Data-driven predictive models for chemical durability of oxide glass under different chemical conditions

Yi Zhang, Aize Li, Binghui Deng and Kelleen K. Hughes

We conducted a comprehensive study to investigate the performance of various machine-learning models in predicting the chemical durability of oxide glasses under different chemical conditions with glass composition as input features, by taking advantage of the large dataset (~1400 datapoints) we have collected. Two typical machine-learning tasks, weight loss regression, and surface appearance change rating classification, were conducted in the study. We successfully made Neural Networks deliver an excellent performance in predicting the weight loss, while Random Forest in classifying the surface appearance change rating. Additionally, feature importance analysis showed that SiO₂, Na₂O, P₂O₅ were the most dominate features for predicting the weight loss, while SiO₂, ZrO₂, CaO were the topmost features for classifying the surface appearance change rating, under acid, HF, and base testing conditions, respectively. We also realized that the trained models fall short of extrapolating data far from the training dataset space even though they exhibit outstanding interpolation performance in some cases. Topology constrained theory fed by structural information from molecular dynamics simulations seems to be a promising approach to address the challenge.

npj Materials Degradation (2020)4:14; https://doi.org/10.1038/s41529-020-0118-x

INTRODUCTION

The chemical durability of oxide glasses has been commonly perceived to be superior to most other materials, which promotes its wide application over the history of human civilization. For instance, many stained glass windows in medieval cathedrals still stand exceptionally well after nearly 1000 years. Nevertheless, glass indeed suffers from gradual dissolution and corrosion when they are exposed to various aqueous solutions, which thwarts its many practical applications. To meet the ever-increasing demands for more advanced glasses exposed to various extreme chemical environments, it is important and urgent to fundamentally understand the glass dissolution kinetics, mechanisms of glasses being attacked by different chemicals, and guidelines to control the chemical durability of glasses.

There have been numerous outstanding research studies conducted and published to unravel the mystery of glass chemical durability over the past few decades. For instance, Pignatelli et al. and Mascaraque et al. correlate the network topology of oxide glasses with their chemical durability, and suggest that it is possible to design new glasses with a specific chemical durability for a determined pH using the topological model. Du et al. investigate the atomic origin of the passivation effect in hydrated silicate glass using reactive molecular dynamics simulations. They propose that the passivation propensity is intrinsically governed by the reorganization of the medium-range order structure of gel upon aging and the formation of small silicate rings that hinder water mobility. Although the insights from these fundamental studies largely boost our understanding of the glass chemical durability, it is still extremely challenging to effectively leverage them to innovate new glass products with advanced properties.

Machine learning (ML) has been increasingly attracting scientific interests in the domain of material science over the past decade. There have been many interesting studies on different materials in literature. For instance, Ziletti et al. applied a deep learning neural network model to automatically classify structures based on the crystal symmetry. Ren et al. successfully leveraged ML modeling and high-throughput experiments to discover three new material glass systems with large glass-forming ability. Different from other conventional physics-based modeling approaches with the aim of understanding specific aspect of material that are otherwise difficult to probe by experiments, ML models have the advantage of handling and comprehending high-dimensional feature space, and thus have great potentials in advancing our understanding of various materials.

As for the chemical durability and other properties of oxide glasses, how to take advantage of the latest advanced ML techniques to directly build up the composition–property relationship map in an end-to-end manner attracts tremendous interest and remains unsolved. As most of the glass properties exhibit highly non-linear relationships with compositions, data-driven ML models theoretically should be a useful tool since they are good at efficiently exploring complex high-dimensional composition space. There have been a few efforts in this field over the past few years. For instance, Deng et al. conducted a comprehensive machine-learning study on predicting the oxide glass density, Young’s modulus, shear modulus, and Poisson’s ratio using glass composition as the input by leveraging the large dataset collected at Corning Incorporated. Liu et al. introduced a topology-informed ML paradigm to predict the dissolution kinetics of silicate glasses given the limited dataset (200 datapoints). The physics-aided ML model is suggested to be capable of improving the model extrapolation ability far from the training sets. The model seems to be very promising since the physics-based topology of the glass is used as an additional feature. Nevertheless, it is built upon a rather simple Na₂O-SiO₂-Al₂O₃ glass system in which the topology constraints are relatively easy to calculate. In many real glass products, the composition can be so complex that practically makes the topology constraints intractable.
To this end, we carried out a comprehensive study to investigate the performance of various ML models to predict the chemical durability of oxide glasses using the glass composition as input features by taking advantage of the large dataset (~1400 datapoints) collected at Corning Incorporated over the past decade. Specifically, the chemical durability in this study is quantified by both the weight loss and glass surface damage severity rating when subjected to different aqueous solutions. Therefore, we are essentially dealing with a ML regression task (weight loss) and a classification task (surface damage rating). Our results showed that the Neural Network model is fully taking advantage of high-dimensional large datasets and exhibits good prediction performance. The feature importance analysis further provides valuable insights into which oxide plays a critical role in controlling the chemical durability of oxide glasses.

**RESULTS**

**Dataset details**

We collected data on the chemical durability of ~1400 oxide glass compositions under various chemical testing conditions that have been conducted over the past decade at Corning Incorporated. Owing to the data availability and quality, we focused on three testing chemical conditions in the study: acid, base, and HF solutions. The specification of each testing condition is detailed in the “Methods” section.

As for the glass composition, it covers a variety of common oxides, including SiO₂, B₂O₃, Al₂O₃, Na₂O, CaO, MgO, K₂O, P₂O₅, BaO, Li₂O, ZnO, SrO, and ZrO₂. These 13 oxides constitute the feature space for the ML models. Note that the composition space also covers some other oxides that are nearly negligible in terms of content, and our ML models therefore do not take them as input features. The detailed dataset size and primary oxide composition range (mol%) are summarized in Table 1.

Note that the composition space used to rank each sample, which are detailed in Fig. 2, together with some example pictures. Figure 3 shows the rating histogram of the dataset. Clearly, there are very limited datapoints of rating 1 and 3 for the acid testing condition, no datapoints of rating 1 and 5 for the base testing condition, and limited datapoints of rating 1, 4, and 5 for the HF-testing condition. Overall, there are two distinct machine-learning tasks in this study: (1) a regression task that correlates glass composition to weight loss, and (2) a classification task that correlates glass composition to surface appearance change. It is also worthy to mention that the weight loss and surface appearance change are not necessarily closely correlated, due to the fact that the former is mainly controlled by the reactivity of glass components with an etchant, while the latter is not only determined by the reactivity but also by the solubility of the etching byproducts in the solution.

**Weight loss regression**

Figure 4(a-1) and (a-2) shows the correlation between predicted and actual weight loss on training dataset under the acid testing condition using trained RF and NN models, respectively. A large majority of the datapoints are centered around the identity function, indicating a decent prediction performance for all the models. In fact, The RF model has a $R^2$ score of 0.87, and while the NN model has a better $R^2$ score of 0.95. Similarly, Fig. 4(b-1) and (b-2) shows the correlation between predicted and actual weight loss on training dataset under the base testing condition using trained RF and NN models, respectively. Clearly, both models show excellent performance with $R^2$ = 0.91 and 0.97 for RF and NN, respectively. Figure 4(c-1) and (c-2) shows the correlation between predicted and actual weight loss on training dataset under the HF-testing condition using trained RF and NN models, respectively. In this scenario, both models have good prediction performances with $R^2$ = 0.92 and 0.94 for RF and NN, respectively. Note that we also trained the KNN model for these datasets. The $R^2$ scores are roughly at the same level as that of the RF model, and therefore, we put all the results from the KNN model in the Supplementary Figs. S1 and S2.

---

### Table 1. Glass weight loss dataset summary.

| Testing condition | Sample size | Composition range (mol%) |
|-------------------|-------------|--------------------------|
|                   | SiO₂ | B₂O₃ | Al₂O₃ | Li₂O | Na₂O | K₂O | CaO | MgO | ZnO | P₂O₅ |
| Acid              | 1410 | 42.6–80.6 | 0–29.4 | 0.02–25.7 | 0–24.2 | 0–23.6 | 0–18.9 | 0–19.8 | 0–19.8 | 0–10.0 |
| Base              | 1462 | 42.6–80.6 | 0–29.4 | 0–25.7 | 0–24.2 | 0–23.6 | 0–18.9 | 0–19.8 | 0–19.8 | 0–10.0 |
| HF                | 1353 | 59.8–74.3 | 0–9.9 | 3.9–18.6 | 0–22.8 | 0–16.3 | 0–18.9 | 0–18.8 | 0–8.5 | 0–3.9 | 0–10.0 |

Details of the three datasets on chemical durability of oxide glasses subjected to acid, base, and HF-solution-testing conditions. Some primary oxide composition ranges are also listed.
Like the weight loss prediction, we also carried out studies on the classification prediction of glass surface appearance change under different chemical conditions. Figure 5(a-1) and (a-2) shows the confusion matrix on the training dataset under the acid testing condition using trained RF and NN models, respectively. The prediction accuracy of RF model reaches as high as 0.96, while the NN model slightly underperforms to reach the accuracy of 0.91. Among those misclassified cases, the misclassification between rating 0 and 2 is mostly dominate (~1.70% for the RF model, ~3.07% for the NN model), followed by the misclassification between rating 2 and 5 (~1.10% for the RF model, ~1.37% for the NN model). Figure 5(b-1) and (b-2) shows the confusion matrix on the training dataset under base testing condition using trained RF and NN models, respectively. The RF model shows a decent prediction accuracy as high of 0.88, while the NN model slightly underperforms to reach the accuracy of 0.83. As for those misclassified cases, the misclassification between rating 0 and 2 (~6.47%) again is most dominate, followed by the misclassification between rating 2 and 3 (~3.70%) for the RF model, while the NN model switches the order with rating 2 and 3 misclassification (~7.99%) running first, followed by rating 0 and 2 misclassification (~5.12%).

Similarly, Fig. 5(c-1) and (c-2) shows the confusion matrix on the training dataset under the HF acid testing condition using trained RF and NN models, respectively. Clearly, the prediction accuracy in this case runs significantly lower for both models. The RF model drops to 0.86, and the NN model drops to 0.80. The careful inspection of those misclassified cases reveals that the misclassification between rating 2 and 3 (~6.72% for the RF model, ~12.18% for the NN model) significantly reduces the prediction accuracy, followed by the misclassification between rating 0 and 2 (~2.94% for the RF model, ~5.04% for the NN model) for both models.

**DISCUSSION**

To further investigate the contribution of each oxide in determining the final prediction of weight loss and surface-damage-rating classification, we calculate the importance score of each oxide that has been implemented in Scikit-learn based on Gini impurity\(^23\) for the RF model. Note that the importance score is purely based on statistics and indicates how closely the target output is correlated with each input feature within the dataset defined composition space. To actually see how each oxide correlates with the prediction response, we further conducted the Prediction Profiler analysis supported by the statistical software JMP\(^24\). Generally, the Prediction Profiler demonstrates how the prediction changes with a single variable. It also recalculates the profiler and prediction in real time as the glass composition changes. Specific to our study, the Prediction Profiler provides a wealth of information about the prediction model with the following useful features: (1) dynamically show how the prediction changes as the glass composition changes; (2) estimate the likely glass composition sets given the target weight loss or glass surface appearance change rating; and (3) gauge the model sensitivity to each variable. It is also worthy to mention that we operate the analysis only at the dataset defined composition domain, and the analysis is insightful when the input

---

**Fig. 2** Glass surface appearance change rating system. The rating based on some visual inspection keywords with some example pictures. The rating indicates the severity of damage resulting from exposure to different chemical solutions.

**Fig. 3** Glass surface appearance change rating histogram. From top to bottom the histograms are for acid (a), base (b), and HF (c) testing conditions.

**Surface appearance change classification**

Like the weight loss prediction, we also carried out studies on the classification prediction of glass surface appearance change under different chemical conditions. Figure 5(a-1) and (a-2) shows the confusion matrix on the training dataset under the acid testing condition using trained RF and NN models, respectively. The prediction accuracy of RF model reaches as high as 0.96, while the NN model slightly underperforms to reach the accuracy of 0.91. Among those misclassified cases, the misclassification between rating 0 and 2 is mostly dominate (~1.70% for the RF model, ~3.07% for the NN model), followed by the misclassification between rating 2 and 5 (~1.10% for the RF model, ~1.37% for the NN model). Figure 5(b-1) and (b-2) shows the confusion matrix on the training dataset under base testing condition using trained RF and NN models, respectively. The RF model shows a decent prediction accuracy as high of 0.88, while the NN model slightly underperforms to reach the accuracy of 0.83. As for those misclassified cases, the misclassification between rating 0 and 2 (~6.47%) again is most dominate, followed by the misclassification between rating 2 and 3 (~3.70%) for the RF model, while the NN model switches the order with rating 2 and 3 misclassification (~7.99%) running first, followed by rating 0 and 2 misclassification (~5.12%).

Similarly, Fig. 5(c-1) and (c-2) shows the confusion matrix on the training dataset under the HF acid testing condition using trained RF and NN models, respectively. Clearly, the prediction accuracy in this case runs significantly lower for both models. The RF model drops to 0.86, and the NN model drops to 0.80. The careful inspection of those misclassified cases reveals that the misclassification between rating 2 and 3 (~6.72% for the RF model, ~12.18% for the NN model) significantly reduces the prediction accuracy, followed by the misclassification between rating 0 and 2 (~2.94% for the RF model, ~5.04% for the NN model) for both models.

**DISCUSSION**

To further investigate the contribution of each oxide in determining the final prediction of weight loss and surface-damage-rating classification, we calculate the importance score of each oxide that has been implemented in Scikit-learn based on Gini impurity\(^23\) for the RF model. Note that the importance score is purely based on statistics and indicates how closely the target output is correlated with each input feature within the dataset defined composition space. To actually see how each oxide correlates with the prediction response, we further conducted the Prediction Profiler analysis supported by the statistical software JMP\(^24\). Generally, the Prediction Profiler demonstrates how the prediction changes with a single variable. It also recalculates the profiler and prediction in real time as the glass composition changes. Specific to our study, the Prediction Profiler provides a wealth of information about the prediction model with the following useful features: (1) dynamically show how the prediction changes as the glass composition changes; (2) estimate the likely glass composition sets given the target weight loss or glass surface appearance change rating; and (3) gauge the model sensitivity to each variable. It is also worthy to mention that we operate the analysis only at the dataset defined composition domain, and the analysis is insightful when the input
features have no or weak interactions, which is the exactly the case for the base and HF chemical conditions.

Figure 6(a) shows the importance score of each oxide calculated by the RF model in predicting the weight loss for glasses subjected to the acid testing condition. Clearly, SiO2 plays the most critical role, followed by B2O3, Na2O, and Al2O3. The essential component in the acid is proton (H+) or hydronium ion (H3O+). Although protons can catalyze the dissociation (breaking) of
Fig. 5  **Surface appearance change classification model performance.** The confusion matrix for classifying glass surface appearance change using trained RF (left) and NN (right) models on the testing dataset for glasses under acid (a), base (b), and HF (c) chemical conditions. Note that there are no available data for the rating 1 and 5 for glasses under the base chemical condition.
siloxane bonds (Si–O–Si) at low proton concentrations, the reaction is very slow. That explains the good acid resistance of vitreous silicate glass and Pyrex® glass (composed primarily of silicate).

Most commercial glasses also contain modifiers and network formers in addition to silicate, such as aluminate, borate, and phosphate. These components, in comparison to silicate, are more susceptible (reactive) to acids. As shown in the following Eqs. 2 and 3, two processes occur sequentially when a glass (alumino-silicate glass) comes into contact with an acid25.

Reaction 1: \[ M^{n+} \text{glass} + nH^+ \text{(sol.)} \rightarrow M^{n+} \text{(sol.)} + nH^+ \text{(glass)} \]  

Reaction 2: \[ \text{Al} (\text{OSi})_4 \text{glass} + 4H^+ \text{(sol.)} \rightarrow \text{Al}^{3+} \text{(sol.)} + 4 \text{HOSi} \text{(glass)} \]

It starts with a diffusion process in which modifiers in the glass are exchanged out by protons (also known as the leaching process), followed by the second process involving the breaking of glass network from aluminate, borate, and phosphate. This second process is called gelation given the fact that it generates a gel layer on the glass surface. As the gel layer has high porosity, the diffusion reaction (Reaction 1) can then go through the gel layer and leach the modifiers present in the deeper interior of the glass. Overall, these two processes are the primary causes for the weight loss of aluminosilicate glass.

Incongruent dissolution has been widely believed to be the dominating corrosion mechanism in acid solutions3. After glass is leached in an acid solution, the residual silica layer holds the integrity of the network structure and provides resistance to further leaching or corrosion to the bulk glass. The network connectivity of the silica gel layer greatly impacts glass acid durability. The better network connectivity or lower O/Si ratio normally leads to a stronger resistance to acid corrosion. However, due to the strong interactions among oxides (see Supplementary Fig. S3), the glass composition effect on the chemical durability will vary significantly depending upon different compositions. Thus, the Prediction Profiler analysis is not shown for this case.

HF solutions are extensively used in glass industry for glass strengthening, thinning, engineering, etc. Congruent dissolution is a well-accepted corrosion mechanism for glass dissolution in HF3. HF primarily attacks Si–O bonds, breaks the glass network, and then consequently all glass components tend to dissolve into the solution with similar etching rates. Figure 7(a) shows the importance score of each oxide calculated by the RF model in predicting the weight loss for glasses under the HF-testing condition. As expected, SiO2 is one of the top contributors to the glass weight loss. In addition, alkaline metals, such as Na2O and K2O, also play important roles in controlling the glass dissolution rate in HF solutions. Figure 7(b) shows the corresponding Prediction Profiler for the oxides. Even though the slopes of the Prediction Profiler might vary slightly with different glass compositions due to the weak interactions among the oxides, the overall evolution trend keeps good consistency. The results show that increasing the content of alkaline earth oxides (i.e., MgO, CaO, SrO, and BaO), Li2O, and ZnO will improve the glass durability in HF. One potential hypothesis is that alkaline earth metals tend to form insoluble etching byproducts in HF acid such as calcium fluorosilicate, which will deposit on the glass surface and prevent further glass dissolution in HF. On the other hand, increasing the amount of Na2O, K2O, B2O3, Al2O3, and SiO2 is
found to reduce glass durability in HF in this case. However, we also want to clarify that the deposition of the etching byproduct on a glass surface can also impact the weight loss or etching rate. For instance, Na$_2$SiF$_6$ is one of the common etching byproducts observed for the alkaline aluminosilicate glass in the HF acid. It tends to precipitate out of the HF solution due to a low solubility limit and consequently deposit on the glass surface, blocking the HF acid from further attacking glass surface with increasing amount of Na$_2$O. In addition, glass chemical durability is determined by both glass composition and glass structure. In a glass composition where it contains high SiO$_2$ contents and alkaline, the generation of NBOs will significantly reduce glass chemical durability. To fully understand glass composition effect on glass chemical durability, glass structure shall be considered.

Similarly, congruent dissolution is believed to be a dominating corrosion mechanism for glass dissolution in NaOH chemistry. Figure 8(a) shows the importance score of each oxide calculated by the RF model in predicting the weight loss for glasses subjected to the base testing condition. Interestingly, P$_2$O$_5$, B$_2$O$_3$, CaO, and SiO$_2$ turn out to be the most important features. Figure 8(b) shows corresponding Prediction Profiler for certain selected oxides. The results show that the glass composition effect on glass base durability is very similar to its effect on the glass chemical durability in HF solutions. All the other glass components show similar effects as they do in the HF chemistry, except that B$_2$O$_3$ shows a non-linear effect and Na$_2$O shows a positive effect on glass durability in base. In other words, increasing the content of alkaline earth oxide (e.g., MgO, CaO, SrO, and BaO), Li$_2$O, and ZnO improves the glass durability, while increasing the amount of K$_2$O, B$_2$O$_3$, Al$_2$O$_3$, and SiO$_2$ is found to reduce the glass durability in base solution. However, we want to draw reader’s attention that some glass compositions in the dataset are not charge balanced. For instance, in the absence of Al$_2$O$_3$ in glass network, the Na$_2$O will primarily attack the SiO$_2$ network, which leads to a worse chemical durability even for SiO$_2$-rich composition as compared to a charge-balance system with less amount of SiO$_2$.

As for classifying the glass surface appearance change under different chemical testing conditions, we also calculate the importance score of each oxide using the RF model and the Prediction Profiler. For the acid testing condition, SiO$_2$ is predominately the most decisive oxide for determining the classification as shown in Fig. 9(a), while the other oxides seem to be relatively trivial features. As mentioned above, SiO$_2$ is very important to glass acid durability. Figure 9(b) further reveals that all the glasses are rated as 0 with the surface showing no defects or optical shifts regardless of the other glass components when the SiO$_2$ content is higher than 70 mol%. However, all the glasses deteriorate to rating 5 with the surface being severely delaminated and heavily leached when the SiO$_2$ content is below 60%. For a glass composition containing SiO$_2$ between 60 mol% and 70 mol%, the surface appearance rating gradually shifts from 4 to 1.

In the base solution, nearly all surface appearances after corrosion is rated between 2 and 3. Rating 2 and 3 represent that glass surfaces develop different levels of haze after corrosion. When a base-solution attacks a glass surface, the formed etching byproducts tend to have a low solubility limit, which leads to considerable deposition on the glass surface and subsequently induce differential etching. As a result, the glass surface is roughened and light is scattered. The higher the roughness, the higher a haze level is achieved. Figure 10(a) shows the importance score of each oxide calculated by the RF model, among which CaO, K$_2$O, MgO, Na$_2$O, Al$_2$O$_3$, and SiO$_2$ are the primary contributors. Figure 10(b) shows the corresponding Prediction Profiler for certain selected oxides. Clearly, we can see that increasing the amount of alkaline earth oxides (e.g., MgO, CaO, and BaO) in a glass composition will lead to more rating transitions from 2 to 3. In other words, the addition of alkaline earth in a glass composition increases surface roughness and the haze level after corrosion. On the other hand, increasing the amount of SiO$_2$, B$_2$O$_3$, Al$_2$O$_3$, Li$_2$O, Na$_2$O, K$_2$O, and ZnO in a glass composition helps to improve the surface appearance in a base solution. This can be explained either by an improved glass durability due to the addition of network formers, or the soluble etching byproducts formed with these components that prohibit the deposition of etching byproducts on glass surface.

Similar to the NaOH chemistry, most glass surface appearances after HF corrosion are rated between 2 and 3, indicating that the main cosmetic defects after HF etching are the glass surface haze level after corrosion. On the other hand, increasing the amount of SiO$_2$, B$_2$O$_3$, Al$_2$O$_3$, Li$_2$O, Na$_2$O, K$_2$O, and ZnO in a glass composition helps to improve the surface appearance in a base solution. This can be explained either by an improved glass durability due to the addition of network formers, or the soluble etching byproducts formed with these components that prohibit the deposition of etching byproducts on glass surface.
roughening and frosty look. The root cause of these defects is nearly the same as that in the NaOH chemistry, that is, the deposition of insoluble etching byproduct on the glass surface. Figure 11(a) shows the importance score of each oxide calculated by the RF model, ZrO\(_2\), Al\(_2\)O\(_3\), K\(_2\)O, B\(_2\)O\(_3\), and SiO\(_2\) are determined to be the main contributors in this case. Since K\(_2\)SiF\(_6\) is considered as a common insoluble etching byproduct in HF chemistry\(^{26}\), increasing the amount of K\(_2\)O in a glass composition leads to an increased rating for surface appearance as shown in Fig. 11(b). Similar effects are also observed for B\(_2\)O\(_3\) and Al\(_2\)O\(_3\). In contrast, when the content of SiO\(_2\) is higher than 68%, it shows that all the glasses are rated as 2 instead of 3, suggesting that increasing the content of SiO\(_2\) in a glass will improve the surface appearance rating in HF. Interestingly, the effect of ZrO\(_2\) on glass surface appearance change is non-linear. Specifically, the presence of small amount of ZrO\(_2\) (i.e., <1.5 mol%) degrades surface appearance in HF, while increasing it to a larger amount (i.e., >1.5 mol%) will improve the surface appearance by dissolving less materials in HF.

Recall that the acid case has the best classification accuracy regardless of the RF and NN model as compared to the other two cases. The above observations on importance score of each oxide might be helpful to explain it. The fact that the acid case has two predominate oxides effectively reduces the dimensionality of dataset space, which on average leads to more training datapoints or high datapoint density for each effective dimension and therefore better prediction accuracy. To the contrary, nearly each oxide in the base and HF-testing conditions plays an unneglectable role, thus the classification still effectively operates in the high-dimensional space with quite sparse datapoint distributions, which results in relatively weak prediction performance. Overall, glass composition effects on weight loss and surface appearance are complicated. It is not only affected by the reactivity of glass components with chemicals, but also the solubility of corrosion byproducts in different chemistries.

As we summarize the results and delve into the discussions, we see that our models are able to deliver very good prediction performance in most cases, while fall short of the prediction accuracy in some other cases. Although the importance score analysis for each oxide provides us some insights into the underlying prediction mechanism and nature of the datasets, we realize that it just touches the tip of the iceberg in the vast sea of puzzles and unknowns. In fact, it increasingly brings up more questions that goes beyond our grasp as we delve into the study. Some of the pressing concerns are listed: (1) Can the not-so-good prediction performance in some cases be simply attributed to the limited datapoints (~1400) or data quality? (2) Why do the top
features differ a lot even within the same chemical testing condition but different prediction tasks? (3) How confident can we extrapolate the model far from the training dataset space even for those models with high prediction or interpolation accuracy?

In fact, many researchers have been putting large efforts trying to answer these questions. The best approach seems to combine the pure data-driven ML algorithms with physics-informed properties. The rational lies in the fact that most state-of-art ML models are good at interpolating data and discovering relationships in high-dimension space, while most physic-based models are superior at extrapolating data in unexplored space. A successful implementation of this “centaurs” strategy is a recent work done by Liu et al.\textsuperscript{22} in which they predict the dissolution kinetics of silicate glasses by topology-informed ML.

The topological constraint theory of glass has been widely and successfully used to correlate the composition and temperature dependence of glass properties\textsuperscript{27–30}. However, the implicit assumption of the theory is that the structure of a given glass is sufficiently known so that the important structural units and associated constraints can be calculated. Unfortunately, the direct experimental approaches (e.g., nuclear magnetic resonance experiments) to characterizing the structure units for complicated multiple-component glass systems are expensive, challenging, and not always available. This is exactly the biggest challenge we are currently facing to leverage the topology constraint theory to enhance the ML model performance in this study.

To address this challenge with the absence of direct experimental measurements for our large dataset, an alternative approach is to resort to molecular dynamics simulations to seek the structural information of glasses. In fact, the recent developments of effective empirical potential for multiple-component glass systems have shown excellent results to capture the structure fingerprints (coordination number, $Q^n$ speciation, ring size distribution, etc.) of oxide glasses as evidenced by the work done by Deng et al.\textsuperscript{31,32}. It is, therefore, promising to move forward through leveraging molecular dynamics simulations, topology constraint theory, and ML models.

**METHODS**

**Chemical durability testing**

All the test glasses were washed with 4 wt\% Semiclean detergent for 2 min at 60 °C using a sonicator. After cleaning, glasses were rinsed with 18 MΩ DI water and dried in an oven at 110 °C for 2 h. The initial weight was measured and recorded after glasses were cooled to room temperature. Three testing conditions were used to evaluate the chemical durability with details listed as follows:

![Fig. 10](image-url) Dependence of glass surface appearance change on glass composition under base-solution testing condition. a The importance score of each oxide calculated by the RF model in predicting the glass surface appearance change. b The Prediction Profiler for certain selected oxides showing the correlation between each oxide and the predicted surface appearance change.
and the predicted surface appearance change. Certain selected oxides showing the correlation between each oxide and the predicted surface appearance change.

Fig. 11 Dependence of glass surface appearance change on glass composition under HF-solution-testing condition. a The importance score of each oxide calculated by the RF model in predicting the glass surface appearance change. b The Prediction Profiler for certain selected oxides showing the correlation between each oxide and the predicted surface appearance change.

(1) Acid: one glass coupon of 1 × 2 inches is soaked in a covered Pyrex tube containing 300 mL of 5 wt% HCl solution. Duplicate tests are pursued for each composition. The test solution is heated using a large water bath (Lindbergh Blue M, model WB1130) with an aluminum cover with 36 holes drilled in it. The test is conducted at 95 °C for 24 h. Afterwards, glasses are quenched and rinsed with 18 MΩ DI water and dried in an oven at 110 °C for at least 30 min. When glasses are cooled, the final glass weight is measured and recorded.

(2) Base: the experimental setup is the same as the acid test. Test samples are soaked in a platinum tube with Teflon watch cover containing 350 mL of 5 wt% NaOH at 95 °C for 6 h. Afterwards, glasses are quenched and rinsed with 18 MΩ DI water and dried in an oven at 110 °C for at least 30 min. When glasses are cooled, the final glass weight is measured and recorded.

(3) HF: the test sample is vertically placed in 220 mL of 10 wt% HF solution for 20 min in a Teflon beaker at room temperature. The final glass weight is measured and recorded after the sample is quenched and rinsed with 18 MΩ DI water and dried in an oven at 110 °C for at least 30 min.

To ensure the consistency of the chemical durability tests, a reference glass is etched in every test to build a control chart. Table 2 summarizes the relative standard deviation (RSD) and numbers of reference glasses used for the acid, base, and HF tests. Clearly, the low RSD value for each testing condition case indicates consistent and robust experimental measurements for our chemical durability tests.

**ML model methodology**

In our previous work on predicting the elastic moduli and density of oxide glasses from composition, we observed that Random Forest (RF) regression, K-Nearest Neighbor (KNN) regression, and artificial Neural Network (NN) regression are consistently the top algorithms that give good performances. Thus, we mainly focus on these three algorithms in this work. The details of each algorithm are available in that work. Some key points with regarding to the model training process, hyperparameter optimization, and reliability testing are listed as follows: (1) The whole dataset is divided into training and testing dataset, respectively; (2) Each feature (oxide composition) is normalized using the formula $x' = \frac{x - \text{min}(x)}{\text{max}(x) - \text{min}(x)}$; (3) The 5-fold cross validation technique is used to optimize the hyperparameters for each model; (4) The mean absolute error (MAE) and categorical cross-entropy are used as the loss function for the weight loss regression and appearance change rating classification, respectively; (5) The training dataset variation is taken into account by randomly sampling 100 different sets for testing. To evaluate the model prediction performance, we mainly use coefficient of determination $R^2$ factor for the weight of loss regression and appearance change rating classification task, respectively; (5) The training dataset variation is taken into account by randomly sampling 100 different sets for testing. To evaluate the model prediction performance, we mainly use coefficient of determination $R^2$ factor for the weight of loss regression and appearance change rating classification task, respectively.

The details of each algorithm are available in that work. Many other details with regarding to model training, complexity, and reliability are available in the Supplementary Tables S1 and Table S2.

**DATA AVAILABILITY**

The datasets used for the machine-learning training in this study are unfortunately not public available at this moment. However, we encourage anyone who is interested in working on this dataset to reach out to us. We are actively seeking help from the community to address some of the challenges mentioned in the work.

**CODE AVAILABILITY**

The study is conducted on the platform of Scikit-learn, which is an open source Python-based platform for conventional machine learning, by leveraging the built-in Random Forest, Neural Network, and K-nearest neighbors models.

Received: 13 January 2020; Accepted: 29 April 2020; Published online: 26 May 2020

**REFERENCES**

1. Scholze, H. Chemical durability of glasses. *J. Non-Crystalline Solids* **52**, 91–103 (1982).
2. Gulbiten, O., Mauro, J. C., Guo, X. & Boratav, O. N. Viscous flow of medieval cathedral glass. *J. Am. Ceram. Soc.* **101**, 5–11 (2018).
3. Frankel, G. S. et al. A comparative review of the aqueous corrosion of glasses, crystalline ceramics, and metals. *NPJ Mater. Degrad.* **2** (2018). https://doi.org/10.1038/s41529-018-0037-2.
4. Conradt, R. Chemical durability of oxide glasses in aqueous solutions: a review. *J. Am. Ceram. Soc.* **91**, 728–735 (2008).

| Table 2. Glass weight loss measurement reliability. |
|---------------------------------------------|
| Acid | Base | HF |
|------|------|----|
| RSD (%) | 2.04 | 1.19 | 3.18 |
| Sample size | 333 | 292 | 160 |

The RSDs calculated for the acid, base, and HF-testing in terms of weight loss.
5. Pignatelli, I., Kumar, A., Bauchy, M. & Sant, G. Topological control on silicates’ dissolution kinetics. Langmuir 32, 4434–4439 (2016).
6. Masaracque, N., Bauchy, M. & Smedskjaer, M. M. Correlating the network topology of oxide glasses with their chemical durability. J. Phys. Chem. B 121, 1139–1147 (2017).
7. Du, T. et al. Atomic origin of the passivation effect in hydrated silicate glasses. NP J Mater. Degrad 3 (2019). https://doi.org/10.1038/s41529-019-0070-9.
8. Butler, K. T., Davies, D. W., Cartwright, H., Iayev, O. & Walsh, A. Machine learning for molecular and materials science. Nature 559, 547–555 (2018).
9. Liu, Y., Zhao, T., Ju, W. & Shi, S. Materials discovery and design using machine learning. J. Materiomics 3, 159–177 (2017).
10. Ziletti, A., Kumar, D., Scheffler, M. & Ghiringhelli, L. M. Insightful classification of crystal structures using deep learning. Nat. Commun. 9, 2775 (2018).
11. Ren, F. et al. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments. Sci. Adv. 4, (2018).
12. Deng, B. & Shi, Y. Dynamic self-assembly of ‘living’ polymeric chains. Chem. Phys. Lett. 668, 14–18 (2017).
13. Deng, B., Luo, J., Harris, J. T. & Smith, C. M. Critical stress map for ZrO2 tetragonal to monoclinic phase transformation in ZrO2-toughened glass-ceramics. Mater. Sci. Eng. A 9, 100548 (2020).
14. Hou, J. et al. Magic auxeticity angle of graphene. Carbon 149, 350–354 (2019).
15. Deng, B., Palermo, E. F. & Shi, Y. Comparison of chain-growth polymerization in solution versus on surface using reactive coarse-grained simulations. Polymer 129, 105–116 (2017).
16. Deng, B., Huang, L. & Shi, Y. Solvent effect on the diffusion of unentangled linear polymer melts. Langmuir 33, 11845–11850 (2017).
17. Deng, B. & Shi, Y. The embrittlement and toughening of metallic glasses from nano-crystallization. J. Appl. Phys. 125, 145101 (2019).
18. Zhou, Q., Shenxi Liu, P. T., Pan, J., Yan, Q. & Zhang., S.-C. Atom2Vec: learning atoms for materials discovery. Proc. Natl Acad. Sci. 115, E6411–E6417 (2018).
19. Liu, H., Fu, Z., Li, Y., Sabri, N. F. A. & Bauchy, M. Balance between accuracy and simplicity in empirical forcefields for glass modeling: Insights from machine learning. J. Non-Crystalline Solids 515, 133–142 (2019).
20. Picklum, M. & Beetz, M. MatCALO: Knowledge-enabled machine learning in materials science. Comput. Mater. Sci. 163, 50–62 (2019).
21. Deng, B. Machine learning on density and elastic property of oxide glasses driven by large dataset. J. Non-Crystalline Solids 529, 119768 (2020).
22. Liu, H. et al. Predicting the dissolution kinetics of silicate glasses by topology-informed machine learning. NPJ Mater. Degrad. 3 (2019).
23. Louppe, G., Wehenkel, L., Sutera, A. & Geurts, P. Understanding variable importance in forests of randomized trees. Proceedings of the 26th International Conference on Neural Information Processing Systems. Curran Associates Inc. 431–439 (2013).
24. Jones, B. & Sall, J. JMP statistical discovery software. Wiley Interdiscip. Rev.: Computational Stat. 3, 188–194 (2011).
25. Paul, A. Chemical durability of glasses; a thermodynamic approach. J. Mater. Sci. 12, 2246–2268 (1977).
26. Barboux, P., Laghzizil, A., Bessaoles, Y., Derouihac, H. & Trouvé, G. Paradoxical crystalline morphology of frosted glass. J. Non-Crystalline Solids 345-346, 137–141 (2004).
27. Mauro, J. C. Topological constraint theory of glass. Am. Ceram. Soc. Bull. 90, 31–37 (2011).
28. Wilkinson, C. J., Zheng, Q., Huang, L. & Mauro, J. C. Topological constraint model for the elasticity of glass-forming systems. J. Non-Crystalline Solids 2, 100019 (2019).
29. Wilkinson, C. J. et al. Topological origins of the mixed alkali effect in glass. J. Phys. Chem. B 123, 7482–7489 (2019).
30. Takeda, W., Wilkinson, C. J., Feller, S. A. & Mauro, J. C. Topological constraint model of high lithium content borate glasses. J. Non-Crystalline Solids 3, 100028 (2019).
31. Deng, L. & Du, J. Development of boron oxide potentials for computer simulations of multicomponent oxide glasses. J. Am. Ceram. Soc. 102, 2482–2505 (2019).
32. Deng, L. & Du, J. Development of effective empirical potentials for molecular dynamics simulations of the structures and properties of boroaluminosilicate glasses. J. Non-Crystalline Solids 453, 177–194 (2016).