Revision 2

Word Count: 8523

# Estimating ferric iron content in clinopyroxene using machine learning models

**Wei-hua Huang[1,2], Yang Lyu[3], Ming-hao Du[4], Can He[5], Shang-de Gao[6], Ren-jun Xu[6], Qun-ke Xia[1,2], J ZhangZhou[1,2]***

[1] School of Earth Sciences, Zhejiang University, Hangzhou 310027, China.

[2] Key Laboratory of Big Data and Deep Earth Resources of Zhejiang Province, Hangzhou 310027, China.

[3] Gemological Institute, China University of Geosciences, Wuhan 430074, China.

[4] School of Geoscience and Info-Physics, Central South University, Changsha 410083, China.

[5] School of Earth Sciences and Engineering, Sun Yat-Sen University, Guangzhou, 510275, China.

[6] Center for Data Science, Zhejiang University, Hangzhou 310058, China.

Corresponding author: J ZhangZhou (zhangzhou333@zju.edu.cn)

## Abstract

Clinopyroxene ferric iron content is an important consideration for garnet-clinopyroxene geothermometry and estimations of water storage in the Earth's interior, but remains difficult and expensive to measure. Here, we develop seven classic algorithms and machine learning methods to estimate $Fe^{3+}/\Sigma Fe$ in clinopyroxene using major element data from electron microprobe analyses. The models were first trained using a large dataset of clinopyroxene $Fe^{3+}/\Sigma Fe$ values determined by Mössbauer spectroscopy and spanning a wide compositional range, with major uncertainties ranging from 0.25 to 0.3 and root-mean-square errors on the test dataset ranging from 0.071 to 0.089. After dividing the entire dataset into three compositional sub-datasets, the machine learning models were trained and compared for each sub-dataset. Our results suggest that ensemble learning algorithms (random forest and Extra-Trees) perform better than principal component analysis-based elastic net polynomial, artificial neural network, artificial neural network ensemble, decision trees, and linear regressions. Using a sub-dataset excluding clinopyroxene in spinel peridotite and omphacite in eclogite, the new models achieved uncertainties of 0.15 to 0.2 and root-mean-square errors on the test dataset ranging from 0.051 to 0.078, decreasing prediction errors by 30–40%. By incorporating compositional data on coexisting spinel, new models for clinopyroxene in spinel peridotite show improved performance, indicating the interaction between spinel and clinopyroxene in spinel peridotite. Feature importance analysis shows $Na^+$, $Ca^{2+}$, and $Mg^{2+}$ to be the most important for predicting $Fe^{3+}$ content, supporting the coupled substitution between $Ca^{2+}$-$M^{2+}$ and $Na^+$-$M^{3+}$ in natural clinopyroxenes. The application of our models to garnet-clinopyroxene geothermometry greatly improves temperature estimates, achieving uncertainties of ±50 °C, compared with uncertainties of ±250 °C using previous models assuming all Fe as $Fe^{2+}$ or calculating $Fe^{3+}$ by charge conservation. Differences in the ferric iron contents, as

calculated using the machine learning models, of clinopyroxenes that did or did not experience hydrogen diffusion during their crystallization from basaltic magma support a redox-driven mechanism for hydrogen diffusion in clinopyroxene.

**Keywords:** Clinopyroxene, $Fe^{3+}/\Sigma Fe$, machine learning, geothermometer, redox

## Introduction

Clinopyroxene is prevalent throughout the deep crust and upper mantle, and its chemical composition carries essential information on metamorphism, partial melting, melt reactions, and deep volatile cycles (Råheim and Green 1974; Rudnick and Fountain 1995; Hirschmann 2000; Xia et al. 2013; Thomson et al. 2016; Beard et al. 2019). In particular, clinopyroxene $Fe^{3+}/\Sigma Fe$ content strongly reflects mantle redox conditions and therefore the stability of carbon-bearing minerals, because the mantle redox state determines whether carbon is present as carbonate, carbonatite melt, diamond, or other reduced phases (Luth and Canil 1993; Dasgupta and Hirschmann 2006; Brey et al. 2008; Frost and McCammon 2008; Rohrbach and Schmidt 2011; Stagno et al. 2013; Zhang et al. 2019). Furthermore, the garnet-clinopyroxene geothermometer, widely applied to garnet-lherzolites, eclogites, and granulites, and based on the reaction (Råheim and Green 1974; Ellis and Green 1979; Ganguly 1979; Saxena 1979; Krogh 1988):

$$\underset{\text{Pyrope}}{\frac{1}{3}\mathrm{Mg_3Al_2Si_3O_{12}}} + \underset{\text{Hedenbergite}}{\mathrm{CaFeSi_2O_6}} \leftrightarrow \underset{\text{Almandine}}{\frac{1}{3}\mathrm{Fe_3Al_2Si_3O_{12}}} + \underset{\text{Diopside}}{\mathrm{CaMgSi_2O_6}}, \qquad (1)$$

assumes iron is present only as ferrous iron; large errors likely result if both clinopyroxene ferrous and ferric iron contents are not considered (McGuire et al. 1989; Canil and O’Neill 1996; Galazka-

Friedman et al. 1998; Schmid et al. 2003; Proyer et al. 2004; Li et al. 2005). Thus, it is important to constrain both the ferric and ferrous iron contents in clinopyroxene.

Moreover, the oxidation of $Fe^{2+}$ to $Fe^{3+}$ has been proposed as an important mechanism of dehydrogenation in clinopyroxenes crystallized from basaltic magmas (Skogby and Rossman 1989; Stalder and Ludwig 2007; Su et al. 2008; Sundvall et al. 2009) following the reaction:

$$\mathrm{OH^-} + \mathrm{Fe^{2+}} = \mathrm{O^{2-}} + \mathrm{Fe^{3+}} + \frac{1}{2}\mathrm{H_2}\,. \tag{2}$$

This reaction is crucial for evaluating the preservation of water in clinopyroxene phenocrysts, which constrains the $H_2O$ contents of basaltic magmas (Wade et al. 2008; Xia et al. 2013; Lloyd et al. 2016; Xu et al. 2019). Thus, the accurate determination of clinopyroxene ferric iron contents greatly impacts constraints of the abundance and distribution of water in Earth's interior.

Although clinopyroxene major element contents, including total iron content, are widely measured by electron microprobe analyses ($n$ = 130,000 in the Georoc database), the ferric and ferrous iron contents are rarely measured separately (<1000) due to technical challenges, cost, and limited access to ferric iron analytical facilities, such as Mössbauer spectroscopy (Dyar et al. 2006), X-ray absorption near edge structure (XANES) spectroscopy (Bajt et al. 1994; Delaney et al. 1998; Kelley and Cottrell 2009; Cottrell and Kelley 2013; Terabayashi et al. 2013), transmission electron microscopy electron energy loss spectroscopy (TEM-EELS; van Aken et al. 1998; van Aken and Liebscher 2002; Rohrbach et al. 2007, 2011; Rzehak et al. 2020) and electron microprobe analyses employing the flank and peak-shift methods (Fialin et al. 2001; Höfer and Brey 2007; Lamb et al. 2012; C. Zhang et al. 2018). Although wet chemistry analyses are widely available, this destructive method requires greater amounts of sample, limiting its application to clinopyroxenes, which commonly account for only small volumes in mineral separates or have compositional heterogeneity. Instead of analytical measurements, the ferric iron contents of minerals are commonly calculated based on charge conservation. However, this method is affected by the accuracy of Si and Na measurements by microprobe. On one hand, Si is a major element with 4+ ionic charge, and a slight deviation in its concentration can have a significant effect on charge balance. On the other hand, precise measurements of Na are challenging due to its low atomic number (Dyar et al. 1989; Canil and O'Neill 1996; Sobolev et al. 1999).

Machine learning (ML) models can reveal complex correlations among elements following a data-driven approach. In recent years, ML methods have been broadly applied in various fields of mineralogy, petrology, and geochemistry (Dyar et al. 2012; Boucher et al. 2015; Petrelli and Perugini 2016; Le Losq et al. 2019; Li et al. 2020a, 2020b; Petrelli et al. 2020; Ptáček et al. 2020; Thomson et al. 2021). In this contribution, we develop and compare the performance of several ML methods to calculate $Fe^{3+}/\Sigma Fe$ in clinopyroxene from microprobe data. To improve ML model performance from the dataset perspective, we divided the clinopyroxene dataset into several compositional sub-datasets. The application of ferric iron contents obtained using our models to geothermometry results are compared to those obtained using the traditional charge balance method. Finally, the application of ferric iron contents is used to test the redox-driven hydrogen diffusion hypothesis in clinopyroxene crystallized from basaltic magma.

## Material and methods

### Data description

The natural clinopyroxene compositional dataset compiled in this study comprises 407 terrestrial samples from 62 publications, including both major element concentrations and

$Fe^{3+}/\Sigma Fe$ ratios (Supplementary Table S1). Synthetic clinopyroxenes ($n$ = 45) were excluded because they mostly have simple compositions or extreme values; e.g., a synthetic diopside sample (Di70_F) reported in Redhammer et al. (2012) has $Fe^{3+}/\Sigma Fe$ = 1.0, which is extremely rare in nature. Because, in principle, the ML training and prediction datasets should overlap, incorporating such simple, non-natural compositions in the training dataset would negatively affect the performance of the model when predicting $Fe^{3+}$ concentrations in natural samples.

Major element concentrations (Si, Ti, Al, Cr, Fe, Mn, Mg, Ca, Na) measured by microprobe were selected as eigenvalues ($X$) for the model input data, and $Fe^{3+}/\Sigma Fe$ ratios obtained by Mössbauer spectroscopy as the 'tag' values ($Y$) representing the model output. Among the analytical methods capable of measuring $Fe^{3+}/\Sigma Fe$, we chose only Mössbauer spectroscopy for two reasons: 1) micro-XANES standards are calibrated by Mössbauer spectroscopy, leading to repetitive counts of Mössbauer spectroscopy data in the training dataset; and 2) ferric iron data from micro-XANES and other methods are significantly less abundant than Mössbauer data.

We used this dataset to train the ML models in two ways. First, the entire dataset ($n$ = 407) was used to train the general models. In a second approach, the dataset was split into three sub-datasets numbered I–III based on mineral species and petrological context: I, diopside and augite in spinel peridotite ($n$ = 127); II, omphacite in eclogite ($n$ = 72); and III, the remaining data ($n$ = 208). The models were trained using each sub-dataset to obtain better accuracy.

As shown in Figure 1, the entire dataset spans a wide range of compositions: 45.32–57.65 wt% $SiO_2$, 0–9.76 wt% $TiO_2$, 0.12–16.37 wt% $Al_2O_3$, 0–29.37 wt% $Cr_2O_3$, 0.67–30.05 wt% FeO, 0–1.97 wt% MnO, 0.01–21.5 wt% MgO, 0.03–24.91 wt% CaO, 0.04–17.26 wt% $Na_2O$, and $Fe^{3+}/\Sigma Fe$ ratios ranging from 0 ($Fe^{2+}$ endmember, hedenbergite) to 1 ($Fe^{3+}$ endmember, aegirine).

Some rare clinopyroxenes, such as esseneite (purely ferric) and spodumene (Fe-free), were excluded.

Most $Fe^{3+}/\Sigma Fe$ ratios reported in the literature ignore the recoil-free fractions of $Fe^{3+}$ and $Fe^{2+}$, which may contribute large errors to $Fe^{3+}/\Sigma Fe$ values measured by Mössbauer spectroscopy (Dyar et al. 2012). To reduce the negative effect of such measurement errors on the tag values in our models, we corrected such values using the recoil-free fractions ($C$) listed in Table 1 as:

$$\frac{A^{3+}}{A^{2+}} = C \times \frac{N^{3+}}{N^{2+}}, \tag{3}$$

where $A$ is the doublet area of $Fe^{3+}$ or $Fe^{2+}$ in the Mössbauer spectrum, and $N$ their corresponding corrected abundances.

## Training and test datasets

We divided the compositional dataset into training and test datasets using the stratified random sampling function in the Scikit-learn library, which produces training sets that reflect, as much as possible, the characteristics of the entire dataset. The training set is used to train the models, whereas the independent test set, which is not involved in model training, is used to estimate the generalization of the trained models by evaluating their performance on a ‘new’ dataset. We selected 80% of the compiled dataset (or sub-dataset) for the training set, leaving 20% for testing. The exception is sub-dataset I, which is much smaller than the other sub-datasets; for this sub-dataset, we selected 90% of the data for training and 10% for testing to allow more data for training the algorithms.

## Machine learning algorithms

We used seven classic algorithms and ML methods that are widely used in regression studies: linear regression, principal component analysis (PCA)-based elastic net polynomial regression, artificial neural network regression, artificial neural network ensemble (ANN ensemble) regression, decision trees regression, random forest regression, and extremely randomized tree (Extra-Trees) regression. Detailed descriptions of the principles of each algorithm are provided in Supplementary Text 1. In addition, because they employ ensemble learning based on bagging, the random forest and Extra-Trees algorithms can prevent overfitting. Therefore, it is worthwhile to test their performance on small datasets such as clinopyroxene composition. Furthermore, these two algorithms provide feature importance information, facilitating the interpretation of their results.

## Performance metrics

Model performance was evaluated using two parameters: the coefficient of determination ($R^2$) and the root-mean-square error (RMSE). To avoid overfitting, we estimated the robustness of the model predictions. $R^2$ is a dimensionless metric of a model's goodness of fit, reflecting the degree to which the input variables explain variations in the output. The formula for $R^2$ is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \quad (4)$$

where $y_i$ denotes the real value, $\bar{y}$ is the mean value of $y_i$, and $\hat{y_i}$ is the predicted value. $R^2$ values range from $-\infty$ to 1, with the best possible score being $R^2 = 1.0$.

RMSE estimates the error between the real and predicted values as the square root of the sum of all errors divided by the number of values:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}. \tag{5}$$

RMSE values closer to zero indicate a better fit.

Considering the small size of our dataset, we performed 10-fold cross-validation to estimate the robustness of each ML algorithm; this method divides the dataset into ten subgroups (nine for training and one for validation) and repeats the process ten times. We repeated the 10-fold cross-validation ten times with various random seeds, resulting in 100 sub-sample sets. Thus, in each 10-fold cross-validation, ten sub-sample sets were created, each consisting of 10% of the total samples. Finally, the $R^2$ distribution of each of the 100 sub-sample sets was analyzed by kernel density estimation, a non-parametric method for estimating the probability density function of a random variable, and which can re-express the discrete histogram with a smooth and continuous curve to make it more accessible for model comparisons.

## Results

### Models trained on the entire dataset

Figure 2a shows the $R^2$ distribution for each model. The random forest and Extra-Trees algorithms showed the best performance, each with $R^2$ distributions characterized by a modal value of 0.89 and a mean RMSE value of 0.078, followed by the polynomial, neural network, and ANN ensemble algorithms with mean RMSE values of 0.080, 0.080, and 0.082, respectively. In contrast, the decision trees algorithm was the least accurate and robust, with very poor performances during some validation iterations. Figure 2b–g shows the details of each ML method by comparing the reference and predicted $Fe^{3+}/\Sigma Fe$ values. Remarkably, the correlations between the reference and predicted values in all model outputs plot along the 1:1 trend, indicating accurate predictions.

The uncertainties of the linear and decision trees models were about ±0.3, whereas the other models showed uncertainties of approximately ±0.25. The RMSE values obtained for the independent test dataset were 0.071 (random forest), 0.077 (Extra-Trees), 0.084 (polynomial), 0.081 (neural network), 0.084 (ANN ensemble), 0.089 (linear), and 0.086 (decision trees), similar to those obtained using the training dataset.

It is crucial to incorporate analytical uncertainties into our models because they determine the reliability of the ML algorithms. Analytical uncertainties on Mössbauer spectroscopy analyses are at least ±5% and increase with decreasing total Fe content (H.L. Zhang et al. 2018). Therefore, analytical errors are more significant for clinopyroxenes with low total Fe contents, which then propagates into the ML models. Indeed, the residual error, i.e., the difference between the predicted and reference values, increased at low Fe contents in all models (Figure 3a–g): the residual errors of the linear and decision trees algorithms increased from 0.2 to 0.3 at Fe contents below 0.4 atoms per formula unit (apfu); those of the polynomial, neural network, and ANN ensemble models increased from 0.15 at >0.4 apfu Fe to about 0.25 at lower Fe contents; in contrast, neither the random forest nor Extra-Trees algorithms were very sensitive to the $Fe^{3+}$ measurement uncertainties (Figure 3f–g). Furthermore, Canil and O'Neill (1996) reported microprobe analytical uncertainties on Si measurements. Indeed, our models inherited Si analytical uncertainties (Figure 3h–n); in contrast to Fe, the residual error increased with increasing Si content. Again, the random forest and Extra-Trees algorithms were the least sensitive to Si measurement uncertainties.

## Models trained on compositional sub-datasets

The $Fe^{3+}/\Sigma Fe$ ratios of certain sample types were poorly predicted by the general models, namely diopside and augite in spinel peridotite (*n* = 127) and omphacite in eclogite (*n* = 72). Therefore, we further trained ML models on these sub-datasets and investigated the feature importance information provided by the random forest and Extra-Trees algorithms.

Figure 4 compares models for diopside and augite in spinel peridotite (sub-dataset I). Linear regression did not achieve a correlation between the reference and predicted values, and simply predicted values around the average composition of the training dataset to achieve a low RMSE (Figure 4a). This is not surprising because the Pearson correlation coefficient shows nearly no correlation between any major element content and $Fe^{3+}/\Sigma Fe$ (below ±0.3), indicating that in these samples, $Fe^{3+}/\Sigma Fe$ does not depend solely on clinopyroxene composition. Previous studies suggest that coexisting spinel might perturb the incorporation of $Fe^{3+}$ in clinopyroxene (Canil and O'Neill 1996; Woodland et al. 2006; Woodland 2009). Therefore, we added the chemical compositions of coexisting spinel into sub-dataset I and trained new models on the combined clinopyroxene and spinel compositional data. The feature importance information returned by the random forest and Extra-Trees algorithms both indicate that the most important feature is Fe content in spinel, with other elemental concentrations in spinel also having moderate contributions to the predictions (Figure 5). Consequently, the performance of these models improved significantly when including coexisting spinel (Figure 4b, d, f), attaining uncertainties of ±0.15 on the linear model and ±0.08 on the random forest and Extra-Trees algorithms.

Because the omphacite-in-eclogite dataset is relatively small (*n* = 72), we do not provide exclusive models for omphacite herein and recommend using the general models to calculate $Fe^{3+}/\Sigma Fe$ in omphacite. Thus, here, we estimate the performance of the general models for omphacite (Figure 6). The maximal uncertainty was within ±0.25, but the correlation between the reference and predicted values was relatively weak, especially for the polynomial, neural network, and ANN ensemble regression algorithms. The random forest and Extra-Trees algorithms again showed the best result. Considering that $Fe^{3+}/\Sigma Fe$ values in omphacite span a wide range (0.11–0.76 in sub-dataset II), the errors on the models are acceptable for omphacite in eclogite (RMSEs for all provided models are below 0.15).

Figure 7 shows that when trained using sub-dataset III, the PCA-based elastic net polynomial, random forest, and Extra-Trees algorithms performed excellently, with modal $R^2$ values of about 0.96, uncertainties below ±0.15, and RMSE values for the test dataset ranging from 0.051 to 0.057. On the other hand, the linear, decision trees, and neural network algorithms performed relatively poorly in accuracy and precision, with uncertainties below ±0.2. Compared to the neural network algorithm, the ANN ensemble regression significantly improved precision, giving robust results without improving accuracy, with uncertainties below ±0.2.

## Discussion

### Evaluation of ML model performances

Here, we compare the different ML models based on the accuracy, uncertainty, and robustness of their predictions. Among the considered models, the random forest and Extra-Trees algorithms show the best performance, with high modal $R^2$ values and low mean RMSE values. Furthermore, these two algorithms are less sensitive to uncertainties on compositions measured by microprobe and/or Mössbauer spectroscopy. In comparison, the polynomial, neural network, and ANN ensemble algorithms show moderate performance, and the ANN ensemble algorithms are more sensitive to measurement uncertainties than the random forest and Extra-Trees algorithms. In particular, the neural network algorithm shows poor performance on the sub-datasets, indicating that the neural network algorithm is not suitable for small datasets. However, the accuracy, precision, and robustness of predictions improve significantly when combining the neural network and bagging algorithms (ANN ensemble algorithms). In contrast, the decision trees algorithm and linear regression show the worst accuracy and robustness, with poor performance during some validation iterations. Therefore, the ensemble learning algorithms are the best suited to the clinopyroxene compositional datasets.

The datasets used to train the ML models also have great impact on their performance. For clinopyroxene in spinel peridotite (sub-dataset I), the ML models' performances improved after incorporating coexisting spinel compositions (Figure 4), consistent with the relative importance of clinopyroxene and spinel compositions to the model predictions. This indicates that the correlation between $Fe^{3+}/\Sigma Fe$ and major element content in clinopyroxenes of sub-dataset I is likely due to spinel-clinopyroxene interaction. Therefore, the ML algorithms will be improved after incorporating a relevant parameter affecting clinopyroxene compositions (see next subsection). In sub-dataset II, the poor model performance was likely due to the complex $Fe^{3+}$ substitution mechanisms in omphacite. Unlike clinopyroxene in spinel peridotite, $Na^{+}$ abundance is strongly correlated with $Fe^{3+}/\Sigma Fe$ in omphacite (Pearson correlation coefficient $r > 0.6$). Furthermore, the small dataset ($n = 72$) spans various $P$-$T$ conditions, some of which may have not reached equilibrium, further contributing to the poor predictive capacity of the ML models for sub-dataset II. After excluding clinopyroxene in spinel peridotite and omphacite in eclogite, ML model performances on sub-dataset III improved compared to the entire dataset ($r = +0.95$, $-0.83$, and $-0.92$ for Na, Mg, and Ca, respectively). This result suggests a general relationship between $Fe^{3+}$ and $Na^{+}$, $Mg^{2+}$, and $Ca^{2+}$. Therefore, ML model performances can be improved with petrological and geological context.

## Cation substitution mechanisms revealed by ML models

Previous studies have proposed cationic substitution mechanisms for $Fe^{3+}$. Redhammer et al. (2000, 2012) reported a coupled $Ca^{2+}$-$Na^{+}$ and $M^{2+}$-$Fe^{3+}$ substitution (where M indicates an ion in the clinopyroxene M1 site) along the join of the hedenbergite-aegirine and diopside-aegirine solid solution series:

$$Fe^{3+\,(M1)} + Na^{+\,(M2)} = Fe^{2+\,(M1)} + Ca^{2+\,(M2)}, \tag{6}$$

$$Fe^{3+\,(M1)} + Na^{+\,(M2)} = Mg^{2+\,(M1)} + Ca^{2+\,(M2)}. \tag{7}$$

Nestola et al. (2007) reported the coupled substitution of Ca and $Fe^{2+}$ for Na and Al along the jadeite-hedenbergite solid solution series and the isovalent substitution of $Fe^{3+}$ for Al along the jadeite-aegirine series:

$$Al^{3+\,(M1)} + Na^{+\,(M2)} = Fe^{2+\,(M1)} + Ca^{2+\,(M2)}, \tag{8}$$

$$Fe^{3+\,(M1)} = Al^{3+\,(M1)}. \tag{9}$$

These coupled substitution mechanisms indicate that the replacement of the original ion in the M1 site by $Fe^{3+}$ is accompanied by substitution between Na and Ca in the M2 site to maintain charge balance.

Based on the feature importance information from the random forest algorithm, Na, Ca, and Mg contents in clinopyroxene are the essential features for predicting $Fe^{3+}/\Sigma Fe$ (Figure 8), in agreement with their strong positive (Na, $r = +0.89$) and negative Pearson correlation coefficients (Mg, $r = -0.78$; Ca, $r = -0.86$) with $Fe^{3+}/\Sigma Fe$ for the entire dataset. The general model indicates that $Fe^{3+}$ is incorporated in clinopyroxene as $NaFe^{3+}Si_2O_6$ (Equations 6, 7). However, clinopyroxene in spinel peridotite shows a poor correlation between Na and $Fe^{3+}$ ($r = +0.09$), with poor model performances unless coexisting spinel is included in the model (Canil and O'Neill 1996; Woodland et al. 2006; Woodland 2009; Hao and Li 2013). As mentioned above, this observation can be explained by reactions between clinopyroxene and coexisting spinel.

To determine the mechanism of $Fe^{3+}$ incorporation in clinopyroxene in spinel peridotite, we analyzed the correlation coefficient matrix for sub-dataset I (Supplementary Table S2) and the feature importance of the model. The Cr and Al contents in clinopyroxene show strong positive correlations with Cr ($r = +0.65$) and Al ($r = +0.70$) contents in spinel, respectively (Figure 9a, b). Similarly, Woodland et al. (2006) suggested that $Fe^{3+}$ partitioning between clinopyroxene and spinel varies with the Cr/Al ratio in spinel, indicating that Cr/Al in spinel may influence $Fe^{3+}/\Sigma Fe$ in clinopyroxene. In addition, Fe has high feature importance scores in both clinopyroxene and spinel. Comparing Fe, Al, and Cr partitioning between spinel and clinopyroxene shows a strong positive correlation between $D^{spl/cpx}_{Fe}$ and $D^{spl/cpx}_{Cr}$, but a poor correlation between $D^{spl/cpx}_{Fe}$ and $D^{spl/cpx}_{Al}$ (Figure 9c, d). These results suggest that $D^{spl/cpx}_{Fe}$ may lead to Cr/Al variation in spinel, in turn affecting $D^{spl/cpx}_{Fe3+}$ and $Fe^{3+}/\Sigma Fe$ in clinopyroxene by the reaction:

$$\underset{\text{Clinopyroxene}}{4CaMg^{2+}Si_2O_6}+\underset{\text{Spinel}}{Fe^{2+}Fe_2^{3+}O_4}+\underset{\text{Spinel}}{Fe^{2+}Al_2^{3+}O_4}=\underset{\text{Clinopyroxene}}{2CaFe^{2+}Si_2O_6}+\underset{\text{Clinopyroxene}}{2CaFe^{3+}Al^{3+}SiO_6}+\underset{\text{olivine}}{2Mg_2^{2+}SiO_4}. \quad (10)$$

However, the effect of $D^{spl/cpx}_{Fe}$ on $Fe^{3+}/\Sigma Fe$ in clinopyroxene should not be the dominant factor controlling $Fe^{3+}$ in clinopyroxene, as reflected by the weak correlation between Fe in clinopyroxene or spinel and $Fe^{3+}/\Sigma Fe$ in clinopyroxene. Furthermore, Na content in clinopyroxene shows moderate to strong negative correlations with Ca ($r = -0.42$) and Mg ($r = -0.64$) in clinopyroxene, indicating the possibility of $NaFe^{3+}Si_2O_4$ incorporation (Equation 7). In addition, the Al content in clinopyroxene was strongly negatively correlated with Si ($r = -0.72$) and Mg ($r = -0.6$) in clinopyroxene, which suggests $Fe^{3+}$ incorporation as $CaFe^{3+}AlSiO_6$:

$$Fe^{3+\,(M1)} + Al^{3+\,(T)} = Mg^{2+\,(M1)} + Si^{4+\,(T)}. \quad (11)$$

We note that all feature importance scores are less than 17.5% (Figure 5), indicating that none of the features are dominant in the system. Therefore, we suggest multiple $Fe^{3+}$ substitution mechanisms between clinopyroxene and spinel in spinel peridotite, without a single mechanism dominating. Furthermore, our results indicate that cation substitution is a mixed process that is effectively quantified by ML models.

## Comparing ML models to charge-balance for geothermometry applications

Figure 10 compares the predicted $Fe^{3+}/\Sigma Fe$ values for both the entire dataset and sub-dataset III (i.e., excluding clinopyroxene in spinel peridotite and omphacite in eclogite) using our polynomial, neural network, ANN ensemble, random forest, and Extra-Trees models to those using the traditional charge conservation method of Droop (1987), a simple general equation for estimating $Fe^{3+}$ content in ferromagnesian oxides and silicate minerals from microprobe analyses:

$$Fe^{3+} = 2G(1 - I/O), \quad (12)$$

where $I$ is the ideal number of cations per formula unit, and $O$ is the observed cation total per $G$ oxygens calculated assuming all iron to be $Fe^{2+}$. Their method generally overestimates $Fe^{3+}/\Sigma Fe$ in clinopyroxene with $Fe^{3+}/\Sigma Fe < 0.4$. By comparing $Fe^{3+}/\Sigma Fe$ values calculated by charge conservation to those obtained by Mössbauer spectroscopy, McGuire et al. (1989) and Canil and O'Neill (1996) showed that the uncertainty on Droop's (1987) method is too significant for practical application. In contrast, the predictions of our models are in good agreement with the true values, and the predicted errors are mainly within ±0.2. Therefore, our models significantly improve both the precision and uncertainty of the traditional method.

## Garnet-clinopyroxene geothermometry

The garnet-clinopyroxene geothermometer has been widely applied to garnet lherzolites, eclogites, and granulites (Råheim and Green 1974; Ellis and Green 1979; Ganguly 1979; Saxena 1979; Powell 1985). Because this geothermometer is based on the exchange of $Fe^{2+}$ and $Mg^{2+}$ between garnet and clinopyroxene, it is important to obtain the $Fe^{2+}$ value from the $Fe^{3+}/\Sigma Fe$ ratio. Whereas the $Fe^{3+}$ content in garnet is negligible at shallow depths (Geiger et al. 1987), that in clinopyroxene is not. Therefore, accurate estimation of $Fe^{3+}$ in clinopyroxene is crucial to the successful application of this geothermometer. Our models address the lack of available methods for accurately calculating $Fe^{3+}$ in clinopyroxene and will contribute to accurate and convenient temperature calculations.

We applied our models to calculate temperatures using the garnet-clinopyroxene geothermometer of Ellis and Green (1979) and compared the results with those calculated from Mössbauer spectroscopy data (Figure 11). We also estimated temperatures by considering all Fe as $Fe^{2+}$ or by predicting $Fe^{2+}$ based on Droop's (1987) charge conservation method. We used the 'WinGrt' program (Yavuz and Yildirim 2020) for the garnet-clinopyroxene geothermometer calculation. The data used here were randomly selected from the test dataset and cover a wide temperature range. Compared to those calculated from Mössbauer spectroscopy data, temperatures are overestimated when assuming all Fe as $Fe^{2+}$, consistent with the results of Galazka-Friedman et al. (1998). The $Fe^{2+}$ contents calculated by Droop's (1987) method also result in large temperature uncertainties. In contrast, the ML models show superior performance; compared to major uncertainties up to about ±250 °C when assuming all Fe as $Fe^{2+}$ or using Droop's (1987) method, the ML models achieved uncertainties below ±50 °C. Temperatures calculated for various samples from Mössbauer spectroscopy data and using different formulations of the garnet-clinopyroxene geothermometer are reported in Table 2. The standard error between different formulas averages to 59 °C, which translates to average uncertainties of ±118 °C (2σ). In contrast, when applying our models to calculate $Fe^{2+}$ content and then temperature, the largest uncertainties are around ±50 °C, within the error of the garnet-clinopyroxene geothermometer. Therefore, our models will improve garnet-clinopyroxene geothermometry calculations.

## Testing the hypothesis of hydrogen diffusion

The oxidation of $Fe^{2+}$ following Equation (2) is commonly regarded as the mechanism of hydrogen diffusion in clinopyroxene (Skogby and Rossman 1989; Bromiley and Keppler 2004; Koch-Müller et al. 2007; Stalder and Ludwig 2007). However, Sundvall et al. (2009) reported that this mechanism may not dominate in Fe-poor clinopyroxenes (<2 wt.% FeO) and that an additional reaction must be considered. These studies focused on a limited number of samples because of the scarcity of Mössbauer analyses. Here, to give a general perspective on the mechanism of hydrogen diffusion in clinopyroxene, we apply our models to calculate $Fe^{3+}/\Sigma Fe$ for a large number of clinopyroxenes in basalts ($n = 109$) and compare the difference between the $Fe^{3+}/\Sigma Fe$ values of dehydrogenated ($n = 31$) and non-dehydrogenated ($n = 78$) samples. These clinopyroxene samples are compiled in Supplementary Table S3; they were collected from several published clinopyroxene samples for which the original literature clearly discussed whether each sample suffered from dehydrogenation. As these clinopyroxenes are from basalts, we applied our models trained on sub-dataset III (i.e., excluding clinopyroxene in spinel peridotite and omphacite in eclogite). We plotted the predicted $Fe^{3+}/\Sigma Fe$ values vs. Fe content for both dehydrogenated and non-dehydrogenated samples in Figure 12a–c and used kernel density estimation to compare the $Fe^{3+}/\Sigma Fe$ values of dehydrogenated and non-dehydrogenated samples, with the RMSE values of the test dataset for each model taken as the model error (Figure 12d–f). Ferric iron content should increase with increasing dehydrogenation if the mechanism of Equation (2) dominates. The distribution of $Fe^{3+}/\Sigma Fe$ values in samples that suffered dehydrogenation is relatively higher than that for samples that did not (Figure 12). Although the differences between the $Fe^{3+}/\Sigma Fe$ values of the dehydrogenated and non-dehydrogenated samples are within the error range of our models, their statistical distribution tends to support the mechanism in Equation (2).

There are two reasons why the application of our models cannot unequivocably support or refuse the mechanism of ferric iron oxidation. First, if the mechanism of Equation (2) dominates, the magnitude of OH loss should equal that of ferric iron gain. However, because there is much less water than iron in clinopyroxene, dehydrogenation may not strongly influence $Fe^{3+}/\Sigma Fe$. Second, except for the oxidation of $Fe^{2+}$ to $Fe^{3+}$, other major element content variations would be limited. Thus, because our models only incorporate major element concentrations, they may not be sensitive enough to distinguish the slight changes in major element contents expected during dehydrogenation. Nonetheless, our investigation implies that major element compositional differences might exist between dehydrogenated and non-dehydrogenated clinopyroxenes, which could be verified by future ML studies.

## Implications

These ML algorithms were trained on a small dataset of 407 samples with known ferric iron contents. As more data becomes available, a more extensive training dataset will improve the performance of the models. These ML models predicting $Fe^{3+}$ in clinopyroxene are useful for estimating clinopyroxene $Fe^{3+}$ content when laboratory measurements are unavailable, whether due to technical or financial reasons. Knowing clinopyroxene $Fe^{3+}$ content also improves the precision of garnet-clinopyroxene geothermometers, offering better constraints and that could be used to provide new insights into metamorphic petrology and mineralogy. Our ML models imply a redox-driven mechanism for hydrogen diffusion in clinopyroxene during their crystallization from basaltic magma, but further work is required to verify or disprove this prediction. Furthermore, similar ML algorithms to predict mineral $Fe^{3+}$ contents could be developed for other $Fe^{3+}$-rich phases such as garnet and bridgmanite. The results of such models will be useful for probing the broader redox distribution of the solid earth. In addition, comprehensive ML investigations can be applied to various other geological questions that require regression, but are poorly fit by simple linear regressions, such as the sulfur content of silicate melts at sulfide saturation (Smythe et al. 2017; Chowdhury and Dasgupta 2020).

## Data Availability Statement

Python scripts for the application of the ML regression models are available on Github at https://github.com/ZJUEarthData/pyro_processor. The $Fe^{3+}$ calculator is not needed to configure the python environment and can be run from the terminal in Windows, Mac OS X, and Linux.

## Acknowledgments

The authors thank Y.H. Zhang, X.M. Zhou, Z.W. Huang, H.H. Wang, S.H. Wei, Z.X. Zhang, W.K. Wang, W.Y. Zhang, W.J. Zhou, W.Q. Liu, X. Chen, L.H. Chen, R.Z. Wang, L.Y. Jin, G.L. Li, Y.D. Lu, L.Z. Zhao, and Y.F. Xu for their assistance in compiling the data and applying the Python codes. We are thankful to Hongluo Zhang and Avishek Rudra for discussions. We are grateful to Robert Dennen for polishing the language of the paper. We highly appreciate insightful reviews by Laura Rzehak and Fang Huang. Finally, the authors acknowledge support from the Key Laboratory of Big Data and Deep Earth Resources of Zhejiang Province and the Fundamental Research Funds for the Central Universities.

## References cited

Ai, Y. (1994) A revision of the garnet-clinopyroxene $Fe^{2+}$-Mg exchange geothermometer. Contributions to Mineralogy and Petrology, 115, 467–473.

Bajt, S., Sutton, S.R., and Delaney, J.S. (1994) X-ray microprobe analysis of iron oxidation states in silicates and oxides using X-ray absorption near edge structure (XANES). Geochimica et Cosmochimica Acta, 58, 5209–5214.

Beard, C.D., Van Hinsberg, V.J., Stix, J., and Wilke, M. (2019) Clinopyroxene/melt trace element partitioning in sodic alkaline magmas. Journal of Petrology, 60, 1797–1823.

Boucher, T.F., Ozanne, M. V., Carmosino, M.L., Dyar, M.D., Mahadevan, S., Breves, E.A., Lepore, K.H., and Clegg, S.M. (2015) A study of machine learning regression methods for major elemental analysis of rocks using laser-induced breakdown spectroscopy. Spectrochimica Acta - Part B Atomic Spectroscopy, 107, 1–10.

Brey, G.P., Bulatov, V.K., Girnis, A. V., and Lahaye, Y. (2008) Experimental melting of carbonated peridotite at 6-10 GPa. Journal of Petrology, 49, 797–821.

Bromiley, G.D., and Keppler, H. (2004) An experimental investigation of hydroxyl solubility in jadeite and Na-rich clinopyroxenes. Contributions to Mineralogy and Petrology, 147, 189–200.

Canil, D., and O'Neill, H.S.C. (1996) Distribution of ferric iron in some upper-mantle assemblages. Journal of Petrology, 37, 609–635.

Chowdhury, P., and Dasgupta, R. (2020) Sulfur extraction via carbonated melts from sulfide-bearing mantle lithologies – Implications for deep sulfur cycle and mantle redox. Geochimica et Cosmochimica Acta, 269, 376–397.

Cottrell, E., and Kelley, K.A. (2013) Redox heterogeneity in Mid-Ocean Ridge basalts as a function of mantle source. Science, 340, 1314 – 1317.

Dasgupta, R., and Hirschmann, M.M. (2006) Melting in the Earth's deep upper mantle caused by carbon dioxide. Nature, 440, 659–662.

De Grave, E., and Van Alboom, A. (1991) Evaluation of ferrous and ferric Mössbauer fractions. Physics and Chemistry of Minerals, 18, 337–342.

Delaney, J.S., Dyar, M.D., Steven, S.R., and Bajt, S. (1998) Redox ratios with relevant resolution: Solving an old problem by using the synchrotron micro-XANES probe. Geology, 26, 139–142.

Droop, G.T.R. (1987) A general equation for estimating $Fe^{3+}$ concentrations in ferromagnesian silicates and oxides from microprobe analyses, using stoichiometric criteria. Mineralogical Magazine, 51, 431–435.

Dyar, M.D., McGuire, A. V., and Ziegler, R.D. (1989) Redox equilibria and crystal chemistry of coexisting minerals from spinel lherzolite mantle xenoliths. American Mineralogist, 74, 969–980.

Dyar, M.D., Agresti, D.G., Schaefer, M.W., Grant, C.A., and Sklute, E.C. (2006) Mössbauer spectroscopy of Earth and planetary materials. Annual Review of Earth and Planetary Sciences, 34, 83–125.

Dyar, M.D., Breves, E.A., Emerson, E., Bell, S.W., Nelms, M., Ozanne, M. V, Peel, S.E., Carmosino, M.L., Tucker, J.M., Gunter, M.E., and others (2012) Accurate determination of ferric iron in garnets by bulk Mössbauer spectroscopy and synchrotron micro-XANES. American Mineralogist, 97, 1726–1740.

Eeckhout, S.G., and De Grave, E. (2003) Evaluation of ferrous and ferric Mössbauer fractions. Part II. Physics and Chemistry of Minerals, 30, 142–146.

Ellis, D.J., and Green, D.H. (1979) An experimental study of the effect of Ca upon garnet-clinopyroxene Fe-Mg exchange equilibria. Contributions to Mineralogy and Petrology, 71, 13–22.

Fialin, M., Wagner, C., Métrich, N., Humler, E., Galoisy, L., and Bézos, A. (2001) $Fe^{3+}/\sum Fe$ vs. FeLα peak energy for minerals and glasses: Recent advances with the electron microprobe. American Mineralogist, 86, 456–465.

Frost, D.J., and McCammon, C.A. (2008) The redox state of earth's mantle. Annual Review of Earth and Planetary Sciences, 36, 389–420.

Galazka-Friedman, J., Bauminger, E.R., and Bakun-Czubarow, N. (1998) Determination of iron oxidation state in omphacites applied to geothermometry of sudetic eclogites. Hyperfine Interactions, 112, 223–226.

Ganguly, J. (1979) Garnet and clinopyroxene solid solutions, and geothermometry based on Fe-Mg distribution coefficient. Geochimica et Cosmochimica Acta, 43, 1021–1029.

Ganguly, J., Cheng, W., and Tirone, M. (1996) Thermodynamics of aluminosilicate garnet solid solution: new experimental data, an optimized model, and thermometric applications. Contributions to Mineralogy and Petrology, 126, 137–151.

Geiger, C.A., Newton, R.C., and Kleppa, O.J. (1987) Enthalpy of mixing of synthetic almandine-grossular and almandine-pyrope garnets from high-temperature solution calorimetry. Geochimica et Cosmochimica Acta, 51, 1755–1763.

Hao, X.-L., and Li, Y.-L. (2013) $^{57}$Fe Mössbauer spectroscopy of mineral assemblages in mantle spinel lherzolites from Cenozoic alkali basalt, eastern China: Petrological applications. Lithos, 156–159, 112–119.

Hirschmann, M.M. (2000) Mantle solidus: Experimental constraints and the effects of peridotite composition. Geochemistry, Geophysics, Geosystems, 1, 2000GC000070.

Höfer, H.E., and Brey, G.P. (2007) The iron oxidation state of garnet by electron microprobe: Its determination with the flank method combined with major-element analysis. American Mineralogist, 92, 873–885.

Kelley, K.A., and Cottrell, E. (2009) Water and the oxidation state of subduction zone magmas. Science, 325, 605–607.

Koch-Müller, M., Abs-Wurmbach, I., Rhede, D., Kahlenberg, V., and Matsyuk, S. (2007) Dehydration experiments on natural omphacites: Qualitative and quantitative characterization by various spectroscopic methods. Physics and Chemistry of Minerals, 34, 663–678.

Krogh, E.J. (1988) The garnet-clinopyroxene Fe-Mg geothermometer — a reinterpretation of existing experimental data. Contributions to Mineralogy and Petrology, 99, 44–48.

Lamb, W.M., Guillemette, R., Popp, R.K., Fritz, S.J., and Chmiel, G.J. (2012) Determination of $Fe^{3+}$/Fe using the electron microprobe: A calibration for amphiboles. American Mineralogist, 97, 951–961.

Lazarov, M., Woodland, A.B., and Brey, G.P. (2009) Thermal state and redox conditions of the Kaapvaal mantle: A study of xenoliths from the Finsch mine, South Africa. Lithos, 112, 913–923.

Le Losq, C., Berry, A.J., Kendrick, M.A., Neuville, D.R., and O'Neill, H.S.C. (2019) Determination of the oxidation state of iron in Mid-Ocean Ridge basalt glasses by Raman spectroscopy. American Mineralogist, 104, 1032–1042.

Li, Y.L., Zheng, Y.F., and Fu, B. (2005) Mössbauer spectroscopy of omphacite and garnet pairs from eclogites: Application to geothermobarometry. American Mineralogist, 90, 90–100.

Li, X., Zhang, C., Behrens, H., and Holtz, F. (2020a) Calculating amphibole formula from electron microprobe analysis data using a machine learning method based on principal components regression. Lithos, 362–363, 105469.

——— (2020b) Calculating biotite formula from electron microprobe analysis data using a machine learning method based on principal components regression. Lithos, 356–357, 105371.

Liu, J. (1998) Assessment of the garnet-clinopyroxene thermometer. International Geology Review, 40, 579–608.

Lloyd, A.S., Ferriss, E., Ruprecht, P., Hauri, E.H., Jicha, B.R., and Plank, T. (2016) An assessment of clinopyroxene as a recorder of magmatic water and magma ascent rate. Journal of Petrology, 57, 1865–1886.

Luth, R.W., and Canil, D. (1993) Ferric iron in mantle-derived pyroxenes and a new oxybarometer for the mantle. Contributions to Mineralogy and Petrology, 113, 236–248.

McGuire, A. V., Dyar, M.D., and Ward, K.A. (1989) Neglected $Fe^{3+}/Fe^{2+}$ ratios-a study of $Fe^{3+}$ content of megacrysts from alkali basalts. Geology, 17, 687–690.

Nakamura, D. (2009) A new formulation of garnet–clinopyroxene geothermometer based on accumulation and statistical analysis of a large experimental data set. Journal of Metamorphic Geology, 27, 495–508.

Nestola, F., Tribaudino, M., Boffa Ballaran, T., Liebske, C., and Bruno, M. (2007) The crystal structure of pyroxenes along the jadeite–hedenbergite and jadeite–aegirine joins. American Mineralogist, 92, 1492–1501.

Nimis, P., Goncharov, A., Ionov, D.A., and McCammon, C. (2015) $Fe^{3+}$ partitioning systematics between orthopyroxene and garnet in mantle peridotite xenoliths and implications for thermobarometry of oxidized and reduced mantle rocks. Contributions to Mineralogy and Petrology, 169, 6.

Petrelli, M., and Perugini, D. (2016) Solving petrological problems through machine learning: the study case of tectonic discrimination using geochemical and isotopic data. Contributions to Mineralogy and Petrology, 171, 81.

Petrelli, M., Caricchi, L., and Perugini, D. (2020) Machine learning Thermo-Barometry: application to clinopyroxene-bearing magmas. Journal of Geophysical Research: Solid Earth, 125, e2020JB020130.

Powell, R. (1985) Regression diagnostics and robust regression in geothermometer/geobarometer calibration: the garnet-clinopyroxene geothermometer revisited. Journal of Metamorphic Geology, 3, 231–243.

Proyer, A., Dachs, E., and McCammon, C. (2004) Pitfalls in geothermobarometry of eclogites: $Fe^{3+}$ and changes in the mineral chemistry of omphacite at ultrahigh pressures. Contributions to Mineralogy and Petrology, 147, 305–318.

Ptáček, M.P., Dauphas, N., and Greber, N.D. (2020) Chemical evolution of the continental crust from a data-driven inversion of terrigenous sediment compositions. Earth and Planetary Science Letters, 539, 116090.

Råheim, A., and Green, D.H. (1974) Experimental determination of the temperature and pressure dependence of the Fe-Mg partition coefficient for coexisting garnet and clinopyroxene. Contributions to Mineralogy and Petrology, 48, 179–203.

Ravna, K. (2000) The garnet–clinopyroxene $Fe^{2+}$–Mg geothermometer: an updated calibration. Journal of Metamorphic Geology, 18, 211–219.

Redhammer, G.J., Amthauer, G., Lottermoser, W., and Treutmann, W. (2000) Synthesis and structural properties of clinopyroxenes of the hedenbergite $CaFe^{2+}Si_2O_6$ - aegirine $NaFe^{3+}Si_2O_6$ solid-solution series. European Journal of Mineralogy, 12, 105–120.

Redhammer, G.J., Tippelt, G., Amthauer, G., and Roth, G. (2012) Structural and $^{57}$Fe Mössbauer spectroscopic characterization of the synthetic $NaFeSi_2O_6$ (aegirine) - $CaMgSi_2O_6$ (diopside) solid solution series. Zeitschrift fur Kristallographie, 227, 396–410.

Rohrbach, A., and Schmidt, M.W. (2011) Redox freezing and melting in the Earths deep mantle resulting from carbon-iron redox coupling. Nature, 472, 209–214.

Rohrbach, A., Ballhaus, C., Golla-Schindler, U., Ulmer, P., Kamenetsky, V.S., and Kuzmin, D. V. (2007) Metal saturation in the upper mantle. Nature, 449, 456–458.

Rohrbach, A., Ballhaus, C., Ulmer, P., Golla-Schindler, U., and Schönbohm, D. (2011) Experimental evidence for a reduced metal-saturated upper mantle. Journal of Petrology, 52, 717–731.

Rudnick, R.L., and Fountain, D.M. (1995) Nature and composition of the continental crust: A lower crustal perspective. Reviews of Geophysics, 33, 267–309.

Rzehak, L.J.A., Rohrbach, A., Vollmer, C., Höfer, H.E., Berndt, J., and Klemme, S. (2020) Ferric-ferrous iron ratios of experimental majoritic garnet and clinopyroxene as a function of oxygen fugacity. American Mineralogist, 105, 1866–1874.

Saxena, S.K. (1979) Garnet-clinopyroxene geothermometer. Contributions to Mineralogy and Petrology, 70, 229–235.

Schmid, R., Wilke, M., Oberhánsli, R., Janssens, K., Falkenberg, G., Franz, L., and Gaab, A. (2003) Micro-XANES determination of ferric iron and its application in thermobarometry. Lithos, 70, 381–392.

Skogby, H., and Rossman, G.R. (1989) $OH^-$ in pyroxene; an experimental study of incorporation mechanisms and stability. American Mineralogist, 74, 1059–1069.

Smythe, D.J., Wood, B.J., and Kiseeva, E.S. (2017) The S content of silicate melts at sulfide saturation: New experiments and a model incorporating the effects of sulfide composition. American Mineralogist, 102, 795–803.

Sobolev, V.N., McCammon, C.A., Taylor, L.A., Snyder, G.A., and Sobolev, N. V. (1999) Precise Mössbauer milliprobe determination of ferric iron in rock-forming minerals and limitations of electron microprobe analysis. American Mineralogist, 84, 78–85.

Stagno, V., Ojwang, D.O., McCammon, C.A., and Frost, D.J. (2013) The oxidation state of the mantle and the extraction of carbon from Earth's interior. Nature, 493, 84–88.

Stalder, R., and Ludwig, T. (2007) OH incorporation in synthetic diopside. European Journal of Mineralogy, 19, 373–380.

Su, W., Zhang, M., Redfern, S.A.T., and Bromiley, G.D. (2008) Dehydroxylation of omphacite of eclogite from the Dabie-Sulu. Lithos, 105, 181–190.

Sundvall, R., Skogby, H., and Stalder, R. (2009) Dehydration-hydration mechanisms in synthetic Fe-poor diopside. European Journal of Mineralogy, 21, 17–26.

Terabayashi, M., Matsui, T., Okamoto, K., Ozawa, H., Kaneko, Y., and Maruyama, S. (2013) Micro-X-ray absorption near edge structure determination of $Fe^{3+}/\Sigma Fe$ in omphacite inclusion within garnet from Dabie eclogite, East-Central China. Island Arc, 22, 37–50.

Thomson, A.R., Walter, M.J., Kohn, S.C., and Brooker, R.A. (2016) Slab melting as a barrier to deep carbon subduction. Nature, 529, 76–79.

Thomson, A.R., Kohn, S.C., Prabhu, A., and Walter, M.J. (2021) Evaluating the formation pressure of diamond-hosted majoritic garnets: A machine learning majorite barometer. Journal of Geophysical Research: Solid Earth, 126, e2020JB020604.

van Aken, P.A., and Liebscher, B. (2002) Quantification of ferrous/ferric ratios in minerals: new evaluation schemes of Fe $L_{23}$ electron energy-loss near-edge spectra. Physics and Chemistry of Minerals, 29, 188–200.

van Aken, P.A., Liebscher, B., and Styrsa, V.J. (1998) Quantitative determination of iron oxidation states in minerals using Fe $L_{2,3}$-edge electron energy-loss near-edge structure spectroscopy. Physics and Chemistry of Minerals, 25, 323–327.

Wade, J.A., Plank, T., Hauri, E.H., Kelley, K.A., Roggensack, K., and Zimmer, M. (2008) Prediction of magmatic water contents via measurement of $H_2O$ in clinopyroxene phenocrysts. Geology, 36, 799–802.

Woodland, A.B. (2009) Ferric iron contents of clinopyroxene from cratonic mantle and partitioning behaviour with garnet. Lithos, 112, 1143–1149.

Woodland, A.B., Kornprobst, J., and Tabit, A. (2006) Ferric iron in orogenic lherzolite massifs and controls of oxygen fugacity in the upper mantle. Lithos, 89, 222–241.

Xia, Q.K., Liu, J., Liu, S.C., Kovács, I., Feng, M., and Dang, L. (2013) High water content in Mesozoic primitive basalts of the North China Craton and implications on the destruction of cratonic mantle lithosphere. Earth and Planetary Science Letters, 361, 85–97.

Xu, Y., Tang, W., Hui, H., Rudnick, R.L., Shang, S., and Zhang, Z. (2019) Reconciling the discrepancy between the dehydration rates in mantle olivine and pyroxene during xenolith emplacement. Geochimica et Cosmochimica Acta, 267, 179–195.

Yavuz, F., and Yildirim, D.K. (2020) WinGrt, a Windows program for garnet supergroup minerals. Journal of Geosciences, 65, 71–95.

Zhang, C., Almeev, R.R., Hughes, E.C., Borisov, A.A., Wolff, E.P., Höfer, H.E., Botcharnikov, R.E., and Koepke, J. (2018) Electron microprobe technique for the determination of iron oxidation state in silicate glasses. American Mineralogist, 103, 1445–1454.

Zhang, H.L., Cottrell, E., Solheid, P.A., Kelley, K.A., and Hirschmann, M.M. (2018) Determination of $Fe^{3+}/\Sigma Fe$ of XANES basaltic glass standards by Mössbauer spectroscopy and its application to the oxidation state of iron in MORB. Chemical Geology, 479, 166–175.

Zhang, Z., Qin, T., Pommier, A., and Hirschmann, M.M. (2019) Carbon storage in Fe-Ni-S liquids in the deep upper mantle and its relation to diamond and Fe-Ni alloy precipitation. Earth and Planetary Science Letters, 520, 164–174.

**Figure 1.** Clinopyroxene compositions: **(a)** Mg-Fe-Ca clinopyroxenes and **(b)** Ca-Na and Na clinopyroxenes. Data sources are listed in Supplementary Table S1. Abbreviations: Wo, wollastonite; En, enstatite; Fs, ferrosilite; Q, Wo + En + Fs; Di, diopside; Hd, hedenbergite; Aug, augite; Pgt, pigeonite; Cen, clinoenstatite; Cfs, clinoferrosilite; Omp, omphacite; Agt, aegirine-augite; Jd, jadeite; Ae, aegirine.

**Figure 2.** Performance of the seven ML algorithms trained on the entire dataset. **(a)** Kernel density estimation of the coefficient of determination ($R^2$) from 10-fold cross-validation. **(b–h)** Correlations between the known $Fe^{3+}$ abundances in the training and test datasets with those predicted by the linear, polynomial, neural network, ANN ensemble, decision trees, random forest, and Extra-Trees algorithms, respectively.

**Figure 3.** Residual error (reference value − predicted value) as a function of **(a–g)** Fe and **(h–n)** Si contents (atoms per formula unit, apfu) in reference clinopyroxenes: **(a, h)** linear regression, **(b,**

**i)** polynomial regression, **(c, j)** neural network regression, **(d, k)** ANN ensemble regression, **(e, l)** decision trees regression, **(f, m)** random forest regression, and **(g, n)** Extra-Trees regression.

**Figure 4.** Comparison of model performance for sub-dataset I, clinopyroxene in spinel peridotite ($n$ = 127). Models were trained using either **(a, c, e)** a dataset of exclusively clinopyroxene compositions (cpx only) or **(b, d, f)** one including both clinopyroxene and coexisting spinel compositions (cpx + spl). Correlations are shown between the known $Fe^{3+}$ abundances in the training and test datasets and those predicted by **(a, b)** linear regression, **(c, d)** random forest regression, and **(e, f)** Extra-Trees regression.

**Figure 5.** Relative feature importance in **(a)** random forest and **(b)** Extra-Trees predictions of $Fe^{3+}/\Sigma Fe$ for clinopyroxene in spinel peridotite (sub-dataset I).

**Figure 6.** Correlations between the known $Fe^{3+}$ abundances in omphacite in eclogite (sub-dataset II) and those predicted by the **(a)** polynomial, **(b)** neural network, **(c)** ANN ensemble, **(d)** random forest, and **(e)** Extra-Trees regressions.

**Figure 7.** Performance of the seven ML algorithms trained on sub-dataset III (i.e., excluding clinopyroxene in spinel peridotite and omphacite in eclogite). **(a)** Kernel density estimation of the coefficient of determination ($R^2$) from 10-fold cross-validation. **(b–h)** Correlations between the known $Fe^{3+}$ abundances in the training and test datasets and those predicted by the linear, polynomial, neural network, ANN ensemble, decision trees, random forest, and Extra-Trees regressions, respectively.

**Figure 8.** Relative feature importance in random forest predictions of $Fe^{3+}/\Sigma Fe$ for clinopyroxene in **(a)** the entire dataset and **(b)** sub-dataset III (excluding clinopyroxene in spinel peridotite and

omphacite in eclogite).

**Figure 9.** Elemental partitioning between clinopyroxene and spinel. Data on **(a)** Cr and **(b)** Al are plotted based on a 12-oxygen formula unit. Partitioning coefficient pairs are plotted in **(c)** $D^{\mathrm{spl/cpx}}{}_{\mathrm{Cr}}$ versus $D^{\mathrm{spl/cpx}}{}_{\mathrm{Fe}}$ and **(d)** $D^{\mathrm{spl/cpx}}{}_{\mathrm{Al}}$ versus $D^{\mathrm{spl/cpx}}{}_{\mathrm{Fe}}$.

**Figure 10.** Comparison of predicted $Fe^{3+}/\Sigma Fe$ values in clinopyroxene using our **(a, b)** polynomial, **(c, d)** neural network, **(e, f)** ANN ensemble, **(g, h)** random forest, and **(i, j)** Extra-Trees methods with those predicted by Droop's (1987) method for **(a, c, e, g, i)** models trained using the entire dataset and **(b, d, f, h, j)** those trained using sub-dataset III (excluding clinopyroxene in spinel peridotite and omphacite in eclogite).

**Figure 11.** Application of our **(a)** polynomial, **(b)** neural network, **(c)** ANN ensemble, **(d)** random forest, and **(e)** Extra-Trees models to the garnet-clinopyroxene geothermometer (Ellis and Green 1979). Red symbols are temperatures obtained considering all Fe as $Fe^{2+}$, black symbols are those using $Fe^{2+}$ contents predicted by our models trained on the entire dataset, and blue symbols are those using $Fe^{2+}$ contents predicted by Droop's (1987) method. The temperatures obtained using these various methods are compared to those calculated directly from Mössbauer spectroscopy data (MS, *y*-axis).

**Figure 12.** Comparison of **(a–c)** $Fe^{3+}/\Sigma Fe$ values as a function of Fe content (atoms per formula unit, apfu) and **(d–f)** kernel density estimations of the distribution of $Fe^{3+}/\Sigma Fe$ values for dehydrogenated and non-dehydrogenated clinopyroxenes predicted by **(a, d)** polynomial, **(b, e)** random forest, and **(c,f)** Extra-Trees regressions.

Table 1. Recoil-free fractions of Fe in clinopyroxene (*C*) at room temperature used to correct literature $Fe^{3+}/\Sigma Fe$ ratios.

| **Species** | ***C*** | **Application to this study** | **Reference** |
|---|---|---|---|
| Di (N) | 1.22 | Di | De Grave and Van Alboom (1991) |
| Hd (N) | 1.17 | Hd | Eeckhout and De Grave (2003) |
| Ae (N) | 1.15 | Ae | Eeckhout and De Grave (2003) |
| Avg | 1.18 | Aug, Pgt, Omp, Jd, Agt | Eeckhout and De Grave (2003) |

Abbreviations: N, natural sample; Avg, average value; Di, diopside; Hd, hedenbergite; Ae, aegirine; Aug, augite; Pgt, pigeonite; Omp, omphacite; Jd, jadeite; Agt, aegirine-augite.

Table 2. Temperatures calculated using various formulations of the garnet-clinopyroxene geothermometer.

| | **Geothermometer formulation** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Sample No.** | **TRG74 (°C)** | **TEG79 (°C)** | **TG79 (°C)** | **TP85 (°C)** | **TK88 (°C)** | **TA94 (°C)** | **TG96 (°C)** | **TL98 (°C)** | **TKR00 (°C)** | **TN09 (°C)** |
| Cpx96[a] | 704 | 693 | 744 | 670 | 647 | 615 | 753 | 625 | 693 | 707 |
| Cpx98[a] | 686 | 673 | 725 | 650 | 626 | 592 | 735 | 603 | 672 | 688 |
| Cpx99[a] | 679 | 666 | 718 | 642 | 618 | 584 | 728 | 595 | 664 | 681 |
| Cpx104[a] | 695 | 684 | 735 | 660 | 637 | 604 | 744 | 614 | 683 | 698 |
| Cpx107[a] | 707 | 697 | 747 | 673 | 651 | 619 | 756 | 628 | 697 | 710 |
| Cpx108[a] | 706 | 695 | 745 | 672 | 649 | 617 | 755 | 627 | 695 | 709 |
| F-8[b] | 1400 | 1256 | 1290 | 1247 | 1243 | 1268 | 1318 | 1293 | 1421 | 1398 |
| F-11[b] | 1423 | 1252 | 1286 | 1241 | 1237 | 1290 | 1313 | 1300 | 1412 | 1414 |
| F-16[b] | 1492 | 1256 | 1307 | 1244 | 1207 | 1310 | 1330 | 1324 | 1441 | 1469 |
| 87-70[c] | 1417 | 1240 | 1265 | 1228 | 1238 | 1297 | 1295 | 1294 | 1384 | 1418 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| BY-19[c] | 1275 | 1297 | 1328 | 1302 | 1302 | 1212 | 1353 | 1250 | 1459 | 1407 |
| 8508-9[c] | 1181 | 1187 | 1233 | 1184 | 1167 | 1064 | 1260 | 1103 | 1314 | 1282 |
| 94M80[d] | 595 | 670 | 677 | 650 | 622 | 614 | 672 | 606 | 630 | 590 |
| 97M32[d] | 513 | 519 | 562 | 495 | 469 | 456 | 560 | 447 | 507 | 505 |
| 944012-11[d] | 678 | 883 | 818 | 870 | 774 | 767 | 827 | 748 | 785 | 679 |

Garnet-clinopyroxene geothermometer formulations: TRG74, Råheim and Green (1974); TEG79, Ellis and Green (1979); TG79, Ganguly (1979); TP85, Powell (1985); TK88, Krogh (1988); TA94, Ai (1994); TG96, Ganguly et al. (1996); TL98, Liu (1998); TKR00, Ravna (2000); TN09, Nakamura (2009).

Data references: [a]Proyer et al. (2004); [b]Lazarov et al. (2009); [c]Nimis et al. (2015); [d]Li et al. (2005).

Figure 1.

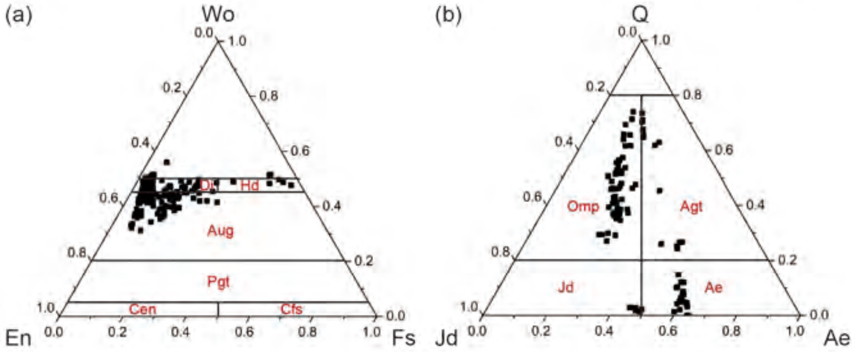

(a)
Wo
0.0
1.0
0.2
0.8
0.4
0.6
0.6
0.4
0.8
0.2
1.0
0.0
Di
Hd
Aug
Pgt
Cen
Cfs
En
0.0
0.2
0.4
0.6
0.8
1.0
Fs
(b)
Q
0.0
1.0
0.2
0.8
0.4
0.6
0.6
0.4
0.8
0.2
1.0
0.0
Omp
Agt
Jd
Ae
Jd
0.0
0.2
0.4
0.6
0.8
1.0
Ae

Figure 2.

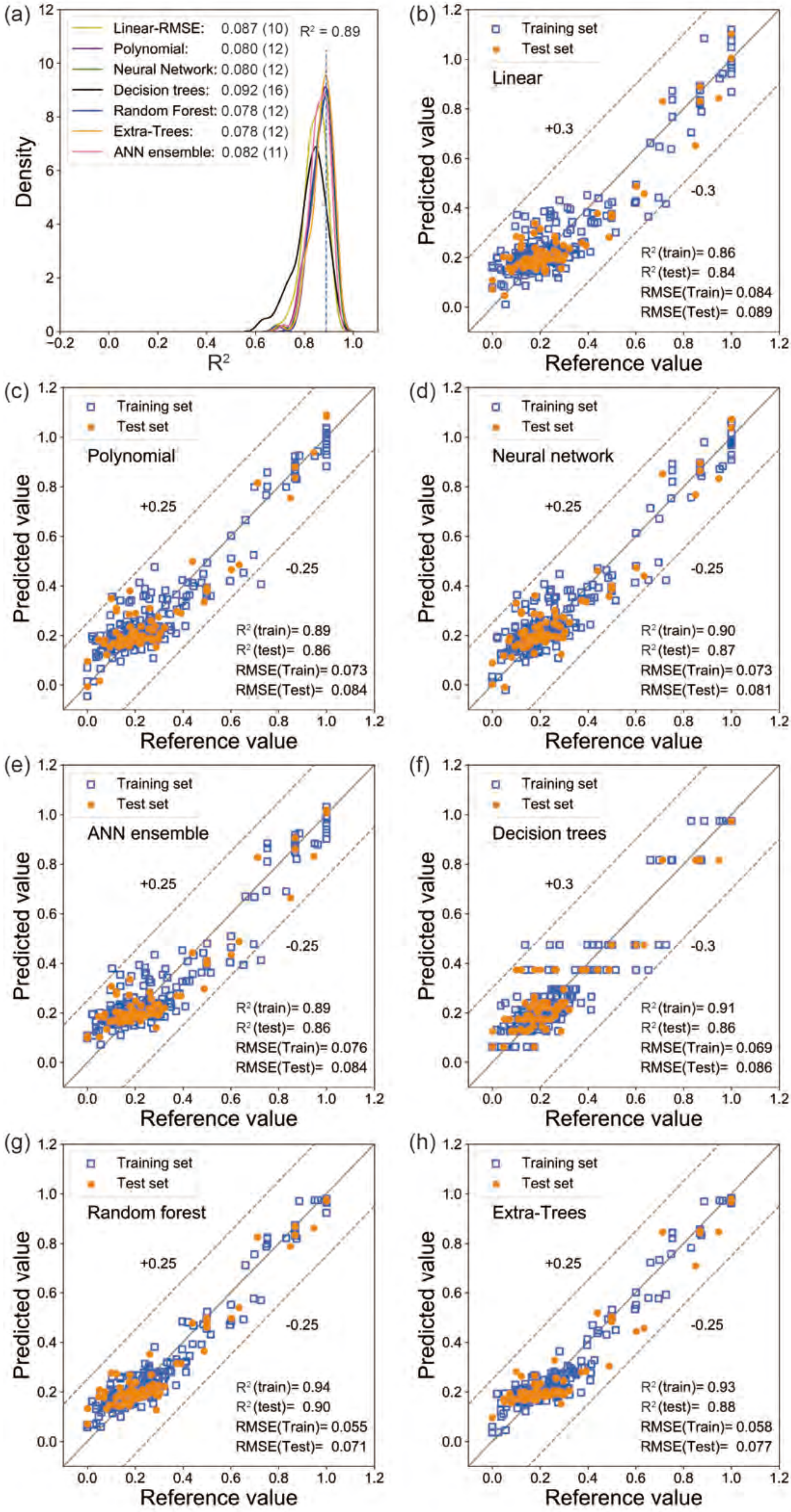

(a)
Linear-RMSE: 0.087 (10)
Polynomial: 0.080 (12)
Neural Network: 0.080 (12)
Decision trees: 0.092 (16)
Random Forest: 0.078 (12)
Extra-Trees: 0.078 (12)
ANN ensemble: 0.082 (11)
$R^2$ = 0.89
Density
$R^2$
(b)
Training set
Test set
Linear
+0.3
-0.3
$R^2$(train)= 0.86
$R^2$(test)= 0.84
RMSE(Train)= 0.084
RMSE(Test)= 0.089
(c)
Polynomial
+0.25
-0.25
$R^2$(train)= 0.89
$R^2$(test)= 0.86
RMSE(Train)= 0.073
RMSE(Test)= 0.084
(d)
Neural network
+0.25
-0.25
$R^2$(train)= 0.90
$R^2$(test)= 0.87
RMSE(Train)= 0.073
RMSE(Test)= 0.081
(e)
ANN ensemble
+0.25
-0.25
$R^2$(train)= 0.89
$R^2$(test)= 0.86
RMSE(Train)= 0.076
RMSE(Test)= 0.084
(f)
Decision trees
+0.3
-0.3
$R^2$(train)= 0.91
$R^2$(test)= 0.86
RMSE(Train)= 0.069
RMSE(Test)= 0.086
(g)
Random forest
+0.25
-0.25
$R^2$(train)= 0.94
$R^2$(test)= 0.90
RMSE(Train)= 0.055
RMSE(Test)= 0.071
(h)
Extra-Trees
+0.25
-0.25
$R^2$(train)= 0.93
$R^2$(test)= 0.88
RMSE(Train)= 0.058
RMSE(Test)= 0.077
Predicted value
Reference value

Figure 3.

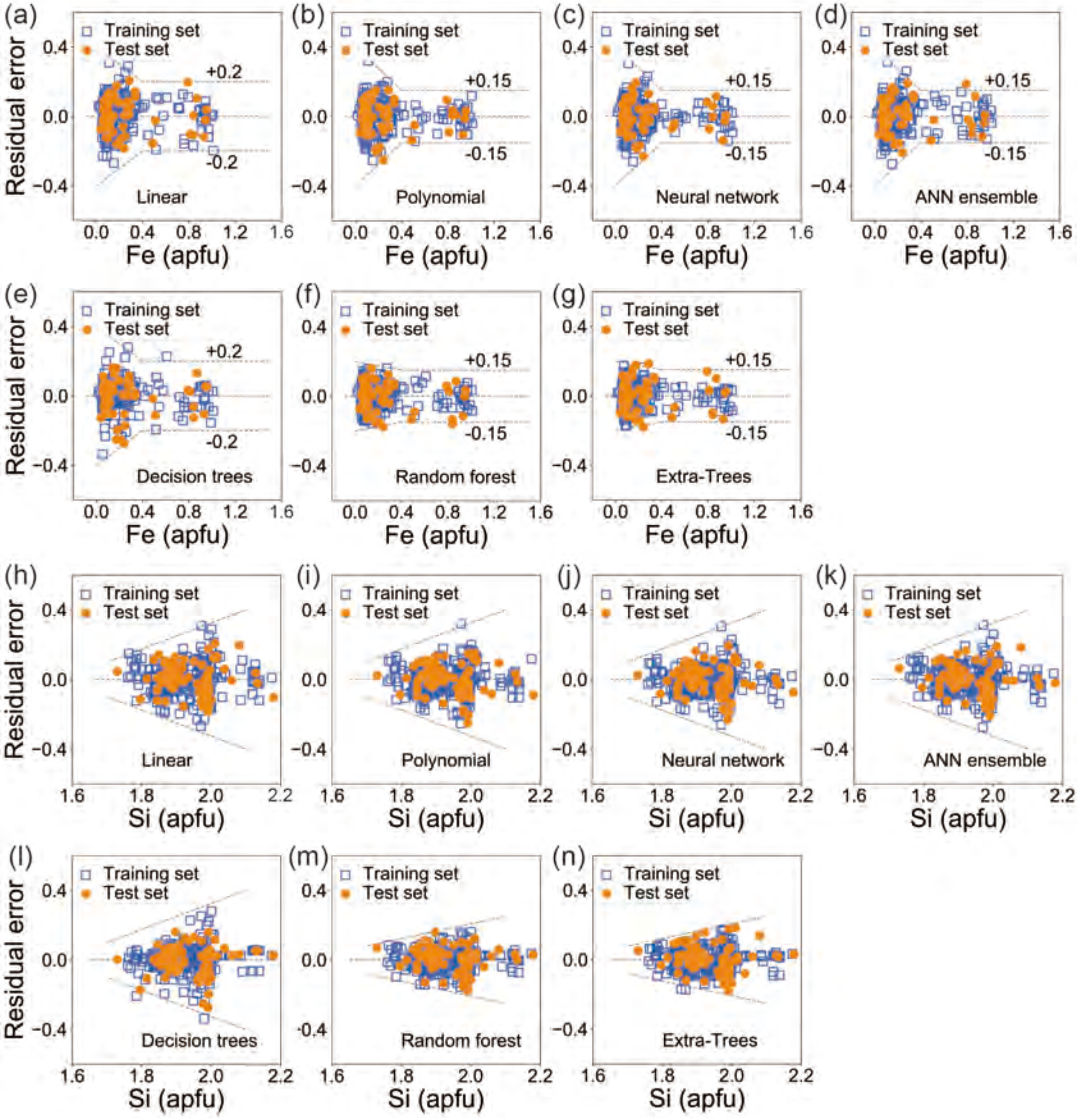

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
□ Training set
● Test set
Residual error
0.4
0.0
−0.4
+0.2
-0.2
+0.15
-0.15
Linear
Polynomial
Neural network
ANN ensemble
Decision trees
Random forest
Extra-Trees
0.0
0.4
0.8
1.2
1.6
Fe (apfu)
1.6
1.8
2.0
2.2
Si (apfu)

Figure 4.

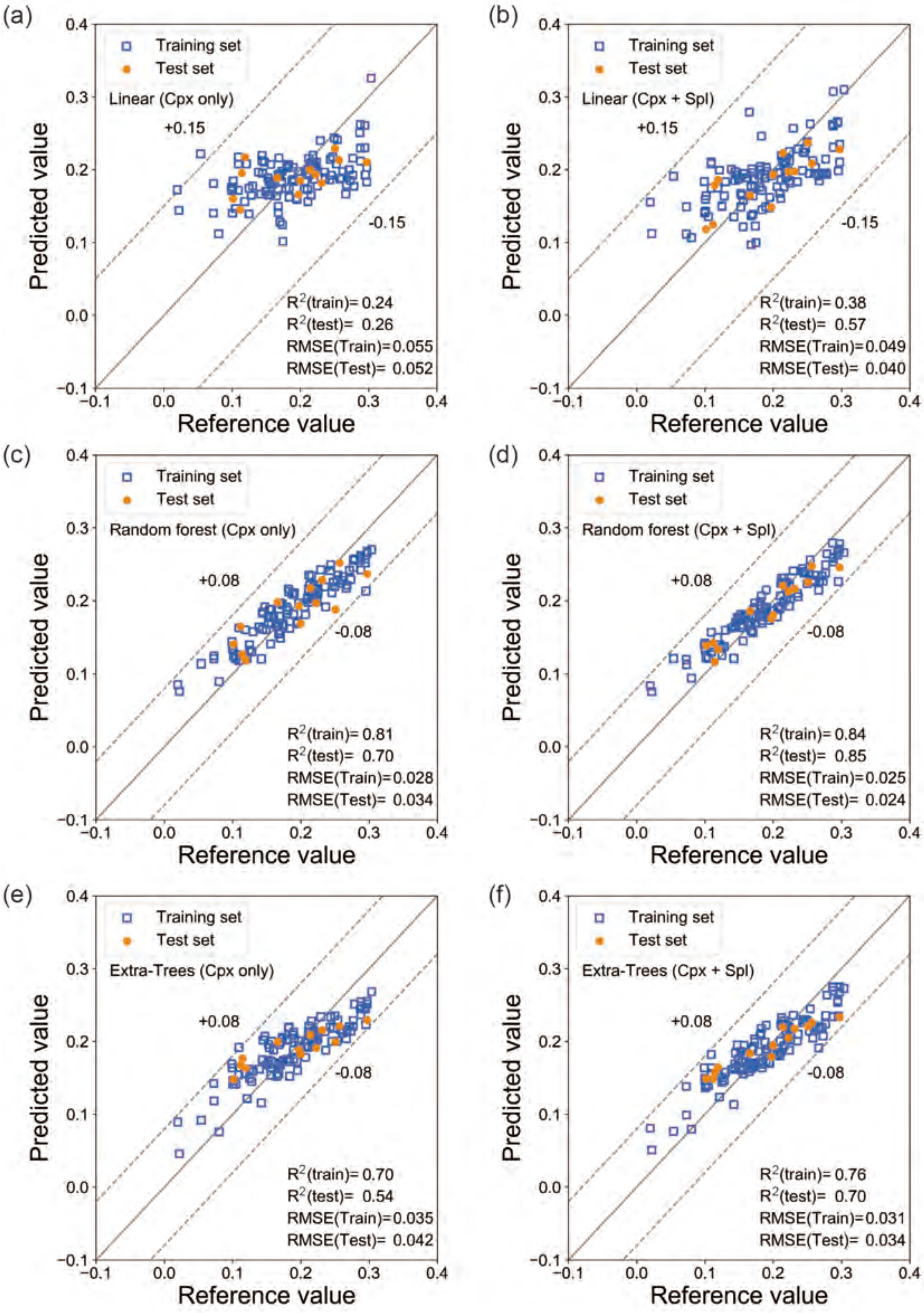

(a)
Training set
Test set
Linear (Cpx only)
+0.15
-0.15
R$^2$(train)= 0.24
R$^2$(test)= 0.26
RMSE(Train)=0.055
RMSE(Test)= 0.052
(b)
Training set
Test set
Linear (Cpx + Spl)
+0.15
-0.15
R$^2$(train)= 0.38
R$^2$(test)= 0.57
RMSE(Train)=0.049
RMSE(Test)= 0.040
(c)
Training set
Test set
Random forest (Cpx only)
+0.08
-0.08
R$^2$(train)= 0.81
R$^2$(test)= 0.70
RMSE(Train)=0.028
RMSE(Test)= 0.034
(d)
Training set
Test set
Random forest (Cpx + Spl)
+0.08
-0.08
R$^2$(train)= 0.84
R$^2$(test)= 0.85
RMSE(Train)=0.025
RMSE(Test)= 0.024
(e)
Training set
Test set
Extra-Trees (Cpx only)
+0.08
-0.08
R$^2$(train)= 0.70
R$^2$(test)= 0.54
RMSE(Train)=0.035
RMSE(Test)= 0.042
(f)
Training set
Test set
Extra-Trees (Cpx + Spl)
+0.08
-0.08
R$^2$(train)= 0.76
R$^2$(test)= 0.70
RMSE(Train)=0.031
RMSE(Test)= 0.034
Predicted value
Reference value
-0.1
0.0
0.1
0.2
0.3
0.4

Figure 5.

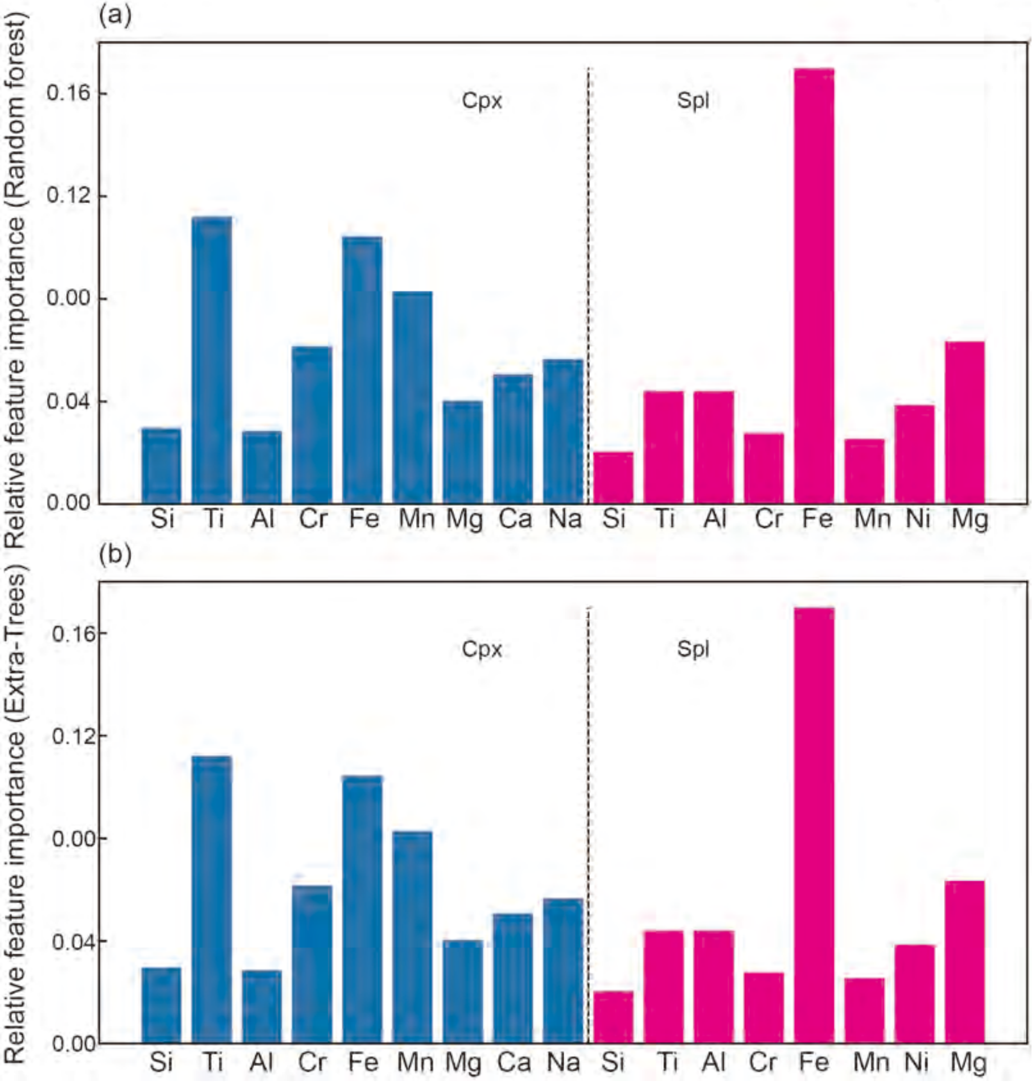

(a)
Relative feature importance (Random forest)
0.16
0.12
0.00
0.04
0.00
Cpx
Spl
Si
Ti
Al
Cr
Fe
Mn
Mg
Ca
Na
Si
Ti
Al
Cr
Fe
Mn
Ni
Mg
(b)
Relative feature importance (Extra-Trees)
0.16
0.12
0.00
0.04
0.00
Cpx
Spl
Si
Ti
Al
Cr
Fe
Mn
Mg
Ca
Na
Si
Ti
Al
Cr
Fe
Mn
Ni
Mg

Figure 6.

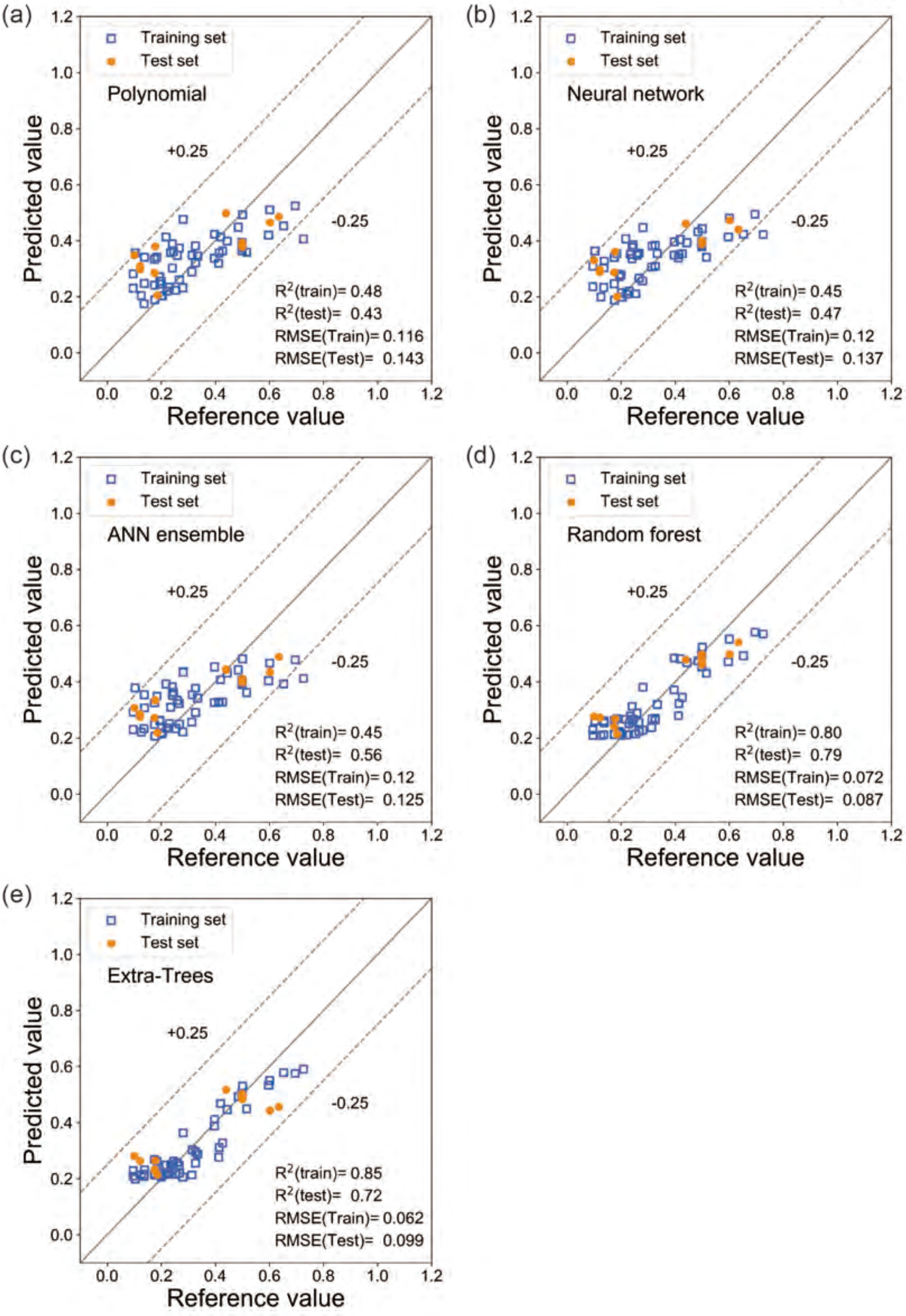

(a)
Training set
Test set
Polynomial
+0.25
-0.25
R$^2$(train)= 0.48
R$^2$(test)= 0.43
RMSE(Train)= 0.116
RMSE(Test)= 0.143
Predicted value
Reference value
(b)
Training set
Test set
Neural network
+0.25
-0.25
R$^2$(train)= 0.45
R$^2$(test)= 0.47
RMSE(Train)= 0.12
RMSE(Test)= 0.137
Predicted value
Reference value
(c)
Training set
Test set
ANN ensemble
+0.25
-0.25
R$^2$(train)= 0.45
R$^2$(test)= 0.56
RMSE(Train)= 0.12
RMSE(Test)= 0.125
Predicted value
Reference value
(d)
Training set
Test set
Random forest
+0.25
-0.25
R$^2$(train)= 0.80
R$^2$(test)= 0.79
RMSE(Train)= 0.072
RMSE(Test)= 0.087
Predicted value
Reference value
(e)
Training set
Test set
Extra-Trees
+0.25
-0.25
R$^2$(train)= 0.85
R$^2$(test)= 0.72
RMSE(Train)= 0.062
RMSE(Test)= 0.099
Predicted value
Reference value
0.0
0.2
0.4
0.6
0.8
1.0
1.2

Figure 7.

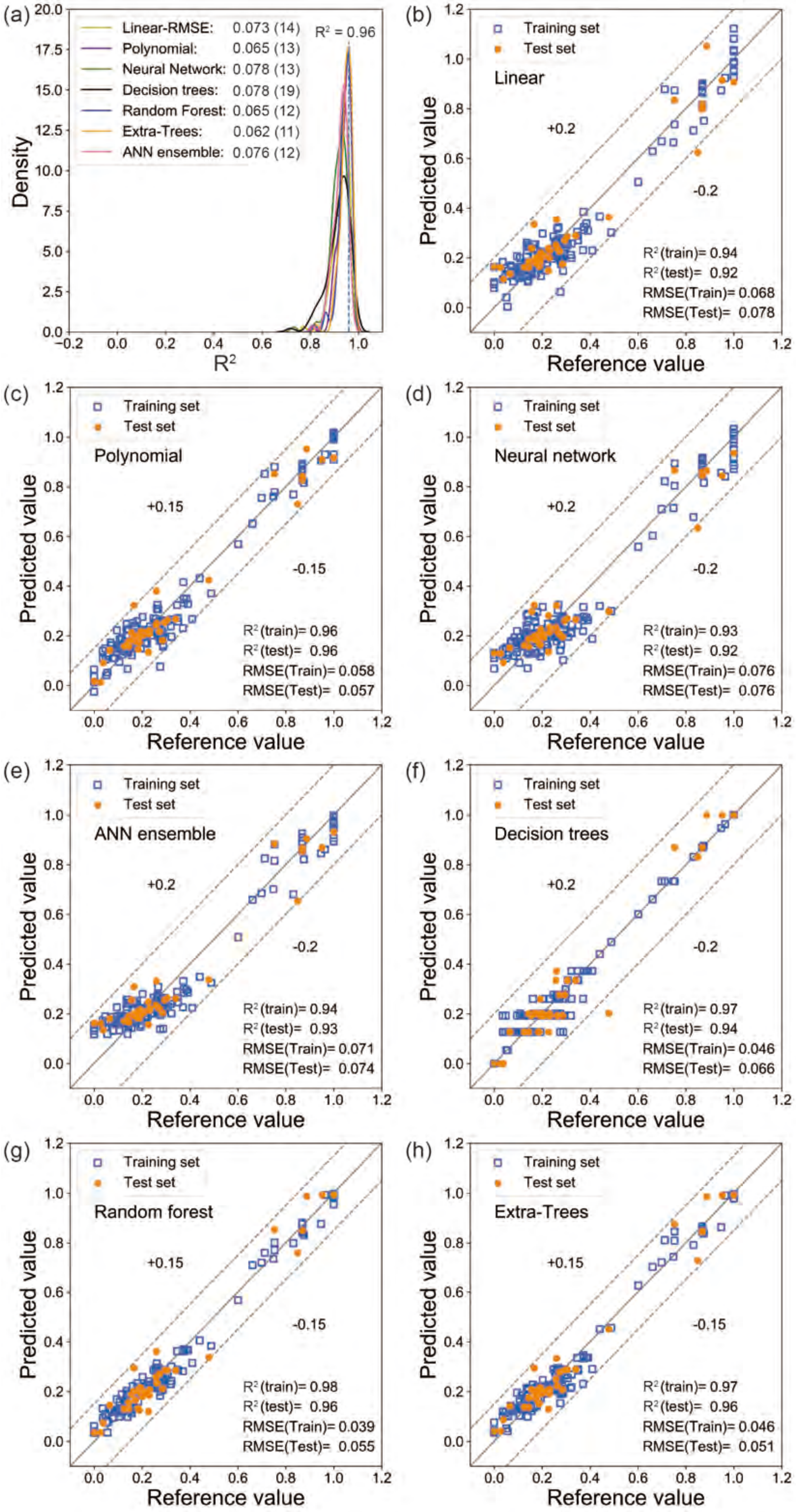

(a)
Density
R²
Linear-RMSE: 0.073 (14)
Polynomial: 0.065 (13)
Neural Network: 0.078 (13)
Decision trees: 0.078 (19)
Random Forest: 0.065 (12)
Extra-Trees: 0.062 (11)
ANN ensemble: 0.076 (12)
R² = 0.96
(b)
Training set
Test set
Linear
+0.2
-0.2
R² (train)= 0.94
R² (test)= 0.92
RMSE(Train)= 0.068
RMSE(Test)= 0.078
(c)
Polynomial
+0.15
-0.15
R² (train)= 0.96
R² (test)= 0.96
RMSE(Train)= 0.058
RMSE(Test)= 0.057
(d)
Neural network
+0.2
-0.2
R² (train)= 0.93
R² (test)= 0.92
RMSE(Train)= 0.076
RMSE(Test)= 0.076
(e)
ANN ensemble
+0.2
-0.2
R² (train)= 0.94
R² (test)= 0.93
RMSE(Train)= 0.071
RMSE(Test)= 0.074
(f)
Decision trees
+0.2
-0.2
R² (train)= 0.97
R² (test)= 0.94
RMSE(Train)= 0.046
RMSE(Test)= 0.066
(g)
Random forest
+0.15
-0.15
R² (train)= 0.98
R² (test)= 0.96
RMSE(Train)= 0.039
RMSE(Test)= 0.055
(h)
Extra-Trees
+0.15
-0.15
R² (train)= 0.97
R² (test)= 0.96
RMSE(Train)= 0.046
RMSE(Test)= 0.051
Predicted value
Reference value

Figure 8.

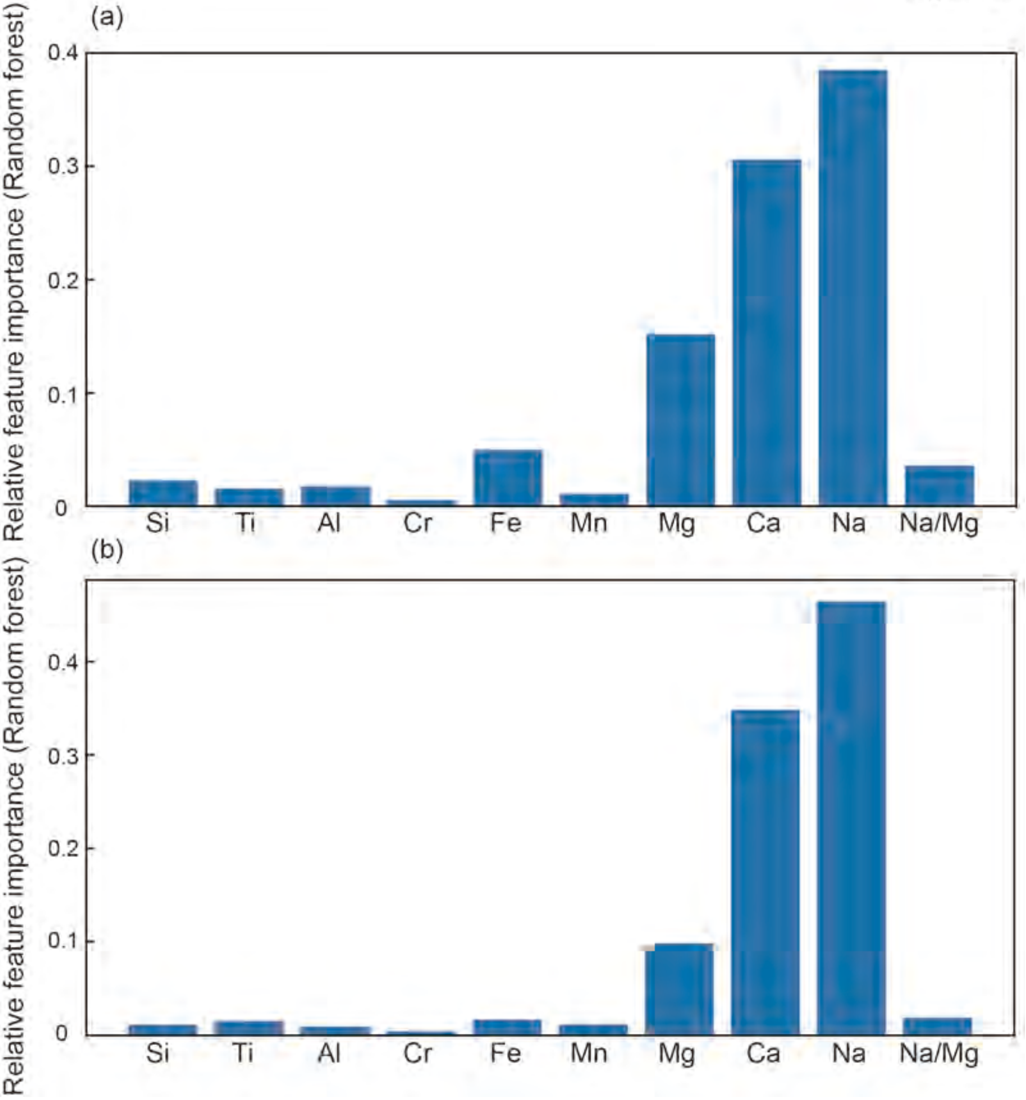

(a)
Relative feature importance (Random forest)
0.4
0.3
0.2
0.1
0
Si
Ti
Al
Cr
Fe
Mn
Mg
Ca
Na
Na/Mg
(b)
Relative feature importance (Random forest)
0.4
0.3
0.2
0.1
0
Si
Ti
Al
Cr
Fe
Mn
Mg
Ca
Na
Na/Mg

Figure 9.

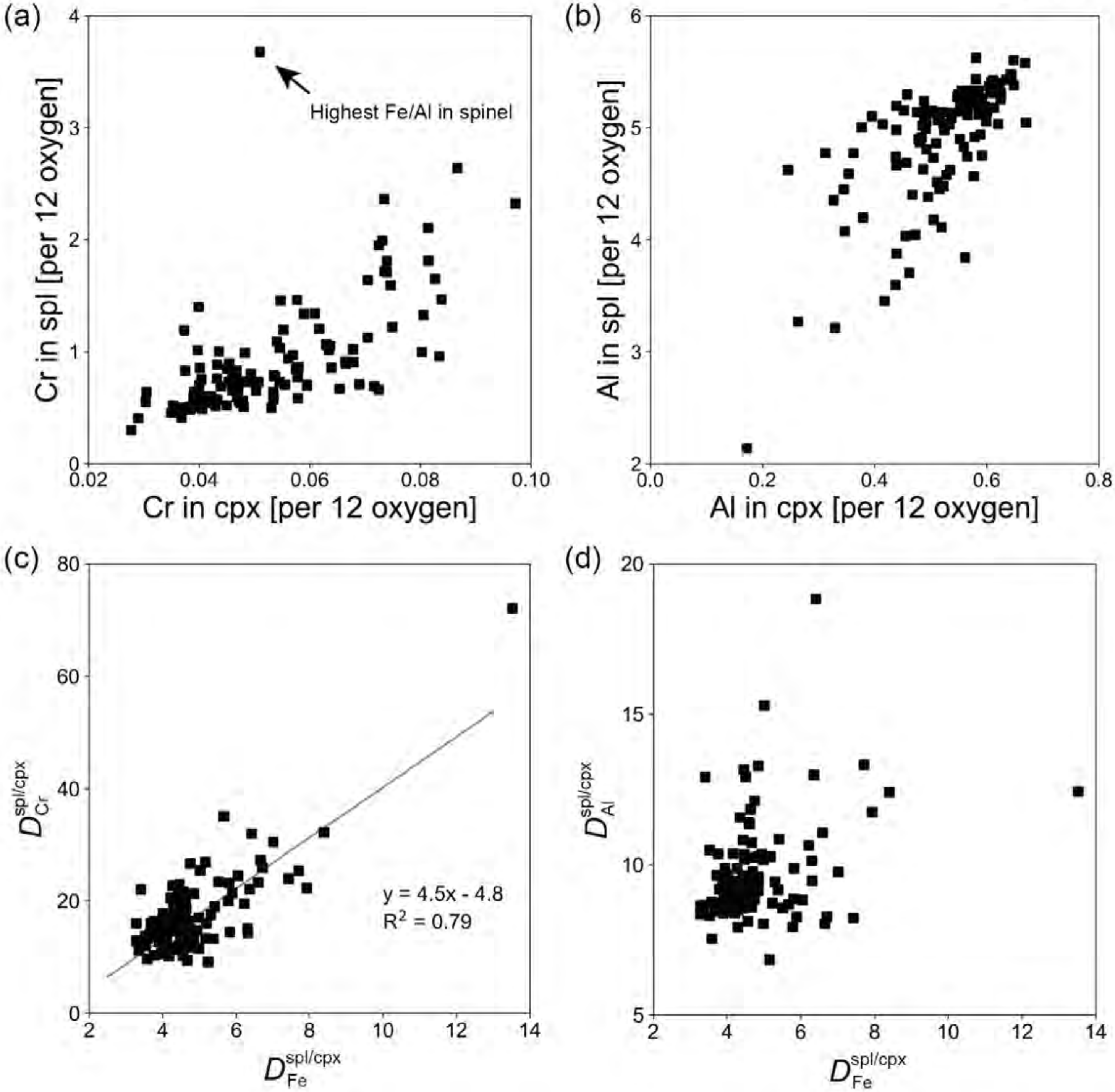

(a)
Highest Fe/Al in spinel
Cr in spl [per 12 oxygen]
Cr in cpx [per 12 oxygen]
0
1
2
3
4
0.02
0.04
0.06
0.08
0.10
(b)
Al in spl [per 12 oxygen]
Al in cpx [per 12 oxygen]
2
3
4
5
6
0.0
0.2
0.4
0.6
0.8
(c)
y = 4.5x - 4.8
R² = 0.79
D_Cr^spl/cpx
D_Fe^spl/cpx
0
20
40
60
80
2
4
6
8
10
12
14
(d)
D_Al^spl/cpx
D_Fe^spl/cpx
5
10
15
20
2
4
6
8
10
12
14

Figure 10.

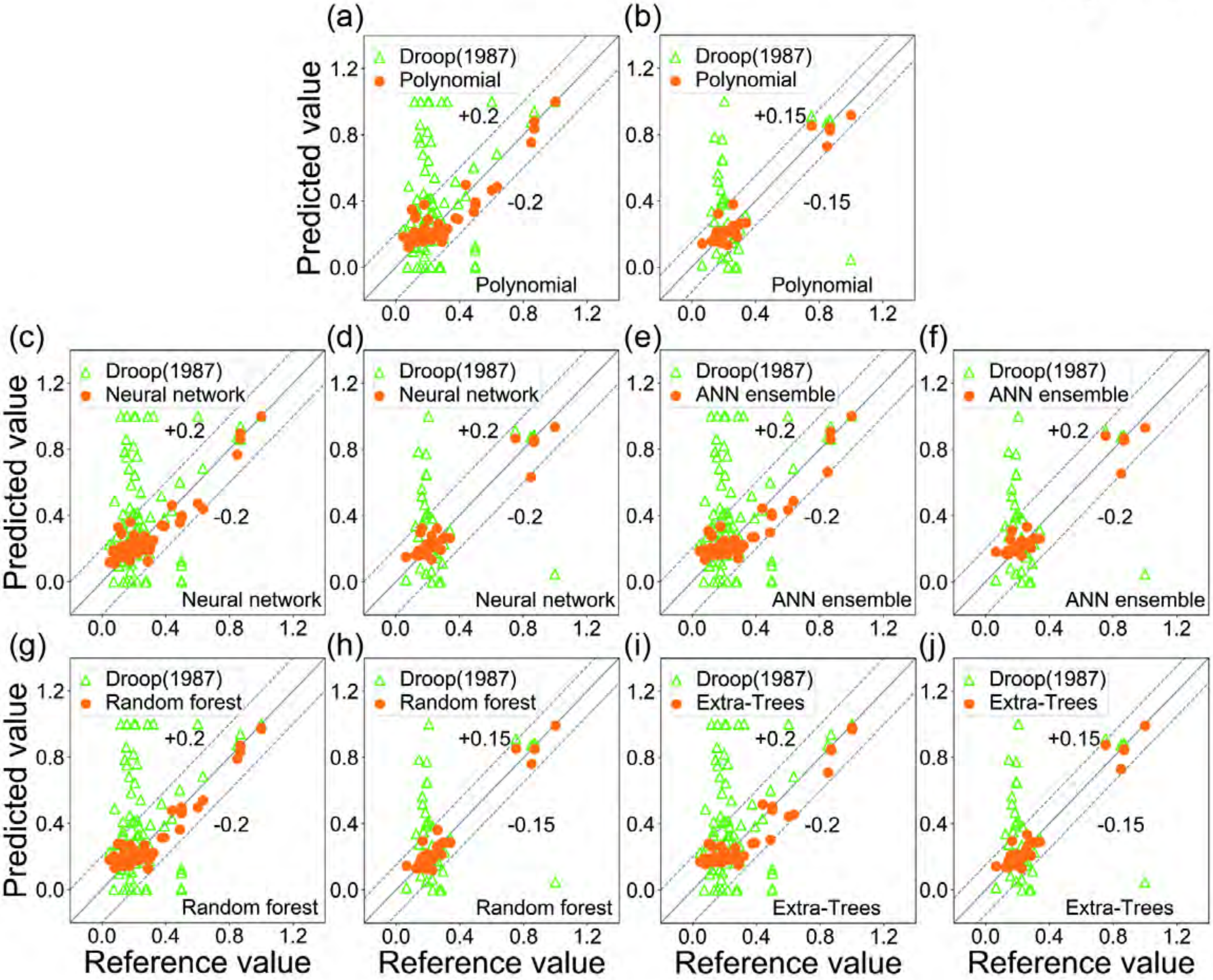

(a)
Droop(1987)
Polynomial
+0.2
-0.2
Polynomial
(b)
Droop(1987)
Polynomial
+0.15
-0.15
Polynomial
(c)
Droop(1987)
Neural network
+0.2
-0.2
Neural network
(d)
Droop(1987)
Neural network
+0.2
-0.2
Neural network
(e)
Droop(1987)
ANN ensemble
+0.2
-0.2
ANN ensemble
(f)
Droop(1987)
ANN ensemble
+0.2
-0.2
ANN ensemble
(g)
Droop(1987)
Random forest
+0.2
-0.2
Random forest
(h)
Droop(1987)
Random forest
+0.15
-0.15
Random forest
(i)
Droop(1987)
Extra-Trees
+0.2
-0.2
Extra-Trees
(j)
Droop(1987)
Extra-Trees
+0.15
-0.15
Extra-Trees
Predicted value
0.0
0.4
0.8
1.2
Reference value
0.0
0.4
0.8
1.2

Figure 11.

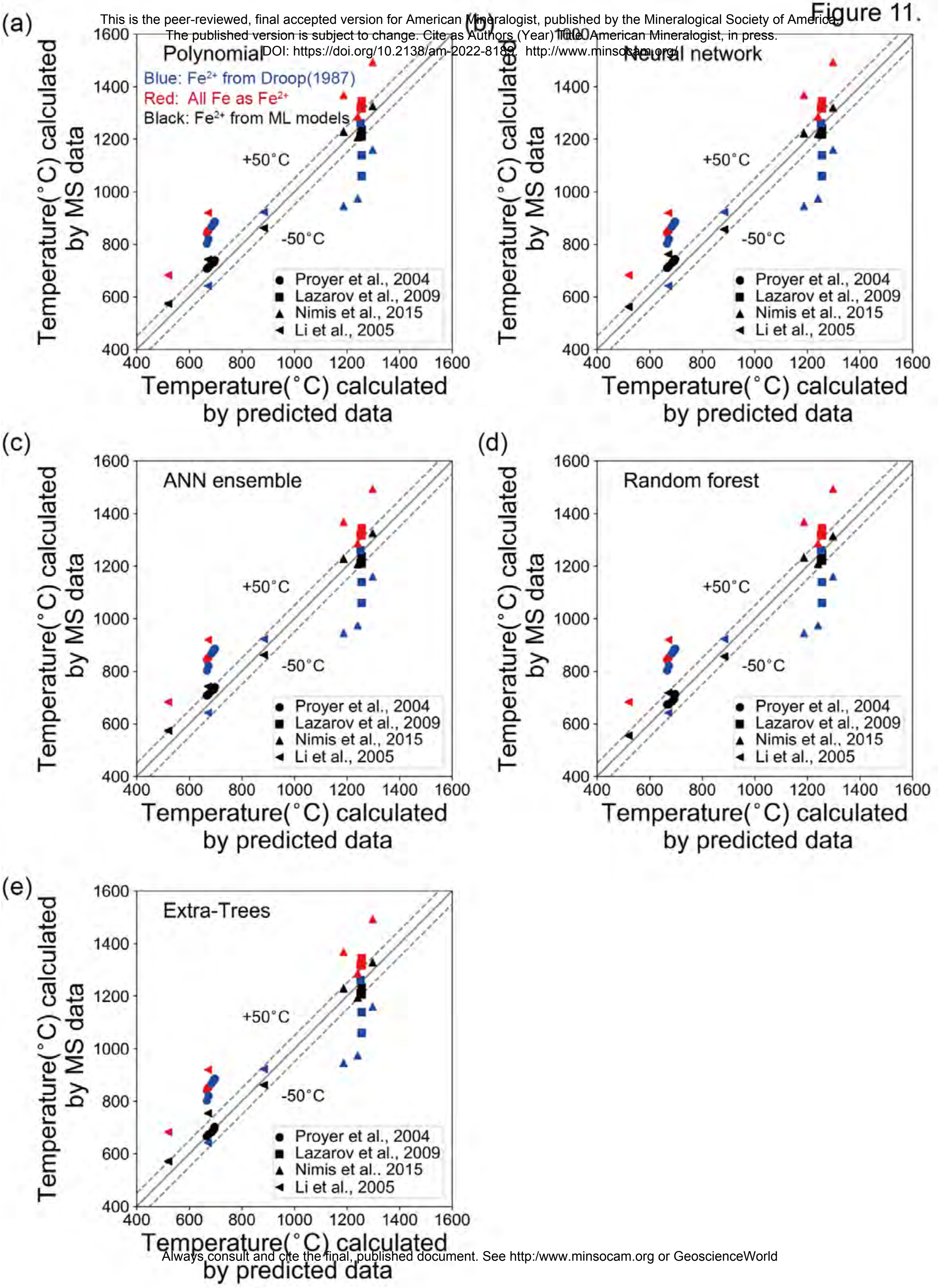
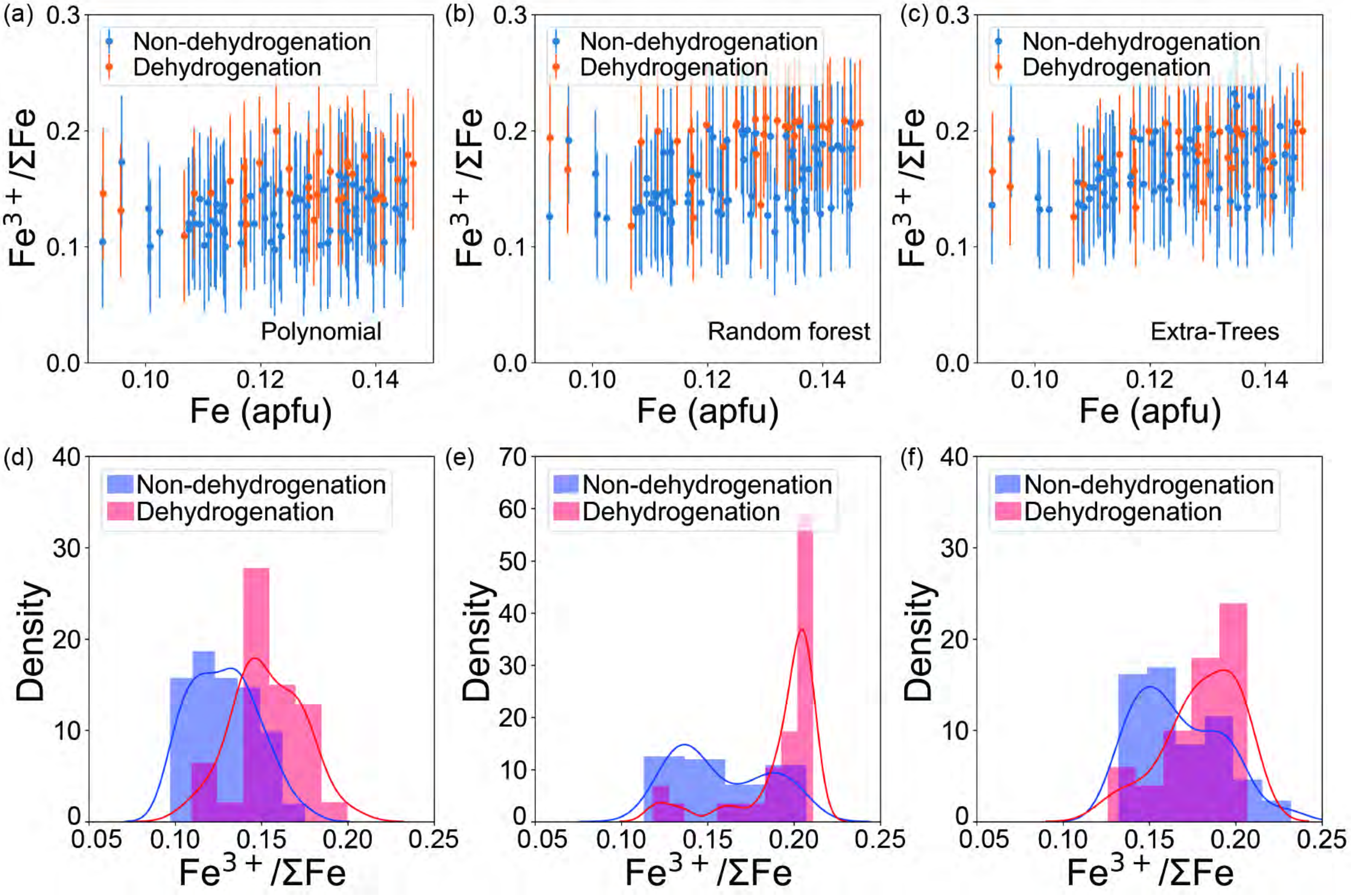

(a)
Polynomial
Blue: Fe²⁺ from Droop(1987)
Red: All Fe as Fe²⁺
Black: Fe²⁺ from ML models
(b)
Neural network
(c)
ANN ensemble
(d)
Random forest
(e)
Extra-Trees
+50°C
-50°C
Proyer et al., 2004
Lazarov et al., 2009
Nimis et al., 2015
Li et al., 2005
Temperature(°C) calculated by MS data
Temperature(°C) calculated by predicted data
400
600
800
1000
1200
1400
1600

Figure 12.


(a)
(b)
(c)
Non-dehydrogenation
Dehydrogenation
Polynomial
Random forest
Extra-Trees
$Fe^{3+}/\Sigma Fe$
0.0
0.1
0.2
0.3
Fe (apfu)
0.10
0.12
0.14
(d)
(e)
(f)
Density
0
10
20
30
40
50
60
70
$Fe^{3+}/\Sigma Fe$
0.05
0.10
0.15
0.20
0.25