Lithos 362–363 (2020) 105469

Contents lists available at ScienceDirect

Lithos

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

Research Article

Calculating amphibole formula from electron microprobe analysis data using a machine learning method based on principal components regression

Xiaoyan Li [a,b], Chao Zhang [a,b,*], Harald Behrens [b], Francois Holtz [b]

[a] State Key Laboratory of Continental Dynamics, Department of Geology, Northwest University, Xi'an 710069, China
[b] Institute of Mineralogy, Leibniz University Hannover, 30167 Hannover, Germany

ARTICLE INFO

Article history:
Received 6 February 2020
Received in revised form 3 March 2020
Accepted 4 March 2020
Available online 10 March 2020

Keywords:
Amphibole formula
Amphibole nomenclature
Machine learning
Principal components regression
Electron microprobe analysis

ABSTRACT

We present a new method for calculating amphibole formula from routine electron microprobe analysis (EMPA) data by applying a principal components regression (PCR)-based machine learning algorithm on reference amphibole data. The reference amphibole data collected from literature are grouped in two datasets, for Li-free and Li-bearing amphiboles respectively, which include Fe^{2+}, Fe^{3+}, OH contents and the ion site assignments determined by single crystal structure refinement. We established two PCR models, for Li-free and Li-bearing amphiboles respectively, by the 10-fold cross validation of training datasets and evaluated by independent test datasets. The results show that our models can successfully reproduce the reference data for most ions with an error less than ±0.01 atom per formula unit (apfu), for Fe^{3+} within an error less than ±0.2 apfu and for ^{W}OH and $^{W}O^{2-}$ with errors less than ±0.3 apfu. The error in estimated $Fe^{3+}/\Sigma Fe$ ratio shows a rough negative dependence on FeO_T content (total iron expressed as FeO), ranging within ±0.3 for amphiboles with $FeO_T \geq 5$ wt% and within ±0.2 for amphiboles with $FeO_T \geq 10$ wt%. Our models are applicable to both ^{W}(OH, F, Cl)-dominant and ^{W}O-dominant amphiboles. It is notable that this method is not suitable for calculating mineral formula of amphiboles that have been affected by deprotonation as a result of secondary oxidation, but it could offer an estimation of initial ^{W}OH prior to the post-formation oxidation. A user-friendly Excel worksheet is provided with two independent PCR models for calculating the formula of Li-free amphibole and Li-bearing amphibole, respectively. An automatic nomenclature function is also provided according to the nomenclature criteria of the 2012 International Mineralogical Association (IMA) report.

1. Introduction

Amphibole is a common rock-forming mineral that can be stable in a variety of geological systems with a wide range of temperature and pressure conditions. Its composition can provide information on the intensive and extensive properties of magma and fluid that are in equilibrium with the amphibole. The composition of amphibole can be used as a thermobarometer to estimate the pressure and temperature of its formation condition (e.g., Putirka, 2016; Ridolfi and Renzulli, 2012). The H_2O content of amphibole and the associated equilibrium mineral assemblage can be applied to estimate the H_2O activity prevailing during its crystallization and the H_2O content of coexisting minerals (Lamb and Popp, 2009; Popp et al., 2006). The OH, F and Cl contents of amphibole can provide information on the volatile content of coexisting melts or fluids (Giesting et al., 2015; Giesting and Filiberto, 2014; Henry and Daigle, 2018; Zhang et al., 2012). The $Fe^{3+}/\Sigma Fe$ ratio and H content of amphibole from mantle xenoliths and volcanic flows are indicative of the redox condition of metasomatic mantle source or the degassing process of the volcanic system (Dyar et al., 1993; McCubbin et al., 2010). All these applications require an accurate measurement of Fe^{3+} and H content and a well-established formula of amphibole. Although modern techniques such as Mössbauer spectroscopy, micro-XANES, SIMS could provide high-quality data for the required Fe^{3+} and H contents in amphibole, their scarcity and high cost have hindered these methods as routine tools for most geologists. The development of the electron microprobe analysis (EMPA) flank method and peak-shift method provides an alternative in-situ method for measuring $Fe^{3+}/\Sigma Fe$ of amphiboles (Lamb et al., 2012; Li et al., 2019), but it is still far from being a routine use in most EMPA laboratories, primarily due to the lack of reference materials. Currently, the calculation of amphibole formula from routine EMPA data is still the most popular method, in spite of difficulties related to the complex structure of amphibole.

* Corresponding author at: State Key Laboratory of Continental Dynamics, Department of Geology, Northwest University, Xi'an 710069, China.
E-mail addresses: zhangchao@nwu.edu.cn, c.zhang@mineralogie.uni-hannover.de (C. Zhang).

https://doi.org/10.1016/j.lithos.2020.105469
0024-4937/

The amphibole supergroup has a general formula $A_{0\text{-}1}B_2C_5T_8O_{22}W_2$, where A = □, Na, K, Ca, Li, Pb^{2+}; B = Na, Ca, Mn^{2+}, Fe^{2+}, Mg, Li; C = Mg, Fe^{2+}, Mn^{2+}, Al, Mn^{3+}, Fe^{3+}, Ti^{4+}, Li; T = Si, Al, Ti, Be^{2+}; W = OH^-, F, Cl, O^{2-} (Hawthorne et al., 2012). The dominant C sites, on which Fe^{3+} resides, are associated with W anions (OH, F, Cl, O) and are made up of three distinct types of octahedrally coordinated sites, 2 M(1), 2 M(2) and 1 M(3) (Hawthorne and Oberti, 2007). The two M(1) sites and one M(3) site are associated with two O(3) W anions in a cis-arrangement and in a trans-arrangement respectively, but the two M(2) site are not associated with the O(3) W anions.

Since amphibole has a complex crystallographic structure that can host a variety of ions, the calculation of mineral formula from routine EMPA data, in which Fe^{2+}, Fe^{3+}, $^{W}OH^-$ and $^{W}O^{2-}$ cannot be measured directly, has long been recognized as a challenging task. Due to the presence of vacancy at the A site, the cation sum of amphibole is not fixed and can vary between 15 and 16 on the anhydrous basis of 23O. Therefore, the most widely accepted methodology is to constrain the subtotal of cations to a specific stoichiometric value and adjusting the Fe^{3+} content to keep the electroneutrality, in which the upper and lower limits of the stoichiometric value would give the corresponding minimum and maximum estimation of Fe^{3+} content respectively. Based on this methodology, several formula calculation protocols could give reasonable results for specific analyses in the literature (Dale et al., 2005; Holland and Blundy, 1994; Leake et al., 1997; Schumacher, 1991; Spear and Kimball, 1984; Yavuz, 2007). However, all these protocols assume that anions on W sites are limited in the form of OH + F + Cl = 2 and ignore the presence of $^{W}O^{2-}$, and thus they usually result in inaccurate Fe^{3+} estimation and are particularly not applicable to the oxy-amphibole group. A recommended method in the 2012 International Mineralogical Association (IMA) report (Hawthorne et al., 2012; Locock, 2014) estimated the $^{W}O^{2-}$ through the relation $^{W}O^{2-}$ = 2Ti, which usually leads to an overestimation if Ti^{4+} does not only resides on M(1) site (see details in the section of Discussion). Estimation of Fe^{3+} in this recommended method also follows the same methodology as previous protocols and gives a mean Fe^{3+} content with minimum and maximum estimates (Hawthorne et al., 2012; Locock, 2014). Recently, Ridolfi et al. (2018) proposed a new method based on multivariate least-square analysis of mass and stoichiometry of selected amphibole from literature data and amphibole endmembers. The mass-normalization procedure can provide estimations of Fe^{3+} through charge balance and $^{W}O^{2-}$ through a regression equation. But this method is only applicable to Li-free amphiboles with low Cl contents (Ridolfi et al., 2018). Furthermore, the new amphibole classification procedure in the 2012 IMA report (Hawthorne et al., 2012) requires an overall site assignment of cations that may reside on different sites, such as Li^+, Fe^{2+}, Mg^{2+} and Mn^{2+} which may be at the B and/or C sites, which however cannot be quantified by the method of Ridolfi et al. (2018).

Indeed, to the best of our knowledge, no method is available currently for calculating amphibole formula that can fulfill the requirements of amphibole classification procedure in the 2012 IMA report. As one step to work on this challenge, we established a new amphibole formula calculation protocol using a machine learning method that can estimate an overall site assignment of amphibole from routine EMPA data. We have previously developed a similar method for calculating biotite formula and the result is promising (Li et al., 2020). In this paper, we describe the method and datasets used for establishing the calculating protocol of amphibole formula, and a user-friendly Excel worksheet executing this method is also provided.

2. Method and dataset

For minerals such as amphibole that have a large number of crystallographic sites, at which a variety of cations and anions could be accommodated, closure and associated multicollinearity are inevitable problems when applying statistical analysis on their chemical data. The performance of closure on mineral chemical data refers to the competition of atomic species for one specific crystallographic position, and charge balance between different crystallographic positions. For example, 7 types of cations could competitively occupy 5 octahedral C sites in amphibole crystal; the whole crystal keeps neutrality which is balanced by all the ions from various sites. The closure phenomenon could exaggerate strong correlation/multicollinearity between the different ions of mineral, which means a pair of predictor variables have a substantial correlation with each other and/or between multiple predictors at once, and therefore can lead to biased estimation when performing normal multivariate analysis (e.g., ordinary least-squares regression method) (Aitchison, 1984; Jolliffe, 2002; Yoo et al., 2014; Young et al., 1997).

In this study, we developed a machine-learning algorithm, principal components regression (PCR), on the reference datasets, to establish a new calculation method for amphibole formula, which could solve the potential closure and multicollinearity in mineral chemical data and successfully produce a reasonable cation/anion-assigned formula from EMPA data. The reference amphibole data are collected into two working datasets, including a Li-free amphibole dataset (n=173) and a Li-bearing amphibole dataset (n = 66) (Supplementary Table 1). As shown in Figs. 1a-d, the collected Li-free amphibole dataset has a wide compositional range, such as 38–53 wt% SiO_2, 0–22 wt% Al_2O_3, 2–30 wt% FeO_T (total iron expressed as FeO), 1–23 wt% MgO, 3–13 wt % CaO, 0–9 wt% Na_2O, 0–0.6 Cl and 0–4.5 wt% F. This dataset also covers a wide range in $Fe^{3+}/\Sigma Fe$ (0–0.9) and $^{W}O^{2-}$ at the O(3) site (0–1.1 apfu, atom per formula unit), which show no apparent correlation (Fig. 1e). For the anions at the O(3) site (Fig. 1f), OH is generally the most abundant one, Cl is commonly at trace level, and F and $^{W}O^{2-}$ are comparatively minor but important ones. Compared to the Li-free amphibole dataset, the Li-bearing amphibole dataset shows different ranges in SiO_2 (49–62 wt%), Al_2O_3 (≤14 wt%), CaO (mostly ≤6 w%), Cl (≤0.02 wt %), but with similar ranges in FeO_T, MgO and Na_2O. In addition, the Li-bearing amphiboles generally have much less $^{W}O^{2-}$ compared to the Li-free amphiboles.

These collected datasets are provided with compositional data analyzed by EMPA, Fe^{3+}/Fe_{tot} ratio analyzed by Mössbauer spectroscopy or micro-XANES, H_2O content analyzed by FTIR, SIMS or C-H-N analysis, and Li content analyzed by SIMS, as well as single crystal structure refinement (SREF) performed by X-ray diffraction. Some rare amphiboles with Pb^{2+} or Li on A site, Mn^{3+} on C site, and Be^{2+} on T site (Hålenius and Bosi, 2012; Iezzi et al., 2004; Moore et al., 1993; Tait et al., 2005) are not considered in this study.

For a general linear regression model defined by eq. $Y = X\beta + \varepsilon$, the process of PCR is composed of two parts: first, finding principal components through principal component analysis (PCA) on matrix X; second, getting regression coefficient β through the regression of matrix Y on the obtained principal components (Jolliffe, 2002; Mevik and Wehrens, 2007). PCA in the first step is a statistical dimension reduction method that is widely used in statistical analysis of compositional data with highly correlated predictors (Abdi and Williams, 2010; Aitchison, 1983; Davis, 2002; Wold et al., 1987). The principal components are a series of linearly independent variables/vectors which are linear combinations of original variables converted through orthogonal transformation (Abdi and Williams, 2010). PCA has been introduced into mineralogy to identify and evaluate independent exchange mechanisms in the mineral composition space (Labotka, 1983; Waters and Charnley, 2002; Young et al., 1997). The obtained principal components, i.e., linear combinations of the multiple cation correlations, represent the combined substitutions that have minimal impact on the crystal structure distortion and are indicative of the underlying physical-chemical control on mineral composition (Labotka, 1983; Waters and Charnley, 2002; Young et al., 1997).

For the linear regression eq. $Y = X\beta + \varepsilon$ in our study, the right-side matrix X represents multi-predictor variables which are the mole contents of elements calculated from EMPA data, and the left-side matrix Y holds multi-response variables which are the site occupations of

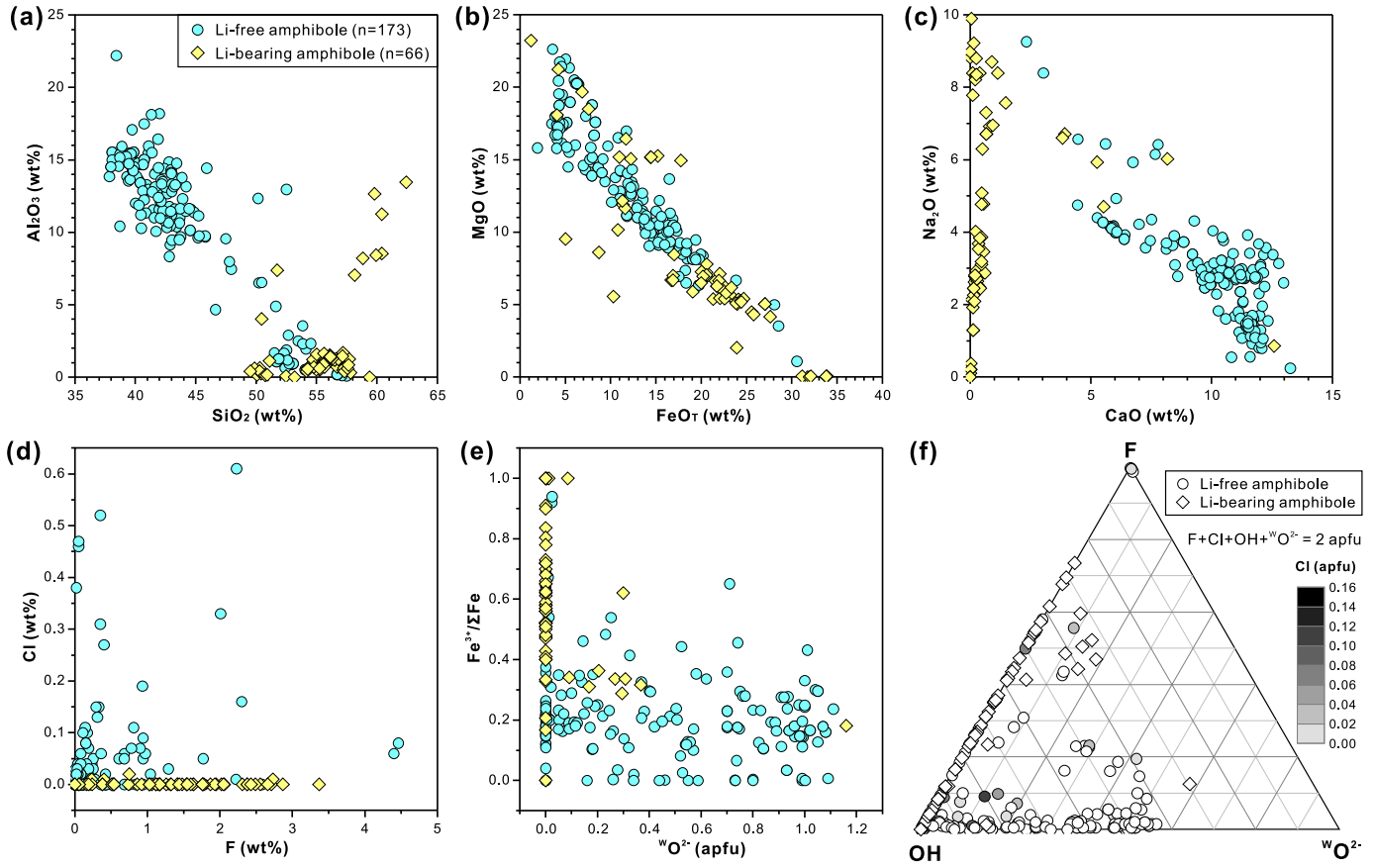

Fig. 1. Representative plots showing the compositional range of collected reference amphiboles used in this study. Two datasets of Li-free and Li-bearing amphiboles are shown separately. (a) SiO_2 vs. Al_2O_3. (b) FeO_T vs. MgO. (c) CaO vs. Na_2O. (d) F vs. Cl. (e) $^{W}O^{2-}$ vs. $Fe^{3+}/\Sigma Fe$. (f) Ternary plot for F, OH, $^{W}O^{2-}$ at the O(3) site. Cl content (≤0.16 apfu) is exclusively indicated by gray scale.

ions obtained from the structure refinement. Each line combined from them stands for a fully analyzed amphibole reference from literature (see details in section Data processing protocol). Because the two matrixes provide the same chemical information of amphibole references in terms of crystal-chemical constraints but in a different form, each principal component capturing the variance of matrix X is correspondingly associated with a component that captures the relevant variance of matrix Y. Therefore, the second regression step, regression of matrix Y on the principal components of matrix X, coincides perfectly with the best application of PCR, which requires that the directions in which the predictors have the largest variation are the exact directions associated with the corresponding response variables (Jolliffe, 2002; Kuhn and Johnson, 2013).

PCR in this study has been performed using the open source GNU R software (R Core Team, 2018) and the PLS package (Mevik et al., 2019). The two datasets of Li-free and Li-bearing amphibole are treated individually in the following procedure, which eventually yield two independent models for Li-free and Li-bearing amphibole. Both datasets are split into a training set (80% in proportion) and a test set (20% in proportion) by simple random sampling (with no replacement) function that is embedded in R software. Therefore, Li-free amphibole dataset has 138 samples in training set and 35 samples in test set while Li-bearing amphibole dataset has 52 samples in training set and 14 samples in test set. The training sets are used to tune the predictive model while performing repeated 10-fold cross-validation to estimate the model performance, and the independent test sets are later employed to evaluate the performance of the model. It is emphasized that the initial regression coefficients obtained by regressing matrix Y on principal components of matrix X have been subsequently transformed back by the PLS package into a coefficient matrix β in the original form of matrix X.

3. Data processing protocol

Step 1. For the collected amphibole EMPA data, molar concentration of ions are calculated from weight percentages (SiO_2, TiO_2, Al_2O3, Cr_2O_3, FeO_T, MnO, MgO, CaO, Na_2O, K_2O, F, Cl, NiO, ZnO, Li_2O), in which FeO_T content is transformed from original Fe_2O_3 and FeO when both of them are available. The major ion concentrations (Si, Ti, Al, Cr, Fe^{2+}, Mn, Mg, Ca, Na, K, F, Cl) without/with Li (when measured) are summarized into two matrices: X_1 (173 by 12) for Li-free amphiboles and X_2 (66 by 13) for Li-bearing amphiboles.

Step 2. The molar proportions of ions at specific sites of amphibole structural formula are collected from literature, including T.Si, T.Al, T.Ti, C·Al, C·Ti, C·Cr, C·Fe^{3+}, C·Fe^{2+}, C·Mg, C·Mn^{2+}, C·Ni, C·Zn, C·Li, B·Mn^{2+}, B·Mg, B·Fe^{2+}, B·Li, B·Ca, B·Na, A.Ca, A.Na, A.K, W.OH, W·F, W·Cl, W·O^{2-}, where prefixes "T.", "C.", "B.", "A." and "W." represent sites where the ions are located at. Parts of these data will form a matrix Y applying for calibration, and all these data will also be used as reference values to evaluate the prediction performance of our model later.

Step 3. The selection criteria of ions into Y matrixes are based on: (1) selecting dominant ions at each crystallographic site (e.g. Si and Al at T sites; Mg, Fe^{2+}, Fe^{3+} and Ti at C sites; K at A site; F and OH at O (3) sites); (2) for ions that can appear at two sites, only one of them is picked (e.g. select C·Mn^{2+} from C·Mn^{2+} and B·Mn^{2+}; select B·Li from C·Li and B·Li). Therefore, for Li-free amphibole, molar proportions of T.Si, T.Al, C·Ti, C·Fe^{3+}, C·Fe^{2+}, C·Mg, C·Mn^{2+}, B·Ca, B·Na, A.K, W.OH, W·F, W·Cl are selected to make up a matrix Y_1 (173 by 13); for Li-

bearing amphibole, T.Si, C·Al, C·Ti, C·Fe^{3+}, C·Fe^{2+}, C·Mg, C·Mn^{2+}, B·Ca, B·Na, A.K, W.OH, W·F and B·Li are selected into a matrix Y_2 (66 by 13). It is notable that the selected ions for Li-free amphibole and Li-bearing amphibole are different due to differences in own chemical properties; for example, the collected Li-bearing amphiboles are mostly Cl-free and Si-rich (^{T}Al-poor), and therefore W·Cl is not selected, and C·Al (instead of T.Al) is selected.

Step 4. Linear equations can be established between X and Y for the two amphibole datasets:

$$Y_1 = X_1\beta_1 + \varepsilon_1 \tag{1}$$

$$Y_2 = X_2\beta_2 + \varepsilon_2 \tag{2}$$

where β_1 and β_2 denote the regression coefficient matrixes obtained through principal component regression (PCR) and ε_1 and ε_2 dominate the vector of random errors with expected values equal to zero. Moreover, each β is composed by the addition of a matrix and a vector, in which the matrix part is made up of the linear coefficients while the vector part represents the intercepts.

Step 5. Once the regression coefficient matrix (β_1 and β_2) obtained by principal component regression has been evaluated and accepted, the predicted matrixes $\hat{Y}_1$ and $\hat{Y}_2$ (symbol hat indicates "predicted") can be calculated through the matrix X and coefficient matrix β.

$$\hat{Y}_1 = X_1\beta_1 \tag{3}$$

$$\hat{Y}_2 = X_2\beta_2 \tag{4}$$

The above two linear equations can be expressed by the following matrix equations respectively:

$$\begin{pmatrix} \widehat{T.Si} & \cdots & \widehat{W.Cl} \\ \vdots & \ddots & \vdots \\ & \cdots & \end{pmatrix}_{173\times 13} = \begin{pmatrix} Si & \cdots & Cl \\ \vdots & \ddots & \vdots \\ & \cdots & \end{pmatrix}_{173\times 12} * \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,13} \\ \vdots & \ddots & \vdots \\ \beta_{12,1} & \cdots & \beta_{12,13} \end{pmatrix}_{12\times 13} + (\beta_{0,1} \ \cdots \ \beta_{0,13})_{1\times 13} \tag{5}$$

$$\begin{pmatrix} \widehat{T.Si} & \cdots & \widehat{W.F} & \widehat{B.Li} \\ \vdots & \ddots & \vdots & \vdots \\ & \cdots & & \end{pmatrix}_{66\times 13} = \begin{pmatrix} Si & \cdots & F & Li \\ \vdots & \ddots & \vdots & \vdots \\ & \cdots & & \end{pmatrix}_{66\times 13} * \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,13} \\ \vdots & \ddots & \vdots \\ \beta_{13,1} & \cdots & \beta_{13,13} \end{pmatrix}_{13\times 13} + (\beta_{0,1} \ \cdots \ \beta_{0,13})_{1\times 13} \tag{6}$$

Consequently, the matrixes $\hat{Y}_1$ and $\hat{Y}_2$ both have 13 linear equations to predict each column of them. This means the 13 site-assigned ions in either Li-free amphibole (T.Si, T.Al, C·Ti, C·Fe^{3+}, C·Fe^{2+}, C·Mg, C·Mn^{2+}, B·Ca, B·Na, A.K, W.OH, W·F, W·Cl) or Li-bearing amphibole (T.Si, C·Al, C·Ti, C·Fe^{3+}, C·Fe^{2+}, C·Mg, C·Mn^{2+}, B·Ca, B·Na, A.K, W.OH, W·F, B·Li) can be calculated directly from EMPA data through these linear equations. For instance, for one specific amphibole EMPA data, the predict T.Si can be calculated through Eq. 7 or Eq. 8, depending on whether Li content is provided or not:

$$\widehat{T.Si} = \beta_{0,1} + \beta_{1,1}Si + \beta_1 Ti + \beta_{3,1}Al + \cdots + \beta_{12,1}Cl \tag{7}$$

$$\widehat{T.Si} = \beta_{0,1} + \beta_{1,1}Si + \beta_1 Ti + \beta_{3,1}Al + \cdots + \beta_{12,1}F + \beta_{13,1}Li \tag{8}$$

Other ions are calculated from analogous equations, and the comparisons of reference values and predicted ones at this step are illustrated in Figs. 2 and 6 for Li-free and Li-bearing amphibole respectively.

Step 6. After getting the predicted value of T.Si ($\widehat{T.Si}$), two correlation coefficients (denoted by γ_1 and γ_2) are calculated for Li-free and Li-bearing amphibole through equations with the same forms:

$$\gamma = \widehat{T.Si}/Si \tag{9}$$

This coefficient is calculated based on Si ratio because silica is a major component in amphibole and can be measured by EMPA with high accuracy.

Subsequently, the predictions of ion mole proportions that are not included in the Y_1 or Y_2 and but only present at one specific crystallographic site, such as C·Cr, C·Ni, C·Zn, and W·Cl (only for Li-bearing amphibole), are calculated through multiplying γ with their mole concentrations derived from EMPA data, such as:

$$\widehat{C.Cr} = \gamma * Cr \tag{10}$$

Step 7. The other ions are calculated according to ideal site occupation (e.g., W·O^{2-}) or to the total amount of specific ion at different sites (e.g., T.Ti and B·Fe^{2+}). The specific calculations are given as the following Eqs. 11–19, in which Eq. 12b and Eq. 13 are only valid for Li-bearing amphiboles. Eq. 19 assumes that total occupation of OH, $^{W}O^{2-}$, F and Cl at the W site equals to 2 apfu.

$$\widehat{T.Ti} = \gamma * Ti - \widehat{C.Ti} \tag{11}$$

$$\widehat{C.Al} = \gamma * Al - \widehat{T.Al} \tag{12a}$$

$$\widehat{T.Al} = \gamma * Al - \widehat{C.Al} \tag{12b}$$

$$\widehat{C.Li} = \gamma * Li - \widehat{B.Li} \tag{13}$$

$$\widehat{B.Fe^{2+}} = \gamma * Fe_T - \widehat{C.Fe^{3+}} - \widehat{C.Fe^{2+}} \tag{14}$$

$$\widehat{B.Mn^{2+}} = \gamma * Mn^{2+} - \widehat{C.Mn^{2+}} \tag{15}$$

$$\widehat{B.Mg} = \gamma * Mg - \widehat{C.Mg} \tag{16}$$

$$\widehat{A.Ca} = \gamma * Ca - \widehat{B.Ca} \tag{17}$$

$$\widehat{A.Na} = \gamma * Na - \widehat{B.Na} \tag{18}$$

$$\widehat{W.O^{2-}} = 2 - \widehat{W.OH} - \widehat{W.F} - \widehat{W.Cl} \tag{19}$$

Step 8. Negative values obtained during the above calculation procedure, which is probably derived from propagated uncertainty, are forced to be zero. The vacancy at A site can be calculated by using the ideal occupation at A site minus all existing cations at this site.

We constructed a user-friendly Excel spreadsheet (Supplementary Table 1) that can be used to perform the above calculation steps and to provide an overall site assignment of cations and anions for amphiboles from an input of routine EMPA data (wt%). The input of NiO and ZnO are optional, and their absence would not affect the final formula quality since these cations have not been used in the regression procedure. For using the model for Li-bearing amphibole, Li_2O concentrations must be provided. The concentrations of 12 major components (i.e., SiO_2, TiO_2, Al_2O_3, Cr_2O_3, FeO_T, MnO, MgO, CaO, Na_2O, K_2O, F, Cl)

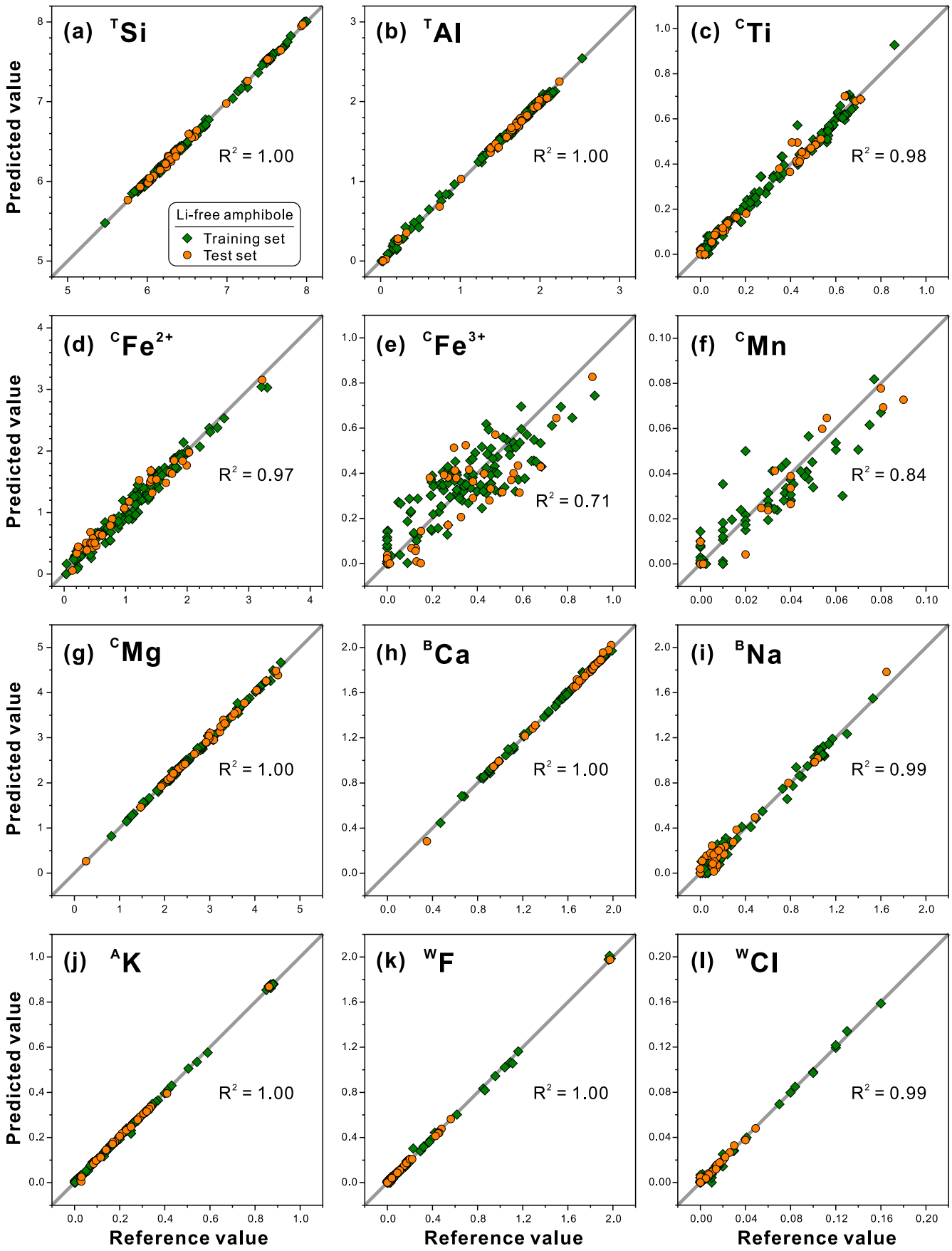

Fig. 2. Comparison between predicted and reference ions for the working dataset of Li-free amphibole. Predictions for these ions are calculated directly from regression coefficients of PCR at Step 5. Plots are in atom per formula unit (apfu).

must be provided; otherwise, the absent value will be considered as zero and this might induce biased estimation of the formula.

4. Discussion

4.1. Calculation reliability and uncertainty

4.1.1. Li-free amphibole

Figs. 2–5 illustrate the comparison of reference values and predicted ones in terms of site assignment for Li-free amphiboles. The site assignments of 12 ions plotted on Fig. 2 and ^{W}OH on Fig. 4a are calculated directly from the regression coefficient equations obtained by PCR (step 5 in above section). Fig. 3 and $^{W}O^{2-}$ on Fig. 4d show the result calculated by steps 6–7 for the other ions. The overall correlations between the predicted value and reference value show a remarkable 1:1 trend, and both data from the training set and the independent test set follow the same distribution behavior, which indicate the high accuracy and the robustness of our model.

The prediction quality is closely related to the quality of original data of ion assignment of reference amphiboles in the literature, the more accuracy the original data, the higher the prediction ability of the model. The majority of ions (^{T}Si, ^{T}Al, ^{C}Mg, ^{C}Cr, ^{C}Ni, ^{B}Ca, ^{A}K, ^{W}F, ^{W}Cl) have a prediction uncertainty less than ±0.01 apfu; ^{C}Ti, $^{C}Fe^{2+}$, ^{C}Al, ^{B}Na and ^{A}Na have a maximum error up to ±0.1 apfu. Comparing with the reference data, our model has a relative low predictability for ^{C}Mn (±0.03 apfu, Fig. 2f), $^{C}Fe^{3+}$ (±0.2 apfu, Fig. 2e) and ^{W}OH (±0.3 apfu, Fig. 4a), which also induce the corresponding low prediction ability for $^{B}Fe^{2+}$ (±0.2 apfu, Fig. 3d) and $^{W}O^{2-}$ (±0.3 apfu, Fig. 4d) according to step 7. The reason for the low predictability for ^{C}Mn and ^{B}Mn (not plotted) is solely due to the extremely low Mn concentration (<0.1 apfu) in the original dataset, which results in a low accuracy for the Mn site assignment in amphibole structure by SREF, and this absolute low accuracy has been inherited by the prediction model for ^{C}Mn as showed by Fig. 2f. However, for the $^{C}Fe^{3+}$ and ^{W}OH, the relatively low prediction quality is propagated from the inconsistencies in the analytical accuracy used in the literature. For example, in the collected dataset, Fe^{3+} content is obtained by Mössbauer spectroscopy or micro-XANES; OH content is measured by FTIR or SIMS. All these techniques have their own analytical uncertainties, e.g. the absolute error of $Fe^{3+}/\Sigma Fe$ analyzed by micro-XANES is ±10–20% (Dyar et al., 2016); the uncertainty in $Fe^{3+}/\Sigma Fe$ determination using Mössbauer spectroscopy is at least ±5% (Dyar, 2002; McCammon et al., 1998; Sobolev et al., 1999). This inherited uncertainty for Fe^{3+} and OH in the original dataset is the major source for the low prediction ability. In addition, the lower the FeO_T concentration in the sample, the lower the accuracy for Fe^{3+} measured by Mössbauer spectroscopy, and this problem is also propagated in the calibration model. This final propagated uncertainty is distinctly displayed by the enlarged discrepancy range in estimated $Fe^{3+}/\Sigma Fe$ ratios towards the low concentration of FeO_T (Fig. 5e). A funnel-shaped trend has been revealed by the range of absolute error of predicted $Fe^{3+}/\Sigma Fe$ varying with FeO_T content, with majority of prediction error falling into ±0.3 if $FeO_T \geq 5$ wt% and further into the ±0.2 if $FeO_T \geq 10$ wt% (Fig. 5e).

4.1.2. Li-bearing amphibole

The comparison of reference values and predicted ones in terms of site assignment for Li-bearing amphiboles are illustrated in Figs. 6–7.

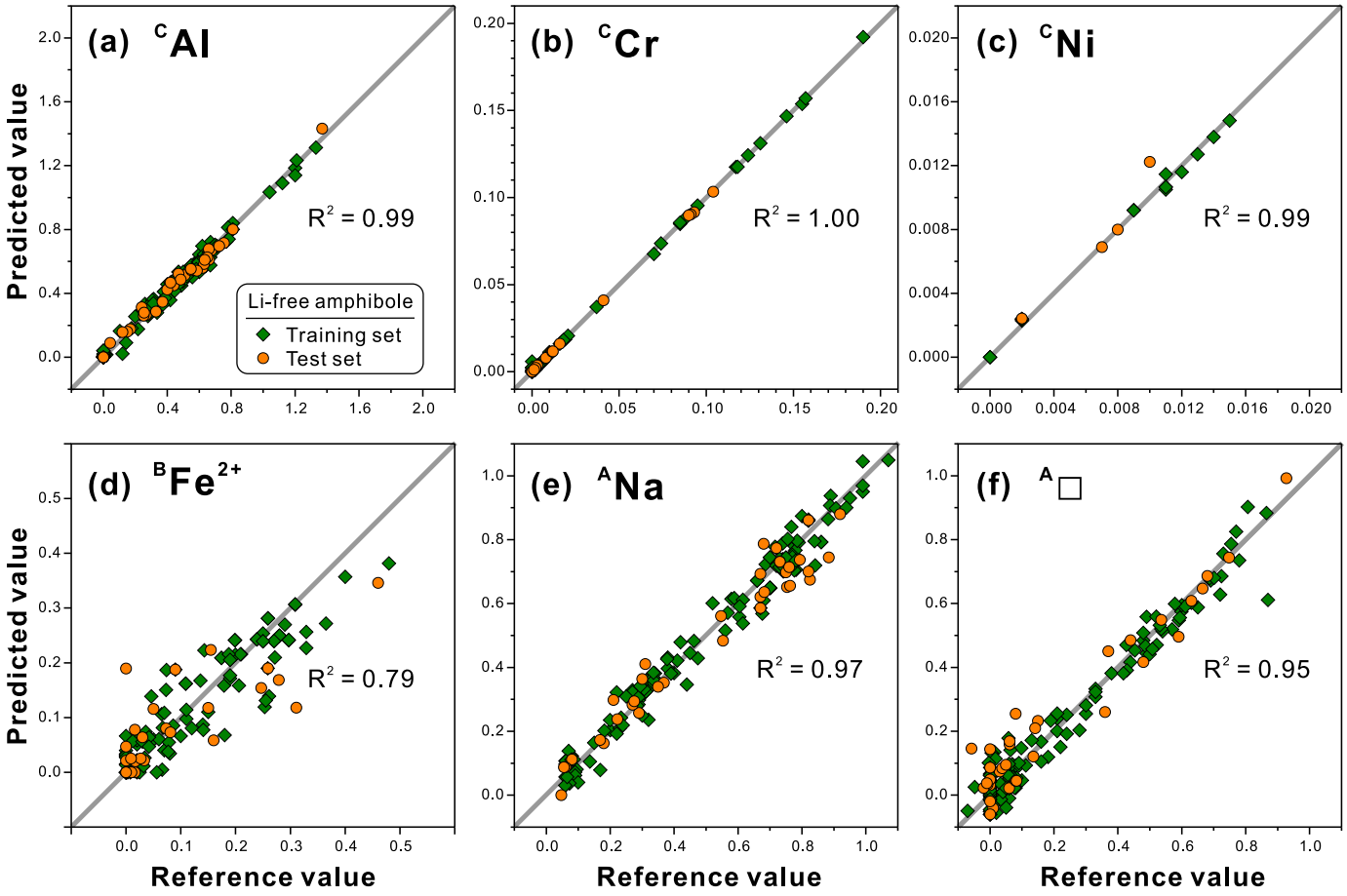

Fig. 3. Comparison between predicted and reference ions for the working dataset of Li-free amphibole. Predictions for these ions are calculated at Step 7 based on constraints from other ions that are calculated directly from regression coefficients of PCR and stoichiometry of amphibole. Plots are in atom per formula unit (apfu).

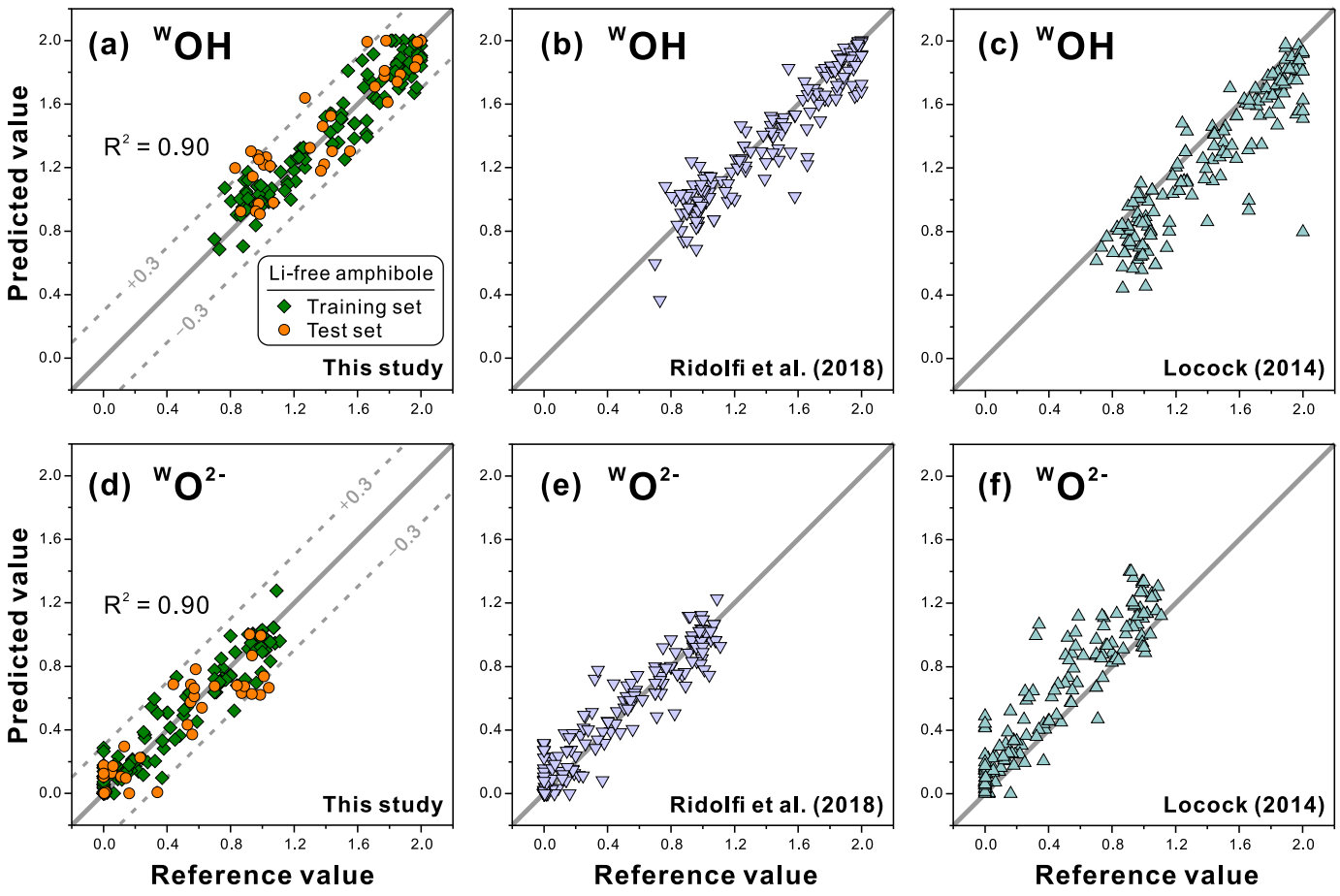

Fig. 4. Comparison between predicted and reference ^{W}OH and $^{W}O^{2-}$ for the working dataset of Li-free amphibole based on different methods. For using the method of this study, ^{W}OH is directly derived from regression coefficients of PCR, and $^{W}O^{2-}$ = 2-F-Cl- ^{W}OH. For using the method of Locock (2014), $^{W}O^{2-} = 2 \times {}^{C}Ti$, and ^{W}OH = 2-F-Cl- $^{W}O^{2-}$. Plots are in atom per formula unit (apfu).

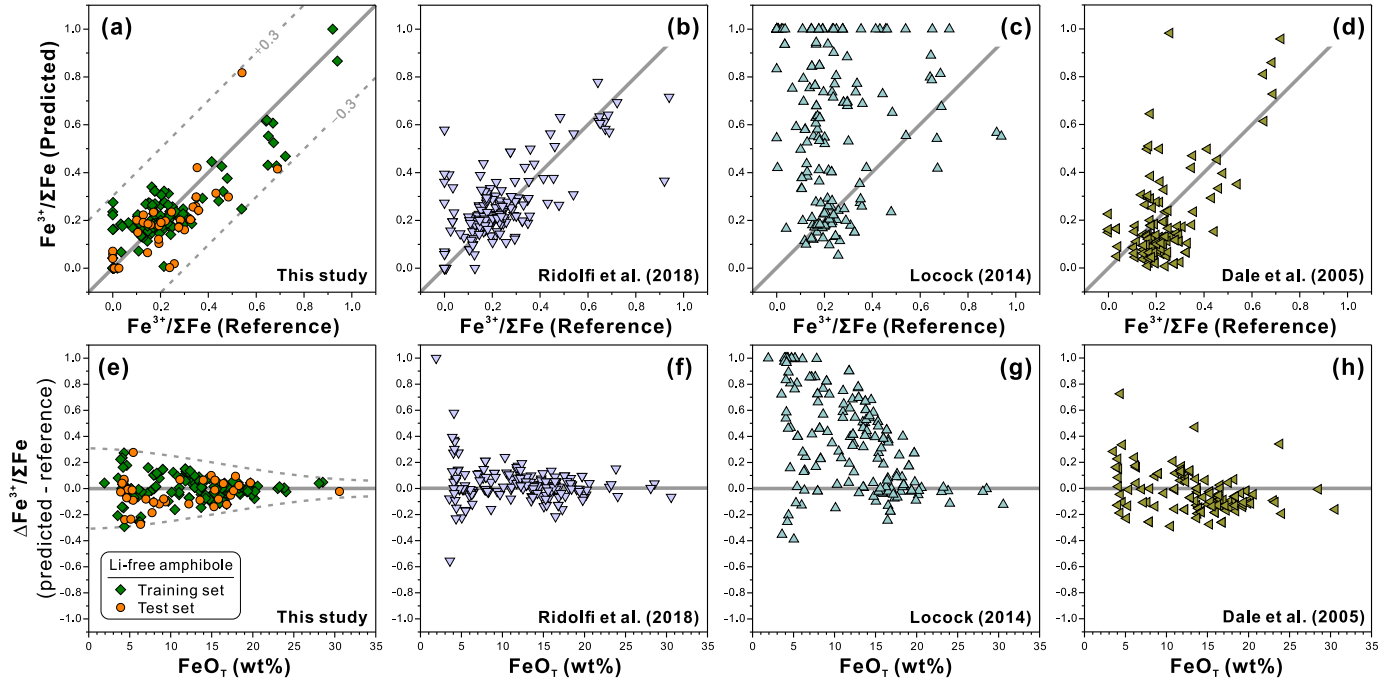

Fig. 5. Upper panels: comparison between predicted and reference $Fe^{3+}/\Sigma Fe$ for the working dataset. Lower panels: correlation between $\Delta Fe^{3+}/\Sigma Fe$ (predicted - reference) and FeO_T content of reference amphibole.

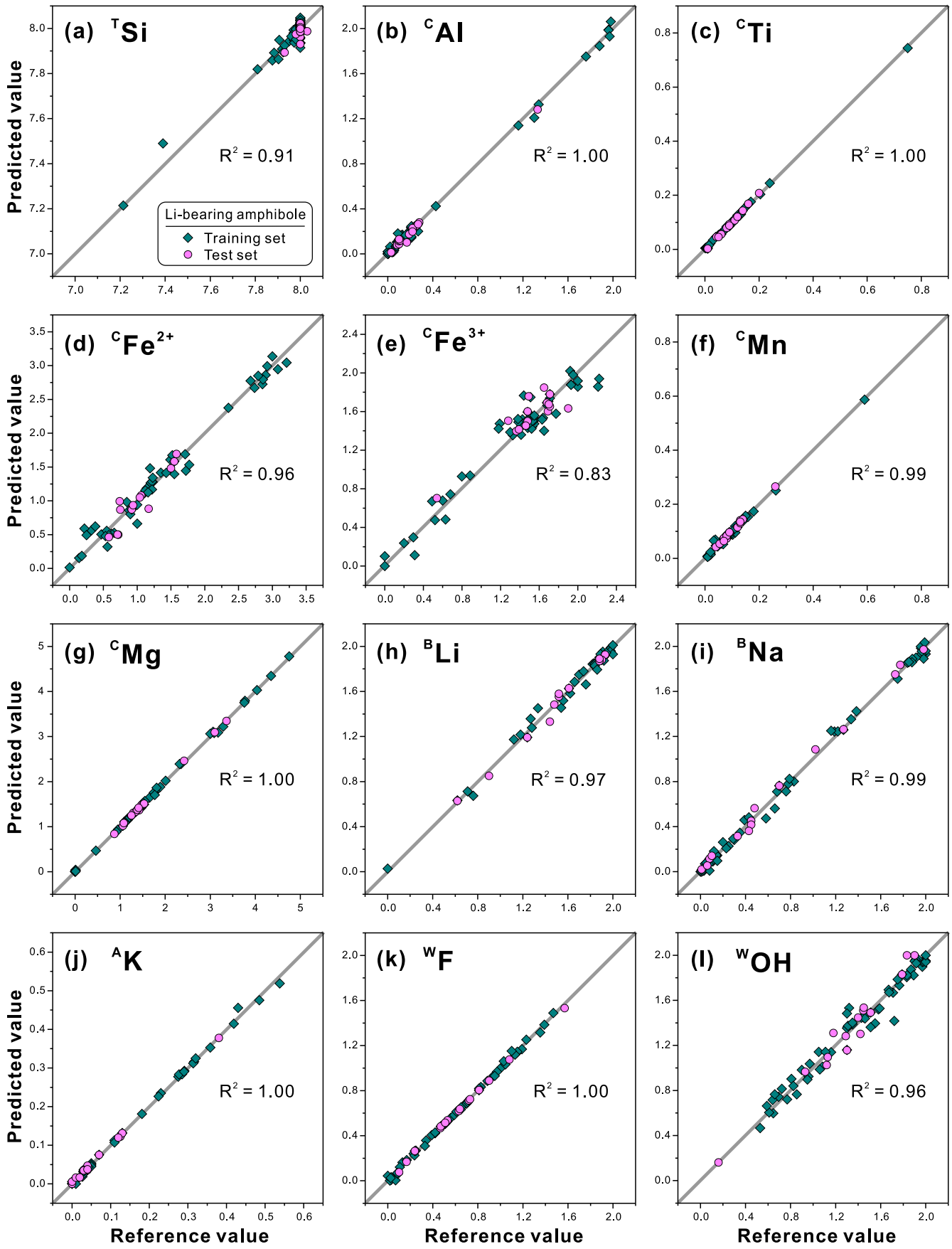

Fig. 6. Comparison between predicted and reference ions for the working dataset of Li-bearing amphibole. Predictions for these ions are calculated directly from regression coefficients of PCR at Step 5. Plots are in atom per formula unit (apfu).

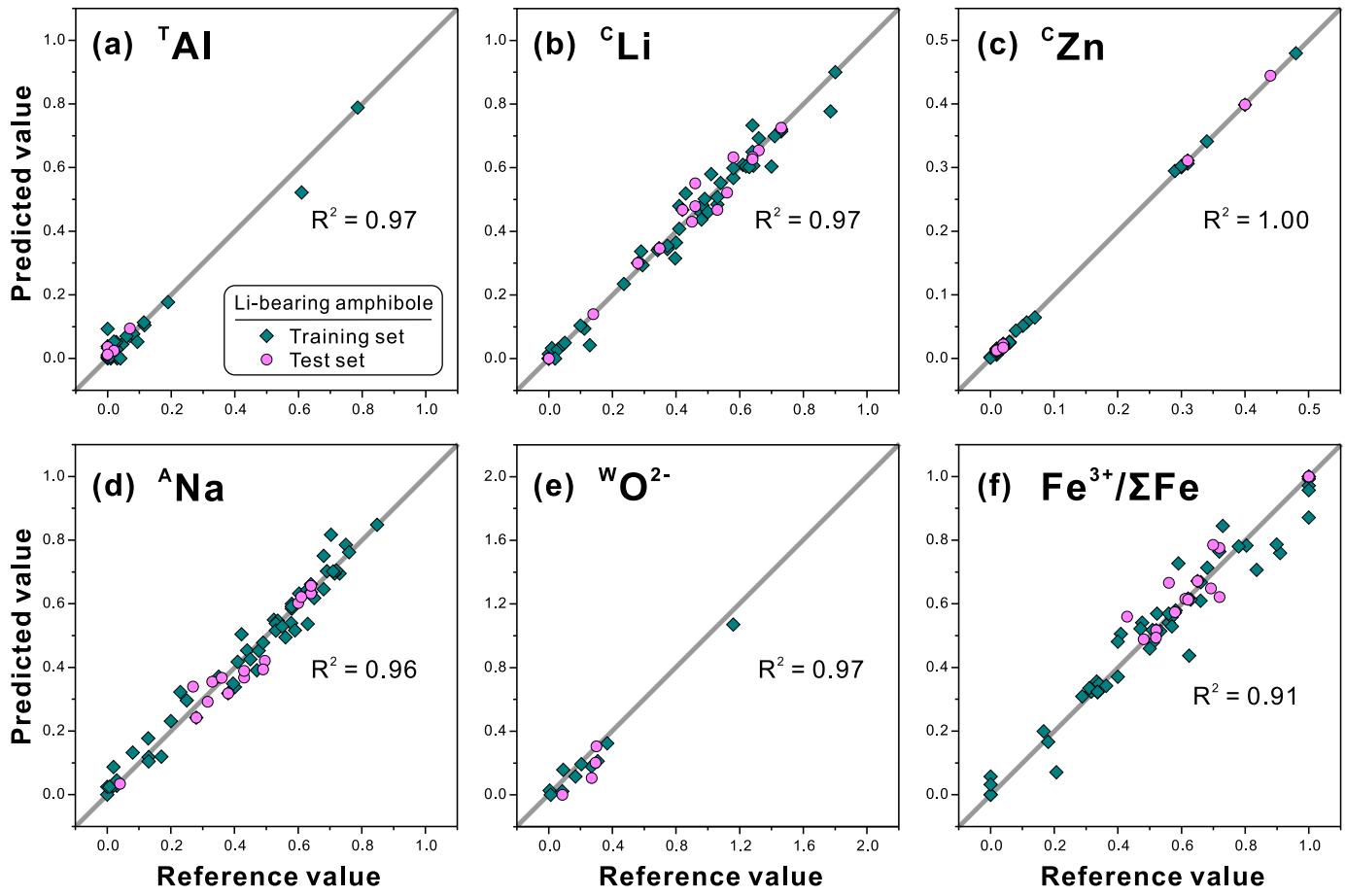

Fig. 7. Comparison between predicted and reference ions for the working dataset of Li-bearing amphibole. Predictions for these ions are calculated at Step 7 based on constraints from other ions that are calculated directly from regression coefficients of PCR and stoichiometry of Li-bearing amphibole. Plots are in atom per formula unit (apfu).

The predicted ion occupations for Li-bearing amphibole follow the 1:1 trend line for the training set as well as the independent test set, which demonstrates the high reliability and robustness of our model. As plotted in Fig. 6 for the ions calculated at Sept 5, the site occupations at ^{C}Ti, ^{C}Mn, ^{C}Mg, ^{B}Ca (not plotted here), ^{A}K, ^{W}F have prediction uncertainties within ±0.01 apfu, and the site occupations at ^{T}Si, ^{C}Al, ^{B}Na and ^{B}Li show prediction uncertainties commonly within ±0.1 apfu, while the site occupations at $^{C}Fe^{2+}$, $^{C}Fe^{3+}$ and ^{W}OH have a maximum prediction uncertainty of ca. ±0.2 apfu. Subsequently, site occupations that are calculated at Step 7 (Fig. 7) show similar prediction uncertainties as for the other ions calculated at Sept 5. The comparison for ^{W}Cl is not shown here due to its extreme low extent. An important outcome is that the predicted $Fe^{3+}/\Sigma Fe$ ratios are in good agreement with the reference values, and the discrepancies are mostly within ±0.15 (Fig. 7f).

4.2. Comparison with other methods

4.2.1. Estimation of ^{W}OH and $^{W}O^{2-}$

Fig. 4 compares the predicted site occupancy of ^{W}OH and $^{W}O^{2-}$ by our model and the results predicted by other two methods for Li-free amphiboles. All these methods have an assumption for the O(3) site that $2 = {}^{W}OH + {}^{W}F + {}^{W}Cl + {}^{W}O^{2-}$. The method used by Locock (2014) for estimating $^{W}O^{2-}$ at O(3) site follows the empirical equation of $^{W}O^{2-} = 2 \times Ti$ as recommended in the 2012 IMA report (Hawthorne et al., 2012). The predicted $^{W}O^{2-}$ by Ridolfi et al. (2018) is built on a regression equation derived via a mass-based calibration procedure. In these two methods, ^{W}OH content is calculated by 2 minus the other three ions. Differently, our model predicts ^{W}OH through a PCR procedure (see above Section), and $^{W}O^{2-}$ is indirectly calculated via Eq. 19.

The predictions of our model and that of Ridolfi et al. (2018) both show good agreements with reference data in terms of ^{W}OH and $^{W}O^{2-}$ (Fig. 4a–b and Figs. 4d–e), with prediction errors within ±0.3 apfu for our model and within ±0.5 apfu for the model of Ridolfi et al. (2018). The result yielded from the method proposed by Locock (2014) is shown in Fig. 4f, and the estimated $^{W}O^{2-}$ is generally significantly overestimated up to 0.8 apfu, which also induces a compensatory underestimation for ^{W}OH (Fig. 4c). The cause for this overestimation is imbed in its intrinsic principal, i.e., the Ti-oxy dehydration mechanism as described in Eq. 20:

$$^{M(1)}Ti^{4+} + 2\,{}^{W}O^{2-} = {}^{M(1)}\left(Mg, Fe^{2+}\right) + 2\,{}^{W}OH^{-} \quad (20)$$

in which M(1) denotes a type of crystallographic site in the amphibole structure (Hawthorne et al., 2012; Hawthorne and Oberti, 2007; Oberti et al., 1992). As described above, C sites are composed of three types of octahedrally coordinated sites 2 M(1), 2 M(2) and 1 M(3). The two M(1) sites and one M(3) site are discriminated by the cis-position of the two coordinated OH groups in the former and trans-position of the two coordinated OH groups in the latter. Therefore, the occurrence of high charge cations (≥3+) at 2 M(1) and/or 1 M (3) sites are responsible for the occurrence of $^{W}O^{2-}$ or dehydrogenation process in amphibole structure (Hawthorne et al., 2012), as commonly believed as the Ti-oxy dehydration/deprotonation mechanism (Eq. 20). However, although Ti preferentially occupies M(1) sites (Oberti et al., 1992), it can also be distorted at M(1), M(2) and M

(3) sites (Gatta et al., 2017; Hawthorne et al., 1998; Tiepolo et al., 1999). The occurrence of Ti at M(3) site does not change the principal of $^{W}O^{2-} = 2^{*}Ti$, but the significant occurrence of Ti at M(2) site for Ti-rich amphibole will obviously cause the overestimation of $^{W}O^{2-}$, because the total Ti at C sites is used in equation $^{W}O^{2-} = 2^{*}Ti$ as an approximation for $^{M(1)}Ti$. Therefore, the relation of $^{W}O^{2-} = 2^{*}Ti$ could not yield a reasonable estimation without a careful structure refinement for site occupation of Ti.

4.2.2. Estimation of $Fe^{3+}/\Sigma Fe$ ratio

Fig. 5 illustrates the $Fe^{3+}/\Sigma Fe$ ratios predicted by our model and by other three methods for Li-free amphiboles. Ridolfi et al. (2018) estimated Fe^{3+} based on a charge balance equation after obtaining estimated $^{W}O^{2-}$ from another independent equation. Locock (2014) used the recommended procedure in the 2012 IMA report (Hawthorne et al., 2012), and calculated a mean value for Fe^{3+} from the minimum and maximum values of estimated Fe^{3+}; the latter are obtained by constraining the sum of cations to a specific value and adjusting the Fe^{3+} amount according to charge balance (Hawthorne et al., 2012). The method used by Dale et al. (2005) was similar to the recommended procedure in the 2012 IMA report, but with more restricted application criteria.

There is a strong tendency of overestimation of the $Fe^{3+}/\Sigma Fe$ ratios estimated by the method of Locock (2014) for Li-free amphibole dataset (Fig. 5c and g). The $Fe^{3+}/\Sigma Fe$ ratios estimated following the method of Dale et al. (2005) are relatively better, and the majority of discrepancies to the reference data are within ±0.3 with a maximum error up to 0.7 (Fig. 5d and h). The application of the method proposed by Ridolfi et al. (2018) yields generally much lower errors for the predicted $Fe^{3+}/\Sigma Fe$ ratios that are mostly within ±0.3 (Fig. 5b), but with a number of low Fe samples ($FeO_T < 5\%$) showing higher errors up to ±0.6 (Fig. 5f). Using the model proposed in this study, for the total 173 amphibole samples from both training set and an independent test set, we obtained better predictions for $Fe^{3+}/\Sigma Fe$ ratio and the resulting errors are within ±0.3 for amphiboles with $FeO_T \geq 5$ wt% and within ± 0.2 for amphiboles with $FeO_T \geq 10$ wt% (Fig. 5a and e).

4.3. Independent application for estimating $Fe^{3+}/\Sigma Fe$ ratio

The recommended nomenclature scheme for amphiboles in the 2012 IMA report (Hawthorne et al., 2012) has received critical comments on the applicability of nomenclature criteria (Evans and Joy, 2014). The main controversial point is that the new nomenclature criteria require accurate $^{C}M^{3+}$ estimations (high-charge cations at C sites); however, the acquirement of accurate Fe^{3+} content is usually difficult to obtain without an independent analysis. In this study, Fe oxidation state can be estimated by the proposed PCR method described above.

In order to test the applicability and the robustness of our model in terms of estimating $Fe^{3+}/\Sigma Fe$ ratio for both Li-free and Li-bearing

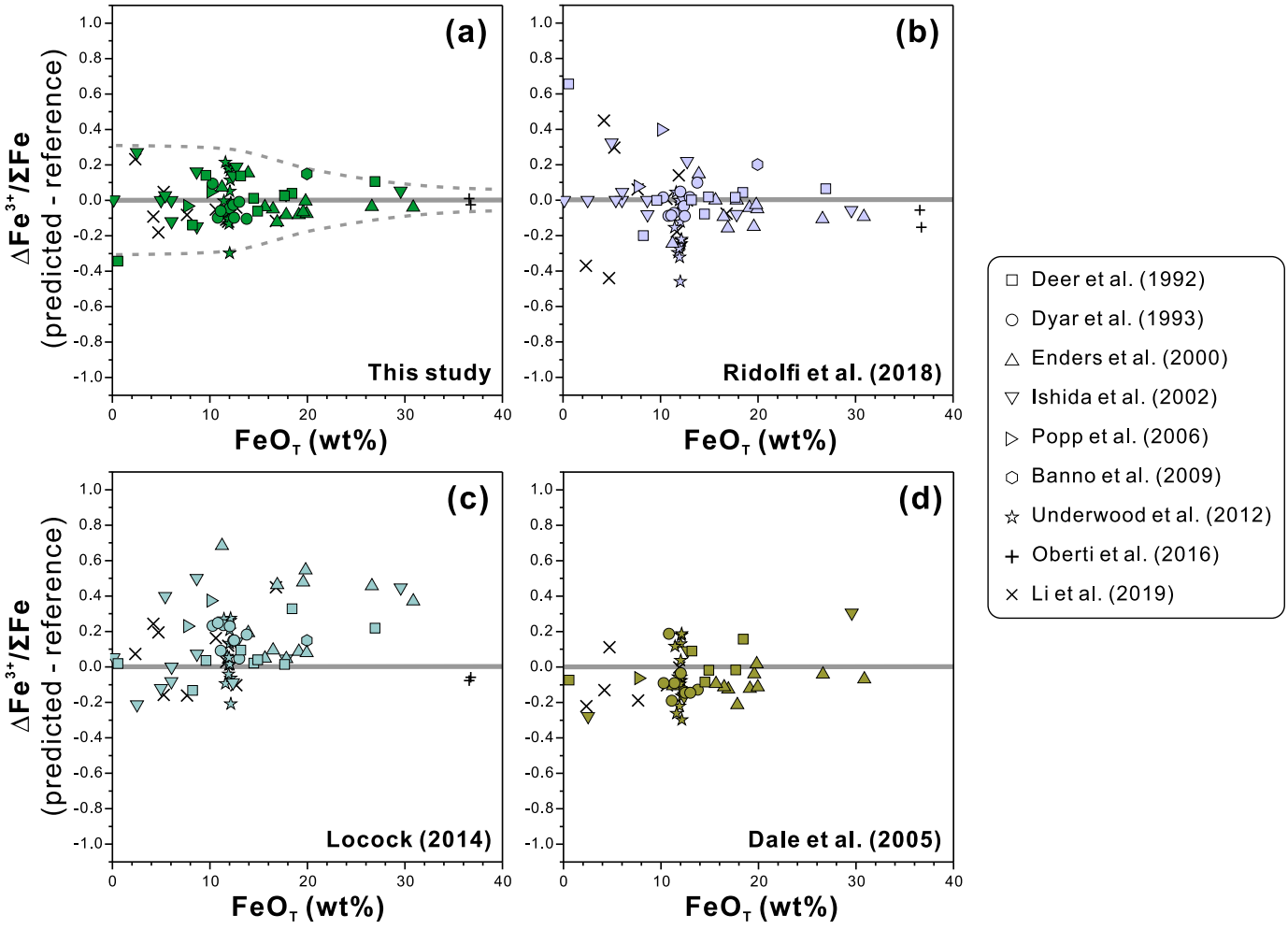

Fig. 8. Correlation between $\Delta Fe^{3+}/\Sigma Fe$ (predicted - reference) and FeO_T content of independent amphibole compositions for which $Fe^{3+}/\Sigma Fe$ were reported in literature. Note that the Li-bearing amphiboles of Enders et al. (2000) that are treated with the method for Li-bearing amphiboles.

amphiboles, we also performed calculations for another independent dataset of amphiboles with $Fe^{3+}/\Sigma Fe$ ratios measured independently that are, however, not involved in the previous evaluation of formula calculation (shown as calculation examples in Supplementary Table 1). It is worth mentioning that Fe^{3+} contents of amphiboles from Enders et al. (2000) are calculated using the model for Li-bearing amphiboles, because these amphiboles contain substantial Li contents. The Fe^{3+} contents of these amphibole have been analyzed by various methods, such as wet chemistry (Deer et al., 1992; Li et al., 2019; Underwood et al., 2012), Mössbauer analysis (Dyar et al., 1993; Enders et al., 2000; Ishida et al., 2002; Oberti et al., 2016; Popp et al., 2006), or adjusted by SREF data (Banno et al., 2009). The errors in estimated $Fe^{3+}/\Sigma Fe$ ratio by our model and the other three models are plotted in Fig. 8, and it is clear that the general prediction uncertainties of our model are consistent with the pattern shown in Fig. 5, yielding prediction errors within ±0.3 for amphiboles with $FeO_T < 15$ wt% and even lower errors for amphiboles with higher FeO_T contents.

4.4. Identification of post-formation oxidation

Fe^{3+} could play a similar role as Ti in the course of Ti-oxy substitution (Hawthorne et al., 2012), which could be described by the following equation:

$$^{M(1)}Fe^{2+} + {}^{W}OH^{-} = {}^{M(1)}Fe^{3+} + {}^{W}O^{2-} + 1/2H_2\,(gas) \quad (21)$$

This Fe-oxidation/deprotonation process is usually regarded as a post-formation process (i.e. post-formation oxidation). Experimental studies on the high-temperature behavior of amphibole under heat-treated conditions provide an opportunity to investigate the deprotonation processes (Della Ventura et al., 2017; Della Ventura et al., 2018a; Della Ventura et al., 2018b; Oberti et al., 2018; Oberti et al., 2019; Popp et al., 2006; Zema et al., 2012). The combination of multiple techniques such as SREF, XANES, Mössbauer, Raman and FTIR measurements performed before and after the heat-treatment have revealed several complex cation and electron migration patterns for the deprotonation process. The iron oxidation process of Fe-rich riebeckite under HT experiment has been observed to involve multiple dynamic processes at different temperatures (Della Ventura et al., 2018b), including the shrinkage of M(1) octahedra accompanying with Fe^{2+} oxidized to Fe^{3+}, the loss of hydrogen at the anionic O(3) site, and the migration/diffusion of hydrogen out of crystal which formed water molecules with oxygen at the crystal surface (Della Ventura et al., 2018a; Della Ventura et al., 2018b). Another study on oxidation process in riebeckite has indicated an alternative cation migration process between variable sites, involving exclusive oxidation of Fe^{2+} to Fe^{3+} at M(1) sites and an exchange process of $^{M(2,3)}Fe^{2+}_{-1}\,{}^{M(1)}Mg_{-1}$ $^{M(1)}Fe^{2+}_{1}\,{}^{M(2,3)}Mg_{1}$ (Oberti et al., 2018). Similar cation migration processes have also been observed in orthorhombic Fe-rich holmquistite, but with a higher complexity due to the presence of Li, because Li migrates from M(4) to M(3) site and Fe^{2+} redistributes among the M(1) and M(3) site with the Fe oxidation occurring only at M(1) site (Oberti et al., 2019).

A Fe-oxidation process involves the oxidation of Fe^{2+} to Fe^{3+}, H^+ and electron diffusion, inter-crystal cation redistribution and crystal structural unit rearrangement. However, only H^+ and electron are generally open to the environmental system during oxidation. By comparing the compositions of the oxidized amphiboles to that of original amphiboles, previous studies show that the exclusive changes are decreasing of H_2O content and increasing of Fe^{3+} content. Dyar et al. (2016) presented several pairs of amphiboles whose compositions were analyzed before and after the heat treatment, and only a significant H_2O loss could be observed when comparing the EMPA data.

Since our method predicts the OH and Fe^{3+} content through solely the EMPA data, only the initial Fe oxidation state and the associated H_2O content could be obtained by our method, which cannot reflect influences from a potential post-formation oxidation process. Therefore, if the H_2O and/or Fe^{3+} contents of oxidized amphibole are analyzed by other independent analytical methods, the significant difference (larger than the prediction error of our model) between the direct analyzed value and value obtained by our method on EMPA data could be used to estimate the amount of released H_2O and/or the possible change of oxygen fugacity between the oxidization stage and the initial crystallization stage. The distinction between the oxidization stage and the initial crystallization stage is usually difficult since there are various possible explanations for the observed Fe^{3+} and H^+ content. The inverse relation between increasing Fe^{3+} and decreasing H^+ content of slightly oxidized hornblende megacrysts from basaltic flows has been related to a partial dehydration/oxidation during magma ascent, and was interpreted to indicate that the mantle source has high H^+ content rather than a high oxygen fugacity (Dyar et al., 1992).

However, by applying the H diffusivity data, the nearly totally dehydrogenated oxy-amphibole may not be explained by the partial H loss during quick ascent but is indicative of its re-equilibrium in an oxidized metasomatic environment that can elevate its $Fe^{3+}/\Sigma Fe$ ratio strongly (Dyar et al., 1993). Indeed, for this kind of oxy-amphiboles that crystalized initially at oxidized conditions rather than being affected by post-formation oxidation, our model estimated consistently high $Fe^{3+}/\Sigma Fe$ ratios as that originally determined by Dyar et al. (1993) (Fig. 8a).

Therefore, using the calculation model for amphibole formula proposed in this study, we can estimate the $Fe^{3+}/\Sigma Fe$ ratio of amphiboles corresponding to the equilibrium conditions at which they were formed, but cannot reflect the changes due to subsequent potential oxidation processes. On the other hand, if a large discrepancy in terms of $Fe^{3+}/\Sigma Fe$ ratio between estimates of our calculation model and independent measurements is observed, we suggest that it is possible that a post-formation oxidation process might have occurred.

4.5. Amphibole nomenclature based on 2012 IMA report

A nomenclature function has been added to the calculation sheet of amphibole formula (Supplementary Table 1). The invoked nomenclature rules are based on the 2012 IMA report for the nomenclature of amphibole supergroup, and the user is recommended to refer to the original publication (Hawthorne et al., 2012) for the detailed explanations of classification criteria.

As stated above, the working datasets of amphiboles in this study comprise only monoclinic amphiboles with low MnO contents (<1 wt %), but neither orthorhombic amphibole nor BeO, Pb_2O and Mn_2O_3 bearing amphibole were involved. Therefore, the calculation method in this study is not applicable to these types of amphibole. Analogously, end-members/rootnames which have formulae dominated by Be^{2+} and Pb^{2+} (joesmithite), Mn^{2+} (e.g. rootname 3, and rootname 11–15) and Mn^{3+} (e.g. dellaventuraite and ungarettiite), and that belong to orthorhombic (e.g. anthophyllite and gedrite) phases, are not implemented in the nomenclature function. Similarly, the prefix "mangani", which indicates the high content of Mn^{3+} at octahedral sites, is also excluded in our nomenclature function.

According to the 2012 IMA amphibole classification criterion, the ^{W}(OH, F, Cl)-dominant amphibole group is divided into subgroups based on the dominant charge-arrangements and on the type of B-group cations (Hawthorne et al., 2012). The cations of ^{B}Ca, ^{B}Na, ^{B}Li and the sum of the other divalent cations $^{B}\Sigma M^{2+}$ ($^{B}\Sigma M^{2+} = {}^{B}Mg + {}^{B}Fe^{2+} + {}^{B}Mn^{2+}$), and their combinations (e.g. $^{B}Li + {}^{B}\Sigma M^{2+}$) are basic constituents whose ratios define the subgroup boundaries. However, the cations Mg^{2+}, Fe^{2+}, Mn^{2+} and Li^+ could reside at the B site and C site, and Ca^{2+} and Na^+ usually reside at the B site and/or A site, which results in difficulties on calculating the accurate proportions of a given ion at distinct sites by using common formula calculation methods from EMPA data without a structure refinement (SREF). It is a challenging task that must be solved when applying the new nomenclature scheme

of IMA, and we provide a practical approach from the perspective of machine learning on the nomenclature issue, for both Li-free and Li-bearing amphiboles that cover a wide compositional range of natural amphiboles.

5. Conclusions

A new machine learning method based on principal components regression for calculating amphibole formula from routine EMPA data has been established. The results of testing on independent datasets confirmed its capability and reliability for both ^{W}(OH, F, Cl)-dominant and ^{W}O-dominant amphiboles. Our method is particularly useful for estimating ^{W}OH, $^{W}O^{2-}$ and $Fe^{3+}/\Sigma Fe$ ratio that cannot be derived directly from EMPA data. The uncertainties in estimated ^{W}OH and $^{W}O^{2-}$ are within ± 0.3 apfu, while the uncertainty in estimated $Fe^{3+}/\Sigma Fe$ ratio is proven within ± 0.3 for amphiboles with $FeO_T \geq 5$ wt% and further less than ± 0.2 for amphiboles with $FeO_T \geq 10$ wt%. The calculation sheet also provides an automatic nomenclature function according to the recommended classification criterion of the 2012 IMA report.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lithos.2020.105469.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This study was supported by German Research Foundation (DFG) (BE 1720/40) and National Natural Science Foundation of China (NSFC) (41902052). We thank He Huang and an anonymous reviewer for their insightful comments and Michael Roden for efficient editorial handling. The Microsoft Excel spreadsheet for calculating amphibole formula from EMPA data is provided as supplementary file online.

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. Wiley Interdisc. Rev. 2, 433–459.
Aitchison, J., 1983. Principal component analysis of compositional data. Biometrika 70, 57–65.
Aitchison, J., 1984. The statistical analysis of geochemical compositions. J. Int. Assoc. Math. Geol. 16, 531–564.
Banno, Y., Miyawaki, R., Matsubara, S., Sato, E., Nakai, I., Matsuo, G.-I., Yamada, S., 2009. Potassic-ferropargasite, a new member of the amphibole group, from Kabutoichiba, Mie Prefecture, Central Japan. J. Mineral. Petrol. Sci. 104, 374–382.
Dale, J., Powell, R., White, R.W., Elmer, F.L., Holland, T.J.B., 2005. A thermodynamic model for Ca–Na clinoamphiboles in Na2O–CaO–FeO–MgO–Al2O3–SiO2–H2O–O for petrological calculations. J. Metamorph. Geol. 23, 771–791.
Davis, J.C., 2002. Statistics and Data Analysis in Geology. Wiley.
Deer, W., Howie, R., Zussman, J., 1992. An Introduction to the Rock-Forming Minerals. Longman.
Della Ventura, G., Susta, U., Bellatreccia, F., Marcelli, A., Redhammer Günther, J., Oberti, R., 2017. Deprotonation of Fe-dominant amphiboles: Single-crystal HT-FTIR spectroscopic studies of synthetic potassic-ferro-richterite. Am. Mineral. 102, 117.
Della Ventura, G., Galdenzi, F., Cibin, G., Oberti, R., Xu, W., Macis, S., Augusto, M., 2018a. Iron oxidation dynamics vs. temperature of synthetic potassic-ferro-richterite: a XANES investigation. Phys. Chem. Chem. Phys. 20, 21764–21771.
Della Ventura, G., Mihailova, B., Susta, U., Guidi, M.C., Marcelli, A., Schlüter, J., Oberti, R., 2018b. The dynamics of Fe oxidation in riebeckite: a model for amphiboles. Am. Mineral. 103, 1103–1111.
Dyar, M.D., 2002. Optical and Mössbauer spectroscopy of iron in micas. Rev. Mineral. Geochem. 46, 313–349.
Dyar, M.D., McGuire, A.V., Mackwell, S.J., 1992. Fe^{3+}/H^{+} and D/H in kaersutites—Misleading indicators of mantle source fugacities. Geology 20, 565–568.
Dyar, M.D., Mackwell, S.J., McGuire, A.V., Cross, L.R., Robertson, J.D., 1993. Crystal chemistry of Fe^{3+} and H^{+} in mantle kaersutite: implications for mantle metasomatism. Am. Mineral. 78, 968–979.
Dyar, M.D., et al., 2016. Use of multivariate analysis for synchrotron micro-XANES analysis of iron valence state in amphiboles. Am. Mineral. 101, 1171–1189.
Enders, M., Speer, D., Maresch, W.V., McCammon, C.A., 2000. Ferric/ferrous iron ratios in sodic amphiboles: Mössbauer analysis, stoichiometry-based model calculations and the high-resolution microanalytical flank method. Contrib. Mineral. Petrol. 140, 135–147.
Evans, B.W., Joy, B.R., 2014. REPLY TO:"Clinoferrogedrite in the contact-metamorphosed Biwabik Iron Formation, northeastern Minnesota: DISCUSSION" by Williams et al. THE CANADIAN MINERALOGIST. Can. Mineral. 52, 921–924.
Gatta, G.D., McIntyre, G.J., Oberti, R., Hawthorne, F.C., 2017. Order of $^{[6]}Ti^{4+}$ in a Ti-rich calcium amphibole from Kaersut, Greenland: a combined X-ray and neutron diffraction study. Phys. Chem. Miner. 44, 83–94.
Giesting, P.A., Filiberto, J., 2014. Quantitative models linking igneous amphibole composition with magma Cl and OH content. Am. Mineral. 99, 852–865.
Giesting, P.A., Schwenzer, S.P., Filiberto, J., Starkey, N.A., Franchi, I.A., Treiman, A.H., Tindle, A.G., Grady, M.M., 2015. Igneous and shock processes affecting chassignite amphibole evaluated using chlorine/water partitioning and hydrogen isotopes. Meteorit. Planet. Sci. 50, 433–460.
Hålenius, U., Bosi, F., 2012. Cation ordering in Pb^{2+}-bearing, Mn^{3+}-rich pargasite from Långban, Sweden. Am. Mineral. 97, 1635–1640.
Hawthorne, F.C., Oberti, R., 2007. Amphiboles: crystal chemistry. Rev. Mineral. Geochem. 67, 1–54.
Hawthorne, F.C., Oberti, R., Zanetti, A., Czamanske, G.K., 1998. The role of Ti in hydrogen-deficient amphiboles: sodic-calcic and sodic amphiboles from Coyote Peak, California. Can. Mineral. 36, 1253–1265.
Hawthorne, F.C., Oberti, R., Harlow, G.E., Maresch, W.V., Martin, R.F., Schumacher, J.C., Welch, M.D., 2012. Nomenclature of the amphibole supergroup. Am. Mineral. 97, 2031–2048.
Henry, D.J., Daigle, N.M., 2018. Chlorine incorporation into amphibole and biotite in high-grade iron-formations: interplay between crystallography and metamorphic fluids. Am. Mineral. 103, 55–68.
Holland, T., Blundy, J., 1994. Non-ideal interactions in calcic amphiboles and their bearing on amphibole-plagioclase thermometry. Contrib. Mineral. Petrol. 116, 433–447.
Iezzi, G., Camara, F., Oberti, R., Della Ventura, G., Pedrazzi, G., Robert, J.-L., 2004. Synthesis, crystal structure and crystal chemistry of ferri-clinoholmquistite, $\square Li_2Mg^3Fe^{3+}2\text{-}Si_8O_{22}(OH)_2$. Phys. Chem. Miner. 31, 375–385.
Ishida, K., Hawthorne, F.C., Ando, Y., 2002. Fine structure of infrared OH-stretching bands in natural and heat-treated amphiboles of the tremolite-ferro-actinolite series. Am. Mineral. 87, 891–898.
Jolliffe, I.T., 2002. Principal Component Analysis. Springer, New York.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York.
Labotka, T.C., 1983. Analysis of the compositional variations of biotite in pelitic hornfelses from northeastern Minnesota. Am. Mineral. 68, 900–914.
Lamb, W.M., Popp, R.K., 2009. Amphibole equilibria in mantle rocks: determining values of mantle a H_2O and implications for mantle H_2O contents. Am. Mineral. 94, 41–52.
Lamb, W.M., Guillemette, R., Popp, R.K., Fritz, S.J., Chmiel, G.J., 2012. Determination of Fe3+/Fe using the electron microprobe: a calibration for amphiboles. Am. Mineral. 97, 951–961.
Leake, B.E., et al., 1997. Nomenclature of amphiboles; report of the Subcommittee on Amphiboles of the International Mineralogical Association Commission on New Minerals and Mineral Names. Eur. J. Mineral. 9, 623–651.
Li, X., Zhang, C., Almeev, R.R., Zhang, X.-C., Zhao, X.-F., Wang, L.-X., Koepke, J., Holtz, F., 2019. Electron probe microanalysis of Fe2+/ΣFe ratios in calcic and sodic-calcic amphibole and biotite using the flank method. Chem. Geol. 509, 152–162.
Li, X., Zhang, C., Behrens, H., Holtz, F., 2020. Calculating biotite formula from electron microprobe analysis data using a machine learning method based on principal components regression. Lithos 356-357, 105371.
Locock, A.J., 2014. An Excel spreadsheet to classify chemical analyses of amphiboles following the IMA 2012 recommendations. Comput. Geosci. 62, 1–11.
McCammon, C., Chinn, I., Gurney, J., McCallum, M., 1998. Ferric iron content of mineral inclusions in diamonds from George Creek, Colorado determined using Mössbauer spectroscopy. Contrib. Mineral. Petrol. 133, 30–37.
McCubbin, F.M., Smirnov, A., Nekvasil, H., Wang, J., Hauri, E., Lindsley, D.H., 2010. Hydrous magmatism on Mars: a source of water for the surface and subsurface during the Amazonian. Earth Planet. Sci. Lett. 292, 132–138.
Mevik, B.-H., Wehrens, R., 2007. The pls package: principal component and partial least squares regression in R. J. Stat. Softw. 18, 1–24.
Mevik, B.-H., Wehrens, R., Liland, K.H., 2019. pls: Partial Least Squares and Principal Component Regression R package version 2.7–1.
Moore, P., Davis, A., Van Derveer, D., Gupta, P.S., 1993. Joesmithite, a plumbous amphibole revisited and comments on bond valences. Mineral. Petrol. 48, 97–113.
Oberti, R., Ungaretti, L., Cannillo, E., Hawthorne, F.C., 1992. The behaviour of Ti in amphiboles; I, four-and six-coordinate Ti in richterite. Eur. J. Mineral. 4, 425–439.
Oberti, R., Boiocchi, M., Zema, M., Della Ventura, G., 2016. Synthetic potassic-ferro-richterite: 1. Composition, crystal structure refinement, and HT behavior by in operando single-crystal X-ray diffraction. Can. Mineral. 54, 353–369.
Oberti, R., Boiocchi, M., Zema, M., Hawthorne, F.C., Redhammer, G.J., Susta, U., Della Ventura, G., 2018. The high-temperature behaviour of riebeckite: expansivity, deprotonation, selective Fe oxidation and a novel cation disordering scheme for amphiboles. Eur. J. Mineral. 30, 437–449.
Oberti, R., Boiocchi, M., Zema, M., 2019. Thermoelasticity, cation exchange and deprotonation in Fe-rich holmquistite: towards a crystal-chemical model for the high-temperature behavior of orthorhombic amphiboles. Am. Mineral. 104, 1829–1839.
Popp, R.K., Hibbert, H.A., Lamb, W.M., 2006. Oxy-amphibole equilibria in Ti-bearing calcic amphiboles: experimental investigation and petrologic implications for mantle-derived amphiboles. Am. Mineral. 91, 54–66.
Putirka, K., 2016. Amphibole thermometers and barometers for igneous systems and some implications for eruption mechanisms of felsic magmas at arc volcanoes. Am. Mineral. 101, 841–858.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ridolfi, F., Renzulli, A., 2012. Calcic amphiboles in calc-alkaline and alkaline magmas: thermobarometric and chemometric empirical equations valid up to 1,130°C and 2.2 GPa. Contrib. Mineral. Petrol. 163, 877–895.

Ridolfi, F., Zanetti, A., Renzulli, A., Perugini, D., Holtz, F., Oberti, R., 2018. AMFORM, a new mass-based model for the calculation of the unit formula of amphiboles from electron microprobe analyses. Am. Mineral. 103, 1112–1125.

Schumacher, J.C., 1991. Empirical ferric iron corrections: necessity, assumptions, and effects on selected geothermobarometers. Mineral. Mag. 55, 3–18.

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

Spear, F., Kimball, K., 1984. RECAMP—A Fortran IV program for estimating Fe3+ contents in amphiboles. Comput. Geosci. 10, 317–325.

Tait, K.T., Hawthorne, F.C., Grice, J.D., Ottolini, L., Nayak, V., 2005. Dellaventuraite, $NaNa_2(MgMn^{3+2}Ti^{4+}Li)Si_8O_{22}O_2$, a new anhydrous amphibole from the Kajlidongri Manganese Mine, Jhabua District, Madhya Pradesh, India. Am. Mineral. 90, 304–309.

Tiepolo, M., Zanetti, A., Oberti, R., 1999. Detection, crystal-chemical mechanisms and petrological implications of [6]Ti4+ partitioning in pargasite and kaersutite. Eur. J. Mineral. 11, 345–354.

Underwood, S., Feeley, T., Clynne, M., 2012. Hydrogen isotope investigation of amphibole and biotite phenocrysts in silicic magmas erupted at Lassen Volcanic Center, California. J. Volcanol. Geotherm. Res. 227, 32–49.

Waters, D., Charnley, N., 2002. Local equilibrium in polymetamorphic gneiss and the titanium substitution in biotite. Am. Mineral. 87, 383.

Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemom. Intell. Lab. Syst. 2, 37–52.

Yavuz, F., 2007. WinAmphcal: a Windows program for the IMA-04 amphibole classification. Geochem. Geophys. Geosyst. 8, Q01004.

Yoo, W., Mayberry, R., Bae, S., Singh, K., He, Q.P., Lillard Jr., J.W., 2014. A study of effects of multicollinearity in the multivariable analysis. Int. J. Appl. Sci. Technol. 4, 9–19.

Young, E., Virgo, D., Popp, R., 1997. Eliminating closure in mineral formulae with specific application to amphiboles. Am. Mineral. 82, 790–806.

Zema, M., Welch, M.D., Oberti, R., 2012. High-T behaviour of gedrite: thermoelasticity, cation ordering and dehydrogenation. Contrib. Mineral. Petrol. 163, 923–937.

Zhang, C., Holtz, F., Ma, C., Wolff, P., Li, X., 2012. Tracing the evolution and distribution of F and Cl in plutonic systems from volatile-bearing minerals: a case study from the Liujiawa pluton (Dabie orogen, China). Contrib. Mineral. Petrol. 164, 859–879.