ASTM INTERNATIONAL

Advances in Civil Engineering Materials

Yu Song,[1,2] Steven Zhang,[1] Kaixin Wang,[1] Chen Jin,[1] Gaurav Sant,[2,3,4,5] and Mathieu Bauchy[1,6]

DOI: 10.1520/ACEM20220024

Interpreting the Strength Activity Index of Fly Ash with Machine Learning

Advances in Civil Engineering Materials

doi:10.1520/ACEM20220024 available online at www.astm.org

Yu Song,[1,2] Steven Zhang,[1] Kaixin Wang,[1] Chen Jin,[1] Gaurav Sant,[2,3,4,5] and Mathieu Bauchy[1,6]

Interpreting the Strength Activity Index of Fly Ash with Machine Learning

Reference

Y. Song, S. Zhang, K. Wang, C. Jin, G. Sant, and M. Bauchy, "Interpreting the Strength Activity Index of Fly Ash with Machine Learning," *Advances in Civil Engineering Materials* https://doi.org/10.1520/ACEM20220024

ABSTRACT

Fly ash from the coal combustion at electric plants is commonly used for partially replacing portland cement in concrete production. Because of the varying nature of the coal source and the different processing protocols, different fly ashes exhibit wide ranges of physical and chemical characteristics, resulting in distinct impacts on concrete strength. Thus far, the most adopted method for assessing a given fly ash is specified by ASTM C618, *Specification for Coal Fly Ash and Raw or Calcined Natural Pozzolan for Use in Concrete*, wherein a series of influential chemical and physical features can be correlated to fly ash's strength activity index (SAI). However, limited knowledge is available on how exactly the individual material attribute affects SAI, so accurately predicting the SAI remains out of reach. Here, we take advantage of recent advances in machine learning to reveal the origins of fly ash's SAI. Leveraging a data set comprising 2,158 fly ash samples, we trained neural network models to predict 28-day SAI based on the sole knowledge of ASTM C618 material attributes. The results demonstrate that SAI is a complex property that does not systematically follow the conventional Class C/F classification. To gain a deeper insight into this matter, we further quantify the influence of each attribute on SAI as captured by the machine learning model.

Keywords

machine learning, waste-to-resource, fly ash, bulk composition, high-throughput screening

Manuscript received March 2, 2022; accepted for publication July 25, 2022; published online September 22, 2022.

[1] Physics of AmoRphous and Inorganic Solids Laboratory (PARISlab), 5731B Boelter Hall, Department of Civil and Environmental Engineering, University of California Los Angeles, 580 Portola Plaza, Los Angeles, CA 90095, USA, https://orcid.org/0000-0001-6218-3234 (Y.S.)

[2] Laboratory for the Chemistry of Construction Materials (LC2), 5731J Boelter Hall, Department of Civil and Environmental Engineering, University of California Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095, USA, https://orcid.org/0000-0002-1124-5498 (G.S.)

[3] Institute for Carbon Management (ICM), University of California, Los Angeles, CA, USA

Introduction

Fly ash is the major by-product of coal combustion as generated from electric power plants. Among all the utilizations of this material, fly ash is extensively used as a supplementary cementitious material (SCM) for partially replacing portland cement in the

[4] Department of Materials Science and Engineering, University of California, Los Angeles, CA, USA

[5] California Nanosystems Institute, University of California, Los Angeles, CA, USA

[6] Institute for Carbon Management (ICM), University of California, Los Angeles, CA, USA (Corresponding author), e-mail: bauchy@ucla.edu, https://orcid.org/0000-0003-4600-0631

production of concrete, which has been broadly adopted over decades.[1–3] Because the production of cement—a key component of concrete materials—is involved with a substantial amount of carbon dioxide emissions (which accounts for roughly 10 % of the total emissions caused by human activities),[4] reducing the usage of cement is thus far the most effective approach for curbing the release of carbon in the concrete industry. In that regard, partially replacing cement with fly ash is among the most important means to fulfill the low-carbon manufacturing of concrete materials. In addition to the environmental benefit, the use of fly ash is further associated with a series of practical advantages for concrete production, such as refining both fresh and hardened properties of concrete and lowering the raw material cost.[5–7] Arguably, fly ash represents one of the most successful cases of waste-to-resource recycling in engineering practices.

Despite the tremendous attractions of using fly ash as an SCM, the addition of unqualified fly ashes can be detrimental to concrete performance, causing diminished strength, delayed setting time, larger shrinkage, incompatibility issues with the chemical additives, etc.[8,9] To minimize the potential adverse effects of the addition of fly ashes in concrete, a list of constraints have been specified by ASTM C618, *Specification for Coal Fly Ash and Raw or Calcined Natural Pozzolan for Use in Concrete*, to avoid unfavorable fly ashes.[10] As the main concern revolves around the impact on concrete strength, because of the abundance of fly ashes in history, the most adopted guideline is established on an empirical rule that the strength active fly ash (known as Class C fly ash) should have more than 18 % calcium oxide (CaO). Although this rule has been widely accepted in practice, previous studies have discussed that it can lead to a huge volume of strength active fly ashes being incorrectly classified.[5,11–14] More recently, given the aggravated conflict between the increasing demand for fly ash in concrete production and the decreasing trend of the fly ash production itself because of the wind-down of coal-based electric plants,[15–17] the crude empirical-based rule for identifying strength active fly ashes has become obsolete. To mitigate the shortage of fly ashes, it is pertinent to develop high-throughput screening methods that can offer an accurate prediction of the strength potential of the raw ashes.

Over the past decades, many studies have proposed more robust testing methods to quantify the chemical reactivity of fly ashes by characterizing their dissolution rate,[18,19] hydration heat,[20] amorphous phases,[21] etc. Thanks to recent advances in computational material science, the chemical reactivity of fly ashes and the mechanical properties of the reaction products have also been successfully predicted based on different theories, such as thermodynamic modeling,[22,23] micromechanical modeling,[24] and topological constraint theory.[18,25] Extensive reviews of such advances in testing and modeling of the fly ash reactivity can be found in the literature.[3,9,26,27] However, though those approaches can offer some critical knowledge on the fundamental mechanisms governing the complex chemical reactions involved in fly ash reactivity, they are generally not well-suited for supporting high-throughput screening approaches. In that regard, a common concern is that those advanced approaches often require accurate quantifications of the crystalline and amorphous phases in fly ash, which makes it practically difficult to implement them at a large scale in real production. In addition, some of the knowledge of the behavior of fly ashes observed under simplified reaction environments may not be transferred to its strength potential in concrete mixtures.

In comparison, it may be more suitable to use the strength activity index (SAI) specified in ASTM C311, *Standard Test Methods for Sampling and Testing Fly Ash or Natural Pozzolans for Use in Portland-Cement Concrete*, and ASTM C618 standards for the fly ash

screening. Admittedly, this approach offers less insight into the chemical reactivity of fly ashes from a fundamental level—for instance, a well-known concern of using SAI to describe the strength potential of fly ash is related to the testing protocols specified by this standard (e.g., 20 % cement replacement by mass and 7 and 28 days of curing time), which may lead to a false-positive indication of the actual reactivity of some inert mineral additives, especially for those with fine particle sizes.[13,14,28] However, the use of SAI has nevertheless remained the most accepted approach for indicating the influence of fly ash on concrete strength in practice.[10,29] Some appealing merits of using SAI are based on the considerations that (i) it offers an intuitive indication of the contribution of a fly ash to the concrete strength, (ii) a huge amount of fly ashes from varying coal sources and production conditions have been systematically evaluated with SAI, which allows us to study the strength potential of fly ashes from the novel perspective of big data, (iii) it is a performance-based criterion, as compared with the prescription-based constraints on material attributes, and (iv) there is no need for the industrial practitioners to adopt either a new standard or a new test method. With that said, a primary obstacle to adopting SAI for high-throughput fly ash screening lies in the time-consuming experimental work needed for lab testing, which is impractical to be applied on a massive scale in daily production.

The new demand for sustainable concrete construction necessitates a more precise and rapid approach to infer the strength activity of fly ashes. To this end, the precursor can be characterizing the material attributes that influence SAI. Because the fly ash's material property alters with the production, processing, and storage conditions,[3,15,30–32] a batch-by-batch characterization is preferred for the proper classification of fly ashes, which requires real-time testing. Across the different characterization methods, X-ray fluorescence (XRF) may be most suitable because of its ease of implementation and the compositional information it provides. Further, using XRF has an innate fit with the established testing standard for fly ash—ASTM C311,[29] wherein XRF is designated for testing the bulk chemical composition of fly ashes. The correlation between the bulk composition and the presence of crystalline and glassy phases in fly ashes has long been recognized since the 1970s.[33–36] The reactivity of a fly ash was found to be primarily governed by that of the amorphous phase (which is analogous to the alkaline earth aluminosilicate glasses), whereas, in contrast, crystalline phases were found to be fairly inert.[2,18] Besides, the reactivity can be further affected by other factors, such as the particle size and the presence of unburnt residues as well as the reaction environment in hydrated cement.[26,27] Previous studies have also investigated the influence of the bulk composition of the major components on the strength activity of fly ashes, such as CaO,[37–41] silicon dioxide (SiO_2),[37–39] aluminum oxide (Al_2O_3),[37,38,44] and ferric oxide (Fe_2O_3).[42] In that regard, network-modifying species (i.e., species like CaO that tend to depolymerize the atomic network of the amorphous phase) generally tend to increase the reaction kinetics.[18,25,43] However, despite the importance of SAI, little effort has been seen in predicting the SAI value directly based on the bulk chemical composition of fly ash.

To promote the optimal use of fly ashes in the design of sustainable concrete materials, it is of special interest to explore the possibility of inferring fly ash's SAI from its fundamental material attributes. In this regard, recent advances in artificial intelligence offer a promising path forward. In particular, machine learning modeling has a unique strength of establishing the prediction without the need for explicit knowledge because the model can be self-driven with the given data for optimizing its prediction.[44] The success of using machine learning for predicting a vast variety of properties of materials has been increasingly reported by recent studies,[45–53] including the determining the reactivity of fly ashes.[54–56] To the best of our knowledge, however, the idea of using machine learning to infer the SAI of fly ash has not been explored by previous research.

This study aims to investigate the use of machine learning for predicting fly ash's strength activity based on the sole knowledge of its bulk material attributes. This study is of special significance for a rapid determination of the SAI of fly ashes. Because this machine learning–based prediction bypasses the need for experimental work, it can greatly improve the efficiency of classifying fly ashes and, therefore, promote the optimal use of fly ashes in concrete production. Note that the strength activity investigated in this study is more of a performance metric of fly ash than a measure of its underlying chemical reactivity. Although the former is fundamentally contingent on the latter, it can be inappropriate to interpret the results the other way around—so cautions should be taken when making inferences to the reactivity of fly ash. To enable a robust machine learning analysis, we first built a large

data set from an industrial database recording the ASTM C618 parameters of a large variety of fly ashes. The curated data set comprises a total of 2,158 samples that correspond to a mix of Class C, F, and N and a small portion of noncompliant fly ashes. The material attributes considered for the SAI prediction include nine chemical contents and one physical property (i.e., fineness). In terms of the machine learning analysis, we developed a neural network model for the SAI prediction. After optimizing this model, its prediction accuracy was evaluated from multiple approaches, and the results indicate that the neural network model is capable of predicting the SAI values with remarkable accuracy. Afterward, we further interrogated the model on how SAI is influenced by each material attribute, whereby novel insights on fly ash's SAI are reversely extracted from the machine learning analysis.

Methodology

THE FLY ASH DATA SET

Data are vital to enable any machine learning analysis. To this end, we curated a data set based on the data provided by a major fly ash manufacturer in the United States. This data set comprises the ASTM C618 test results of 2,158 fly ash samples ranging over a large variety of chemical compositions, conditions, and sources. These samples broadly cover Class C, F, and N fly ashes as well as a small portion of noncompliant ashes. A total of 10 features are adopted in the curated data set for predicting SAI, which are summarized in **Table 1**. In contrast, as the focus of this study is investigating the influence of fly ash material attributes on its strength potential, some performance-based parameters (e.g., water demand and autoclave expansion) and oxide features that are present in low quantities phosphorous oxide (P_2O_5) and titanium dioxide (TiO_2) are not considered in the data set. In addition, the density of fly ash is not accounted for because it is correlated to the composition of the major oxides in **Table 1**.

In terms of the variation of the 10 predictive features within this data set (see **Table 1**), their distributions can represent the compositional variation of the different types of fly ashes used in concrete production. Based on ASTM C618-19,[10] the nine chemical features (mass fractions of the oxides in the bulk material) and one physical feature (fineness, mass fraction of particles retained on the 45-μm sieve) are considered. On average, the nine compositional features considered herein as input sum up to 97.4 % of the total mass of the fly ashes. Based on the chemical requirements of ASTM C618—i.e., the contents of $SiO_2 + Al_2O_3 + Fe_2O_3$, CaO, sulfur trioxide ($SO_3$), moisture, and loss on ignition (LOI)[10]—459 (21.3 %), 1,578 (73.1 %), and 89 (4.1 %) samples in this data set are classified as Class C, F, and N fly ashes, respectively. The remaining 32 (1.5 %) samples are noncompliant as per ASTM C618.

As for the target of the machine learning prediction, the 28-day SAI varies between 74 and 114, wherein the percentage here indicates how the 28-day strength of a mortar with 20 % comprising fly ash compares with that of a reference mortar with no replacement.[10] Among all, about a quarter of the samples exhibit an SAI that is larger than 100. This fraction roughly echoes with the fraction of Class C fly ashes that are presented in this data set.

TABLE 1

Distribution of the ASTM C618 parameters in the fly ash data set involved in this study

Statistics		SiO_2, %	Al_2O_3, %	Fe_2O_3, %	SO_3, %	CaO, %	MgO, %	TotalAlk,[a] %	Moist, %	LOI, %	Fineness, %	28 d SAI
Percentile	0th	24.6	9.2	1.5	0.0	0.8	0.5	0.5	0.0	0.0	3.2	74.0
	25th	42.4	18.1	4.9	0.6	6.0	1.9	1.6	0.1	0.3	16.9	90.0
	50th	51.7	19.3	5.9	0.7	11.3	2.8	2.1	0.1	0.6	20.0	94.0
	75th	56.7	21.2	7.5	1.3	17.7	4.5	2.9	0.1	1.8	23.7	99.0
	100th	66.8	29.5	28.4	18.2	43.2	10.3	9.6	0.6	13.1	47.1	114.0
Mean		49.7	19.6	7.1	1.2	12.9	3.2	2.3	0.1	1.3	20.2	94.6
Standard deviation		8.9	3.0	3.7	1.6	8.0	1.7	1.1	0.1	1.7	5.2	6.6

Note: [a] Total alkali, which equals $Na_2O + 0.658K_2O$.

Although we observe some level of correlation between bulk and amorphous phase compositions, no direct, obvious transformation/correlation rule exists. From the perspective of machine learning, the large range of compositions and SAI values featured in this data set is key to enabling the training of a robust, comprehensive predictive model, i.e., a model that can be applied to a wide range of fly ashes.

MACHINE LEARNING MODELING

To model the SAI of fly ash based on the features prescribed in the section "The Fly Ash Data Set," we conduct the machine learning analysis by using artificial neural networks (hereinafter, neural network). As one of the most frequently used machine learning models, the neural network is known for its efficiency in learning the pattern embedded in the data set (and therefore, good prediction accuracy) and its flexibility in handling different types of data and machine learning tasks (such as the regression task in this study).[57]

In terms of the implementation, we build the neural network model on PyTorch, which is a high-performance machine learning library in Python.[58] The design of the neural network herein generally follows the pipeline established in a previous study.[54] Given the size of the fly ash data set and the number of features and based on our previous machine learning tasks,[53,59] the neural network model is built with two layers of hidden neurons, which contain 10 and 6 artificial neurons, respectively. To ensure the model's learning efficiency, our modeling scheme also includes (i) adding Rectified Linear Unit activation and Batch Normalization between the hidden layers and (ii) preventing the model from overfitting on the noise in the data by using L_2 Norm for weight decay as well as (iii) using a small batch size of 128 for data loading when training the model. In addition, we couple the Adam optimizer in PyTorch with the mean square error loss function for the model optimization.

As for the actual modeling optimization, we use 85 % of the total samples in the fly ash data set to train the neural network while holding the other 15 % for evaluating the prediction accuracy of the trained model (as a blind test on the 15 % never-seen samples). Within the 85 % training set, 20 % of the samples are selected for determining the optimal hyperparameters (i.e., the model parameters that need to be preset before the actual optimization),[57] which is typically known as the validation set. To ensure a fair split of samples between the 85 % training set and the 15 % test set (as well as the selection of the validation set) such that the feature distribution is not biased in either set, we implement stratification sampling for dividing the data set, using technical details that have been adopted in several studies.[45,54,63] To obtain the best accuracy for predicting fly ash SAI,[60] we carry out a grid search on the training set samples to determine the optimal hyperparameters of the neural network model, which include learning rate and weight decay. To this end, the searching of these two hyperparameters is set within their common ranges (10^{-1} to 10^{-4}).[61] Based on the results averaged from 10 stratified train-test splits, the optimal learning rate and weight decay are determined as 1×10^{-3} and 7×10^{-3}, respectively.

For the grid search of optimal hyperparameters, the prediction accuracy is primarily assessed based on the coefficient of determination (R^2), where the result is 1 for a perfect prediction without any error) of the validation set samples, whereas root mean square error (RMSE) and mean absolute percentage error (MAPE) are also calculated to support the model evaluation.[44] It should be noted that neural networks do not require a specific configuration of their network structure, settings, and hyperparameters to yield a model of good accuracy, so the previously stated modeling scheme and parameters only serve as a reference and other configurations may also be plausible.

SHAPLEY ADDITIVE EXPLANATIONS ANALYSIS FOR MODEL INTERPRETATION

In comparison with the prediction accuracy, the interpretability of a machine learning model is arguably equally or even more important for material research because it opens up the possibility of "reverse learning" from artificial intelligence (and a meaningful knowledge extraction thereof). Recent advances in the field of artificial intelligence have seen a series of methods for inferring the true relationship as learned by a machine learning model, such as through specific schemes of data sampling and shuffling.[62,63] Herein, after obtaining the optimal

neural network model for predicting SAI, SHapley Additive exPlanations (SHAP) analysis is implemented for interpreting the trained model.[61]

In simple terms, SHAP analysis is a statistical approach that assesses the marginal contributions of the input features (e.g., material attributes of a fly ash sample) to an outcome that is resulted from their combined effects (e.g., fly ash's SAI value), in which the marginal contribution is presented as Shapley values.[64] Note that the Shapley value can be either positive or negative, corresponding to the direction of the marginal contribution —this is, either increasing or decreasing the value of the outcome. In a previous study, SHAP analysis has been demonstrated as a powerful method to understand the role of major oxides in fly ash to the rigidity of its atomic network.[54] In this study, SHAP analysis is used to reveal the data patterns from the trained neural network model regarding (i) how SAI is jointly determined by the 10 features considered for modeling, (ii) the importance of each feature on the model prediction, and (iii) the isolated effect of each feature on SAI. Further details about the implementation of SHAP analysis are provided in the section "Model Interpretations."

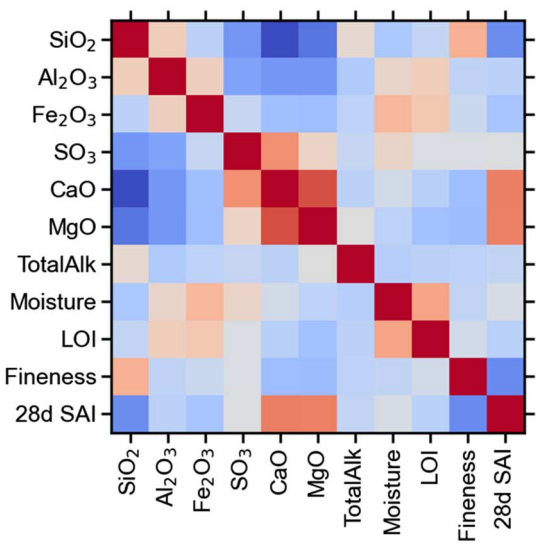
Results

CORRELATIONS BETWEEN THE FEATURES AND SAI

To obtain an overall evaluation of the nature of the fly ash data set, we first determine the correlation between each pair of inputs and output based on a direct linear fitting. The correlations are normalized and presented as a correlation heat map in **figure 1**, where the highest level of correlation is displayed in dark red color. Here, because each entry is fully correlated to itself, all correlations along the −45° diagonal line are equal to the highest value, 1. We find that CaO and magnesium oxide (MgO) both exhibit a strong positive correlation with the 28-day SAI. This echoes the fact that the content of CaO is typically used for discriminating Class C from F fly ashes[10] and that these network-modifying oxides tend to increase the reactivity of fly ashes,[18,25,43] which, in turn, further improve their strength activity. For this pair of oxides, the high correlation between CaO and MgO also indicates that these oxides tend to accompany each other. In contrast, SiO_2 and fineness both exhibit a negative correlation with the SAI. The negative correlation between SiO_2 and SAI is in agreement with the fact that silicon atoms act as network formers in glasses and tend to increase the topological rigidity of the glass atomic network within fly ashes, thereby making fly ashes less chemically reactive and thus reducing their strength activity.[37,44] In turn, the negative correlation between fineness and SAI arises from the fact that chemical reactivity is typically proportional to

FIG. 1

Correlation matrix between all the input features and the 28-day SAI (28d SAI) output for the samples in the fly ash data set. The correlation of each pair is color-coded from dark red (positive) to dark blue (negative), based on linear fitting. Note that TotalAlk equals $Na_2O + 0.658K_2O$.



the specific surface area of fly ashes. As for the other features, the correlation analysis from figure 1 also suggests that Fe_2O_3, Al_2O_3, Total Alkali (TotalAlk), and LOI all have slightly negative influences on the 28-day SAI (following a roughly decreasing order). No obvious effect can be observed from this analysis for SO_3 and moisture.

It is also worth noting that a negative correlation exists between CaO and SiO_2 (also Al_2O_3), which indicates that the influence of the major oxides in fly ash on SAI is further intricated by the fact that increasing the content of one oxide unavoidably leads to content changes of the other oxides. This issue highlights a major challenge of isolating the sole effect of a feature on SAI. This point will be further addressed in the section "Model Interpretations."

To gain a more intuitive view of the correlations in the fly ash data set, we further inspect the correlation between each input feature and the 28-day SAI, as displayed in figure 2. Note that the plots in figure 2 are color-coded based on the density of the scatters, and the feature ranges are truncated to focus on the most populated region. In general, a rough agreement can be established between the results shown in figures 1 and 2 in which the positive or negative feature correlations to 28-day SAI in figure 1 can be perceived from figure 2. An interesting finding from figure 2 is that the data distributions of some features appear to be fairly heterogeneous or discrete. Upon inspecting the notable clusters, we find it unintuitive to identify any first-order correlations for the majority of the noticeable clusters in the feature-SAI plots, whereas a few identifiable correlations are marked as follows: (i) the samples featuring high CaO content (>18 %) but very low SAI values (see fig. 2*E*) have very high SO_3 contents (of around 10 %); (ii) MgO has an isolated cluster at low SAI values (see fig. 2*F*) wherein those samples systematically exhibit low Al_2O_3 (around 50 %) but high SiO_2 contents (around 50 %); (iii) in the case of TotalAlk, a clear bifurcation is observed between 1.5 and 2.5 % (see fig. 2*G*), wherein the lower cluster can be correlated with very low SO_3 contents (merely about 0.5 %); (iv) fineness shows a diverged trend between 15 and 25 % (see fig. 2*J*), and, in comparison, the samples in the low-SAI cluster have low CaO contents (below 7 %). The presence of some of these clusters might be due to the different coal sources and processing protocols in fly ash production.[3,38,65]

Furthermore, it is important to note that the exact feature correlations in figure 2 are not clear enough for establishing a direct mapping from each of the input features (or simple feature combinations) to the 28-day SAI. For example, when the CaO content equals 15 %, the 28-day SAI can vary significantly between 85 (i.e., largely inactive) and 103 (i.e., active). Similarly, such a wide variation is also seen in the case of fineness because the SAI exhibits a great variability at fixed fineness values. From the standpoint of maximizing the strength efficiency of fly ashes, the aforementioned observation indeed illustrates the difficulty of predicting fly ash's SAI based on the sole knowledge of the basic material attributes.

MODEL PREDICTION OF THE 28-DAY SAI

Based on the pipeline established in the section "Machine Learning Modeling," we obtained an optimized neural network model for predicting the 28-day SAI with the sole inputs of the 10 input features as prescribed by ASTM C618. The prediction performance of this model on the 85 % training set and the 15 % test set (this is, the holdout samples that are used to assess the true model accuracy when generalizing to new samples) is compared in figure 3, where the scatters are color-coded to indicate the density of the distribution. It can be seen from figure 3*A* that the majority of the training samples are distributed along the line of equality between the true SAI values of 85 and 100, suggesting a good prediction accuracy of the neural network model for fly ash samples of relatively low strength activity (e.g., Class F fly ashes). At the less populated region of an even lower SAI value (<85), the model shows a tendency to overpredict SAI. This is likely due to the fact that there is only a small portion of samples with such a low SAI value (see Table 1), so the model cannot get sufficiently trained to offer an accurate prediction for that region, as also reported in several other machine learning–related studies.[45,53,63,66] For the strength reactive fly ashes with an SAI higher than 100, although the model predictions appear to be slightly tilted clockwise, the SAI values of most samples are still accurately estimated. Note that the model does not predict any SAI value higher than about 107 (see fig. 3*A*). This is likely a consequence of the very limited number of fly

FIG. 2

Correlations between the individual input features and the 28-day SAI: (*A*) SiO_2, (*B*) Al_2O_3, (*C*) Fe_2O_3, (*D*) SO_3, (*E*) CaO, (*F*) MgO, (*G*) Total Alkali, (*H*) Moisture, (*I*) Loss on Ignition, and (*J*) Fineness. In these plots, the local density of the scatters increases from blue to red. The circled regions in several of these plots are data clusters identified to correlate with the presence of other features (see text).

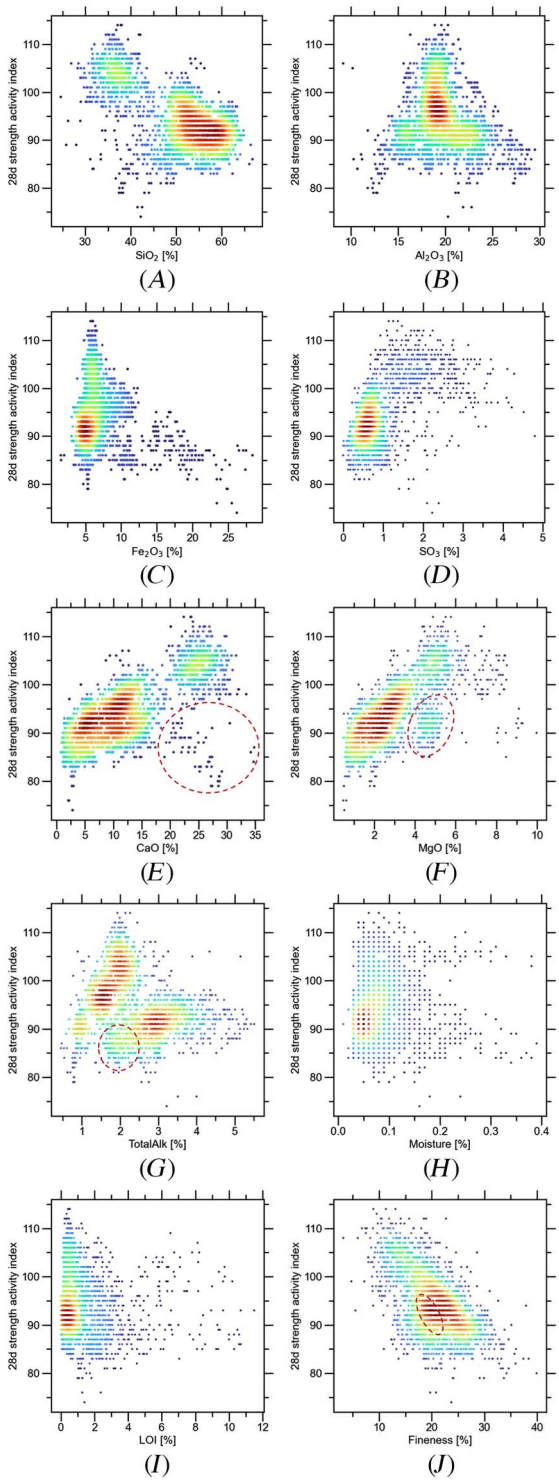
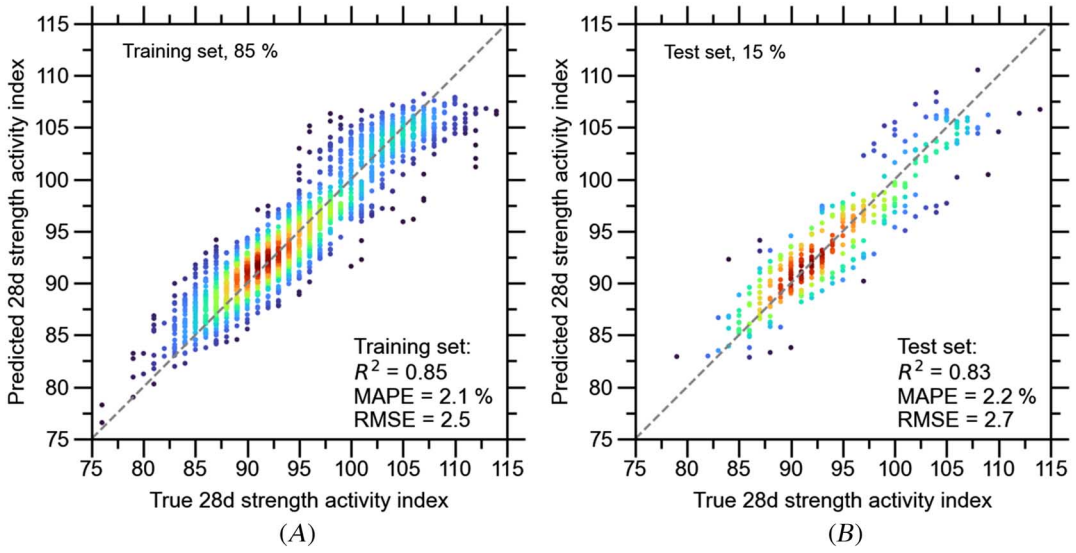


FIG. 3 Predicted vs. actual 28-day SAI of the samples of the (*A*) training set (i.e., for training the neural network model) and (*B*) test set (i.e., for evaluating the ability of the trained model to generalize on new samples). The dashed line in these plots indicates perfect agreement.



ash samples showing very high SAI values so that the model is unable to learn how to properly predict SAI in this regime.

Compared with the training set, the model achieves fairly closed prediction performance on the test set samples. This observation is suggested by the similar overall distribution of the data points in **figure 3*B*** and is further supported by the comparable accuracy metrics (i.e., R^2, RMSE, and MAPE) between the two sets. Taking MAPE of the test set as an example (see **fig. 3*B***), the model has an average prediction error of merely 2.2 %, which is rather impressive. Overall, this evaluation demonstrates that the neural network model does not show a sign of either underfit (i.e., cannot predict the true SAI well enough) or overfit (i.e., cannot achieve the same performance on the test set). More importantly, because the test set samples are never involved in the model training process, the results indicate that the neural network model has a strong ability to offer robust SAI prediction for new fly ash samples. Altogether, the various pieces of evidence presented in this section support that our model can offer a robust prediction of fly ash's SAI, and this model is especially applicable for predicting fly ashes within an SAI range from approximately 85 to 105—which is well representative for fly ashes that are compliant to ASTM C618.

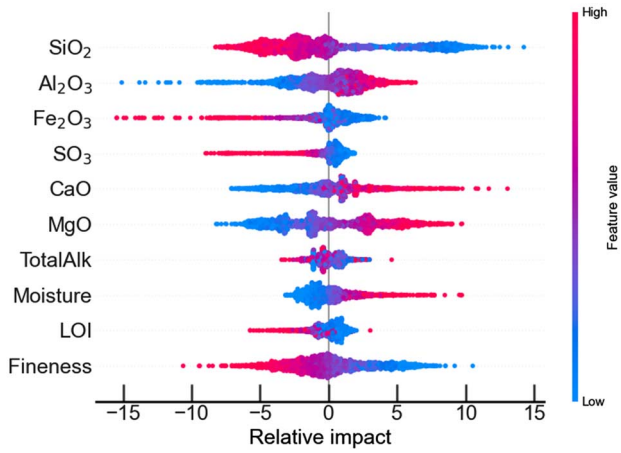
MODEL INTERPRETATIONS

To further investigate the data patterns that are learned by the machine learning model, we then proceed to interpret the optimized neural network model presented in the last section. Machine learning models are conventionally well known for their abilities to make accurate predictions yet lack transparency for people to rationalize how the input features are mapped to the output prediction (i.e., a black-box model), with neural networks as a typical example. Thus, the challenge of interpretation constitutes a major hindrance to the broader adoption of machine learning analyses in material research, which needs to be established on explicit evidence and knowledge.

Herein, we adopted the SHAP to decode the relationship between the 10 input features and fly ash's 28-day SAI that is hidden in the latent space of our neural network model.[61] Through computing and comparing the Shapley value of the complex interactions between the features involved in the model predictions, the SHAP analysis aims at assessing the marginal contribution of each feature to the prediction of a machine learning model.

FIG. 4

Summary of the SHAP analysis comparing the relative impact (i.e., marginal contribution) of each input feature to the 28-day SAI, based on all 2,158 samples in the fly ash data set. For ease of understanding, the relative impact is aligned to the same scale of SAI.



In short, a positive Shapley value correlates to a positive marginal contribution that increases the value of the model prediction from its average over the entire data set (stated another way, a positive relative impact on the model prediction) and vice versa.[61,64] Further details about the implementation of SHAP are given in the section "SHapley Additive exPlanations Analysis for Model Interpretation."

Here, we first compared the Shapley values (i.e., relative impact to 28-day SAI) for all 10 input features, as summarized in **figure 4**. In this figure, the scatters on each row correspond to the distribution of the relative impact from a feature over all the samples in the fly ash data set, in which the relative impact gradually shifts from negative to positive from left to right. For comparison, the values of each feature are normalized to enable the scatters of all features to be displayed on the same color scale, from minimum (in blue) to maximum (in pink); in addition, the variation of the local thickness in each cluster corresponds to the distribution of the feature values between the extremities. In general, the horizontal span of the clusters of each feature corresponds to its range of impact, meaning that impactful features are expected to exhibit wider spreads. In that regard, the oxides that are most influential to 28-day SAI include, in roughly descending order, SiO_2, CaO (as well as MgO), Al_2O_3, and Fe_2O_3, which cover the main composition of fly ashes. For the instance of SiO_2, this analysis suggests that the 28-day SAI systematically decreases as the SiO_2 content becomes higher in fly ashes (moving from the blue to the pink end), whereas only a small portion of the samples in this data set are seen with relatively low SiO_2 contents (also seen in **Table 1**). In comparison, this analysis indicates that a higher content of CaO generally indicates a larger 28-day SAI.

A somewhat surprising finding from this SHAP summary plot is that the physical attribute, fineness, appears to be more influential than any single oxide content, where a clear and consistent increase of the relative impact on SAI is observed as fineness becomes progressively smaller (i.e., fly ash particle size decreases). This implies that, regardless of the chemical composition of fly ash, reducing the particle size can be a generic yet effective approach to improve its strength activity. This point is consistent with the long-established understanding of the effect of fly ash's fineness on concrete strength, wherein finer fly ash particles are expected to be more reactive because of their higher specific surface area available for the pozzolanic reaction.[40,67–69] In that regard, mechanical treatments such as vibratory milling have been proven to be an effective approach to enhancing the reactivity of fly ash,[70] although reducing the fineness also tends to increase the material processing cost and water demand for the concrete bleeding process.[71,72]

To gain a more detailed model interpretation, we then narrow the analysis down to the marginal contribution belonging to each feature. This is fulfilled by plotting the actual feature value (not normalized) against its

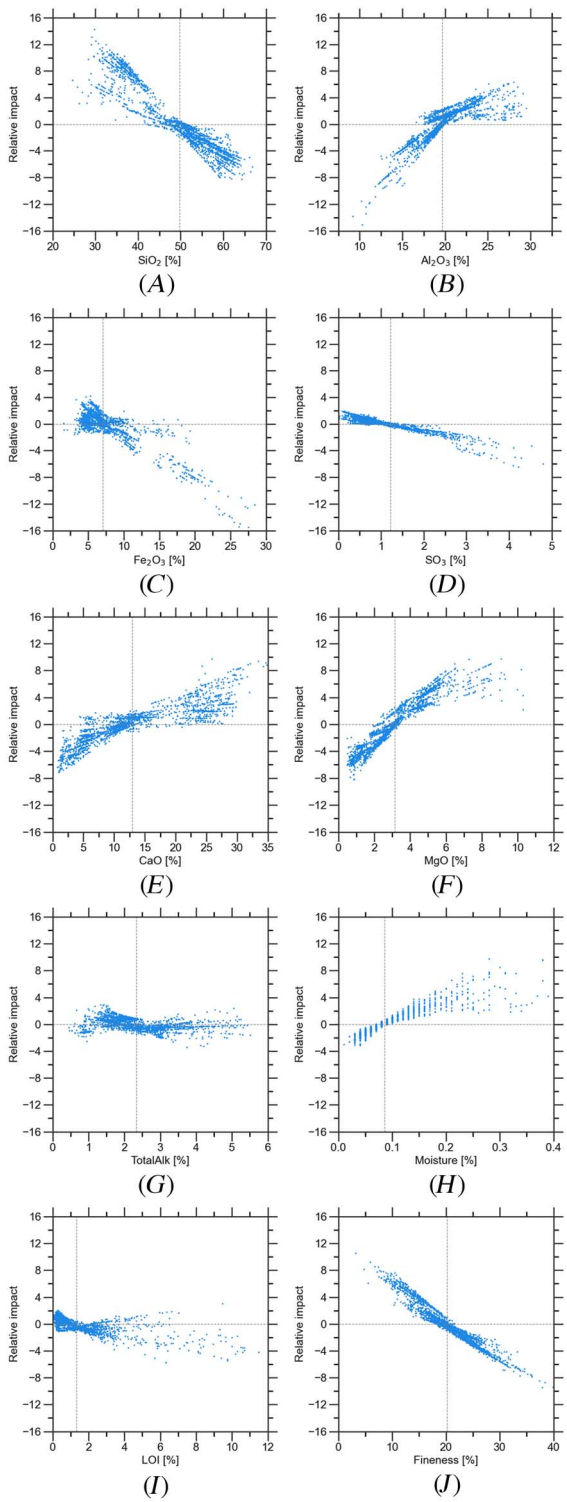
Shapley values corresponding to each sample in the fly ash data set. The feature-specific SHAP results are displayed in **figure 5**. Unlike the direct feature-property correlation that treats each feature independently (see **fig. 2**), the SHAP plots in **figure 5** present the influence of each feature, which considers the synergistic effect between the feature of interest and all the other features. This is valuable because a standalone feature may be trivial by itself, but it becomes much more influential when combined with some other features—which is a more reasonable indication of its true influence.

Compared with the feature correlations in **figure 2**, we first observe that the SHAP analysis from **figure 5** provides more consistent feature trends, which implies a successful extraction of the feature's marginal contribution. In terms of the features that exhibit notable effects from **figure 2**, their SHAP results generally suggest the same trends, as judged from the slope of the scatters. For example, increasing fineness or the content of SiO_2 or Fe_2O_3 tends to reduce the 28-day SAI of fly ashes, but larger SAI values can be expected if the CaO or MgO content becomes larger, which is in line with the established studies.[37–41] On the other hand, the effects of several weaker features become more phenomenal based on SHAP. Taking LOI as an example, **figure 5*I*** indicates that this feature should have a slightly negative influence on SAI, which is expected because LOI is mostly attributed to the fact that unburnt impurities tend to slow down the reaction of fly ash and increase the water demand.[73] Similarly, a clearer influence is also observed on moisture. In contrast, the machine learning model does not learn a clear effect of TotalAlk, where its marginal contribution seems to be barely changed with the content variation.

The largest difference is found on SO_3, where it shows a positive correlation in **figure 2*D*** yet a negative effect is obtained based on SHAP in **figure 5*D***. Although the effect of SO_3 on SAI is less prominent as compared with the other major oxides, previous studies have reported mixed observations on its influence on the strength of concrete containing fly ashes. It is generally expected that SO_3 imposes setting retardation on concrete mixtures and a high SO_3 content typically weakens the hardened cement paste because of the increased risks of sulfate attack and expansion.[74,75] However, it has also been observed that the final strength of a self-cementitious fly ash hydration system can be improved upon increasing SO_3 content.[76] In practice, ASTM C618 restricts a maximum SO_3 content of 5 % in fly ashes,[10] whereas the same prescription is also given for portland cement by ASTM C150, *Specification for Portland Cement*.[77] Thus, it should be reasonable to expect a decline of SAI at an elevated level of SO_3 content in fly ash. In that regard, the direct correlation misleadingly suggests a contrary trend; however, the true effect of SO_3 is nevertheless revealed by the model interpretation. As for the misleading trend suggested by the direct correlation, it may be attributed to the fact that SO_3 has a positive correlation with CaO (see the correlation map in **fig. 1**). Consequently, the true effects of SO_3 are obscured by its correlation with CaO. This issue represents a typical challenge involved in data interpretation. In this sense, it is encouraging that the right pattern is captured by the machine learning model and can be further interpreted into meaningful insights that can advance our knowledge of fly ash materials from a novel perspective.

The aforementioned observation may also be applicable to Al_2O_3, where the direct correlation suggests it has an optimal content at around 19 % (see **fig. 2*B***) while the machine learning model captures that the 28-day SAI increases with Al_2O_3 monotonically. The latter agrees well with our previous observation that the presence of Al_2O_3 in fly ash tends to increase its strength activity by reducing the rigidity of the atomic network.[54] Another interesting observation for Al_2O_3 is that this oxide exhibits two distinct slopes at low and high contents, where **figure 5*B*** shows that Al_2O_3 is more influential on SAI when its content is less than about 20 %. This could be related to the different roles played by Al_2O_3 in modifying the atomic structure when comparing aluminum-poor peralkaline silicate glasses and aluminum-rich peraluminous silicate glasses materials (both of which apply to fly ashes).[37,38,44,78] Although this finding needs to be further verified with experiments, it provides intriguing insights regarding the real influence of Al_2O_3 on SAI. Further investigation on this point may help develop new activation strategies of fly ashes, such as blending fly ashes with different Al_2O_3 contents.

FIG. 5



Feature-based SHAP analysis showing the relative impact (i.e., marginal contribution) of each input feature to the 28-day SAI, where the feature values are based on the actual scale: (*A*) SiO_2, (*B*) Al_2O_3, (*C*) Fe_2O_3, (*D*) SO_3, (*E*) CaO, (*F*) MgO, (*G*) Total Alkali, (*H*) Moisture, (*I*) Loss on Ignition, and (*J*) Fineness. To assist the comparison, the vertical and horizontal dash lines indicate the mean value of each feature and the zero marginal contribution, respectively. For ease of understanding, the relative impact is aligned to the same scale of SAI.

Conclusions

In this work, we investigate a large fly ash data set with machine learning techniques. To this end, we build a neural network model to predict the fly ash's 28-day SAI based on the sole input of 10 material attributes as specified in ASTM C618. Further, we implement SHAP analysis to extract insights from the "black-box" neural network model. The major findings of this study are remarked as follows:

- Fly ash's SAI cannot be accurately inferred based on its direct correlation with any single ASTM C618 attribute. Although the CaO content is found to be the most positively correlated to 28-day SAI, the conventional rule of using CaO content (>18 %) to differentiate the strength active fly ash (i.e., SAI >100) does not hold for many samples in the fly ash data set. In fact, a considerable portion of strength active samples exhibit low CaO contents down to 10 %.
- Our optimized neural network model achieves remarkable accuracy in predicting the 28-day SAI. This is demonstrated from its predictions for the test set samples (i.e., the holdout samples that are never involved in the model training), wherein the average prediction error is merely 2.2 %. To the best of our knowledge, this is the first time that a machine learning model is reported to be capable of predicting SAI.
- The neural network model is successfully decoded for interpreting the influence of each feature on SAI as captured by machine learning. A broad agreement is observed between machine learning and established studies on the major features (e.g., CaO, SiO_2, and fineness). Further, the effects of the minor features are clarified.

Overall, it is rather encouraging to find that machine learning can achieve high accuracy for predicting SAI and further offer meaningful insights on how SAI is synergistically determined by the individual material attributes. From a practical perspective, this approach provides a superior way for determining the strength activity of fly ashes, which can significantly promote the high-throughput screening of fly ash for fulfilling the goal of sustainable concrete construction. More importantly, this study offers a paradigm of knowledge extraction from artificial intelligence for the broader adoption of machine learning techniques in material research.

ACKNOWLEDGMENTS

The authors acknowledge the financial support for this research by the US Department of Transportation through the Federal Highway Administration (Grant #693JJ31950021) and the US National Science Foundation (DMREF: 1922167).

References

1. K. L. Aughenbaugh, T. Williamson, and M. C. G. Juenger, "Critical Evaluation of Strength Prediction Methods for Alkali-Activated Fly Ash," *Materials and Structures* 48, no. 3 (March 2015): 607–620, https://doi.org/10.1617/s11527-014-0496-z
2. T. Oey, J. Timmons, P. Stutzman, J. W. Bullard, M. Balonis, M. Bauchy, and G. Sant, "An Improved Basis for Characterizing the Suitability of Fly Ash as a Cement Replacement Agent," *Journal of the American Ceramic Society* 100, no. 10 (October 2017): 4785–4800, https://doi.org/10.1111/jace.14974
3. G. Xu and X. Shi, "Characteristics and Applications of Fly Ash as a Sustainable Construction Material: A State-of-the-Art Review," *Resources, Conservation and Recycling* 136 (September 2018): 95–109, https://doi.org/10.1016/j.resconrec.2018.04.010
4. A. Hasanbeigi, L. Price, and E. Lin, "Emerging Energy-Efficiency and CO_2 Emission-Reduction Technologies for Cement and Concrete Production: A Technical Review," *Renewable and Sustainable Energy Reviews* 16, no. 8 (October 2012): 6220–6238, https://doi.org/10.1016/j.rser.2012.07.019
5. J. M. Paris, J. G. Roessler, C. C. Ferraro, H. D. DeFord, and T. G. Townsend, "A Review of Waste Products Utilized as Supplements to Portland Cement in Concrete," *Journal of Cleaner Production* 121 (May 2016): 1–18, https://doi.org/10.1016/j.jclepro.2016.02.013
6. C. R. Shearer, "The Productive Reuse of Coal, Biomass and Co-fired Fly Ash" (PhD diss., Georgia Institute of Technology, 2014), http://hdl.handle.net/1853/52298
7. A. J. Saraber, "Fly Ash from Coal and Biomass for Use in Concrete: Origin, Properties and Performance" (PhD diss., Delft University of Technology, 2017), https://doi.org/10.4233/uuid:2113768d-6b00-4b13-91f0-d338dac9e47a

8. S. Tsimas and A. Moutsatsou-Tsima, "High-Calcium Fly Ash as the Fourth Constituent in Concrete: Problems, Solutions and Perspectives," *Cement and Concrete Composites* 27, no. 2 (February 2005): 231–237, https://doi.org/10.1016/j.cemconcomp.2004.02.012
9. T. Hemalatha and A. Ramaswamy, "A Review on Fly Ash Characteristics – Towards Promoting High Volume Utilization in Developing Sustainable Concrete," *Journal of Cleaner Production* 147 (March 2017): 546–559, https://doi.org/10.1016/j.jclepro.2017.01.114
10. *Specification for Coal Fly Ash and Raw or Calcined Natural Pozzolan for Use in Concrete*, ASTM C618-19 (West Conshohocken, PA: ASTM International, approved January 1, 2019), https://doi.org/10.1520/C0618-19
11. A. R. Pourkhorshidi, M. Najimi, T. Parhizkar, F. Jafarpour, and B. Hillemeier, "Applicability of the Standard Specifications of ASTM C618 for Evaluation of Natural Pozzolans," *Cement and Concrete Composites* 32, no. 10 (November 2010): 794–800, https://doi.org/10.1016/j.cemconcomp.2010.08.007
12. G. P. Gava and L. R. Prudêncio, "Pozzolanic Activity Tests as a Measure of Pozzolans' Performance. Part 1," *Magazine of Concrete Research* 59, no. 10 (December 2007): 729–734, https://doi.org/10.1680/macr.2007.59.10.729
13. D. P. Bentz, A. Durán-Herrera, and D. Galvez-Moreno, "Comparison of ASTM C311 Strength Activity Index Testing Versus Testing Based on Constant Volumetric Proportions," *Journal of ASTM International* 9, no. 1 (2011): 1–7, https://doi.org/10.1520/JAI104138
14. P. Suraneni, L. Burris, C. R. Shearer, and R. D. Hooton, "ASTM C618 Fly Ash Specification: Comparison with Other Specifications, Shortcomings, and Solutions," *ACI Materials Journal* 118, no. 1 (2021): 157–167, https://doi.org/10.14359/51725994
15. *Coal Combustion Product (CCP) Production & Use Survey Report* (Denver, CO: American Coal Ash Association, 2018).
16. M. Schneider, M. Romer, M. Tschudin, and H. Bolio, "Sustainable Cement Production—Present and Future," *Cement and Concrete Research* 41, no. 7 (July 2011): 642–650, https://doi.org/10.1016/j.cemconres.2011.03.019
17. E. Benhelal, G. Zahedi, E. Shamsaei, and A. Bahadori, "Global Strategies and Potentials to Curb CO_2 Emissions in Cement Industry," *Journal of Cleaner Production* 51 (July 2013): 142–161, https://doi.org/10.1016/j.jclepro.2012.10.049
18. T. Oey, A. Kumar, I. Pignatelli, Y. Yu, N. Neithalath, J. W. Bullard, M. Bauchy, and G. Sant, "Topological Controls on the Dissolution Kinetics of Glassy Aluminosilicates," *Journal of the American Ceramic Society* 100, no. 12 (December 2017): 5521–5527, https://doi.org/10.1111/jace.15122
19. W. Bumrongjaroen, I. Muller, J. Schweitzer, and R. A. Lvingston, "Application of Glass Corrosion Tests to the Reactivity of Fly Ash" (paper presentation, 2007 World of Coal Ash [WOCA], Covington, KY, May 7–10, 2007).
20. K. L. Lin, K. S. Wang, T. Y. Lee, and B. Y Tzeng, "The Hydration Characteristics of MSWI Fly Ash Slag Present in C3S," *Cement and Concrete Research* 33, no. 7 (July 2003): 957–964, https://doi.org/10.1016/S0008-8846(02)01002-5
21. P. T. Durdziński, C. F. Dunant, M. B. Haha, and K. L. Scrivener, "A New Quantification Method Based on SEM-EDS to Assess Fly Ash Composition and Study the Reaction of Its Individual Components in Hydrating Cement Paste," *Cement and Concrete Research* 73 (July 2015): 111–122, https://doi.org/10.1016/j.cemconres.2015.02.008
22. D. Glosser, P. Suraneni, O. B. Isgor, and W. J. Weiss, "Estimating Reaction Kinetics of Cementitious Pastes Containing Fly Ash," *Cement and Concrete Composites* 112 (September 2020): 103655, https://doi.org/10.1016/j.cemconcomp.2020.103655
23. K. Bharadwaj, R. M. Ghantous, F. Sahan, O. B. Isgor, and W. J. Weiss, "Predicting Pore Volume, Compressive Strength, Pore Connectivity, and Formation Factor in Cementitious Pastes Containing Fly Ash," *Cement and Concrete Composites* 122 (September 2021): 104113, https://doi.org/10.1016/j.cemconcomp.2021.104113
24. B. Pichler and C. Hellmich, "Upscaling Quasi-brittle Strength of Cement Paste and Mortar: A Multi-scale Engineering Mechanics Model," *Cement and Concrete Research* 41, no. 5 (May 2011): 467–476, https://doi.org/10.1016/j.cemconres.2011.01.010
25. M. Bauchy, "Deciphering the Atomic Genome of Glasses by Topological Constraint Theory and Molecular Dynamics: A Review," *Computational Materials Science* 159 (March 2019): 95–102, https://doi.org/10.1016/j.commatsci.2018.12.004
26. Z. Li, G. Xu, and X. Shi, "Reactivity of Coal Fly Ash Used in Cementitious Binder Systems: A State-of-the-Art Overview," *Fuel* 301 (October 2021): 121031, https://doi.org/10.1016/j.fuel.2021.121031
27. J. Skibsted and R. Snellings, "Reactivity of Supplementary Cementitious Materials (SCMs) in Cement Blends," *Cement and Concrete Research* 124 (October 2019): 105799, https://doi.org/10.1016/j.cemconres.2019.105799
28. R. D. Kalina, S. Al-Shmaisani, R. D. Ferron, and M. C. G. Juenger, "False Positives in ASTM C618 Specifications for Natural Pozzolans," *ACI Materials Journal* 116, no. 1 (2019): 165–172, https://doi.org/10.14359/51712243
29. *Standard Test Methods for Sampling and Testing Fly Ash or Natural Pozzolans for Use in Portland-Cement Concrete* (Superseded), ASTM C311/C311M-16 (West Conshohocken, PA: ASTM International, approved December 15, 2016), https://doi.org/10.1520/C0311_C0311M-16
30. W. Zhang, A. Noble, X. Yang, and R. Honaker, "A Comprehensive Review of Rare Earth Elements Recovery from Coal-Related Materials," *Minerals* 10, no. 5 (May 2020): 451, https://doi.org/10.3390/min10050451
31. W. Franus, M. M. Wiatros-Motyka, and M. Wdowin, "Coal Fly Ash as a Resource for Rare Earth Elements," *Environmental Science and Pollution Research* 22, no. 12 (June 2015): 9464–9474, https://doi.org/10.1007/s11356-015-4111-9
32. P. Liu, R. Huang, and Y. Tang, "Comprehensive Understandings of Rare Earth Element (REE) Speciation in Coal Fly Ashes and Implication for REE Extractability," *Environmental Science & Technology* 53, no. 9 (May 2019): 5369–5377, https://doi.org/10.1021/acs.est.9b00005

33. D. M. Roy, K. Luke, and S. Diamond, "Characterization of Fly Ash and Its Reactions in Concrete," *MRS Online Proceedings Library (OPL)* 43, no. 1 (December 1984): 3–20, https://doi.org/10.1557/PROC-43-3
34. P. K. Mehta, "Influence of Fly Ash Characteristics on the Strength of Portland-Fly Ash Mixtures," *Cement and Concrete Research* 15, no. 4 (July 1985): 669–674, https://doi.org/10.1016/0008-8846(85)90067-5
35. G. J. McCarthy, F. P. Glasser, D. M. Roy, and S. Diamond, eds., *Fly Ash and Coal Conversion By-Products: Characterization, Utilization and Disposal III* (Pittsburgh, PA: Materials Research Society, 1987).
36. R. T. Hemmings, E. E. Berry, G. J. McCarthy, and F. P. Glaser, eds., *Fly Ash and Coal Conversion By-Products: Characterization, Utilization and Disposal V* (Pittsburgh, PA: Materials Research Society, 1989).
37. M. Bauchy, "Structural, Vibrational, and Elastic Properties of a Calcium Aluminosilicate Glass from Molecular Dynamics Simulations: The Role of the Potential," *The Journal of Chemical Physics* 141, no. 2 (July 2014): 024507, https://doi.org/10.1063/1.4886421
38. K. L. Aughenbaugh, P. Stutzman, and M. C. G. Juenger, "Identifying Glass Compositions in Fly Ash," *Frontiers in Materials* 3 (2016): 1, https://doi.org/10.3389/fmats.2016.00001
39. R. Szabo and G. Mucsi, "Effect of SiO_2, Al_2O_3 and Na_2O Content and Fly Ash Fineness on the Structure and Mechanical Properties of Fly Ash Based Geopolymer," *Reciklaza i odrzivi razvoj* 12 (2019): 61–68.
40. Š. Slanička, "The Influence of Fly Ash Fineness on the Strength of Concrete," *Cement and Concrete Research* 21, nos. 2–3 (March–May 1991): 285–296, https://doi.org/10.1016/0008-8846(91)90010-F
41. R. T. Hemmings and E. E. Berry, "On the Glass in Coal Fly Ashes: Recent Advances," *MRS Online Proceedings Library* 113, no. 1 (December 1987): 3–38, https://doi.org/10.1557/PROC-113-3
42. E. Gerasimova, "The Effect of Fe_2O_3 on the Mechanical Properties of the Polymer Modified Cement Containing Fly Ash," *Procedia Engineering* 150 (2016): 1553–1557, https://doi.org/10.1016/j.proeng.2016.07.110
43. I. Pignatelli, A. Kumar, M. Bauchy, and G. Sant, "Topological Control on Silicates' Dissolution Kinetics," *Langmuir* 32, no. 18 (May 2016): 4434–4439, https://doi.org/10.1021/acs.langmuir.6b00359
44. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, et al., "Scikit-Learn: Machine Learning in Python," *Journal of Machine Learning Research* 12 (2011): 2825–2830.
45. K. M. Jablonka, D. Ongari, S. M. Moosavi, and B. Smit, "Big-Data Science in Porous Materials: Materials Genomics and Machine Learning," *Chemical Reviews* 120, no. 16 (August 2020): 8066–8129, https://doi.org/10.1021/acs.chemrev.0c00004
46. K. Guo, Z. Yang, C.-H. Yu, and M. J. Buehler, "Artificial Intelligence and Machine Learning in Design of Mechanical Materials," *Materials Horizons* 8 (2021): 1153–1172, https://doi.org/10.1039/D0MH01451F
47. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, "Machine Learning in Materials Informatics: Recent Applications and Prospects," *npj Computational Materials* 3, no. 1 (2017): 54, https://doi.org/10.1038/s41524-017-0056-5
48. H. Liu, Z. Fu, K. Yang, X. Xu, and M. Bauchy, "Machine Learning for Glass Science and Engineering: A Review," *Journal of Non-crystalline Solids* 557 (April 2021): 119419, https://doi.org/10.1016/j.jnoncrysol.2019.04.039
49. Y. Song, Z. Huang, C. Shen, H. Shi, and D. A. Lange, "Deep Learning-Based Automated Image Segmentation for Concrete Petrographic Analysis," *Cement and Concrete Research* 135 (September 2020): 106118, https://doi.org/10.1016/j.cemconres.2020.106118
50. B. A. Young, A. Hall, L. Pilon, P. Gupta, and G. Sant, "Can the Compressive Strength of Concrete Be Estimated from Knowledge of the Mixture Proportions?: New Insights from Statistical Analysis and Machine Learning Methods," *Cement and Concrete Research* 115 (January 2019): 379–388.
51. N. M. Anoop Krishnan, S. Mangalathu, M. M. Smedskjaer, A. Tandia, H. Burton, and M. Bauchy, "Predicting the Dissolution Kinetics of Silicate Glasses Using Machine Learning," *Journal of Non-crystalline Solids* 487 (May 2018): 37–45.
52. J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent Advances and Applications of Machine Learning in Solid-State Materials Science," *npj Computational Materials* 5, no. 1 (2019): 83, https://doi.org/10.1038/s41524-019-0221-0
53. B. Ouyang, Y. Song, Y. Li, F. Wu, H. Yu, Y. Wang, G. Sant, and M. Bauchy, "Predicting Concrete's Strength by Machine Learning: Balance between Accuracy and Complexity of Algorithms," *ACI Materials Journal* 117, no. 6 (2020): 125–133, https://doi.org/10.14359/51728128
54. Y. Song, K. Yang, J. Chen, K. Wang, G. Sant, and M. Bauchy, "Machine Learning Enables Rapid Screening of Reactive Fly Ashes Based on Their Network Topology," *ACS Sustainable Chemistry & Engineering* 9, no. 7 (February 2021): 2639–2650, https://doi.org/10.1021/acssuschemeng.0c06978
55. K. T. Nguyen, Q. D. Nguyen, T. A. Le, J. Shin, and K. Lee, "Analyzing the Compressive Strength of Green Fly Ash Based Geopolymer Concrete Using Experiment and Machine Learning Approaches," *Construction and Building Materials* 247 (June 2020): 118581, https://doi.org/10.1016/j.conbuildmat.2020.118581
56. H. Naseri, H. Jahanbakhsh, F. Moghadas Nejad, and A. Golroo, "Developing a Novel Machine Learning Method to Predict the Compressive Strength of Fly Ash Concrete in Different Ages," *AUT Journal of Civil Engineering* 4, no. 4 (December 2020): 3, https://doi.org/10.22060/ajce.2019.16124.5569
57. J. Heaton, *Introduction to Neural Networks with Java* (Chesterfield, MO: Heaton Research, Inc., 2008).
58. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Advances in Neural Information Processing Systems* 32 (Red Hook, NY: Curran Associates, Inc., 2019), 7994–8005.

59. B. Ouyang, Y. Song, Y. Li, F. Wu, H. Yu, Y. Wang, Z. Yin, X. Luo, G. Sant, and M. Bauchy, "Using Machine Learning to Predict Concrete's Strength: Learning from Small Datasets," *Engineering Research Express* 3, no. 1 (March 2021): 015022, http://doi.org/10.1088/2631-8695/abe344
60. O. Demir-Kavuk, M. Kamada, T. Akutsu, and E.-W. Knapp, "Prediction Using Step-Wise L1, L2 Regularization and Feature Selection for Small Data Sets with Large Number of Features," *BMC Bioinformatics* 12, no. 1 (2011): 412, https://doi.org/10.1186/1471-2105-12-412
61. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems 30*, ed. I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Red Hook, NY: Curran Associates, Inc., 2017), 4765–4774.
62. L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," in *2018 IEEE Fifth International Conference on Data Science and Advanced Analytics (DSAA)* (New York: IEEE, 2018), 80–89, https://doi.org/10.1109/DSAA.2018.00018
63. D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine Learning Interpretability: A Survey on Methods and Metrics," *Electronics* 8, no. 8 (August 2019): 832, https://doi.org/10.3390/electronics8080832
64. A. E. Roth, ed., *The Shapley Value: Essays in Honor of Lloyd S. Shapley* (Cambridge, UK: Cambridge University Press, 1988).
65. M. Chelberg, "The Effect of Fly Ash Chemical Composition on Compressive Strength of Fly Ash Portland Cement Concrete" (master's thesis, Ohio State University, 2019), http://rave.ohiolink.edu/etdc/view?acc_num=osu1555611247091087
66. P. Ziolkowski and M. Niedostatkiewicz, "Machine Learning Techniques in Concrete Mix Design," *Materials* 12, no. 8 (April 2019): 1256, https://doi.org/10.3390/ma12081256
67. L. Kan, R. Shi, and J. Zhu, "Effect of Fineness and Calcium Content of Fly Ash on the Mechanical Properties of Engineered Cementitious Composites (ECC)," *Construction and Building Materials* 209 (June 2019): 476–484, https://doi.org/10.1016/j.conbuildmat.2019.03.129
68. Z. Giergiczny and A. Werynsk, "Influence of Fineness of Fly Ashes on Their Hydraulic Activity," in *Third International Conference Proceedings, Fly Ash, Silica Fume, Slag, and Natural Pozzolans in Concrete* (Farmington Hills, MI: American Concrete Institute, 1989), 97–116.
69. Ş. Yazici and H. Ş. Arel, "Effects of Fly Ash Fineness on the Mechanical Properties of Concrete," *Sadhana* 37, no. 3 (June 2012): 389–403, https://doi.org/10.1007/s12046-012-0083-3
70. S. Kumar and R. Kumar, "Mechanical Activation of Fly Ash: Effect on Reaction, Structure and Properties of Resulting Geopolymer," *Ceramics International* 37, no. 2 (March 2011): 533–541, https://doi.org/10.1016/j.ceramint.2010.09.038
71. C. Jaturapitakkul, K. Kiattikomol, V. Sata, and T. Leekeeratikul, "Use of Ground Coarse Fly Ash as a Replacement of Condensed Silica Fume in Producing High-Strength Concrete," *Cement and Concrete Research* 34, no. 4 (April 2004): 549–555, https://doi.org/10.1016/S0008-8846(03)00150-9
72. B. Felekoğlu, S. Türkel, and H. Kalyoncu, "Optimization of Fineness to Maximize the Strength Activity of High-Calcium Ground Fly Ash – Portland Cement Composites," *Construction and Building Materials* 23, no. 5 (May 2009): 2053–2061, https://doi.org/10.1016/j.conbuildmat.2008.08.024
73. A. Fernández-Jiménez and A. Palomo, "Characterisation of Fly Ashes. Potential Reactivity as Alkaline Cements," *Fuel* 82, no. 18 (December 2003): 2259–2265, https://doi.org/10.1016/S0016-2361(03)00194-7
74. A. Ríos, M. González, C. Montes, J. Vásquez, and J. Arellano, "Assessing the Effect of Fly Ash with a High SO_3 Content in Hybrid Alkaline Fly Ash Pastes (HAFAPs)," *Construction and Building Materials* 238 (March 2020): 117776, https://doi.org/10.1016/j.conbuildmat.2019.117776
75. F. Zunino, D. P. Bentz, and J. Castro, "Reducing Setting Time of Blended Cement Paste Containing High-SO_3 Fly Ash (HSFA) Using Chemical/Physical Accelerators and by Fly Ash Pre-washing," *Cement and Concrete Composites* 90 (July 2018): 14–26, https://doi.org/10.1016/j.cemconcomp.2018.03.018
76. G. Sheng, Q. Li, J. Zhai, and F. Li, "Self-Cementitious Properties of Fly Ashes from CFBC Boilers Co-firing Coal and High-Sulphur Petroleum Coke," *Cement and Concrete Research* 37, no. 6 (June 2007): 871–876, https://doi.org/10.1016/j.cemconres.2007.03.013
77. *Specification for Portland Cement*, ASTM C150 (West Conshohocken, PA: ASTM International, approved July 1, 2022), https://doi.org/10.1520/C0150_C0150M-22
78. A. K. Varshneya, *Fundamentals of Inorganic Glasses* (Sheffield, UK: Society of Glass Technology, 2013).