Construction and Building Materials 336 (2022) 127557

Contents lists available at ScienceDirect

Construction and Building Materials

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

Predicting compressive strength of alkali-activated systems based on the network topology and phase assemblages using tree-structure computing algorithms

Rohan Bhat [a], Taihao Han [a], Sai Akshay Ponduru [a], Arianit Reka [a,b], Jie Huang [c], Gaurav Sant [d], Aditya Kumar [a,*]

[a] Department of Materials Science and Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[b] Faculty of Natural Sciences and Mathematics, University of Tetova, Tetovo, Republic of North Macedonia
[c] Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[d] Civil and Environmental Engineering, University of California, Los Angeles, Los Angeles, CA 90095, USA

ARTICLE INFO

Keywords:
Alkali-activated system
Compressive strength
Topological constraint theory
Thermodynamic simulation
Machine learning

ABSTRACT

Alkali-activated system is an environment-friendly, sustainable construction material utilized to replace ordinary Portland cement (OPC) that contributes to 9% of the global carbon footprint. Moreover, the alkali-activated system has exhibited superior strength at early ages and better corrosion resistance compared to OPC. The current state of analytical and machine learning models cannot produce highly reliable predictions of the compressive strength of alkali-activated systems made from different types of aluminosilicate-rich precursors owing to substantive variation in the chemical compositions and reactivity of these precursors. In this study, a random forest model with two constraints (i.e., topological network and thermodynamic constraints) is employed to predict the compressive strength of alkali-activated systems made from 26 aluminosilicate-rich precursors and distinct processing parameters. Results show that once the model is rigorously trained and optimized, the RF model can yield *a priori*, high-fidelity predictions of the compressive strength in relation to the physicochemical properties of aluminosilicate-rich precursors; processing parameters; and constraints. The topological network constraint provides the chemostructural properties and reactivity of the aluminosilicate-rich precursors. Whereas the thermodynamic constraint estimates the phase assemblages at different degrees of reaction of the aluminosilicate-rich precursors. Finally, the correlations between topological network constraint; phase assemblage; and compressive strength are demonstrated. When the topological network constraint equals 3.4, the alkali-activated systems can achieve their optimal compressive strength.

1. Introduction

Ordinary Portland cement (OPC) is a fundamental material for infrastructure development. The global production of OPC is 4 billion tons per year with a growth rate of 80 million tons per year [1]. This gigantic production presents a tremendous energy consumption (11 exajoule per year [2]) and 9% of the global CO_2 emission [3]. Calcination of limestone above 1450 °C contributes at least 50% of CO_2 release in OPC manufacture, where sustainability cannot be improved through clean energy [4]. Another shortcoming faced by OPC is susceptibility to degradation, which results in additional OPC consumption in the form of repair and reconstruction. With the increasing demand of infrastructures, an alternative cementitious material – more sustainable and durable than OPC – is urgently needed. Alkali-activated system (including binder; mortar; and concrete) – also named as geopolymer system – is a rising candidate to replace OPC [5,6]. The alkali-activated system can be made from aluminosilicate-rich precursors such as fly ash; ground granulated blast-furnace slag; rice husk ash; and other urban waste ashes [6,7]. Although they are made from recycled materials, the alkali-activated systems exhibit stronger strength at early ages and better corrosion resistance than OPC [7–9]. The replacement of OPC by the alkali-activated system would reduce 73% of CO_2 emission and 43%

* Corresponding author at: Department of Materials Science and Engineering, Missouri University of Science and Technology, B49 McNutt Hall, 1400 N Bishop, Rolla, MO 65409, United States.
E-mail address: kumarad@mst.edu (A. Kumar).

https://doi.org/10.1016/j.conbuildmat.2022.127557
Received 21 February 2022; Received in revised form 31 March 2022; Accepted 15 April 2022

of energy consumption in the cement and construction industry [10].

Alkali-activated systems are made from amorphous aluminosilicate-rich precursors and alkali activators (e.g., sodium silicate; sodium hydroxide; potassium silicate; etc.) through a geopolymerization reaction [5,8]. The alkali activator significantly increases the concentration of cations and OH^- in the solution, resulting in an acceleration of leaching of aluminate and silicate ions from surfaces of aluminosilicate-rich precursors. After this, sodium aluminosilicate hydrate (N-A-S-H) and calcium sodium aluminosilicate hydrate (C-N-A-S-H) gels are precipitated as a result of the geopolymerization reaction. (*N.B.:* $N = Na_2O$; $A = Al_2O_3$; $C = CaO$; $S = SiO_2$; $H = H_2O$). The dominant product of Ca-rich precursor is C-N-A-S-H; otherwise, the dominant product is N-A-S-H. Conventional (two-part) alkali-activated systems are produced by mixing aluminosilicate-rich precursors; alkali activator solutions; and additional water [11]. Due to the danger of handling large amounts of the viscous and corrosive alkali activator solutions, one-part alkali-activated systems have been widely researched upon. One-part alkali-activated systems are synthesized by the addition of water to dry premixture, prepared from aluminosilicate-rich precursors and solid alkali-activators with or without calcination [8].

Reactivity of aluminosilicate-rich precursors is a vital parameter that affects fresh and hardened properties of alkali-activated systems [12,13]. The highly reactive aluminosilicate-rich precursors set in a short time and yield high compressive strength. The reactivity of the aluminosilicate-rich precursors is associated with chemostructural properties of precursors, especially the amorphous phases accounting for almost 50–90% in aluminosilicate-rich precursors [14,15]. Previous studies [15,16] have shown that the topological constraint theory (TCT) is the key enabler to evaluate the reactivity of aluminosilicate-rich precursors. The TCT is developed based on the topology of the atomic structure of amorphous materials and their macroscopic properties [17]. To be specific, TCT reduces the complexity of chemostructural properties of amorphous materials into a singular constraint [*number of constraints* (n_c)] [18]. This parameter quantitatively evaluates the reactivity of aluminosilicate-rich precursors by solely relying on the major components (i.e., CaO; SiO_2; and Al_2O_3).

Compressive strength is one of the most important mechanical properties of alkali-activated systems, giving a general idea about the performance regarding the quality of alkali-activated systems [19,20]. For alkali-activated systems to achieve the targeted compressive strength as well as other mechanical properties, researchers must invest a significant amount of time in exploring mixture design and processing parameters [21]. The laboratory exploring work, however, is costly and labor-intensive. Thus, being able to predict the compressive strength is essential especially in saving time and cost. Consequently, many researchers have worked on various analytical models to predict the compressive strength of alkali-activated systems. Le et al. [22] have developed a modified Feret's model based on the mixture design of alkali-activated systems and standard compressive strength at 28-day. This model, however, cannot properly predict the compressive strength of the systems without aggregate [23]. An analytical model for compressive strength of alkali-activated systems made from a fly ash has been developed by Beluah et al. [24]. However, the shortcoming is that a calibration is required when the model is applied to a new database due to a limited number of data-records used for the model development. Jonbi and Fulazzaky [25] have related compressive strength with the age of alkali-activated systems. This study cannot produce reliable predictions of compressive because of exclusion of some influential variables. This study also reinforces that analytical models cannot account for all influential variables and variables without understanding their contributions on properties.

Machine learning (ML), a data-driven artificial intelligence, can overcome the abovementioned shortcomings of current analytical models. In the past decades, many studies [19,26–30] have applied machine learning models to predict different properties of alkali-activated systems. Zhang et al. [29] have shown regression- and tree-based ML models to produce reliable predictions of compressive strength of alkali-activated systems in relation to mixture design and chemistry of precursors. Gomma et al. [28] have demonstrated the reliable performance of random forest (RF) model on predicting fresh and hardened properties of alkali-activated systems related to mixture design and various processing conditions. Lahoti et al. [30] have used RF, Naïve Bayes, and k-nearest neighbor models to predict mechanical properties of alkali-activated systems with a small database. Although previous studies have shown that machine learning is a promising tool to predict properties of alkali-activated systems, some knowledge gaps still exist. First, previous studies only include one or two types of aluminosilicate-rich precursors in their databases. The performance of ML models on multiple (more than two) aluminosilicate-rich precursors has not been valid. Next, as aforesaid, the reactivity of the aluminosilicate-rich precursors is a critical parameter that influences the properties of the alkali-activated systems. The parameter related to the reactivity has not been included in previous studies. Finally, ML models find the underlying structure between the mixture design and properties, all without understanding mechanisms. This leads to a critical problem that predictions as produced by ML may violate fundamental material laws. Such violation significantly undermines the credibility of results from ML. However, this problem can be solved by applying thermodynamic constraints to ML models. Han et al. [31] have shown that the thermodynamic constraint regulates the RF model to predict properties of OPC systems, where the accuracy is better than models without constraints.

Gibbs Energy Minimization Software (GEMS) is a tool for geochemical modeling [32,33], which can be employed to conduct thermodynamic simulations of alkali-activated systems. More specifically, GEMS simulations and the degree of reaction of the aluminosilicate-rich precursors can produce equilibrium phase assemblages and microstructure for alkali-activated systems. Several studies [34,35] have shown that the phase assemblages from GEMS simulations accurately replicate phases appearing in real alkali-activated systems. In the thermodynamic simulations, the expected products for alkali-activated systems are: C-N-A-S-H; N-A-S-H; stratlingite; gibbsite; monocarboaluminate hydrate; hydrogarnet; and amorphous zeolites. With the understanding of phase assemblages at different degrees of reaction of aluminosilicate-rich precursors, microstructures and properties of the alkali-activated systems can be correlated. This would help researchers understand the mechanisms behind properties of alkali-activated systems.

The study presented here aims to overcome the abovementioned limitations of current analytical and ML models by advancing network and thermodynamic constraints that enhance the prediction performance of ML models on the compressive strength of alkali-activated systems. Herein, the RF model is employed to find correlations between compressive strength and inputs (i.e., mixture design and processing parameters of alkali-activated systems). Such correlations are utilized to produce *a priori*, reliable predictions of compressive strength in new alkali-activated systems. The network constraint (developed from chemostructural properties of aluminosilicate-rich precursors based on the topological constraint theory) and thermodynamic constraint (obtained from thermodynamic simulations) are employed to regulate the RF model, which ensures that predictions do not violate fundamental material laws. Furthermore, the RF model links the compressive strength with phase assemblages of alkali-activated systems. This leads to an investigation of compressive strength behavior in relation to microstructures reacted from aluminosilicate-rich precursors with different chemostructural properties.

2. Modeling method

In this study, a RF model is employed to predict the compressive strength of alkali-activated systems. First, the model is trained and tested without any constraints. Second, a network constraint (i.e.,

number of constraints; described in **Section 2.1**) is applied to the model to produce more accurate predictions. Third, an additional thermodynamic constraint (*volume fraction of products*; described in **Section 2.2**) regulates the network-constrained model to ensure non-violation of fundamental material laws. Prediction performances of three stages are rigorously appraised by comparing their prediction against measured values. Fig. 1 demonstrates the framework of the RF model with constraints. As can be seen, the model used in this study unites the network and thermodynamic constraints with the RF model to achieve optimal predictions of compressive strength of alkali-activated systems. The purpose of using three parallel predictions is to investigate the optimal prediction strategy (one that treads the balance between ease-of-use/simplicity and accuracy) for alkali-activated systems. The optimal strategy can be directly applied to the database without other two stages. A detailed description of the RF model can be found in **Section 2.3**.

2.1. Network constraint

Herein, a brief description of TCT is given below, and details can be found elsewhere [16]. The major components (i.e., SiO_2, CaO, and Al_2O_3) of the aluminosilicate-rich precursors are utilized to derive the *number of constraints* based on the topological constraint theory (TCT). The compositions of the aluminosilicate-rich precursors are considered as $(CaO)_x(Al_2O_3)_y(SiO_2)_{1-x-y}$, where x and y are normalized molar fractions. Aluminosilicate-rich precursors contain radial bonding-stretching (BS) and angular bond-bending (BB) constraints [16,18,36]. BS exists between two bonded atoms. BB exists when atoms have fixed interatomic angles. There are six types of atoms that have constraints: Si atom; Al atom (four-fold and five-fold); tricluster oxygen atom (TO); free oxygen atoms (FO); bridging atom (BO); and non-bridging atom (NBO).

In the aluminosilicate-rich precursors, Si and Al atoms are the network formers, forming tetrahedral structures. O atoms acting as a bridge to connect Si/Al tetrahedrons are BOs [37,38]. The Ca atom is the network modifier existing between Si/Al tetrahedral molecular as the interstitial site [37,38]. The O atom connected with Ca atom and Si/Al tetrahedral molecular is NBO. FO is an O atom that only connects with a Ca atom. TO is the O atom acting as the charge compensating atom for the Al tetrahedron. Depending on the CaO and Al_2O_3 contents, the network structure of aluminosilicate materials can be divided into three regimes: fully depolymerized ($y-x \le -\frac{2}{3}$); partially depolymerized ($-\frac{2}{3} \le y-x \le 0$); and fully polymerized ($0 \le y-x$) [16]. In the fully depolymerized regime, Ca atom is the dominant composition. All Si and Al tetrahedrons isolate from each other. The excess Ca atoms result in the formation of NBOs and FOs. In the partially depolymerized regime, Si atom is the dominant composition. The Ca atoms connect with Si/Al tetrahedrons and compensate charge of Al tetrahedrons. This regime includes both BOs and NBOs. In fully polymerized regime, Al atom is the dominant composition. Due to insufficient Ca atoms, the charge of Al atoms cannot be compensated to form four-fold tetrahedral units, where over-coordinated Al atoms (five-fold) are formed. This regime includes both BOs and TOs.

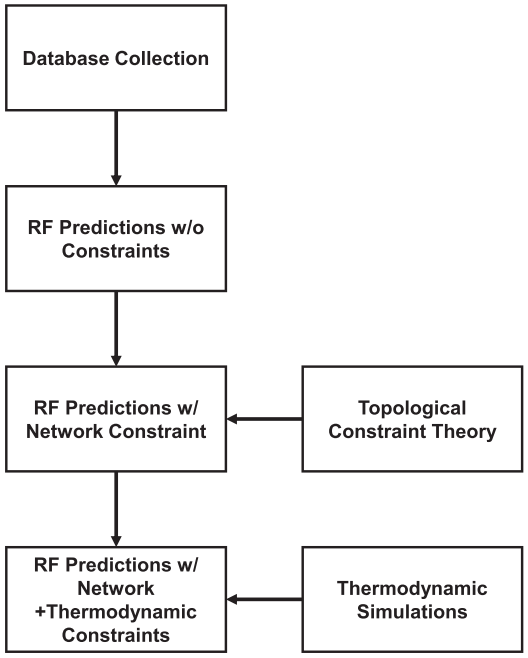

Fig. 1. Schematic of the RF model to predict the compressive strength of alkali-activated systems by incorporating network and thermodynamic constraints. Network constraint is developed from the topological constraint theory. Thermodynamic constraint is acquired from phase assemblages of alkali-activated systems.

Ca atom is the charge-compensating ion; thus, it is excluded from topological constraints. But it creates 1 BS constrain with a connected NBO [39]. Ca atom connected to a FO can form 1 BS constraint. Previous studies [16,36,39–41] have shown that Si atoms form tetrahedral structure at five fixed angles in the amorphous phases, which contains 4 BS and 5 BB constraints. Additionally, Si tetrahedron creates 1 BB constraint with BO. Four-fold Al atom has 4 BS constraints and 5 BB constraints. Five-fold atom only has 5 BS constraints. TO has 3 BB constraints. The *number of constraints* (n_c) corresponding to each regime is shown in **Eq. 1**. Additional emphasis is given to that the aluminosilicate-rich precursors in this study fall into partially depolymerized and fully polymerized regimes.

$$n_c = \frac{11+y-10x}{3-2x+2y} \text{(Fully depolymerized)} \tag{1a}$$

$$n_c = \frac{11+10y-10x}{3-2x+2y} \text{(Partially depolymerized)} \tag{1b}$$

$$n_c = \frac{11+13y-13x}{3-2x+2y} \text{(Fully polymerized)} \tag{1c}$$

2.2. Thermodynamic modeling

GEMS [32,33] is employed to produce thermodynamic simulations and phase assemblages of alkali-activated systems in relation to their mixture design [i.e., chemical composition of binders (aluminosilicate-rich precursors; alkali-activators; and water)] and processing conditions (i.e., temperature). The quality of the thermodynamic simulation results is dependent upon the accuracy and completeness of the input properties of the substances and phases, which can typically be found within literature and thermodynamic databases. The thermodynamic data for aqueous species and numerous solids are acquired from the PSI-GEMS thermodynamic database, while solubility products for relevant phases are extracted from the Cemdata 18 [42,43] and zeolite 20 [44]. An extended Debye-Huckel calculation is utilized by the software to estimate the activity coefficients of the aqueous species. The assumption that the aqueous phase is dominated by NaOH is made. The average ion size and parameter for common short-range interactions of charged are 3.31 A and 0.098 kg/mol, respectively. All phase assemblages are simulated at 1 atm and 20-to-85 °C.

Phase assemblages obtained from the thermodynamic simulations are shown in Fig. 2. The figure reveals volume of all reactants and products at increasing degrees of reaction of the aluminosilicate-rich precursors. In Fig. 2, it has been observed that GEMS can estimate phase assemblages of alkali-activated systems made from multiple aluminosilicate-rich precursors at different water-to-solid ratio. The phases shown herein are in agreement with previous studies [20,34,45]. The main phases of alkali-activated systems include: C-N-A-S-H; N-A-S-

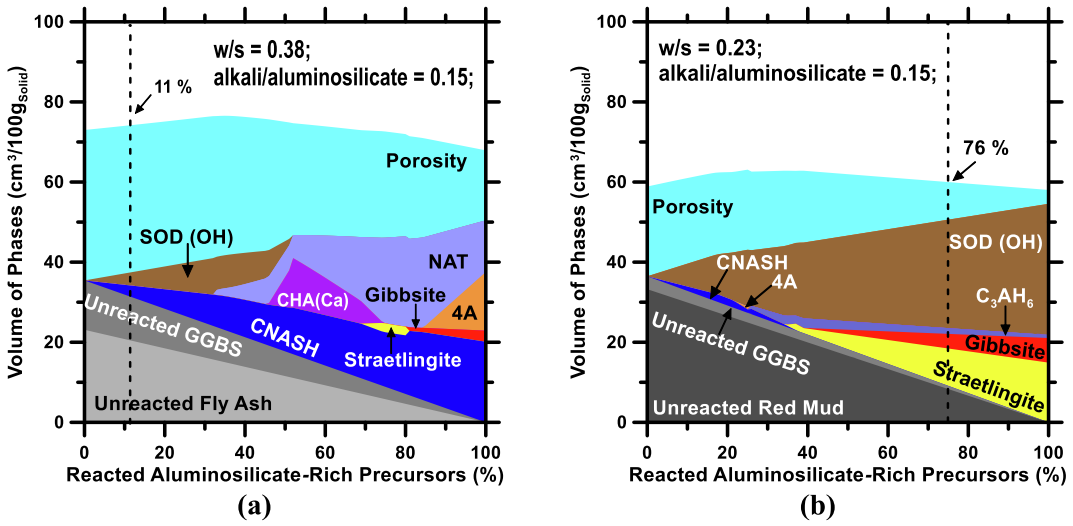

Fig. 2. Equilibrium phase assemblage, simulated using GEMS, of alkali-activated systems made from **(a)** fly ash and ground granulated blast-furnace slag (GGBS); and **(b)** red mud and ground granulated blast-furnace slag (GGBS). The vertical dashed line presents the phase assemblage at the targeted degree of reaction of aluminosilicate-rich precursors estimated from the compressive strength. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

H (Na-chabazite; zeolite 4A; hydroxysodalite; and natrolite); amorphous Si; Ca-chabazite; Gibbsite; straetlingite; and C_3AH_6. The phase assemblage is an important outcome because researchers can use it to estimate the volume fraction of products and porosity in any given aluminosilicate-rich precursor at the targeted degree of reaction. The compressive strength can be roughly estimated from porosity and volume fraction of products because they provide a basic solid-to-solid connectivity within the alkali-activated systems [46–48]. In order to obtain accurate phase assemblages, it is essential to identify the degree of reaction of the aluminosilicate-rich precursors, which can be estimated directly from compressive strength at any given age. The method to determine the degree of reaction is explained in **Section 3.0**. The volume fraction of products is employed as the thermodynamic constraint to regulate the RF model.

2.3. Random forest (RF) model

RF model is a tree-structure model developed from the classification-and-regression trees (CARTs) model with bagging technique [49,50]. During the training, the model parallelly constructs hundreds of independent CARTs from bootstraps. Each tree grows through binary splits in a recursive fashion until the terminal nodes reach a "near-homogenous" state. The model allows each CART to grow to its maximum size without pruning and smoothing. When new data is applied to the RF model, the model averages the predictions from all trees to produce the final output. Compared to other ML models, the RF model is different in the sense that two-stage randomization [50,51] is employed during the growth of CARTs. The first randomization is that the bootstrap used to construct a tree randomly selects data from the original dataset. The second stage involves that a randomly chosen subset of variables, instead of using all variables, is utilized to ascertain the best scenario for each split. Due to the two-stage randomization, trees in the forest decorrelate with each other, resulting in a reduction of variance errors [52,53]. To achieve the RF model's best performance, 10-fold cross-validation (CV) method [54,55] and the grid-search method [28,56] are employed to determine the optimal hyperparameters for the RF model. In this study, 300 trees and 4 splits at each node are for all predictions.

To train and validate the RF model, the database (shown in **Section 3.0**) is split into two non-overlapping datasets: training dataset and testing dataset. The training dataset contain 75% of randomly selected data-records form the parent database, and the remaining 25% data-records form the testing dataset. The training dataset is used to train the RF model rigorously. The testing dataset is used to quantitatively evaluate the prediction performance of the RF model through 5 statistical parameters – Pearson correlation coefficient (R); coefficient of determination (R^2); mean absolute percentage error (*MAPE*); root mean squared error (*RMSE*); and mean absolute error (*MAE*). Mathematical functions for these statistical parameters were detailed in our previous studies [54,57,58].

2.4. Database Collection

Data pertaining to compressive strength in 520 alkali-activated systems [including pure binder (aluminosilicate-rich precursors + alkali activator + water); mortar (binder + fine aggregate); and concrete (binder + fine aggregate + coarse aggregate)] were collected from 15 studies [59–73]. Majority of the alkali-activated systems were synthesized by fly ash and ground granulated blast furnace slag, and the remaining systems were synthesized from fly ash sinking spherical beads, red mud, metakaolin, silica fume, and rice husk ash. An alkali-activated system can contain at the most three different aluminosilicate-rich precursors. Not all the densities of the aluminosilicate-rich precursors were illustrated in previous studies. In our database, densities of the aluminosilicate-rich precursors are calculated based on the density of CaO (3.34 g/cm^3); Na_2O (2.27 g/cm^3); SiO_2 (amorphous, 2.20 g/cm^3) [74]; and Al_2O_3 (amorphous, 2.32 g/cm^3) [75,76] with respect to their normalized mass fractions. Other compositions (e.g., MgO; K_2O; etc.) are minor, and their products have negligible contribution to the compressive strength [7]. Therefore, they are excluded from the density calculations and thermodynamic simulations. It is worth noting that this database includes not only two-part (aluminosilicate-rich precursors mixed with an alkali-activator solution) but also one-part (aluminosilicate-rich precursors and solid alkali-activator mixed with water) alkali-activated systems.

Here, compressive strength is used as a direct indicator to calculate the degree of reaction of the aluminosilicate-rich precursors. Previous studies [77,78] have demonstrated that compressive strength of cementitious materials and the degree of reaction of binder exhibits a linear correlation. The degree of reaction of the aluminosilicate-rich precursors is employed to locate the phase assemblage of the binder in thermodynamic simulations. The volume fraction of products obtained

from the phase assemblages act as a thermodynamic constraint to regulate and improve predictions as produced by the RF model. A pre-processing is required before the degree of reaction is correlated with the compressive strength. In this database, the compressive strength was tested from binder; mortar; and concrete. Fu et al. [79] have demonstrated that compressive strength is improved by increasing binder content and decreasing aggregate content. To unify the degree of reaction for all specimens in our database, it is important to normalize measured compressive strength by a standard specimen composition (i.e., a standard binder-to-aggregate ratio in volume). This is because inconsistent binder-to-aggregate ratios result in inaccurate correlations between compressive strength and the degrees of reaction of the aluminosilicate-rich precursors. The compressive strength of pure binder systems is directly used to determine the degree of reaction. For the remaining alkali-activated systems, the volume ratio of binder-to-aggregate is calculated, and most systems fall into the 0.4-to-0.6 range. Thus, 0.5 of binder-to-aggregate ratio is selected as the standard composition to normalize compressive strength. Here, the normalized compressive strength ($CS_{normalized}$) is calculated on the basis of **Eq. (2)**. $CS_{measured}$ is compressive strength measured from experiment; $r_{measured}$ is the volume ratio of binder-to-aggregate of each specimen; and $r_{standard}$ is the volume ratio of binder-to-aggregate of the standard specimen, which is 0.5. The reactivity (maximum degree of reaction) of the aluminosilicate-rich precursors depends on the crystallinity, where a highly-crystalline aluminosilicate-rich precursors results in a low reactivity and thus yielding a low compressive strength in the alkali-activated systems [12]. Previous studies [14,15,80] have revealed that the amorphous content of the aluminosilicate-rich precursors ranges from 40% to 90%. This indicates that the maximum degree of reaction of the aluminosilicate-rich precursors cannot exceed 90%. In our study, we assume that 80% of the aluminosilicate-rich precursors have reacted in the specimen with the highest compressive strength (80 MPa). The degree of reaction of the aluminosilicate-rich precursors for the remaining specimens is computed as the proportion of 80% based on the fraction of the compressive strength to maximum compressive strength.

$$CS_{normalized} = \frac{CS_{measured}}{\frac{r_{measured}}{r_{standard}}} \quad (2)$$

The database used for ML models contains 10 inputs and 1 output. The inputs include: normalized mass of aluminosilicate material; solid NaOH; solid Na_2SiO_3; fine aggregate; coarse aggregate; and water ($\%_{mass}$); curing temperature (°C); age (days); type (unitless; 1 = one-part and 2 = two-part). The output is compressive strength (MPa), which is the measured value obtained from the literature. It is worth noting that NaOH and Na_2SiO_3 are alkali activators. The *number of constraints* of the aluminosilicate-rich precursors (unitless) and volume fraction of products (unitless) are applied to the model as additional inputs when the network and thermodynamic constraints are applied. It should be noted that the *number of constraints* of the aluminosilicate-rich precursors is calculated by adding up the number of constraints of each aluminosilicate-rich precursor with respect to its mass fraction in all aluminosilicate-rich precursors. Statistical parameters pertaining to inputs and output are itemized in Table 1. The database used in this study is provided in **Supplementary Information**.

3. Results and discussion

3.1. Thermodynamic simulations

Based on GEMS simulations, we find that across all alkali-activated systems resulted in a near-linear correlation (Fig. 3) with the compressive strength against the volume fraction of products. The reason to evaluate this relationship is that the compressive strength of any alkali-activated system is correlated with the extent of reaction of the aluminosilicate-rich precursors, which, in turn, dictates the volume fraction of the products. It should be noted that a mathematical function that interprets the correlation between compressive strength and volume fraction of products in 520 alkali-activated systems is shown in the legend. This mathematical function allows scientists to promptly examine the compressive strength of any given alkali-activated system by using its volume fraction of products as the sole input.

It is worth pointing out that the RF model can directly produce predictions of the volume fraction of products in alkali-activated systems, without calculating them based on phase assemblages from thermodynamic simulations. This requires the RF model to be trained with a new output, which consists of the same input variables shown in **Section 3.0**, but the volume of products as the output. Once the RF model is trained, it can predict the volume fraction of products (Fig. 4) in new alkali-activated systems with respect to their mixture design. The predicted volume fraction of products can also be used as the thermodynamic constraint (the same constraint obtained from the phase assemblage) to regulate the RF model to predict the compressive strength. However, in this study, all thermodynamic constraints are derived from the GEMS simulations.

Table 1
Statistical parameters pertaining to 10 inputs; 2 constraints; and 1 output (bold) of 520 alkali activated systems in the compressive strength database.

Attribute	Unit	Min.	Max.	Mean	Std. Dev.
Aluminosilicate Material Content	$\%_{mass}$	9.451	73.76	26.25	16.64
Fine Aggregate Content	$\%_{mass}$	0	66.67	32.33	16.99
Coarse Aggregate Content	$\%_{mass}$	0	79.55	27.33	23.07
Na_2SiO_3 Content	$\%_{mass}$	0	17.243	4.420	3.600
NaOH Content	$\%_{mass}$	0	3.194	0.857	0.821
Superplasticizer Content	$\%_{mass}$	0	0.427	0.075	0.141
Water Content	$\%_{mass}$	2.685	25.92	8.395	4.073
Curing Temperature	°C	20	85	36.47	22.94
Curing time	Days	1	28	12.84	10.75
Type	unitless	1	2	–	–
Number of Constraints	unitless	3.092	4.155	3.634	0.297
Volume Fraction of Product	unitless	0.023	0.6921	0.214	0.136
Compressive Strength	**MPa**	**5.152**	**79.82**	**27.16**	**16.80**

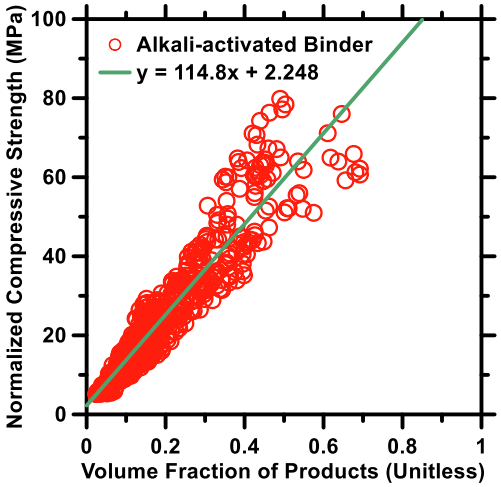

Fig. 3. A linear correlation between compressive strength and volume fraction of products in 520 alkali-activated systems used in this study. A mathematical function qualitatively estimating this correlation is shown in the legend.

3.2. Machine learning predictions

Fig. 5 demonstrates the predictions of the compressive strength of alkali-activated systems produced by the RF model without constraints

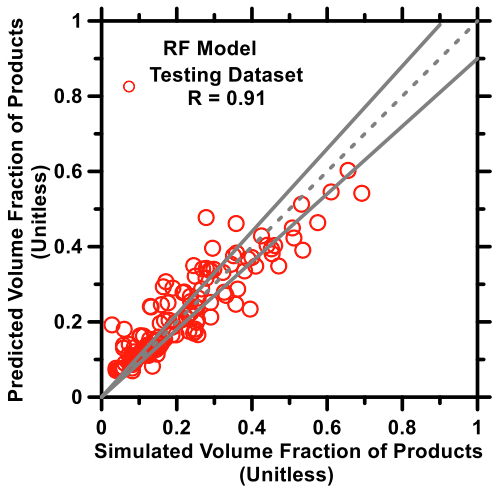

Fig. 4. Predictions of volume fraction of products as produced by the RF model against values obtained from thermodynamic simulations. Pearson correlation coefficient (R) of predicted results is shown in the legend. The dash line indicates ideal prediction, and solid lines represent a ± 10% error bound.

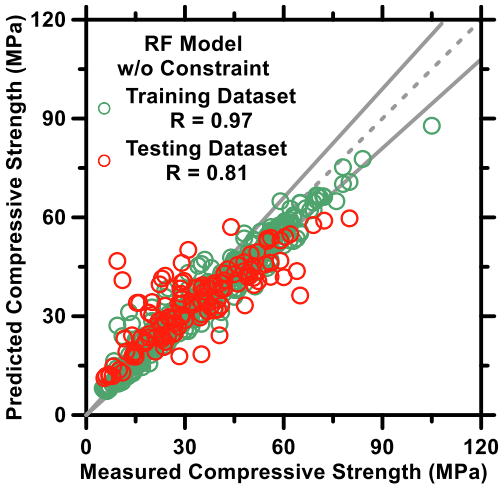

Fig. 5. Predictions of compressive strength of alkali activated systems as produced by the RF model without any constraints compared against measured compressive strength. Pearson correlation coefficient (R) of predicted results is shown in the legend. The dash line indicates ideal prediction, and solid lines represent a ± 10% error bound.

against the measured compressive strength. Predictions for both training and testing datasets are shown in the figure. The five statistical parameters (mentioned in **Section 2.3**) pertaining to the performance of the RF model on the testing dataset are itemized in Table 2.

As shown in Fig. 5 and Table 2, the accuracy of predictions produced by the RF model is moderate. R^2 is 0.63 and *RMSE* is 9.752 MPa. Although previous studies [19,26,28,81,82] have demonstrated that standalone ML models can yield accurate predictions of compressive strength of alkali-activated systems, it is expected that a standalone model cannot produce highly accurate predictions for the database used in this study. This is because our database is much more complex than databases used in previous studies [19,26,28,81,82]. First, both one-part and two-part alkali-activated systems are included in the database. In general, due to heat generated from the dissolution of solid activators, one-part alkali-activated systems are more reactive than two-part systems [83]. Therefore, one-part alkali-activated systems demonstrate higher compressive strength at early ages and a shorter setting time compared to two-part systems. Second, the chemical composition of the aluminosilicate-rich precursors in previous studies were simple, only containing a few different compositions. However, there are 26 different compositions of the aluminosilicate-rich precursors in this study. Literature [9,20,84] has shown that chemical compositions of the aluminosilicate-rich precursors significantly affects the compressive strength of alkali-activated systems, where high Al_2O_3 and SiO_2 contents lead to binders with high compressive strength. Third, the database in this study contains a large-scale of water-to-solid ratio from 0.18-to-0.50. The geopolymerization reaction is incomplete without sufficient water; in contrast, excessive water creates an additional porosity that can reduce compressive strength. Finally, the database consists of a large range of curing temperatures, from 20-to-85 °C. Elevating the curing temperature increases the extent and rate of the reaction through escalating mesopore volume and surface area for nuclei of geopolymerization reaction [85]. This accelerates setting time accompanied by higher compressive strength at early ages.

Table 2
The statistical parameters pertaining the performance of the RF model with or without network and thermodynamic constraints on predictions of alkali-activated systems' compressive strength against the testing dataset.

Model Name	R Unitless	R^2 Unitless	MAE MPa	MAPE %	RMSE MPa
RF	0.8095	0.6553	6.854	29.39	9.433
RF + Network Constraint	0.8841	0.7816	5.956	24.56	7.868
RF + Network Constraint + Thermodynamic Constraint	0.9619	0.9253	3.649	13.45	4.865

As abovementioned, due to the complexity of the database used in this study, it is important to further calibrate the RF model to obtain better predictions of compressive strength. Towards this, the topological network constraint (described in **Section 2.1**) and volume fraction of products from thermodynamic simulations (described in **Section 2.2**) are used to provide information of fundamental material laws to guide the RF model. Fig. 6**a** demonstrates the predictions of compressive strength of alkali-activated systems produced by the RF model with the network constraint. Fig. 6**b** shows the predictions of compressive strength of alkali-activated systems produced by the RF model with the network and thermodynamic constraints. Predictions for both training and testing datasets are shown in the figures. The corresponding prediction errors of the testing datasets are summarized in Table 2.

As can be seen in Fig. 6 and Table 2**,** the RF model produces accurate predictions for alkali-activated systems' compressive strength with network and thermodynamic constraints. R^2 and *RMSE* of the predictions with the network constraint are 0.78 and 7.868 MPa, respectively. R^2 and *RMSE* of the predictions with the network and thermodynamic constraints are 0.93 and 4.865 MPa, respectively. The prediction error is even smaller than the standard deviation (5 MPa) of the compressive strength measurement [86]. Based on these values, it is clear that the predictions with constraints are superior compared to predictions produced by the standalone RF model, which is expected. Unlike the standalone model without knowing the chemical composition of the aluminosilicate-rich precursors, the network constraint provides chemostructural properties (e.g., the quantities of NBOs; BOs; FOs; and TOs and strength of chemical bonds) of aluminosilicate-rich precursors to guide the RF model. This information can serve as the proxy for the reactivity of aluminosilicate-rich precursors. The highly reactive aluminosilicate-rich precursors react fast with the alkali activators and water resulting in higher compressive strength at early ages and an overall higher degree of reaction. Thus, critical information from the network constraint helps the RF model to catch the trend between inputs and compressive strength. The thermodynamic constraint refines predictions of compressive strength further. As shown in Fig. 3, compressive strength exhibits a strong correlation with the thermodynamic constraint (volume fraction of products). Therefore, the thermodynamic constraint restricts predictions produced by the RF model in a small and reliable range. Furthermore, the thermodynamic constraint prohibits

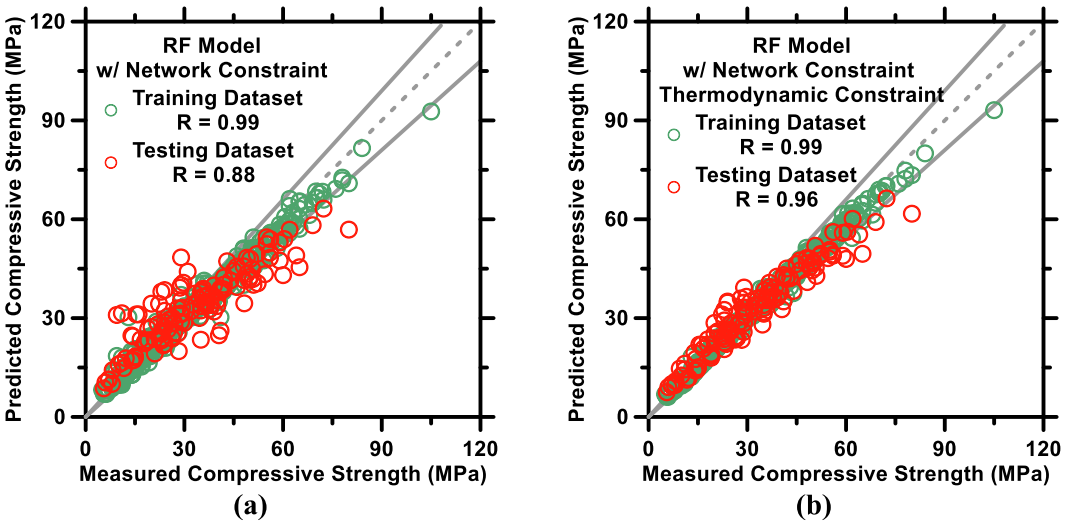

Fig. 6. Predictions of compressive strength of alkali activated systems as produced by the RF model with (a) network constraint and (b) network + thermodynamic constraints compared against measured compressive strength. Pearson correlation coefficient (R) of predicted results is shown in the legend. The dash line indicates ideal prediction and solid lines represent a ± 10% error bound.

the RF model from violating fundamental material laws.

3.3. The role of chemical compositions of aluminosilicate materials

Previous sections have shown that the RF model with constraints can produce reliable predictions of compressive strength. However, because machine learning models are a "black-box", they cannot interpret mechanisms behind compressive strength and mixture design. This section reveals the influence of the chemical composition of the aluminosilicate-rich precursors and the microstructure of alkali-activated systems on the compressive strength through a relationship between the *number of constraints*, phase assemblages, and compressive strength. The simulations are conducted with 0.35 of water-to-solid ratio; 0.14 of alkali-to-fly ash ratio; 0.27 of NaOH-to-$NaSiO_2$ ratio; and 20 °C. The fly ash composition is $(CaO)_{0.4-x}(Al_2O_3)_x(SiO_2)_{0.6}$, and the degree of reaction is 60%. Such parameters are chosen because most data-records in the database have similar parameters. The simulated results for fully polymerized and partially depolymerized are shown in Fig. 7. As expected, the main phases are C-N-A-S-H and N-A-S-H (Na-Chabazite) [9,45]. For both regimes, at $n_c = 3.4$, the compressive strength has the highest value. This is because the trade-off between C-N-A-S-H and N-A-S-H phases reaches the optimal state. Specifically, before $n_c = 3.4$, alkali-activated systems do not contain enough network forming atoms to form a high strength gel-like polymeric structure.

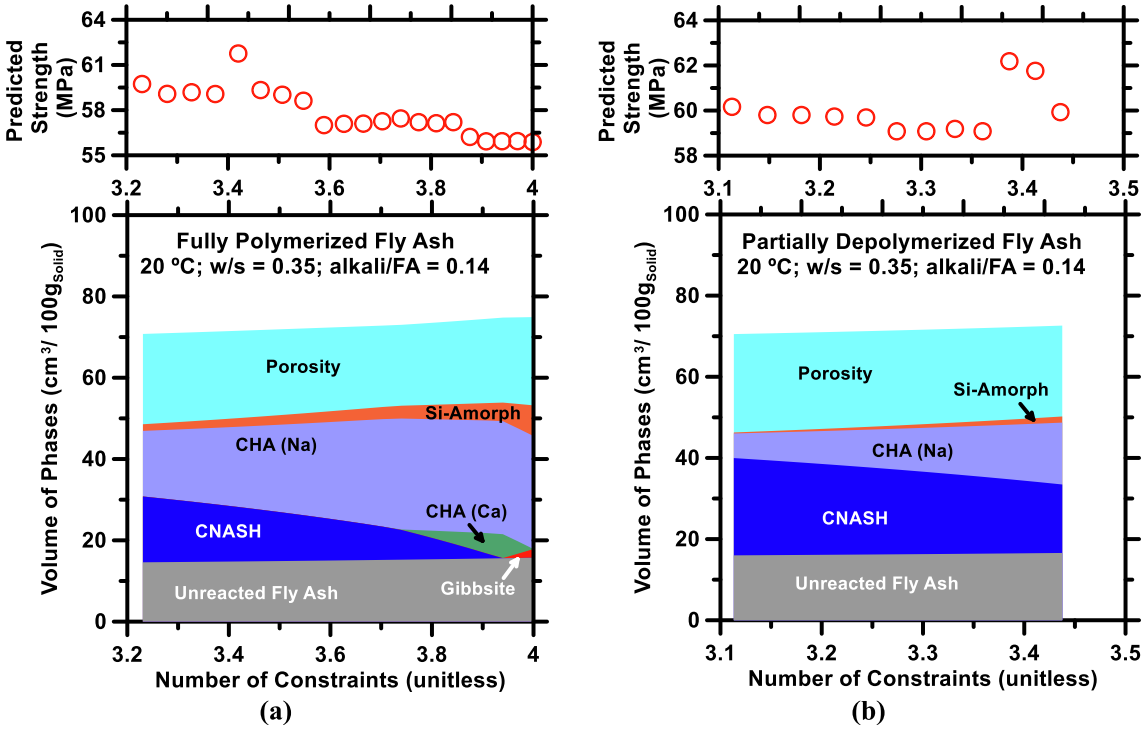

Fig. 7. Phase assemblages (bottom) of alkali-activated systems simulated from thermodynamic modeling and compressive strength (top) predicted from the RF model corresponding to network structure (*number of constraints*) of **(a)** fully polymerized fly ash; and **(b)** partially depolymerized fly ash. The parameters for thermodynamic simulations are shown in the legend.

However, after $n_c = 3.4$, the reactivity of fly ash diminishes due to excess network former atoms.

Fig. 7**a** demonstrates the phase assemblage and compressive strength for fully polymerized fly ash. The Al_2O_3 varies from 0.2 ($n_c = 3.23$) to 0.4 ($n_c = 4.00$). With increasing Al_2O_3 and decreasing CaO, it can be observed that the volume of C-N-A-S-H decrease monotonically. At $n_c = 3.7$, Ca-chabazite starts forming. Most of the Na_2O has formed stable phases (Na-Chabazite) with Al_2O_3, and insufficient CaO forces more Al_2O_3 to get into a Ca-rich structure, where producing 1 mol Ca-chabazite consumes 2 mol of Al_2O_3. Later CaO is too low to form a Ca-rich structure, thus gibbsite starts forming. With decreasing volume of C-N-A-S-H, the compressive strength values tend to plummet. This reinforces that C-N-A-S-H is the major component to provide strength in alkali-activate systems [7,8,20]. Ca-chabazite and gibbsite barely contribute to compressive strength. Fig. 7**b** shows phase assemblage and compressive strength for partially depolymerized fly ash. The Al_2O_3 varies from 0.19 ($n_c = 3.23$) to 0.08 ($n_c = 4.00$). It is clearly observed that volume of C-N-A-S-H decrease with increasing *number of constraints*. Compared to fully polymerized fly ash, partially depolymerized fly ash produces more C-N-A-S-H. Therefore, the compressive strength of partially depolymerized fly ash is higher, even though there is more porosity in the structure. Partially depolymerized fly ashes also form lesser amorphous Si phase because they have sufficient CaO and Al_2O_3 to form C-N-A-S-H and N-A-S-H phases.

4. Conclusions

This paper employed the RF model to predict the compressive strength of alkali-activated systems made from seven types (26 different compositions) of aluminosilicate-rich precursors. In order to improve prediction performances, thermodynamic (volume fraction of products of alkali-activated systems) and network (*number of constraints* of the aluminosilicate-rich precursors) constraints were utilized to regulate the RF model. The correlations between the phase assemblages, *number of constraints*, and compressive strength were revealed and explained. It is worthwhile to point out that this is the first study that employs ML models with constraints to predict the compressive strength of alkali-activated systems in relation to complex compositions of the aluminosilicate-rich precursors and wide processing parameters.

The database for alkali-activated systems contained 520 unique data-records with 10 inputs variables. The parent database was split into training and testing datasets to train the model and evaluate the performance. The prediction accuracy (R = 0.81) of RF model without any constraint is acceptable. The prediction accuracy (R = 0.88) of RF model with network constraint is moderate. The prediction accuracy (R = 0.96) of RF model with network and thermodynamic constraints is highly reliable. Results sufficiently reveal that the network and thermodynamic constraints provide necessary chemostructural information and material laws to enhance the performance of the RF model. The well-trained RF model with constraints was used to establish correlations between microstructure of alkali-activated systems; reactivity of aluminosilicate-rich precursors; and compressive strength. When the *number of constraints* was equal to 3.4, the alkali-activated systems achieved the optimal compressive strength. The C-N-A-S-H phase is the major phase that provide this high compressive strength.

In conclusion, the predictions shown in this study can still be improved with a larger and more diverse database. In the future, except for Al_2O_3, CaO, SiO_2, other chemical compositions of the aluminosilicate-rich precursors can be accounted in the phase assemblage. This is towards understanding the influence of minor components and their products on compressive strength. However, this paper presents a new pathway to understanding microstructure-property correlations. Meanwhile, the ML models can provide a preview of properties of alkali-activated systems made from new aluminosilicate-rich precursors without performing experiments.

CRediT authorship contribution statement

Rohan Bhat: . **Taihao Han:** . **Sai Akshay Ponduru:** . **Arianit Reka:** Conceptualization. **Jie Huang:** . **Gaurav Sant:** Conceptualization. **Aditya Kumar:** Conceptualization.

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.

Acknowledgement

This study was financially supported by: the Leonard Wood Institute (LWI: W911NF-07-2-0062); the National Science Foundation (NSF-CMMI: 1661609; NSF-CMMI: 1932690; NSF-DMR: 2034856); and the Federal Highway Administration (Award no: 693JJ31950021); and the Fulbright Visiting Scholar Program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.conbuildmat.2022.127557.

References

[1] C. Shi, A.F. Jiménez, A. Palomo, New cements for the 21st century: The pursuit of an alternative to Portland cement, Cem. Concr. Res. 41 (2011) 750–763, https://doi.org/10.1016/j.cemconres.2011.03.016.
[2] J.S. Damtoft, J. Lukasik, D. Herfort, D. Sorrentino, E.M. Gartner, Sustainable development and climate change initiatives, Cem. Concr. Res. 38 (2008) 115–127, https://doi.org/10.1016/j.cemconres.2007.09.008.
[3] E. Gartner, H. Hirao, A review of alternative approaches to the reduction of CO_2 emissions associated with the manufacture of the binder phase in concrete, Cem. Concr. Res. 78 (2015) 126–142, https://doi.org/10.1016/j.cemconres.2015.04.012.
[4] Cement Industry Energy and CO_2 Performance, World Business Council for Sustainable Development. Conches-Geneva (2009). https://www.wbcsd.org/Projects/Cement-Sustainability-Initiative/Resources/Getting-the-Numbers-Right.
[5] J. Davidovits, Geopolymer Chemistry and Applications, 4th Ed, Institute Geopolymer, 2015.
[6] J.L. Provis, S.A. Bernal, Geopolymers and related alkali-activated materials, Annu. Rev. Mater. Res. 44 (2014) 299–327, https://doi.org/10.1146/annurev-matsci-070813-113515.
[7] B. Singh, G. Ishwarya, M. Gupta, S.K. Bhattacharyya, Geopolymer concrete: A review of some recent developments, Constr. Build. Mater. 85 (2015) 78–90, https://doi.org/10.1016/j.conbuildmat.2015.03.036.
[8] T. Luukkonen, Z. Abdollahnejad, J. Yliniemi, P. Kinnunen, M. Illikainen, One-part alkali-activated materials: A review, Cem. Concr. Res. 103 (2018) 21–34, https://doi.org/10.1016/j.cemconres.2017.10.001.
[9] A.A. Mohammed, H.U. Ahmed, A. Mosavi, Survey of mechanical properties of geopolymer concrete: A comprehensive review and data analysis, Materials. 14 (2021) 4690, https://doi.org/10.3390/ma14164690.
[10] C. Meyer, The greening of the concrete industry, Cem. Concr. Compos. 31 (2009) 601–605, https://doi.org/10.1016/j.cemconcomp.2008.12.010.
[11] P. Duxson, A. Fernández-Jiménez, J.L. Provis, G.C. Lukey, A. Palomo, J.S.J. van Deventer, Geopolymer technology: the current state of the art, J. Mater. Sci. 42 (2007) 2917–2933, https://doi.org/10.1007/s10853-006-0637-z.
[12] K.U. Ambikakumari Sanalkumar, M. Lahoti, E.-H. Yang, Investigating the potential reactivity of fly ash for geopolymerization, Constr. Build. Mater. 225 (2019) 283–291, https://doi.org/10.1016/j.conbuildmat.2019.07.140.
[13] N.K. Lee, G.H. An, K.T. Koh, G.S. Ryu, Improved Reactivity of fly ash-slag geopolymer by the addition of silica fume, Adv. Mater. Sci. Eng. 2016 (2016) e2192053, https://doi.org/10.1155/2016/2192053.
[14] C. Li, H.H. Sun, L.T. Li, Glass phase structure of blast furnace slag, Adv. Mater. Res. 168–170 (2011) 3–7, https://doi.org/10.4028/www.scientific.net/AMR.168-170.3.
[15] Y. Song, K. Yang, J. Chen, K. Wang, G. Sant, M. Bauchy, Machine learning enables rapid screening of reactive fly ashes based on their network topology, ACS Sustainable Chem. Eng. 9 (2021) 2639–2650, https://doi.org/10.1021/acssuschemeng.0c06978.
[16] K. Yang, Y. Hu, Z. Li, N.M.A. Krishnan, M.M. Smedskjaer, C.G. Hoover, J.C. Mauro, G. Sant, M. Bauchy, Analytical model of the network topology and rigidity of calcium aluminosilicate glasses, J. Am. Ceram. Soc. 104 (2021) 3947–3962, https://doi.org/10.1111/jace.17781.

[17] J.C. Phillips, Topology of covalent non-crystalline solids I: Short-range order in chalcogenide alloys, J. Non-Cryst. Solids 34 (1979) 153–181, https://doi.org/10.1016/0022-3093(79)90033-4.

[18] M. Bauchy, Deciphering the atomic genome of glasses by topological constraint theory and molecular dynamics: A review, Comput. Mater. Sci. 159 (2019) 95–102, https://doi.org/10.1016/j.commatsci.2018.12.004.

[19] K.T. Nguyen, Q.D. Nguyen, T.A. Le, J. Shin, K. Lee, Analyzing the compressive strength of green fly ash based geopolymer concrete using experiment and machine learning approaches, Constr. Build. Mater. 247 (2020), 118581, https://doi.org/10.1016/j.conbuildmat.2020.118581.

[20] C. Ng, U.J. Alengaram, L.S. Wong, K.H. Mo, M.Z. Jumaat, S. Ramesh, A review on microstructural study and compressive strength of geopolymer mortar, paste and concrete, Constr. Build. Mater. 186 (2018) 550–576, https://doi.org/10.1016/j.conbuildmat.2018.07.075.

[21] B. Panda, M.J. Tan, Experimental study on mix proportion and fresh properties of fly ash based geopolymer for 3D concrete printing, Ceram. Int. 44 (2018) 10258–10265, https://doi.org/10.1016/j.ceramint.2018.03.031.

[22] H.-B. Le, Q.-B. Bui, L. Tang, Geopolymer recycled aggregate concrete: From experiments to empirical models, Materials 14 (2021) 1180, https://doi.org/10.3390/ma14051180.

[23] J. De Brito, R. Kurda, P. Raposeiro da Silva, Can we truly predict the compressive strength of concrete without knowing the properties of aggregates? Appl. Sci. 8 (2018) 1095, https://doi.org/10.3390/app8071095.

[24] B.M.M.R. Sudhir, S. Chen, S. Rai, D. Jain, An empirical model for geopolymer reactions involving fly ash and GGBS, Adv. Mater. Sci. Eng. 2022 (2022), https://doi.org/10.1155/2022/8801294.

[25] J. Jonbi, M.A. Fulazzaky, Modeling the water absorption and compressive strength of geopolymer paving block: An empirical approach, Measurement 158 (2020), 107695, https://doi.org/10.1016/j.measurement.2020.107695.

[26] M.M. Yadollahi, A. Benli, R. Demirboğa, Prediction of compressive strength of geopolymer composites using an artificial neural network, Mater. Res. Innov. 19 (2015) 453–458, https://doi.org/10.1179/1433075X15Y.0000000020.

[27] D. Dao, S. Trinh, H.-B. Ly, B. Pham, Prediction of compressive strength of geopolymer concrete using entirely steel slag aggregates: Novel hybrid artificial intelligence approaches, Appl. Sci. 9 (2019) 1113, https://doi.org/10.3390/app9061113.

[28] E. Gomaa, T. Han, M. ElGawady, J. Huang, A. Kumar, Machine Learning to predict properties of fresh and hardened alkali-activated concrete, Cem. Concr. Compos. 115 (2021), 103863, https://doi.org/10.1016/j.cemconcomp.2020.103863.

[29] L.V. Zhang, A. Marani, M.L. Nehdi, Chemistry-informed machine learning prediction of compressive strength for alkali-activated materials, Constr. Build. Mater. 316 (2022), 126103, https://doi.org/10.1016/j.conbuildmat.2021.126103.

[30] M. Lahoti, P. Narang, K.H. Tan, E.-H. Yang, Mix design factors and strength prediction of metakaolin-based geopolymer, Ceram. Int. 43 (2017) 11433–11441, https://doi.org/10.1016/j.ceramint.2017.06.006.

[31] T. Han, S.A. Ponduru, R. Cook, J. Huang, G. Sant, A. Kumar, A deep learning approach to design and discover sustainable cementitious binders: Strategies to learn from small databases and develop closed-form analytical models, Front. Mater. 8 (2022), https://doi.org/10.3389/fmats.2021.796476.

[32] D.A. Kulik, T. Wagner, S.V. Dmytrieva, G. Kosakowski, F.F. Hingerl, K. V. Chudnenko, U.R. Berner, GEM-Selektor geochemical modeling package: revised algorithm and GEMS3K numerical kernel for coupled simulation codes, Comput. Geosci. (2012), https://doi.org/10.1007/s10596-012-9310-6.

[33] T. Wagner, D.A. Kulik, F.F. Hingerl, S.V. Dmytrieva, GEM-Selektor geochemical modeling package: TSolMod library and data interface for multicomponent phase models, Can. Mineral. 50 (2012) 1173–1195, https://doi.org/10.3749/canmin.50.5.1173.

[34] R.J. Myers, B. Lothenbach, S.A. Bernal, J.L. Provis, Thermodynamic modelling of alkali-activated slag cements, Appl. Geochem. 61 (2015) 233–247, https://doi.org/10.1016/j.apgeochem.2015.06.006.

[35] Q. Wan, F. Rao, S. Song, R.E. García, R.M. Estrella, C.L. Patiño, Y. Zhang, Geopolymerization reaction, microstructure and simulation of metakaolin-based geopolymers at extended Si/Al ratios, Cem. Concr. Compos. 79 (2017) 45–52, https://doi.org/10.1016/j.cemconcomp.2017.01.014.

[36] T. Oey, A. Kumar, I. Pignatelli, Y. Yu, N. Neithalath, J.W. Bullard, M. Bauchy, G. Sant, Topological controls on the dissolution kinetics of glassy aluminosilicates, J. Am. Ceram. Soc. 100 (2017) 5521–5527, https://doi.org/10.1111/jace.15122.

[37] A.K. Varshneya, Fundamentals of Inorganic Glasses, Second Edition, Society of Glass Technology, UK, 2013.

[38] J.E. Shelby, R.S. of C (Great Britain), Introduction to Glass Science and Technology, Royal Society of Chemistry, 2005.

[39] M. Bauchy, M.J. Abdolhosseini Qomi, C. Bichara, F.-J. Ulm, R.-J.-M. Pellenq, Nanoscale structure of cement: Viewpoint of rigidity theory, J. Phys. Chem. C. 118 (2014) 12485–12493, https://doi.org/10.1021/jp502550z.

[40] M. Bauchy, M. Micoulaut, Atomic scale foundation of temperature-dependent bonding constraints in network glasses and liquids, J. Non-Cryst. Solids 357 (2011) 2530–2537, https://doi.org/10.1016/j.jnoncrysol.2011.03.017.

[41] T. Oey, K.F. Frederiksen, N. Mascaraque, R. Youngman, M. Balonis, M. M. Smedskjaer, M. Bauchy, G. Sant, The role of the network-modifier's field-strength in the chemical durability of aluminoborate glasses, J. Non-Cryst. Solids 505 (2019) 279–285, https://doi.org/10.1016/j.jnoncrysol.2018.11.019.

[42] B. Lothenbach, T. Matschei, G. Möschner, F.P. Glasser, Thermodynamic modelling of the effect of temperature on the hydration and porosity of Portland cement, Cem. Concr. Res. 38 (2008) 1–18, https://doi.org/10.1016/j.cemconres.2007.08.017.

[43] T. Matschei, B. Lothenbach, F.P. Glasser, Thermodynamic properties of Portland cement hydrates in the system $CaO–Al_2O_3–SiO_2–CaSO_4–CaCO_3–H_2O$, Cem. Concr. Res. 37 (2007) 1379–1410, https://doi.org/10.1016/j.cemconres.2007.06.002.

[44] B. Ma, B. Lothenbach, Synthesis, characterization, and thermodynamic study of selected K-based zeolites, Cem. Concr. Res. 148 (2021), 106537, https://doi.org/10.1016/j.cemconres.2021.106537.

[45] X. Ke, Y. Duan, Coupling machine learning with thermodynamic modelling to develop a composition-property model for alkali-activated materials, Compos. B Eng. 216 (2021), 108801, https://doi.org/10.1016/j.compositesb.2021.108801.

[46] D. Lootens, D.P. Bentz, On the relation of setting and early-age strength development to porosity and hydration in cement-based materials, Cement Concrete, Composites. 68 (2016) 9–14, https://doi.org/10.1016/j.cemconcomp.2016.02.010.

[47] O.A. Mohamed, A review of durability and strength characteristics of alkali-activated slag concrete, Materials. 12 (2019) 1198, https://doi.org/10.3390/ma12081198.

[48] K. Gijbels, Y. Pontikes, P. Samyn, S. Schreurs, W. Schroeyers, Effect of NaOH content on hydration, mineralogy, porosity and strength in alkali/sulfate-activated binders from ground granulated blast furnace slag and phosphogypsum, Cem. Concr. Res. 132 (2020), 106054, https://doi.org/10.1016/j.cemconres.2020.106054.

[49] L. Breiman, Bagging predictors, Machine Learn. 24 (1996) 123–140, https://doi.org/10.1007/BF00058655.

[50] L. Breiman, Random forests, Machine Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[51] A. Liaw, M. Wiener, Classification and Regression by RandomForest, 2001.

[52] Gãś. Biau, L. Devroye, Gâą Lugosi, Consistency of random forests and other averaging classifiers, J. Machine Learn. Res. 9 (2008) 2015–2033.

[53] X. Chen, H. Ishwaran, Random forests for genomic data analysis, Genomics 99 (2012) 323–329, https://doi.org/10.1016/j.ygeno.2012.04.003.

[54] R. Cook, J. Lapeyre, H. Ma, A. Kumar, Prediction of compressive strength of concrete: A critical comparison of performance of a hybrid machine learning model with standalone models, ASCE J. Mater. Civil Eng. 31 (2019) 04019255, https://doi.org/10.1061/(ASCE)MT.1943-5533.0002902.

[55] C. Schaffer, Selecting a classification method by cross-validation, Machine Learn. 13 (1993) 135–143, https://doi.org/10.1007/BF00993106.

[56] J. Lapeyre, T. Han, B. Wiles, H. Ma, J. Huang, G. Sant, A. Kumar, Machine learning enables prompt prediction of hydration kinetics of multicomponent cementitious systems, Sci. Rep. 11 (2021) 3922, https://doi.org/10.1038/s41598-021-83582-6.

[57] T. Han, N. Stone-Weiss, J. Huang, A. Goel, A. Kumar, Machine learning as a tool to design glasses with controlled dissolution for application in healthcare industry, Acta Biomater. 107 (2020) 286–298, https://doi.org/10.1016/j.actbio.2020.02.037.

[58] R. Cook, C.M. Keitumetse, M.B. Hayat, A. Kumar, L. Alagha, Prediction of flotation performance of sulfide minerals using an original hybrid machine learning model, Eng. Rep. 12 (2020), e12167, https://doi.org/10.1002/eng2.12167.

[59] S. Zhou, C. Ma, G. Long, Y. Xie, A novel non-Portland cementitious material: Mechanical properties, durability and characterization, Constr. Build. Mater. 238 (2020), 117671, https://doi.org/10.1016/j.conbuildmat.2019.117671.

[60] S. Guo, C. Ma, G. Long, Y. Xie, Cleaner one-part geopolymer prepared by introducing fly ash sinking spherical beads: Properties and geopolymerization mechanism, J. Cleaner Prod. 219 (2019) 686–697, https://doi.org/10.1016/j.jclepro.2019.02.116.

[61] J. Wongpa, K. Kiattikomol, C. Jaturapitakkul, P. Chindaprasirt, Compressive strength, modulus of elasticity, and water permeability of inorganic polymer concrete, Mater. Des. 31 (2010) 4748–4754, https://doi.org/10.1016/j.matdes.2010.05.012.

[62] M. Dong, M. Elchalakani, A. Karrech, Development of high strength one-part geopolymer mortar using sodium metasilicate, Constr. Build. Mater. 236 (2020), 117611, https://doi.org/10.1016/j.conbuildmat.2019.117611.

[63] S. Haruna, B.S. Mohammed, M.M.A. Wahab, M.S. Liew, Effect of paste aggregate ratio and curing methods on the performance of one-part alkali-activated concrete, Constr. Build. Mater. 261 (2020), 120024, https://doi.org/10.1016/j.conbuildmat.2020.120024.

[64] S. Haruna, B. Mohammed, M.M.A. Wahab, A. Al-Fakih, Effect of aggregate-binder proportion and curing technique on the strength and water absorption of fly ash-based one-part geopolymer mortars, IOP Conference Series: Materials Science and Engineering. 1101 (2021) 012022. https://doi.org/10.1088/1757-899X/1101/1/012022.

[65] S. Singh, M.U. Aswath, R.V. Ranganath, Effect of mechanical activation of red mud on the strength of geopolymer binder, Constr. Build. Mater. 177 (2018) 91–101, https://doi.org/10.1016/j.conbuildmat.2018.05.096.

[66] P.O. Awoyera, M.S. Kirgiz, A. Viloria, D. Ovallos-Gazabon, Estimating strength properties of geopolymer self-compacting concrete using machine learning techniques, J. Mater. Res. Technol. 9 (2020) 9016–9028, https://doi.org/10.1016/j.jmrt.2020.06.008.

[67] A. Rafeet, R. Vinai, M. Soutsos, W. Sha, Guidelines for mix proportioning of fly ash/GGBS based alkali activated concretes, Constr. Build. Mater. 147 (2017) 130–142, https://doi.org/10.1016/j.conbuildmat.2017.04.036.

[68] H.-Y. Zhang, J.-C. Liu, B. Wu, Mechanical properties and reaction mechanism of one-part geopolymer mortars, Constr. Build. Mater. 273 (2021), 121973, https://doi.org/10.1016/j.conbuildmat.2020.121973.

[69] C. Ma, G. Long, Y. Shi, Y. Xie, Preparation of cleaner one-part geopolymer by investigating different types of commercial sodium metasilicate in China, J. Cleaner Prod. 201 (2018) 636–647, https://doi.org/10.1016/j.jclepro.2018.08.060.

[70] K.-H. Yang, J.-K. Song, A.F. Ashour, E.-T. Lee, Properties of cementless mortars activated by sodium silicate, Constr. Build. Mater. 22 (2008) 1981–1989, https://doi.org/10.1016/j.conbuildmat.2007.07.003.
[71] M. Olivia, H. Nikraz, Properties of fly ash geopolymer concrete designed by Taguchi method, Mater. Des. 36 (2012) 191–198, https://doi.org/10.1016/j.matdes.2011.10.036.
[72] R. Vinai, A. Rafeet, M. Soutsos, W. Sha, The role of water content and paste proportion on physico-mechanical properties of alkali activated fly ash–GGBS concrete, J. Sust. Metall. 2 (2016) 51–61, https://doi.org/10.1007/s40831-015-0032-6.
[73] G. Fang, W.K. Ho, W. Tu, M. Zhang, Workability and mechanical properties of alkali-activated fly ash-slag concrete cured at ambient temperature, Constr. Build. Mater. 172 (2018) 476–487, https://doi.org/10.1016/j.conbuildmat.2018.04.008.
[74] C. Graf, A. Silica, Amorphous, in: Kirk-Othmer Encyclopedia of Chemical Technology, American Cancer Society, 2018, pp. 1–43, 10.1002/0471238961.0113151823010404.a01.pub3.
[75] C. Århammar, A. Pietzsch, N. Bock, E. Holmström, C.M. Araujo, J. Gråsjö, S. Zhao, S. Green, T. Peery, F. Hennies, S. Amerioun, A. Föhlisch, J. Schlappa, T. Schmitt, V. N. Strocov, G.A. Niklasson, D.C. Wallace, J.-E. Rubensson, B. Johansson, R. Ahuja, Unveiling the complex electronic structure of amorphous metal oxides, Proc. Natl. Acad. Sci. 108 (2011) 6355–6360, https://doi.org/10.1073/pnas.1019698108.
[76] K. Koski, J. Hölsä, P. Juliet, Properties of aluminium oxide thin films deposited by reactive magnetron sputtering, Thin Solid Films 339 (1999) 240–248, https://doi.org/10.1016/S0040-6090(98)01232-2.
[77] X. Chen, E. Kim, P. Suraneni, L. Struble, Quantitative correlation between the degree of reaction and compressive strength of metakaolin-based geopolymers, Materials. 13 (2020) 5784, https://doi.org/10.3390/ma13245784.
[78] N. Shafiq, Degree of hydration and compressive strength of conditioned samples made of normal and blended cement system, KSCE J. Civ. Eng. 15 (7) (2011) 1253–1257.
[79] T.C. Fu, W. Yeih, J.J. Chang, R. Huang, The influence of aggregate size and binder material on the properties of pervious concrete, Advances in Materials Science and Engineering. 2014 (2014) e963971. https://doi.org/10.1155/2014/963971.
[80] R.T. Chancey, P. Stutzman, M.C.G. Juenger, D.W. Fowler, Comprehensive phase characterization of crystalline and amorphous phases of a Class F fly ash, Cem. Concr. Res. 40 (2010) 146–156, https://doi.org/10.1016/j.cemconres.2009.08.029.
[81] D. Dao, H.-B. Ly, S. Trinh, T.-T. Le, B. Pham, Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete, Materials. 12 (2019) 983, https://doi.org/10.3390/ma12060983.
[82] M.A. Khan, A. Zafar, F. Farooq, M.F. Javed, R. Alyousef, H. Alabduljabbar, M. I. Khan, Geopolymer concrete compressive strength via artificial neural network, adaptive neuro fuzzy interface system, and gene expression programming with k-fold cross validation, Front. Mater. 8 (2021) 66, https://doi.org/10.3389/fmats.2021.621163.
[83] T. Suwan, M. Fan, Effect of manufacturing process on the mechanisms and mechanical properties of fly ash-based geopolymer in ambient curing temperature, Mater. Manuf. Processes 32 (2017) 461–467, https://doi.org/10.1080/10426914.2016.1198013.
[84] O.K. Wattimena, D. Hardjito Antoni, A review on the effect of fly ash characteristics and their variations on the synthesis of fly ash based geopolymer, AIP Conf. Proc. 1887 (2017), 020041, https://doi.org/10.1063/1.5003524.
[85] Sindhunata, J.S.J. van Deventer, G.C. Lukey, H. Xu, Xu, Effect of curing temperature and silicate concentration on fly-ash-based geopolymerization, Ind. Eng. Chem. Res. 45 (10) (2006) 3559–3568.
[86] E. Nawy, Concrete Construction Engineering Handbook, CRC Press, 2008, 10.1201/9781420007657.