donza, H., Menendez, G., Cabrera, O., Irassar, E., 2003. Limestone filler cement in low w/c concrete: A rational use of energy. Cem. Concr. Res. 33 (6), 865–871.

C150M-20, A.C., 2020. Standard Specification for Portland Cement. ASTM International, West Conshohocken, PA.

Chou, J.-S., Chiu, C.-K., Farfoura, M., Al-Taharwa, I., 2011. Optimizing the prediction accuracy of concrete compressive strength based on a comparison of data-mining techniques. J. Comput. Civil Eng. 25 (3), 242–253.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Support-vector networks. Machine Learn. 20 (3), 273–297.

Dadsetan, S., Bai, J., 2017. Mechanical and microstructural properties of self-compacting concrete blended with metakaolin, ground granulated blast-furnace slag and fly ash. Constr. Build. Mater. 146, 658–667.

DeRousseau, M., Kasprzyk, J., Srubar III, W., 2018. Computational design optimization of concrete mixtures: A review. Cem. Concr. Res. 109, 42–53.

Dinakar, P., Sethy, K.P., Sahoo, U.C., 2013. Design of self-compacting concrete with ground granulated blast furnace slag. Mater. Des. 43, 161–169.

Elchalakani, M., Basarir, H., Karrech, A., 2017. Green concrete with high-volume fly ash and slag with recycled aggregate and recycled water to build future sustainable cities. J. Mater. Civ. Eng. 29 (2), 04016219.

Elemam, W.E., Abdelraheem, A.H., Mahdy, M.G., Tahwia, A.M., 2020. Optimizing fresh properties and compressive strength of self-consolidating concrete. Constr. Build. Mater. 249, 118781.

EN197, B., 1, 2011. Cement Composition, specifications and conformity criteria for common cements. British Standards Institution (BSI).

Esen, Y., Orhan, E., 2016. Investigation of the effect on the physical and mechanical properties of the dosage of additive in self-consolidating concrete. KSCE J. Civ. Eng. 20 (7), 2849–2858.

Farooq, F., Ahmed, W., Akbar, A., Aslam, F., Alyousef, R., 2021a. Predictive modeling for sustainable high-performance concrete from industrial wastes: A comparison and optimization of models using ensemble learners. J. Cleaner Product. 292, 126032.

Farooq, F., Czarnecki, S., Niewiadomski, P., Aslam, F., Alabduljabbar, H., Ostrowski, K. A., Śliwa-Wieczorek, K., Nowobilski, T., Malazdrewicz, S., 2021b. A comparative study for the prediction of the compressive strength of self-compacting concrete modified with fly ash. Materials 14 (17), 4934.

Farooq, F., Nasir Amin, M., Khan, K., Rehan Sadiq, M., Faisal Javed, M., Aslam, F., Alyousef, R., 2020. A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). Applied Sciences 10 (20), 7330.

Gill, A.S., Siddique, R., 2017. Strength and micro-structural properties of self-compacting concrete containing metakaolin and rice husk ash. Constr. Build. Mater. 157, 51–64.

Güneyisi, E., Gesoğlu, M., 2011. Properties of self-compacting portland pozzolana and limestone blended cement concretes containing different replacement levels of slag. Mater. Struct. 44 (8), 1399–1410.

Güneyisi, E., Gesoğlu, M., Özbay, E., 2010. Strength and drying shrinkage properties of self-compacting concretes incorporating multi-system blended mineral admixtures. Constr. Build. Mater. 24 (10), 1878–1887.

Harrisson, A.M., 2019. Constitution and Specification of Portland Cement. Lea's Chem. Cement Concr., 87–155.

Heidaripanah, A., Nazemi, M., Soltani, F., 2017. Prediction of resilient modulus of lime-treated subgrade soil using different kernels of support vector machine. Int. J. Geomech. 17 (2), 06016020.

Hoang, N.-D., Pham, A.-D., Nguyen, Q.-L., Pham, Q.-N., 2016. Estimating compressive strength of high performance concrete with Gaussian process regression model. Adv. Civil Eng. 2016, 1–8.

Houst, Y.F., Bowen, P., Perche, F., Kauppi, A., Borget, P., Galmiche, L., Le Meins, J.-F., Lafuma, F., Flatt, R.J., Schober, I., Banfill, P.F.G., Swift, D.S., Myrvold, B.O., Petersen, B.G., Reknes, K., 2008. Design and function of novel superplasticizers for more durable high performance concrete (superplast project). Cem. Concr. Res. 38 (10), 1197–1209.

Ilg, M., Plank, J., 2019. Synthesis and properties of a polycarboxylate superplasticizer with a jellyfish-like structure comprising hyperbranched polyglycerols. Ind. Eng. Chem. Res. 58 (29), 12913–12926.

Ilyas, I., Zafar, A., Javed, M.F., Farooq, F., Aslam, F., Musarat, M.A., Vatin, N.I., 2021. Forecasting strength of CFRP confined concrete using multi expression programming. Materials 14 (23), 7134.

Javed, M.F., Amin, M.N., Shah, M.I., Khan, K., Iftikhar, B., Farooq, F., Aslam, F., Alyousef, R., Alabduljabbar, H., 2020. Applications of gene expression programming and regression techniques for estimating compressive strength of bagasse ash based concrete. Crystals 10 (9), 737.

Kannan, V., 2018. Strength and durability performance of self compacting concrete containing self-combusted rice husk ash and metakaolin. Constr. Build. Mater. 160, 169–179.

Kannan, V., Ganesan, K., 2014a. Mechanical properties of self-compacting concrete with binary and ternary cementitious blends of metakaolin and fly ash. J. South Afr. Inst. Civil Eng. 56 (2), 97–105.

Kannan, V., Ganesan, K., 2014b. Synergic effect of pozzolanic materials on the structural properties of self-compacting concrete. Arab. J. Sci. Eng. 39 (4), 2601–2609.

L'Hôpital, E., Lothenbach, B., Le Saout, G., Kulik, D., Scrivener, K., 2015. Incorporation of aluminium in calcium-silicate-hydrates. Cem. Concr. Res. 75, 91–103.

Le, H.T., Ludwig, H.-M., 2016. Effect of rice husk ash and other mineral admixtures on properties of self-compacting high performance concrete. Mater. Des. 89, 156–166.

Lee, J., Lee, T., 2020. Durability and engineering performance evaluation of CaO content and ratio of binary blended concrete containing ground granulated blast-furnace slag. Appl. Sci. 10 (7), 2504.

Li, L.-L., Zhao, X., Tseng, M.-L., Tan, R.R., 2020. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J. Cleaner Prod. 242, 118447.

Li, M., Hao, H., Shi, Y., Hao, Y., 2018. Specimen shape and size effects on the concrete compressive strength under static and dynamic tests. Constr. Build. Mater. 161, 84–93.

Liu, M., 2010. Self-compacting concrete with different levels of pulverized fuel ash. Constr. Build. Mater. 24 (7), 1245–1252.

Liu, Y., Song, Y., Keller, J., Bond, P., Jiang, G., 2017. Prediction of concrete corrosion in sewers with hybrid Gaussian processes regression model. RSC Adv. 7 (49), 30894–30903.

Long, G., Gao, Y., Xie, Y., 2015. Designing more sustainable and greener self-compacting concrete. Constr. Build. Mater. 84, 301–306.

Meko, B., Ighalo, J.O., Ofuyatan, O.M., 2021. Enhancement of self-compactability of fresh self-compacting concrete: A review. Clean. Mater. 1, 100019.

Moosberg-Bustnes, H., Lagerblad, B., Forssberg, E., 2004. The function of fillers in concrete. Mater. Struct. 37 (2), 74–81.

Nafees, A., Javed, M.F., Khan, S., Nazir, K., Farooq, F., Aslam, F., Musarat, M.A., Vatin, N.I., 2021. Predictive modeling of mechanical properties of silica fume-based green concrete using artificial intelligence approaches: MLPNN, ANFIS, and GEP. Materials 14 (24), 7531.

Nanthagopalan, P., Santhanam, M., 2012. An empirical approach for the optimisation of aggregate combinations for self-compacting concrete. Mater. Struct. 45 (8), 1167–1179.

Nikbin, I.M., Beygi, M.H.A., Kazemi, M.T., Vaseghi Amiri, J., Rabbanifar, S., Rahmani, E., Rahimi, S., 2014. A comprehensive investigation into the effect of water to cement ratio and powder content on mechanical properties of self-compacting concrete. Constr. Build. Mater. 57, 69–80.

Niknezhad, D., Kamali-Bernard, S., Mesbah, H.-A., 2017. Self-compacting concretes with supplementary cementitious materials: shrinkage and cracking tendency. J. Mater. Civ. Eng. 29 (7).

Okamura, H., Ozawa, K., Ouchi, M., 2000. Self-compacting concrete. Structural concrete 1(1), 3-17.

Onyelowe, K.C., Ebid, A.M., Nwobia, L.I., 2021. Predictive models of volumetric stability (durability) and erodibility of lateritic soil treated with different nanotextured bio-ashes with application of loss of strength on immersion; GP, ANN and EPR performance study. Cleaner Materials 1, 100006.

Pacewska, B., Wilińska, I., 2020. Usage of supplementary cementitious materials: advantages and limitations. J. Therm. Anal. Calorim. 142 (1), 371–393.

Pal, S., Mukherjee, A., Pathak, S., 2003. Investigation of hydraulic activity of ground granulated blast furnace slag in concrete. Cem. Concr. Res. 33 (9), 1481–1486.

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environ. Modell. Software 79, 214–232.

Rahla, K.M., Mateus, R., Bragança, L., 2019. Comparative sustainability assessment of binary blended concretes using Supplementary Cementitious Materials (SCMs) and Ordinary Portland Cement (OPC). J. Cleaner Prod. 220, 445–459.

Sadrossadat, E., Basarir, H., 2019. An evolutionary-based prediction model of the 28-day compressive strength of high-performance concrete containing cementitious materials. Adv. Civil Eng. Mater. 8 (3), 484–497.

Sadrossadat, E., Basarir, H., Karrech, A., Durham, R., Fourie, A., Bin, H., 2019. The optimization of cemented hydraulic backfill mixture design parameters for different strength conditions using artificial intelligence algorithms, International Symposium on Mine Planning & Equipment Selection. Springer, 219–227.

Sadrossadat, E., Basarir, H., Karrech, A., Elchalakani, M., 2021. Multi-objective mixture design and optimisation of steel fiber reinforced UHPC using machine learning algorithms and metaheuristics. Eng. Comput., 1–14.

Sadrossadat, E., Basarir, H., Luo, G., Karrech, A., Durham, R., Fourie, A., Elchalakani, M., 2020. Multi-objective mixture design of cemented paste backfill using particle swarm optimisation algorithm. Miner. Eng. 153, 106385.

Sadrossadat, E., Soltani, F., Mousavi, S.M., Marandi, S.M., Alavi, A.H., 2013. A new design equation for prediction of ultimate bearing capacity of shallow foundation on granular soils. J. Civil Eng. Manage. 19 (sup1), S78–S90.
Saleh Ahari, R., Erdem, T.K., Ramyar, K., 2015. Permeability properties of self-consolidating concrete containing various supplementary cementitious materials. Constr. Build. Mater. 79, 326–336.
Siddique, R., Aggarwal, P., Aggarwal, Y., Gupta, S., 2008. Modeling properties of self-compacting concrete: support vector machines approach. Comput. Concr. 5 (5), 123–129.
Siddique, R., Bennacer, R., 2012. Use of iron and steel industry by-product (GGBS) in cement paste and mortar. Resour. Conserv. Recycl. 69, 29–34.
Turk, K., Karatas, M., Gonen, T., 2012. Effect of Fly Ash and Silica Fume on compressive strength, sorptivity and carbonation of SCC. KSCE J. Civ. Eng. 17 (1), 202–209.
Uysal, M., 2012. The influence of coarse aggregate type on mechanical properties of fly ash additive self-compacting concrete. Constr. Build. Mater. 37, 533–540.
Uysal, M., Sumer, M., 2011. Performance of self-compacting concrete containing different mineral admixtures. Constr. Build. Mater. 25 (11), 4112–4120.
Vapnik, V.N., 1999. An overview of statistical learning theory. IEEE Trans. Neural Networks 10 (5), 988–999.
Vivek, S.S., Dhinakaran, G., 2017. Fresh and hardened properties of binary blend high strength self compacting concrete. Eng. Sci. Technol. Internat. J. 20 (3), 1173–1179.
Xie, T., Ali, M.S.M., Elchalakani, M., Visintin, P., 2021. Modelling fresh and hardened properties of self-compacting concrete containing supplementary cementitious materials using reactive moduli. Constr. Build. Mater. 272, 121954.
Xie, T., Visintin, P., 2018. A unified approach for mix design of concrete containing supplementary cementitious materials based on reactivity moduli. J. Cleaner Prod. 203, 68–82.
Young, B.A., Hall, A., Pilon, L., Gupta, P., 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. Cem. Concr. Res. 115, 379–388.
Zhang, J., Ma, G., Huang, Y., Aslani, F., Nener, B., 2019. Modelling uniaxial compressive strength of lightweight self-compacting concrete using random forest regression. Constr. Build. Mater. 210, 713–719.
Zhang, N., Xiong, J., Zhong, J., Leatham, K., 2018. Gaussian process regression method for classification for high-dimensional data with limited samples. In: 2018 Eighth International Conference on Information Science and Technology (ICIST). IEEE, pp. 358–363.
Zhao, H., Sun, W., Wu, X., Gao, B., 2015. The properties of the self-compacting concrete with fly ash and ground granulated blast furnace slag mineral admixtures. J. Cleaner Prod. 95, 66–74.