Estimating ferric iron content in clinopyroxene using machine learning models

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

\textsuperscript{1} School of Earth Sciences, Zhejiang University, Hangzhou 310027, China.
\textsuperscript{2} Key Laboratory of Big Data and Deep Earth Resources of Zhejiang Province, Hangzhou 310027, China.
\textsuperscript{3} Gemological Institute, China University of Geosciences, Wuhan 430074, China.
\textsuperscript{4} School of Geoscience and Info-Physics, Central South University, Changsha 410083, China.
\textsuperscript{5} School of Earth Sciences and Engineering, Sun Yat-Sen University, Guangzhou, 510275, China.
\textsuperscript{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$^{3+}$/ΣFe in clinopyroxene using major element data from electron microprobe analyses. The models were first trained using a large dataset of clinopyroxene Fe$^{3+}$/ΣFe 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):

$$\frac{1}{3}\text{Mg}_3\text{Al}_2\text{Si}_3\text{O}_{12} + \text{CaFeSi}_2\text{O}_6 \leftrightarrow \frac{1}{3}\text{Fe}_3\text{Al}_2\text{Si}_3\text{O}_{12} + \text{CaMgSi}_2\text{O}_6,$$

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:

$$\text{OH}^- + \text{Fe}^{2+} = \text{O}^{2-} + \text{Fe}^{3+} + \frac{1}{2}\text{H}_2 .$$

This reaction is crucial for evaluating the preservation of water in clinopyroxene phenocrysts, which constrains the H$_2$O 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$^{3+}$/ΣFe 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$^{3+}$/ΣFe 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$^{3+}$/ΣFe = 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$^{3+}$/ΣFe ratios obtained by Mössbauer spectroscopy as the ‘tag’ values ($Y$) representing the model output. Among the analytical methods capable of measuring Fe$^{3+}$/ΣFe, 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$_2$O$_3$, 0–29.37 wt% Cr$_2$O$_3$, 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$_2$O, and Fe$^{3+}$/ΣFe 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 $\text{Fe}^{3+}/\Sigma\text{Fe}$ ratios reported in the literature ignore the recoil-free fractions of $\text{Fe}^{3+}$ and $\text{Fe}^{2+}$, which may contribute large errors to $\text{Fe}^{3+}/\Sigma\text{Fe}$ 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:

$$\frac{A_{3+}}{A_{2+}} = C \times \frac{N_{3+}}{N_{2+}},$$

where $A$ is the doublet area of $\text{Fe}^{3+}$ or $\text{Fe}^{2+}$ 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:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$ (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:
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 $\text{Fe}^{3+}/\Sigma\text{Fe}$ 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$^{3+}$/ΣFe 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\(^{3+}/\Sigma\text{Fe}\) (below ±0.3), indicating that in these samples, Fe\(^{3+}/\Sigma\text{Fe}\) 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\(^{3+}/\Sigma\text{Fe}\) 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\(^{3+}\)/ΣFe 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 $\text{Fe}^{3+}/\Sigma\text{Fe}$ 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 $\text{Fe}^{3+}$ substitution mechanisms in omphacite. Unlike clinopyroxene in spinel peridotite, $\text{Na}^+$ abundance is strongly correlated with $\text{Fe}^{3+}/\Sigma\text{Fe}$ 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 $\text{Fe}^{3+}$ and $\text{Na}^+$, $\text{Mg}^{2+}$, and $\text{Ca}^{2+}$. 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:

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

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

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:

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

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

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$^{3+}$/ΣFe (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$^{3+}$/ΣFe for the entire dataset. The general model indicates that Fe$^{3+}$ is incorporated in clinopyroxene as NaFe$^{3+}$Si$_2$O$_6$ (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$^{3+}$/ΣFe 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 $D_{\text{spl/cpx}}^{\text{Fe}}$ and $D_{\text{spl/cpx}}^{\text{Cr}}$, but a poor correlation between $D_{\text{spl/cpx}}^{\text{Fe}}$ and $D_{\text{spl/cpx}}^{\text{Al}}$ (Figure 9c, d). These results suggest that $D_{\text{spl/cpx}}^{\text{Fe}}$ may lead to Cr/Al variation in spinel, in turn affecting $D_{\text{spl/cpx}}^{\text{Fe}^{3+}}$ and Fe$^{3+}$/ΣFe in clinopyroxene by the reaction:

\[
4\text{CaMg}^{2+}\text{Si}_2\text{O}_6 + \text{Fe}^{2+}\text{Fe}_2\text{O}_4 + \text{Fe}^{2+}\text{Al}_2\text{O}_4 = 2\text{CaFe}^{2+}\text{Si}_2\text{O}_6 + 2\text{CaFe}^{3+}\text{Al}^{3+}\text{SiO}_6 + 2\text{Mg}^{2+}\text{SiO}_4. \tag{10}
\]

Clinopyroxene Spinel Spinel Clinopyroxene Clinopyroxene olivine

However, the effect of $D_{\text{spl/cpx}}^{\text{Fe}}$ on Fe$^{3+}$/ΣFe 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$^{3+}$/ΣFe 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$^{3+}\text{Si}_2\text{O}_4$ 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$_6$:

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

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:

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

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 to assess the reliability of the models and the robustness of the hypothesis that Fe$^{2+}$ oxidation is the dominant mechanism of hydrogen diffusion in clinopyroxene.
non-dehydrogenated samples in Figure 12a–c and used kernel density estimation to compare the
Fe\(^{3+}/\Sigma\text{Fe}\) 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\(^{3+}/\Sigma\text{Fe}\) values in samples that suffered dehydrogenation is relatively higher than
that for samples that did not (Figure 12). Although the differences between the Fe\(^{3+}/\Sigma\text{Fe}\) 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 unequivocally 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\(^{3+}/\Sigma\text{Fe}\).
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\textsuperscript{3+} in clinopyroxene are useful for estimating clinopyroxene Fe\textsuperscript{3+} content when laboratory measurements are unavailable, whether due to technical or financial reasons. Knowing clinopyroxene Fe\textsuperscript{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\textsuperscript{3+} contents could be developed for other Fe\textsuperscript{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\textsuperscript{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

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


**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, clinoferrasilite; 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$^{3+}$/ΣFe 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$^{3+}$/ΣFe 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) $D_{\text{spl/cpxCr}}$ versus $D_{\text{spl/cpxFe}}$ and (d) $D_{\text{spl/cpxAl}}$ versus $D_{\text{spl/cpxFe}}$.

Figure 10. Comparison of predicted Fe$^{3+}$/ΣFe 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$^{3+}$/ΣFe values as a function of Fe content (atoms per formula unit, apfu) and (d–f) kernel density estimations of the distribution of Fe$^{3+}$/ΣFe 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$^{3+}$/ΣFe ratios.

<table>
<thead>
<tr>
<th>Species</th>
<th>C</th>
<th>Application to this study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Di (N)</td>
<td>1.22</td>
<td>Di</td>
<td>De Grave and Van Alboom (1991)</td>
</tr>
<tr>
<td>Hd (N)</td>
<td>1.17</td>
<td>Hd</td>
<td>Eeckhout and De Grave (2003)</td>
</tr>
<tr>
<td>Ae (N)</td>
<td>1.15</td>
<td>Ae</td>
<td>Eeckhout and De Grave (2003)</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>TRG74 (°C)</th>
<th>TEG79 (°C)</th>
<th>TG79 (°C)</th>
<th>TP85 (°C)</th>
<th>TK88 (°C)</th>
<th>TA94 (°C)</th>
<th>TG96 (°C)</th>
<th>TL98 (°C)</th>
<th>TKR00 (°C)</th>
<th>TN09 (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cpx96a</td>
<td>704</td>
<td>693</td>
<td>744</td>
<td>670</td>
<td>647</td>
<td>615</td>
<td>753</td>
<td>625</td>
<td>693</td>
<td>707</td>
</tr>
<tr>
<td>Cpx98a</td>
<td>686</td>
<td>673</td>
<td>725</td>
<td>650</td>
<td>626</td>
<td>592</td>
<td>735</td>
<td>603</td>
<td>672</td>
<td>688</td>
</tr>
<tr>
<td>Cpx99a</td>
<td>679</td>
<td>666</td>
<td>718</td>
<td>642</td>
<td>618</td>
<td>584</td>
<td>728</td>
<td>595</td>
<td>664</td>
<td>681</td>
</tr>
<tr>
<td>Cpx104a</td>
<td>695</td>
<td>684</td>
<td>735</td>
<td>660</td>
<td>637</td>
<td>604</td>
<td>744</td>
<td>614</td>
<td>683</td>
<td>698</td>
</tr>
<tr>
<td>Cpx107a</td>
<td>707</td>
<td>697</td>
<td>747</td>
<td>673</td>
<td>651</td>
<td>619</td>
<td>756</td>
<td>628</td>
<td>697</td>
<td>710</td>
</tr>
<tr>
<td>Cpx108a</td>
<td>706</td>
<td>695</td>
<td>745</td>
<td>672</td>
<td>649</td>
<td>617</td>
<td>755</td>
<td>627</td>
<td>695</td>
<td>709</td>
</tr>
<tr>
<td>F-8b</td>
<td>1400</td>
<td>1256</td>
<td>1290</td>
<td>1247</td>
<td>1243</td>
<td>1268</td>
<td>1318</td>
<td>1293</td>
<td>1421</td>
<td>1398</td>
</tr>
<tr>
<td>F-11b</td>
<td>1423</td>
<td>1252</td>
<td>1286</td>
<td>1241</td>
<td>1237</td>
<td>1290</td>
<td>1313</td>
<td>1300</td>
<td>1412</td>
<td>1414</td>
</tr>
<tr>
<td>F-16b</td>
<td>1492</td>
<td>1256</td>
<td>1307</td>
<td>1244</td>
<td>1207</td>
<td>1310</td>
<td>1330</td>
<td>1324</td>
<td>1441</td>
<td>1469</td>
</tr>
<tr>
<td>87-70c</td>
<td>1417</td>
<td>1240</td>
<td>1265</td>
<td>1228</td>
<td>1238</td>
<td>1297</td>
<td>1295</td>
<td>1294</td>
<td>1384</td>
<td>1418</td>
</tr>
</tbody>
</table>

Always consult and cite the final, published document. See http://www.minsocam.org or GeoscienceWorld
<p>| | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BY-19&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1275</td>
<td>1297</td>
<td>1328</td>
<td>1302</td>
<td>1302</td>
<td>1212</td>
<td>1353</td>
<td>1250</td>
<td>1459</td>
<td>1407</td>
<td></td>
</tr>
<tr>
<td>8508-9&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1181</td>
<td>1187</td>
<td>1233</td>
<td>1184</td>
<td>1167</td>
<td>1064</td>
<td>1260</td>
<td>1103</td>
<td>1314</td>
<td>1282</td>
<td></td>
</tr>
<tr>
<td>94M80&lt;sup&gt;d&lt;/sup&gt;</td>
<td>595</td>
<td>670</td>
<td>677</td>
<td>650</td>
<td>622</td>
<td>614</td>
<td>672</td>
<td>606</td>
<td>630</td>
<td>590</td>
<td></td>
</tr>
<tr>
<td>97M32&lt;sup&gt;d&lt;/sup&gt;</td>
<td>513</td>
<td>519</td>
<td>562</td>
<td>495</td>
<td>469</td>
<td>456</td>
<td>560</td>
<td>447</td>
<td>507</td>
<td>505</td>
<td></td>
</tr>
<tr>
<td>944012-11&lt;sup&gt;d&lt;/sup&gt;</td>
<td>678</td>
<td>883</td>
<td>818</td>
<td>870</td>
<td>774</td>
<td>767</td>
<td>827</td>
<td>748</td>
<td>785</td>
<td>679</td>
<td></td>
</tr>
</tbody>
</table>


Data references: <sup>a</sup>Proyer et al. (2004); <sup>b</sup>Lazarov et al. (2009); <sup>c</sup>Nimis et al. (2015); <sup>d</sup>Li et al. (2005).
Figure 3.

(a) Training set, Test set
(b) Training set, Test set
(c) Training set, Test set
(d) Training set, Test set

Linear
Polynomial
Neural network
ANN ensemble

(e) Training set, Test set
(f) Training set, Test set
(g) Training set, Test set

Decision trees
Random forest
Extra-Trees

(h) Training set, Test set
(i) Training set, Test set
(j) Training set, Test set
(k) Training set, Test set

Linear
Polynomial
Neural network
ANN ensemble

(l) Training set, Test set
(m) Training set, Test set
(n) Training set, Test set

Decision trees
Random forest
Extra-Trees
Figure 4.

(a) Linear (Cpx only)
- Training set
- Test set
- $R^2$ (train) = 0.24
- $R^2$ (test) = 0.26
- RMSE (Train) = 0.055
- RMSE (Test) = 0.052

(b) Linear (Cpx + Spl)
- Training set
- Test set
- $R^2$ (train) = 0.38
- $R^2$ (test) = 0.57
- RMSE (Train) = 0.049
- RMSE (Test) = 0.040

(c) Random forest (Cpx only)
- Training set
- Test set
- $R^2$ (train) = 0.81
- $R^2$ (test) = 0.70
- RMSE (Train) = 0.028
- RMSE (Test) = 0.034

(d) Random forest (Cpx + Spl)
- Training set
- Test set
- $R^2$ (train) = 0.84
- $R^2$ (test) = 0.85
- RMSE (Train) = 0.025
- RMSE (Test) = 0.024

(e) Extra-Trees (Cpx only)
- Training set
- Test set
- $R^2$ (train) = 0.70
- $R^2$ (test) = 0.54
- RMSE (Train) = 0.035
- RMSE (Test) = 0.042

(f) Extra-Trees (Cpx + Spl)
- Training set
- Test set
- $R^2$ (train) = 0.76
- $R^2$ (test) = 0.70
- RMSE (Train) = 0.031
- RMSE (Test) = 0.034

Always consult and cite the final, published document. See http://www.minsocam.org or GeoscienceWorld
Figure 5.

(a) Cpx

Relative feature importance (Random forest)

Si  Ti  Al  Cr  Fe  Mn  Mg  Ca  Na

(b) Spl

Relative feature importance (Extra-Trees)

Si  Ti  Al  Cr  Fe  Mn  Mg  Ca  Na

Always consult and cite the final, published document. See http://www.minsocam.org or GeoscienceWorld
Figure 6.
Figure 9.

(a) Cr in spl [per 12 oxygen] vs Cr in cpx [per 12 oxygen].

(b) Al in spl [per 12 oxygen] vs Al in cpx [per 12 oxygen].

(c) $D_{Cr}^{spl/cpx}$ vs $D_{Fe}^{spl/cpx}$.

(d) $D_{Al}^{spl/cpx}$ vs $D_{Fe}^{spl/cpx}$.

Highest Fe/Al in spinel.

$y = 4.5x - 4.8$

$R^2 = 0.79$
Figure 10.
Figure 12.