Elsevier Editorial System(tm) for Computers & Geosciences

Manuscript Draft

Manuscript Number: CAGEO-D-15-00083R3

Title: Precursors predicted by artificial neural networks for mass balance calculations: quantifying hydrothermal alteration in volcanic rocks

Article Type: Research paper

Keywords: hydrothermal alteration; immobile elements; mineral exploration; neural networks; multi-layer perceptron; mass balance calculations.

Corresponding Author: Dr. Réal Daigneault,

Corresponding Author's Institution: Université du Québec à Chicoutimi

First Author: Sylvain Trépanier

Order of Authors: Sylvain Trépanier; Lucie Mathieu, Ph.D.; Réal Daigneault; Stéphane Faure, Ph.D.

Abstract: This study proposes an artificial neural networks-based method for predicting the unaltered (precursor) chemical compositions of hydrothermally altered volcanic rock. The method aims at predicting precursor's major components contents (SiO_2, FeOT, MgO, CaO, Na_2O, K_2O). The prediction is based on ratios of elements generally immobile during alteration processes; i.e. Zr, TiO_2, Al_2O_3, Y, Nb, Th, and Cr, which are provided as inputs to the neural networks. Multi-layer perceptron neural networks were trained on a large dataset of least-altered volcanic rock samples that document a wide range of volcanic rock types, tectonic settings and ages. The precursors thus predicted are then used to perform mass balance calculations. Various statistics were calculated to validate the predictions of precursors' major components, which indicate that, overall, the predictions are precise and accurate. For example, rank-based correlation coefficients were calculated to compare predicted and analysed values from a least-altered test dataset that had not been used to train the networks. Coefficients over 0.87 were obtained for all components, except for Na_2O (0.77), indicating that predictions for alkali might be less performant. Also, predictions are performant for most volcanic rock compositions, except for ultra-K rocks. The proposed method provides an easy and rapid solution to the often difficult task of determining appropriate volcanic precursor compositions to rocks modified by hydrothermal alteration. It is intended for large volcanic rock databases and is most useful, for example, to mineral exploration performed in complex or poorly known volcanic settings. The method is implemented as a simple C++ console program.

Januray 2016,

Dear editor Jef Caers,

Many thanks for your interest for our article entitled “Mass balance calculations for hydrothermally altered volcanic rocks using precursors predicted by artificial neural networks”. My co-authors and I have carefully read the reviewers’ comments and made the appropriate modifications to the manuscript.

The reviewer made three comments in his last review. His first comment is answered hereafter, the other two led use to modify our manuscript. We would like to take this opportunity to thank this anonymous reviewer for helping us improve our manuscript.

We hope that the manuscript is now mature and has reached the high quality required for publication in your journal.

Summary of modifications:

- Modifications have been made on lines 201-202, 211-214 and lines 321-325
- A table has been added (Table 3)

Kind regards,

Réal Daigneault and co-authors,

Centre d'études sur les Ressources minérales (CERM)
Université du Québec à Chicoutimi (UQAC)
555 Boulevard de l'université, Chicoutimi, Québec, G7H2B1, Canada

Below, detailed answers are given to each comments formulated by the reviewer. The answers are provided in the following format:

- *In italic, reviewer's comments are reproduced.*
- Below, you may find my answer to each comment.
- *Line numbers: in italic, correspond to the lines from the initial manuscript.*
- **Line numbers: in bold, correspond to the lines from the new manuscript.**

Comments of reviewer #1:

Why are oxides being used? It is stoichiometrically reasonable to use cation values. When mass balance calculations are carried out, they are based on cations, not oxides. This may be a nomenclature issues, but it is important to be consistent.

Answer: Indeed, we could have used cations, but we choose to use oxides because chemical analyses are generally provided as oxides, so the whole procedure was thought and calibrated for oxides. Note that the use of cations would not have modified our procedure, and that this is only a choice that we made early when we developed our procedure. In this respect, we followed authors such as Barrett and MacLean (1994), for example, who also use oxides in their procedure.

My reference to groups was intended to recognize that they are testing their methodology with a number of "rock types", which I interpreted to be, a priori, groups. Whilst I recognize that there is a continuum of rock types, it is useful to use groups, as it refines the methodology. The authors refer to the application to a wide range of volcanic rock types, which in this reviewer's opinion is a recognition that the data are composed of groups. I suggest that the authors add a statement, that although they refer to a range of volcanic rock types, which are groups, their methodology is applied across a continuum of compositions.

Answer: A statement has been added **lines 201-202** to include the reviewer's suggestion in the manuscript.

I accept the authors statements and revisions about the compositional nature of the data. It is important to discuss that the data are compositional and, if exceptions are being made, they must be justified. A detailed explanation **(should?)** *be given that describes the tests that they carried out to demonstrate that the use of ratios yielded similar results.*

Answer: Many thanks for accepting the explanations that we provided in our previous review. Following the reviewer's comments, detailed explanation on the test performed with ratios has been added to the manuscript (see **lines 211-214** and **lines 321-325**). Also, **table 3** has been added to demonstrate that both tests yield similar results.

Highlights

Method for predicting the fresh precursor to hydrothermally altered volcanic rocks.

Precursors are predicted using multi-layer perceptron neural networks.

Precursor compositions (major oxides) are predicted from immobile elements ratios.

Chemical method for quantifying alteration from large datasets of volcanic rocks.

New chemical method developed for grassroot or regional mineral exploration.

*Manuscript
Click here to download Manuscript: Manuscript_BM_MODIF-FOLLOWED.docx Click here to view linked References

Precursors predicted by artificial neural networks for mass balance calculations: quantifying hydrothermal alteration in volcanic rocks

Sylvain Trépanier [a, b], Lucie Mathieu [a], Réal Daigneault [c*], Stéphane Faure [a]

[a] CONSOREM - Mineral Exploration Research Consortium, 555 Boul. de l'Université, Chicoutimi, Canada, G7H 2B1

[b] Redevance Aurifères Osisko Ltd., Exploration Osisko Baie James – 300 rue Saint-Paul (office 200), Québec, Québec, G1K 7R1.

[c] Centre d'études sur les Ressources minérales (CERM) - Université du Québec à Chicoutimi, 555 Boul. de l'Université, Chicoutimi, Canada, G7H 2B1

* Corresponding author. E-mail address: real_daigneault@uqac.ca,
Tel.:1-418-545-5011. (R. Daigneault)

Keywords: hydrothermal alteration, immobile elements, mineral exploration, neural networks, multi-layer perceptron, mass balance calculations.

Abstract

This study proposes an artificial neural networks-based method for predicting the unaltered (precursor) chemical compositions of hydrothermally altered volcanic rock. The method aims at predicting precursor's major components contents (SiO_2, FeO^T, MgO, CaO, Na_2O, K_2O). The prediction is based on ratios of elements generally immobile during alteration processes; i.e. Zr, TiO_2, Al_2O_3, Y, Nb, Th, and Cr, which are provided as inputs to the neural networks. Multi-layer perceptron neural networks were trained on a large dataset of least-altered volcanic rock samples that document a wide range of volcanic rock types, tectonic settings and ages. The precursors thus predicted are then used to perform mass balance calculations. Various statistics were calculated to validate the predictions of precursors' major components, which indicate that, overall, the predictions are precise and accurate. For example, rank-based correlation coefficients were calculated to compare predicted and analysed values from a least-altered test dataset that had not been used to train the networks. Coefficients over 0.87 were obtained for all components, except for Na_2O (0.77), indicating that predictions for alkali might be less performant. Also, predictions are performant for most volcanic rock compositions, except for ultra-K rocks. The proposed method provides an easy and rapid solution to the often difficult task of determining appropriate volcanic precursor compositions to rocks modified by hydrothermal alteration. It is intended for large volcanic rock databases and is most useful, for example, to mineral exploration performed in complex or poorly known volcanic settings. The method is implemented as a simple C++ console program.

Formatted: Superscript

1. Introduction

Grassroots, or regional mineral exploration, necessitates the use of powerful and simple methods to interpret large amounts of chemical data. The interpretation and quantification of hydrothermal alterations in volcanic terrains, for instance, is particularly important as many economic substances are accumulated by the circulation of hydrothermal fluids in areas with magmatic activity, and because altered rocks are one of the main vectors used to explore for such mineralisations.

Geochemical studies of hydrothermally altered rocks aim at determining the amount of elements gained and lost during alteration of an initially fresh rock (i.e. the precursor), by performing mass balance calculations while assuming that one or more elements is immobile during alteration (see Gresens, 1967). Gresens' approach (Gresens, 1967) has been extensively and successfully used (MacLean and Kranidiotis, 1987; Shriver and MacLean, 1993; Cadieux et al., 2006; to name a few).

A major question that arises when using such mass balance equations concerns the choice of an appropriate precursor for the studied altered sample. Precursor's compositions are usually obtained by analysing "fresh" samples (i.e. samples that lack alteration minerals) that have been geologically related to the altered sample using detailed knowledge of the local geology (e.g. Grant, 1986; Grant, 2005). Such approaches are possible only for detailed deposit-scale studies, if fresh rocks exist and if our level of knowledge of the area render their identification possible. Even so, usually nothing proves that the fresh sample is an exact match for the precursor to the altered sample, and these difficulties can make mass balance calculations un-reliable.

45 At a more regional scale, on the other hand, these difficulties can be in-surmountable.
46 Thus, orebody targeting using chemical datasets more often rely on alteration indexes
47 (ratio of major elements), which values are sensitive to the precursor's compositions and
48 could provide miss-leading indications on the intensity of alteration (see discussion in
49 Trépanier et al., in press).
50 Mass balance calculations would be more reliable if a suitable precursor composition to
51 an altered rock could be determined directly from the chemical characteristics of the
52 altered sample, thus providing a precursor adapted to this sample.
53 This paper proposes a method for estimating the composition in SiO_2, CaO, MgO, FeO^T,
54 K_2O, and Na_2O of fresh precursors to altered volcanic rocks, using ratios of commonly
55 analysed least mobile elements (i.e. Zr, TiO_2, Al_2O_3, Y, Nb, Th, and Cr) and a regression
56 method (i.e. artificial neural network). The method is provided as supplemental material,
57 as a simple C++ program that calculates precursor compositions (this study) prior
58 performing mass balance calculations (using published methods).
59

Formatted: Superscript

60 **2. Methodology**

61 **2.1 Description of the method**

62 In this contribution, a method to predict fresh precursors from the chemical composition
63 (i.e. immobile elements content) of altered volcanic rocks is presented. Predictions are
64 performed for magmatic rocks only because relatively simple relationships, which are
65 controlled by magmatic processes, exist between the trace and major elements content of
66 such rocks. For example, Zr and Si are both incompatible elements that tend to increase

during fractional crystallisation, making Zr a proxy for Si (see Winchested and Floyd, 1977).

Also, precursor predictions are performed using ratio of elements considered immobile in regard of hydrothermal processes. Ratios are used because the value of immobile elements ratios is not modified by mass or volume changes and because ratios are unlikely to be strongly fractionated by hydrothermal processes. Instead, such ratios mostly vary between volcanic rocks of different types and affinities, which originate from various sources and different conditions of partial melting, fractional crystallisation and other processes. Because ratios of immobile elements are insensitive to alteration and can reflect geodynamic settings and differentiation, they have been extensively used by discrimination diagrams (e.g. Winchester and Floyd, 1977; Barrett and MacLean, 1994), and their use is here extended to predict major elements compositions of the fresh precursors to altered volcanic rocks.

In order to relate immobile element ratios to the major elements of volcanic rocks, the typical composition of least-altered volcanic rocks is documented using a large dataset of samples of various volcanic rock types from different tectonic settings and ages (see section 3.1; **Figures 1 and 2**) and on which predictions are performed using a method able to solve regression problems. Since this prediction implies a large variety of rocks, non-linear solutions can be expected. Tests were initially performed with various methods, e.g. multiple regressions, but non produced results as good as these obtained with neural networks. For this reason, a multi-layer perceptron neural network was trained on a large dataset of least-altered volcanic rocks and used to predict precursor compositions.

2.2 Immobile elements ratios

Hydrothermal fluids modify the mineral assemblages (alteration) and chemical composition (metasomatism) of rocks, with various minerals and elements being more or less susceptible to be modified or displaced. Here, the immobile elements used by the proposed method are immobile in most mineralised contexts and are commonly analysed, so that the method can be applied to datasets typically used by exploration geologists.

Elements relatively immobile during hydrothermal alteration, weathering and low-grade metamorphic processes are characterized by ions that are of intermediate ionic potential (0.03-0.1 ppm^{-1}; Pearce, 1996). From the elements considered immobile by various authors (Cann, 1970; Pearce and Norry, 1979; Hill et al., 2000; Kurtz et al., 2000; Hastie et al., 2007), seven elements were selected as inputs to the neural networks (i.e. Zr, Al, Ti, Nb, Cr, Y, and Th) because, according to Pearce (1996), these elements are sufficient to properly characterize most basalts and, presumably, most volcanic rocks. These elements have indeed been used to characterize various magmatic processes, such as: 1) the degree of differentiation (Winchester and Floyd, 1977; Pearce and Norry, 1979); 2) the alkalinity of rocks (Pearce and Cann, 1973); 3) the geodynamic context (Pearce and Cann, 1973; Wood, 1980; Meschede, 1986; Schandl and Gorton, 2002; Hastie et al, 2007; Ross and Bédard, 2009); 4) the depth of partial melting and/or the composition of the source (Drummond and Defaut, 1990; Petford and Atherton, 1996; Pearce, 2008); etc.

The selected immobile elements are considered relatively immobile in many, but not all, hydrothermal processes, as various parameters may increase the mobility of these elements (Hynes, 1980; McCulloch and Gamble, 1991; Pearce, 1996; Hill et al., 2000).

For example, the mobility of Zr, Y, Ti, and Nb has been documented from oceanic crust samples (Cann, 1970), in rocks altered by fluorine-enriched fluids (Salvi et al., 2000; Jiang et al., 2005) and in various mineralised areas (Finlow-Bates and Stumpfl, 1981; van Dongen et al., 2010). However, Zr, Al, Ti, Nb, Cr, Y, and Th are considered to be sufficiently immobile in most cases; as such, the method developed here should be applicable to most mineralised contexts.

Even so, it is recommended to test for the immobility of Zr, Al, Ti, Nb, Cr, Y, and Th prior performing the precursors' prediction described below, by using, for example, binary plots (see Cann, 1970) or Pearce Element Ratio diagrams (Pearce, 1968; Stanley and Russell, 1989) adapted to the study of hydrothermal processes (Stanley and Madeisky, 1994). Also, in certain well constrained contexts, it might be possible to sample a fresh precursor and the immobility of elements could be verified using the isocon diagram of Grant (1986, 2005). Study of element's variabilities could also be performed (Schiano et al., 1993).

2.3 Artificial neural networks definition

Artificial neural networks are numerical models made of simple processing elements named neurons (Zurada, 1992). Neurons are interconnected by links, to which values (i.e. weights) are attributed. Weights describe the strength of the connections between neurons. Neural networks are adaptive; i.e. "knowledge" is gained through a process of training and learning from examples. After training, the "knowledge" is stored by the weight values. Also, the learning process can be supervised, if the input examples already contain the desired outputs.

Many neural network topologies exist. Here, multilayer perceptrons (MLP) are supervised networks that were selected because of their capacity to approximate nearly any function (Hornik et al., 1989). The first layer, called the input layer, distributes the input data (i.e. chemical analyses of volcanic rocks in our case) to the network. The MLP's input layer is made of several neurons, i.e. one for each input variable. Processing (i.e. additions, multiplications and other non-linear transformations; **Figure 3**) is performed by one or more hidden layers of neurons. The final, predicted results, is calculated by the output layer, which contains an amount of neurons that is dependent on the number of target values to be modelled by the network.

An MLP is trained using successive cycles. During each cycle, training examples are fed to the network, which gradually modifies the weights in order to diminish the sum of the squared error between desired (actual) and output (modelled) values. The network's predictions thus improve at each cycle but excessive learning can lead to overfitting, which should be avoided as an overfitted network has started to learn the noise, or random component, of the input dataset and cannot be generalised. To avoid overfitting and to evaluate the performance of a network, examples are often randomly separated into three subsets: 1) the *training set*; 2) the *cross-validation set*; and 3) the *test set.*

The *training set* is used to train the network. The *cross-validation set* is used to monitor the sum of squared errors between the desired and modelled outputs and to avoid overfitting (see early stopping procedure; Wang et al., 1994). The *test set*, on the other hand, is used to evaluate the generalisation performance of the network once the training is completed. In addition, networks are usually trained several times using random initial network weights to avoid local minima in the approximated function.

3. Data processing

3.1 Least-altered volcanic rocks dataset

Relationships between major and immobile elements are here evaluated with neural networks using a large chemical dataset of modern volcanic rocks from around the World. Post-Archean to modern rocks are used because they are well constrained samples for which names, geodynamic settings and freshness can be confidently established; i.e. characteristics that would be harder to determine for more ancient samples. Besides, there might be large differences between ancient (Archean) and modern geodynamics and bulk chemistry of magmas (e.g. Bédard, 2006), but this does not necessarily mean that the relationships between major and immobile elements, in magmas, changed between the Archean and more recent times; i.e. if it is considered that partial melting, fractional crystallisation and other processes operate in similar ways at all times. Also, volcanic rocks are used because a large dataset is needed, and because the highest quality and largest dataset of magmatic rocks available, i.e. the Georoc web database (Sarbas, 2008; Georoc, 2011), documents mostly volcanic rocks.

The Georoc dataset was thus used and processed as follows. First, all volcanic samples from the "Precompiled Files --> Rocks" page were downloaded; i.e. a total of 110 724 samples of volcanic rocks. Then, the following criteria were used to filter the dataset.

1. Only samples with major elements, LOI, Zr, Cr, Nb, Y, and Th analysed were retained.
2. Only samples with Nb >= 8 ppm, or with Nb < 8 ppm but precise to the tenth decimal place, were kept.

3. Samples described as altered in the Georoc dataset were rejected (see codes M, E, T; Georoc, 2011).
4. Samples (picrites, komatiites, meimechites, and boninites excepted) with LOI > 2.5% (i.e. approximately H_2O+ > 2% and CO_2 > 0.5% that characterises most altered volcanic rocks according to the IUGS volcanic rock classification; Le Maitre et al., 1989) were rejected.
5. Samples (picrites, picrobasalts, komatiites, meimechites, and boninites excepted) plotting outside the "fresh field" of a modified Hughes (1972) diagram (i.e. diagram with a "fresh field" expanded toward the right to include fresh ultra-K rocks) were rejected (**Figure 1**).
6. Samples with a Chemical Index of Alteration (CIA; Nesbitt, 2003) over 52.5 (i.e. 52.5 was used instead of 50 to retain slightly peraluminous igneous rocks) were discarded.
7. To ensure that a single oversampled volcanic unit is not overrepresented in the dataset, a maximum of 25 samples per publications has been conserved, using random subsampling.
8. The Georoc dataset is also strongly skewed toward basalts, basaltic andesites, and to a lesser extent, andesites. For neural networks processing, more balanced inputs are preferable. The dataset was thus subsampled to ensure that no single rock type (as defined by IUGS; Le Maitre et al., 1989) represents over 10% of the final dataset. Note that these rock types, or groups, are no longer used hereafter, as the methodology is applied across a continuum of compositions.

The final dataset, considered to represent fresh rocks, contains 5 810 samples (**Figure 2a)** from various tectonic settings (**Figure 2b**).

3.2 Artificial neural networks processing

The Georoc database's samples have then been used to train multi-layers supervised networks (i.e. MPL, see section 2.3). Prior performing the training, the dataset was randomly split into three subsets: 1) the *training set*, to which 2 905 samples (50%) were allocated; 2) the *cross-validation set* received 1 453 samples (25%); and 3) the *test set* received the remaining 1 452 samples (25%). Also, tests were performed using ~~log-~~ transformed data (centered log-ratios) and un-transformed, compositional, data. As both tests produced similar results (see section 5.1), only un-transformed data were ~~used to train the MPL~~retained.

The inputs fed to the MLP are ratios of immobile elements and outputs (i.e. the components to be predicted) are major mobile elements. Before processing, the chemical analyses were recalculated on a LOI-free basis and input chemical elements were normalised. Once normalised, the values of each variable is comprised between 0 and 1 and, as such, the units of input data (wt% or ppm) is unimportant.

Then, the neural networks were trained using the Levensberg-Marquardt supervised algorithm (NeuroSolutions 6.11 for excel software; NeuroDimension, 2012). The MLP networks used here contain an input layer made of 21 neurons (i.e. number of possible ratios between the immobile elements), one hidden layer and uses an hyperbolic tangent transfer function. To determine the optimal number of hidden neurons for the hidden layer of each network, five models that contain 2, 4, 6, 8, and 10 hidden neurons, respectively, were tested. Each test consisted in 3 runs, or trainings, of the networks and

the most performant network (i.e. network with minimum sum of squared errors on the cross-validation set of all variations and runs) was kept (see **Table 1**).

Also, for each of the output components that are to be predicted (i.e. SiO_2, CaO, FeO^T, MgO, K_2O, Na_2O, Zr, and TiO_2), separate networks were trained. Note that the predicted Zr and TiO_2 values are used later, to perform mass balance calculations. Once trained with the *training* and cross-*validation sets*, and validated using the *test set* (see section 5), the networks were used to predict precursor's compositions for new samples, and these precursors were used to perform mass balance calculations (see section 4).

Eventually and to evaluate new samples (see sections 6 and 7), an additional constrain was added to the model. It should be noted that trained neural networks provide valid predictions of precursor's compositions if and only if all the ratios of immobile elements of a sample to be evaluated are within the ranges of the *training set's* ratios. On the other hand, if the ratios of a sample fall outside the range of values of the *training set*, the prediction would be invalid and it should not be attempted. Similarly, the predicted values are expected to fall within the range of the *training set's* values, otherwise the prediction should be regarded as invalid. Valid ranges for immobile elements ratios and for predicted values were recorded from the networks and embedded in the software (see section 4), in order to filter unsuitable samples and invalid outputs.

4. Description of the software

4.1 Software

This study aims at modelling the chemical composition of fresh precursors to altered rocks, in order to facilitate mass balance calculations. The outputs of the neural networks

Formatted: Superscript

(i.e. the results of the training process) can however be difficult to use as such, as the software used to train the neural networks (NeuroDimension, 2012) provides outputs files with a ".NSW" extension, which are simple ASCII text files that record the network's structure, parameters and weights.

To facilitate access to these results, they were embedded in a C++ program that is able to receive chemical analyses from new samples (input), that predicts the precursors' compositions and that performs mass balance calculation (see section 4.2). This program provides users with the results of these calculations; i.e. precursors' compositions as well as mass changes values, which are provided in percent (relative mass change) and grams (absolute mass change).

4.2 Mass balance calculations

Once precursor values are predicted, mass balances can be calculated using the equation of Grant (1986, 2005), which is a rewritten form of Gresens' (1967) equation. Grant's (1986) equation defines a line called the "isocon" that passes through the origin of a multi-components graphic. The slope of the isocon yields the overall mass change between the fresh and altered sample, and this slope is constrained using elements immobile during the alteration process. Here, the slope value is obtained from the average of the slopes (Leitch and Lentz, 1994) calculated from TiO_2 and Zr (see equation 1). Absolute mass changes, in grams, are then calculated for 100 grams of precursor (see equation 2) and relative mass changes are calculated in percent (see equation 3).

(1) $M^P/M^A = (C_{TiO2}{}^A / C_{TiO2}{}^P + C_{Zr}{}^A / C_{Zr}{}^P) / 2$

where C_{TiO2} is the concentration of $TiO_{2;}$ C_{Zr} is the concentration of Zr; "A" refers to the altered rock and "P" to the precursor rock as calculated by the neural networks. M^P and M^A are the equivalent masses before and after alteration. M^P/M^A is the slope of the isocon used in subsequent mass balance calculations.

(2) $\Delta C_i = C_i^A * M^A/M^P - C_i^P$

(3) $\Delta C_i / C_i^P = 100 * (M^A/M^P)(C_i^A / C_i^P) - 1$

where *Ci* is the concentration of element "i"; M^A/M^P is the inverse of the slope of the isocon calculated using equation 3; ΔC_i is the change in concentration of element "i" in grams per 100 grams of precursor.

5. Neural network validations

Prior to using the method to predict new sample's precursors, the neural networks' outputs were validated. The consequences that Zr, Y, Nb, Cr, Th, Al, and Ti mobility might have on the prediction's accuracy are also tested. Overall, the proposed approach appears accurate, precise, and may tolerate a limited mobility of the immobile elements used by the model.

5.1 Evaluation of prediction's accuracy and precision

Once the training was completed, the accuracy and precision of the predictions were evaluated by feeding the *test set* to the trained networks and by comparing the predicted and actual values, using median errors (for accuracy) as well as, for precision, rank-based correlation coefficients, standard deviation of errors and normalized mean square error values (NMSE; Neurodimension, 2012) (see equations 4 and 5).

(4) $MSE = \sum^{N} (Actual\ value - Predicted\ value)^2$

(5) $NMSE = MSE / Variance\ of\ the\ actual\ values$

where N is the number of samples in the training set, MSE is the mean square error, and $NMSE$ is the normalized mean square error.

First, the *training set* was fed to the networks and the outputs were compared to the analysed values using scatter plots (**Figure 4**). Overall, the points on the scatterplots tend to lie equally on either side of the 1:1 lines for the entire ranges of values, indicating that the prediction method is accurate and precise, in particular for TiO_2 and Zr, but less so for K_2O and Na_2O. Also, the model appears generally less precise for high values than for average or low values, especially for the predictions of Na_2O, Zr, TiO_2, and MgO (**Figure 4**).

Then, median errors between predicted and analysed values were examined and found to be near 0 for all components and for both subalkaline and alkaline rocks (i.e. models are accurate) (**Table 2**). Standard deviation of error values are also low (**Table 2**), indicating that the method is precise (**Table 2**).

Then, global rank-based correlation coefficients between predicted and analysed values were calculated (**Table 2**). Global correlations are over 0.93 for most components (i.e. the predictions are precise), except for Na_2O (0.77) and CaO (0.88), and are especially elevated for TiO_2 and Zr (0.99). Isocon evaluations, which are based on TiO_2 and Zr values, are thus likely to be precise (see section 4.2). Also, note that correlations are generally slightly higher for subalkaline rocks than for alkaline rocks, with the maximum correlations' difference being 0.09 (for K_2O; **Table 2**).

Eventually, NMSE were calculated, which are found to be the lowest for TiO_2 and Zr, and the highest for Na_2O and K_2O (**Table 2**). NMSE values are also generally higher for alkaline rocks, except for FeO^T and TiO_2. Overall, the values of NMSE are generally low, validating the precision and accuracy of the models.

As mentioned in section 3.2, the procedure described above uses un-transformed data but has also been performed using log-transformed data divided by the immobile element Al_2O_3. The rank-based correlation coefficients for both tests are similar except for SiO_2 (**Table 3**) and only the results of the test performed with un-transformed data were implemented by the C++ program provided as additional material to this contribution.

5.2 Effects of the mobility of an immobile element on major oxide predictions

Precursor compositions are predicted using ratios of elements considered immobile, but that might be mobile under certain conditions (see section 2.2). In this section, the effect that such mobility might have on the precision and accuracy of predictions is modelled.

For each immobile element, one sample from the *test set*, for which the immobile element considered has approximately median value, was selected. The value of the considered immobile element of the selected samples was then increased and decreased by 33%, while the values of the other immobile elements were left untouched (**Table ~~3~~4**); i.e. the ratios of immobile elements were modified. The resulting data were fed to the neural networks, which provided predictions reported in **Table ~~3~~4**.

This simple model of element's mobility indicates that moderate enrichments or depletions of Y, Cr, and Nb do not significantly modify the predicted values. Changes in Zr and TiO_2 moderately affect the predictions for FeO^T, MgO, and CaO (**Table ~~3~~4**). On

Formatted: Superscript

Formatted: Superscript

the other hand, changes in Zr and TiO_2 strongly affect Zr and TiO_2 predictions, respectively. As these elements are used to evaluate the isocon (see section 4.2), it is recommended to evaluate the extent of their mobility prior using the method presented here (see section 2.2).

6. Method validation using case studies

In this section, several fresh volcanic rocks, which were not included in the Georoc database used to train the neural network, are used to validate the approach (see case studies 1 to 3). Then, altered rocks from several mineralised areas (i.e. volcanogenic massive sulphide, or VMS, and porphyry deposits) are compiled, their precursors compositions are predicted and mass balance calculations are performed (see case studies 4 to 6).

6.1 Validation using fresh volcanic rocks

Case study 1 – Unaltered dacite samples from the 2004-2006 and 1986 eruptions of Mount Saint-Helens, Washington, USA, were compiled from Pallister et al. (2008). The major elements composition of these samples was then predicted using the neural networks, which accurately predicts most components (**Figure 5**), except for K_2O values, which are moderately overestimated by 0.6 to 0.8% (**Figure 5f**), and SiO_2 values, which are underestimated by ~2% (**Figure 5a).**

Case study 2 – Ten samples of unaltered mafic volcanic rocks (alkaline basalts and hawaiites) from the Heimaey volcanic centre, Iceland, were compiled from Mattsson and Oskarsson (2005). Their analysed and predicted values are shown in **Figure 6.** The

predictions are generally good but systematic errors are observed for some elements, such as CaO (errors up to 1.5%). Also, overestimations of 2%, 0.1% to 0.4% and 1.5% are observed for SiO_2, K_2O, and MgO. Conversely, underestimations of 0.05 to 0.15% and 0.2 to 1% are observed for TiO_2 and Na_2O values. Overall, these errors on the predictions are minor and the method reproduces satisfactorily the compositions in major elements of the input rocks.

Case study 3 – Nineteen samples of mildly to strongly alkaline volcanic rocks from the Middle Latin Valley volcanic field, Italy, were compiled from Boari et al. (2009). These samples are described as devoid of alteration minerals and are named melilite-bearing kamafugitic rocks, plagioclase leucitites and leucitites, shoshonites, and calc-alkaline rocks (Boari et al., 2009). Predicted values for these samples show that neural networks perform very poorly for nearly all components (except for Zr) (see **Figure 7**), indicating that the method cannot be applied to ultra-K rocks. This is, however, a minor limitation as alkaline rocks in general, and ultra-K rocks in particular, represent small volumes of the crust.

6.2 Case study 4: Rhyodacite from the Phelps Dodge VMS deposit

Twenty-two samples of altered rhyodacites from the Phelps Dodge VMS deposit, located in the Matagami Camp within the Archean Abitibi Greenstone Belt, in Québec, were compiled from MacLean and Kranidiotis (1987). These samples were selected because of their well-documented mass changes in Si, Mg, Fe, and other elements calculated by comparing each sample to the least altered, likely chloritised, sample #18 (see MacLean and Kranidiotis, 1987).

Major element contents of the 22 samples were predicted using the neural network solution. Also, mass balance calculations using the Grant's method (Grant, 1986), sample #18 as a fresh precursor and Al-Ti as immobile elements, were performed, following earlier similar calculations (see MacLean and Kranidiotis, 1987; Barrett and MacLean, 1994). Results of these calculations are displayed using a $Fe_2O_3^T$+MgO vs SiO_2 diagram similar to this used by Barrett and MacLean (1994) (**Figure 8**). On this diagram, the least-altered sample #18 is identified and the evolution of the alteration in three stages, as interpreted by MacLean and Kranidiotis (1987), is identified: 1) chloritisation (with Fe > Mg) and silicification; 2) silica leaching; and 3) destruction of Fe-chlorite (with residual Mg-chlorite) (**Figure 8**).

Formatted: Superscript

Using the $Fe_2O_3^T$+MgO vs SiO_2 diagram, we observe that mass changes calculated with a unique precursor (i.e. sample #18; **Figure 8a**) are well reproduced by these calculated using precursors predicted by neural networks (**Figure 8b**). Furthermore, neural network predictions confirm that sample #18 is altered (chloritized).

Formatted: Superscript

6.3 Case study 5: Neves Corvo VMS deposit

Twenty samples of altered felsic to intermediate volcanic rocks from the Neves Corvo VMS deposit, located in the Iberian pyrite belt, in Spain, were compiled from Relvas et al. (2006). The hydrothermal alteration is zoned, with a chlorite-enriched core enveloped by K-sericite and Na-sericite halos (Relvas et al., 2006). The method presented here was applied to samples from these three alteration zones and results are shown by **Figure 9**.

The calculations confirm that the chlorite-enriched samples have gained Fe-Mg and lost alkali (see chloritisation). Also, the sericite-enriched samples have not gained Fe-Mg and

409 have lost less alkali than the chloritized rocks; i.e. these rocks have mostly endure an
410 acidic alteration, which is consistent with the acidic character of Neves Corvo deposit's
411 hydrothermal fluids (see Relvas et al., 2006). Thus, mass balance calculations enable a
412 clear distinction between sericitised and chloritised samples (see **Figure 9**), pointing
413 toward agreement between the calculation performed here and the petrology of the
414 samples, as described by Relvas et al. (2006).
415
416 **6.4 Case study 6: porphyry deposits**
417 A large amount (n= 565) of samples of altered felsic to intermediate volcanic and
418 intrusive rocks, collected in the vicinity of porphyry deposits, have been published by
419 Dilles (2012) (see also Cohen, 2011; Cohen et al., 2011). The petrology of alteration
420 minerals contained in these samples has been studied in details, leading to their
421 classification into the following groups: 1) albite; 2) albite, K-feldspar, sericite; 3)
422 plagioclase; 4) plagioclase, K-feldspar; 5) sericite-albite; and 6) sericite (see Dilles, 2012
423 and references included). Most of these alterations correspond to alkali gains (i.e.
424 albitisation and K-feldspar alterations), locally accompanied by acidic alteration (i.e.
425 sericitisation).
426 Neural networks calculations were performed on extrusive and intrusive rocks because,
427 while the neural networks were trained using volcanic rocks only, there is no reason for
428 the relationships between trace and major elements to be different in these two types of
429 rocks. Predictions were thus obtained for all the compiled samples, prior calculating mass
430 changes, and results are shown by **Figure 10**. Calculations performed here indicate that
431 albite- or plagioclase-enriched samples have gained Na (albitisation), that sericitised

samples have gained K and lost Na, and that the other samples have gained variable amounts of Na and K (**Figure 10**). Overall, a correlation is observed between the Na and K mass changes calculated here and petrographic observations of Dilles (2012).

8. Discussion and conclusions

This study proposes a method for predicting the composition in major elements of precursors to hydrothermally altered volcanic rocks, using artificial neural networks. To facilitate its utilisation by the reader, the method is embedded in a C++ console program that also performs mass balance calculations, allowing for the evaluation of alteration, in terms of type and magnitude, from large datasets of volcanic rocks.

Neural networks used here "learned" to predict the composition in major elements of volcanic rocks using ratios of immobile elements from a large dataset of fresh volcanic rocks compiled from the Georoc database. These predictions were checked for their accuracy and precision, which appear to be generally very good, except for uncommon volcanic rocks, i.e. ultra-K rocks (see section 6.3), for which the method presented here should not be applied.

Also, evaluation of the precision of the method assumes that the dataset used to train the neural networks contains samples completely devoid of alteration. However, as alteration may initiate as early as during the cooling of a lava flow, it is reasonable to assume that the used dataset still contains slightly altered samples. As alteration affects the most mobile elements, i.e. Na_2O, CaO, and K_2O, it may explain a part of the imprecision of the neural networks predictions. The rest of the imprecisions is likely inherited from the neural network inability to perfectly predict the proper precursor to every volcanic rock.

When the method is used on samples from mineralised areas, another source of inaccuracy of the predictions might arise, which is related to the mobility of the immobile elements used to constrain the model. To evaluate this problem, a "stress test" was performed (see section 5.2), which shows that the predictions are robust and tolerate minor to moderate mobility of the immobile elements. However, if Zr and TiO_2 are mobile, the precursor prediction will be little affected, but the isocon will be erroneous, leading to inaccurate mass balance calculations. Note that, however, these elements are particularly immobile in most hydrothermal systems (e.g. Finlow-Bates and Stumpfl, 1981), except in particular settings (Jiang et al., 2005). As such, the method presented here should be operational in most contexts, but should be used with care in unusual areas, where there is reasons to doubt the immobility of Al, Cr, Y, Th, or Nb, and especially of Ti and Zr.

Also, the method was calibrated on a volcanic rocks dataset and should thus be applied to volcanic rocks only. However, chemical differences between extrusive and intrusive rocks are generally limited. In addition, there is no reason for the relationship between mobile and immobile elements to be different in both types of rocks. As such, the method should be applicable to magmatic rocks in general (see section 7.3), as long as they are neither ultra-K rocks, nor rocks that do not have extrusive equivalents (i.e. cumulate and pegmatite). Note that the applicability of the method likely extends to metamorphic rocks with magmatic protoliths, as metamorphism does not modify the chemistry of rocks, except for volatiles. Thus, only sedimentary rocks and paragneisses should not be evaluated with this method.

In conclusion, the method presented here appears sufficiently precise and accurate to be used in most contexts to quantify alterations from altered magmatic rocks. It should be used on large datasets that document a large variety of volcanic units, when the volcanic stratigraphy is poorly constrained, etc. As such, the method is most helpful in grassroots or regional mineral exploration contexts, but could also be used to evaluate alterations quantified using other methods, and is thus viewed as complementary of existing methods.

Acknowledgments

The authors wish to address special thanks to the editor and reviewers of the manuscript, Jef Caers and David Lentz, as well as an anonymous reviewer. This project was supported by Canada Economic Development, the Ministère de l'Énergie et des Ressources naturelles du Québec, the Conférence régionale des élus Saguenay-Lac-Saint-Jean and companies members of the Consorem. The authors warmly thank their colleague Silvain Rafini for constructive discussions on this project, Geneviève Boudrias for editing the original Consorem report on this work (report available online at http://www.consorem.ca/), Judit Ozoray for correcting this manuscript's spelling and the various users of the non-public LithoModeleur software, that integrates the method presented here since April 2008.

Field Code Changed

References

Barrett, T.J., MacLean, W.H, 1994. Chemostratigraphy and hydrothermal alteration in exploration for VHMS deposits in greenstones and younger volcanic rocks. In:

D.R. Lentz (Ed.), Alteration and alteration processes associated with ore-forming systems. Geological Association of Canada, short course notes 11, p. 433-465.

Bédard, J.H., 2006. A catalytic delamination-driven model for coupled genesis of Archaean crust and sub-continental lithospheric mantle. Geochimica et Cosmochimica Acta 70(5), p. 1188-1214.

Boari, E., Tommasini, S., Laurenzi, M.A., Conticelli, S., 2009. Transition from ultrapotassic kamafugitic to sub-alkaline magmas: Sr, Nd and Pb isotope, trace element and ^{40}Ar-^{39}Ar Age data from the Middle Latin Valley Volcanic Field, Roman Magmatic Province, Central Italy. Journal of Petrology 50, 1327-1357.

Cadieux, A.-M., Dubé, B., Williamson, K., Malo, M., Twomey, T., 2006. Characterization of hydrothermal alterations at the Red Lake mine, northwestern Ontario. Geological Survey of Canada, Current Research 2006-C2, 14p.

Cann, J.R., 1970. Rb, Sr, Y, Zr and Nb in some ocean-floor basaltic rocks. Earth Planet. Science Letters 10(1), 7-11.

Cohen, J.F., 2011. Mineralogy and geochemistry of hydrothermal alteration at the Ann-Mason porphyry copper deposit, Nevada: Comparison of large-scale ore exploration techniques to mineral chemistry. Unpublished Master thesis, Oregon State University, 580 p.

Cohen, J.F., Dilles, J.H., Tosdal, R.M., Halley, S., 2011. Compositional variations in hydrothermal white micas and chlorites in a porphyry Cu system at Yerington. Abstracts with Programs - Geological Society of America, May, 43(4), 63-64.

521 Dilles, J.H., 2012. Footprints of porphyry Cu deposits: Vectors to the hydrothermal
522 center using mineral mapping and lithogeochemistry. Final Technical Report for
523 USGS MRERP Grant Award number G10AP00052, 599 p.
524 Finlow-Bates, T., Stumpfl, E.F., 1981. The behaviour of so-called immobile elements in
525 hydrothermally altered rocks associated with volcanogenic submarine exhalative
526 ore deposits. Mineralium Deposita 16, 319-328.
527 GEOROC - Geochemistry of Rocks of the Oceans and Continents. 2011. Max-Planck-
528 Institut für Chemie. http://georoc.mpch-mainz.gwdg.de/georoc/Start.asp, [accessed
529 30 November 2011].
530 Grant, J.A., 1986. The Isocon Diagram - A simple solution to gresens's rquation for
531 metasomatic alteration. Economic Geology 81, 1976-1982.
532 Grant, J.A., 2005. Isocon analysis: a brief review of the method and applications. Physics
533 and Chemistry of the Earth, Parts A/B/C 30, 17, 997-1004.
534 Gresens, R.L., 1967. Composition-volume relationships of metasomatism. Chemical
535 Geology 2, 47-55.
536 Hastie, A.R., Kerr, A.C., Pearce, J.A., Mitchell, S.F., 2007. Classification of altered
537 volcanic island arc rocks using immobile trace elements: development of the Th–
538 Co discrimination diagram. Journal of Petrology 48(12), 2341-2357.
539 Hill, I.G., Worden, R.H., Meighan, I.G., 2000. Yttrium: The Immobility-mobility
540 transition during basaltic weathering. Geology 28, 923-926.
541 Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are
542 universal approximators. Neural Networks 4, 251-257.

Hughes, C.J., 1972. Spilites, keratophyres and the igneous spectrum. Geological Magazine 109, 515-527.

Hynes, A., 1980. Carbonitization and mobility of Ti, Y, and Zr in Ascot formation metabasalts, S.E. Quebec. Contributions to Mineralogy and Petrology 75, 79-87.

Jiang, S.-Y., Wang, R.-C., Xu, X.-S., Zhao, K.-D., 2005. Mobility of high field strength elements (HFSE) in magmatic-, metamorphic-, and submarine-hydrothermal systems. Physics and Chemistry of the Earth 30, 1020-1029.

Kurtz, A.C., Derry, L.A., Chadwick, O.A., Alfano, M.J., 2000. Refractory element mobility in volcanic soils. Geology 28, 683-686.

LeBas, M.J., Le Maitre, R.W., Streckeisen, A., Zanettin, B., 1986. A chemical classification of volcanic rocks based on the total alkali-silica diagram. Journal of Petrology 27, 745-750.

Leitch, C.H.B., Lentz, D.R., 1994. The Gresens approach to mass balance constraints of alteration systems: methods, pitfall, examples. In Lentz, D.R. (Ed), Alteration and Alteration Processses associated with ore-forming systems. Geological Association of Canada, short course notes 11, pp. 161-192.

Le Maitre, R.W., Bateman, P., Dudek, A., Keller, J., Lameyre, J., Le Bas, M.J., Sabine, P.A., Schmid, R., Sorensen, H., Streckeisen, A., Woolley, A.R., Zanettin, B., 1989. A classification of igneous rocks and glossary of terms: recommendations of the International Union of Geological Sciences subcommission on the systematics of igneous rocks. Blackwell Scientific Publications, Oxford, 206 p.

Formatted: English (United Kingdom)

MacLean, W.H., Kranidiotis, P., 1987. Immobile elements as monitors of mass transfer in hydrothermal alteration: Phelps Dodge massive sulphide deposit, Matagami, Québec. Economic Geology 82, 951-962.

Mattsson, H.B., Oskarsson, N., 2005. Petrogenesis of alkaline basalts at the tip of a propagating rift: Evidence from the Heimaey volcanic centre, south Iceland. Journal of Volcanology and Geothermal Research 147, 245-267.

McCulloch, M.T., Gamble, J.A., 1991. Geochemical and geodynamical constraints on subduction zone magmatism. Earth and Planetary Science Letters 102, 358-374.

Meschede, M., 1986. A method of discriminating between different types of mid-ocean ridge basalts and continental tholeiites with the Nb-Zr-Y diagram. Chemical geology 56(3), 207-218.

Nesbitt, H.W., 2003. Petrogenesis of siliciclastic sediments and sedimentary rocks. In: Lentz, D.R. (Ed.), Geochemistry of sediments and sedimentary rocks: evolutionary considerations to mineral deposits-forming environments. Geological Association of Canada, GeoText 4, pp.39-51.

NeuroDimension, 2012. NeuroSolutions for Excel 6.11. http://www.neurosolutions.com/. [Accessed June 20th, 2012].

Pallister, J.S., Thornber, C.R., Cashman, K.V., Clynne, M.A., Lowers, H.A., Mandeville, C.W., Brownfield, I.K., Meeker, G.P., 2008. Petrology of the 2004-2006 Mount St.Helens lava dome - Implications for magmatic plumbing and eruption triggering. In: Sherrod, D.R., Scott, W.E., Stauffer P.H. (Eds.), A volcano rekindled: the renewed eruption of Mount St. Helens, 2004-2006. U.S. Geological Survey Professional Paper 1750, pp.647-702.

Field Code Changed

587 Pearce, T.H., 1968. A contribution to the theory of variation diagrams. Contributions to
588 Mineralogy and Petrology 19(2):142-157.
589 Pearce, J.A., 1996. A user's guide to basalt discrimination diagrams. In: Wyman, D.A.
590 (Ed.), Trace element geochemistry of volcanic rocks: applications for massive
591 sulphide exploration. Geological Association of Canada, short course notes 12,
592 pp.79-113.
593 Pearce, J.A., 2008. Geochemical fingerprinting of oceanic basalts with applications to
594 ophiolite classification and the search for Archean oceanic crust. Lithosphere
595 100(1), 14-48.
596 Pearce, J.A., Cann, J.R., 1973. Tectonic setting of basic volcanic rocks determined using
597 trace element analyses. Earth and Planetary Science Letters 19, 290-300.
598 Pearce, J.A., Norry, M.J., 1979. Petrogenetic implications of Ti, Zr, Y, and Nb variations
599 in volcanic rocks. Contributions to Mineralogy and Petrology 69(1), 33-47.
600 Petford, N., Atherton, M., 1996. Na-rich partial melts from newly underplated basaltic
601 crust: the Cordillera Blanca Batholith, Peru. Journal of Petrology 37(6), 1491-1521.
602 Relvas, J.M., Barriga, F.J., Ferreira, A., Noiva, P.C., Pacheco, N., Barriga, G. 2006.
603 Hydrothermal alteration and mineralization in the Neves-Corvo volcanic-hosted
604 massive sulfide deposit, Portugal. I. Geology, mineralogy, and geochemistry.
605 Economic Geology 101(4), 753-790.
606 Ross, P.S., Bédard, J.H., 2009. Magmatic affinity of modern and ancient subalkaline
607 volcanic rocks determined from trace-element discriminant diagrams. Canadian
608 Journal of Earth Sciences 46(11), 823-839.

Salvi, S., Fontan, F., Monchoux, O., Williams-Jones, A.E., Moine, B., 2000. Hydrothermal mobilization of high field strength elements in alkaline igneous systems: Evidence from the Tamazeght Complex (Morocco). Economic Geology 95, 559-576.

Sarbas, B., 2008. The GEOROC database as part of a growing geoinformatics network. In: Brady, S.R., Sinha, A.K., Gundersen, L.C. (Ed.), Geoinformatics 2008, Data to Knowledge, Proceedings. U.S. Geological Survey Scientific Investigations Report 2008-5172, p. 42-43.

Schandl, E.S., Gorton, M.P., 2002. Application of high field strength elements to discriminate tectonic settings in VMS environments. Economic Geology 97(3), 629-642.Schiano, P., Dupré, B., Lewin, E., 1993. Application of element concentration variability to the study of basalt alteration (Fangataufa atoll, French Polynesia). Chemical Geology 104(1), 99-124.

Shriver, N.A., MacLean, W.H., 1993. Mass, volume and chemical changes in the alteration zone at the Norbec mine, Noranda, Québec. Mineralium Deposita 28, 157-166.

Stanley, C.R., Madeisky, H.E., 1994. Lithogeochemical exploration for hydrothermal ore deposits using Pearce element ratio analysis: Alteration and alteration processes associated with ore forming systems. Geological Association of Canada Short Course Notes 11, 193-211

Trépanier S., Mathieu L., Daigneault R., In press. CONSONORM_LG: new normative minerals and alteration indexes for low-grade metamorphic rocks. Economic Geology.

Formatted: English (United Kingdom)

632 Van Dongen, M., Weinberg, R.F., Tomkins, A.G. (2010) REE-Y, TI, and P
633 remobilization in magmatic rocks by hydrothermal alteration during Cu-Au deposit
634 formation. Economic Geology 105(4), 763-776.
635 Wang, C., Venkatesh, S.S., Judd, J.S., 1994. Optimal stopping and effective machine
636 complexity in learning. In: Cowan, J.D., Tesauro, G., Alspector J. (Eds.), Advances
637 in neural information processing systems (Denver, 1993), Morgan Kaufmann, San
638 Francisco, pp. 303-310.
639 Winchester, J.A., Floyd, P.A., 1977. Geochemical discrimination of magma series and
640 their differentiation products using immobile elements. Chemical Geology 20, 325-
641 343.
642 Wood, D.A., 1980. The application of a Th-Hf-Ta diagram to problems of
643 tectonomagmatic classification and to establishing the nature of crustal
644 contamination of basaltic lavas of the British Tertiary Volcanic Province. Earth and
645 Planetary Science Letters 50(1), 11-30.
646 Zurada, J. M., 1992. Introduction to artificial neural systems. West Publishing Company,
647 St. Paul, Minnesota, 758p.
648

649 **Figure captions**

650 **Figure 1.** Modified Hughes' diagram (Hughes, 1972) on which the Georoc dataset is
651 represented. This diagram has been used to remove K-Na altered samples from the initial
652 dataset. Black triangles highlight kept samples and grey crosses correspond to discarded
653 samples.

654

655 **Figure 2**. Pie plots displaying several characteristics of the samples of the least-altered
656 dataset (n = 5 810). The displayed characteristics are: (a) the proportions of rock types
657 (see IUGS geochemical classification of volcanic rocks; Le Maitre et al., 1989); and (b)
658 the tectonic settings, as identified in the Georoc database.

659

660 **Figure 3.** Sketch showing calculations performed at an individual neuron within a
661 multilayer perceptron (MLP) network. The neuron's input is calculated by multiplying
662 the input signal (x_i) by the weight of the links (w_i). Inputs are summed and then
663 transformed by a non-linear function, in this case the hyperbolic tangent function. The
664 result corresponds to neuron's output.

665

666 **Figure 4.** Binary scatterplots of predicted vs analysed values, for the *test set* (n = 1452).
667 The 1:1 line is shown as dotted lines. The plots are available for the following elements:
668 (a) SiO_2; (b) FeO^T; (c) MgO; (d) CaO; (e) Na_2O; (f) K_2O; (g) TiO_2; and (h) Zr.

Formatted: Superscript

669

670 **Figure 5.** Binary scatterplots of predicted vs analysed values, for samples from Mount
671 Saint-Helens, USA (data from Pallister et al., 2008). The 1:1 line is shown as dotted lines.

672 The plots are available for the following elements: (a) SiO_2; (b) FeO^T; (c) MgO; (d) CaO;
673 (e) Na_2O; (f) K_2O; (g) TiO_2; and (h) Zr.
674
675 **Figure 6.** Binary scatterplots of predicted vs analysed values, for samples from Heimaey
676 volcanic zone, Iceland (data from Mattsson and Oskarsson, 2005). The 1:1 line is shown
677 as dotted lines. The plots are available for the following elements: (a) SiO_2; (b) FeO^T; (c)
678 MgO; (d) CaO; (e) Na_2O; (f) K_2O; (g) TiO_2; and (h) Zr.
679
680 **Figure 7.** Binary scatterplots of predicted vs analysed values, for samples from Middle
681 Latin Valley, Roman Volcanic Province, Italy (data from Boari et al., 2009). The 1:1 line
682 is shown as dotted lines. The plots are available for the following elements: (a) SiO_2; (b)
683 FeO^T; (c) MgO; (d) CaO; (e) Na_2O; (f) K_2O; (g) TiO_2; and (h) Zr.
684
685 **Figure 8.** Absolute mass changes, in grams, calculated for the rhyodacite rocks that host
686 the Phelps Dodge deposits (samples from MacLean and Kranidiotis, 1987) and displayed
687 in $Fe_2O_3^T$+MgO vs SiO_2 diagrams inspired from Barrett and MacLean (1994). The mass
688 changes have been calculated (a) using sample #18 as a fresh precursor (see MacLean
689 and Kranidiotis, 1987); and (b) using precursors predicted by neural networks. Also,
690 principal alteration paths are shown by arrows.
691
692 **Figure 9.** Absolute mass changes, in grams, calculated for altered rocks from the Neves
693 Corvo deposit (samples from Relvas et al., 2006). The calculation was performed using

Formatted: Superscript
Formatted: Superscript
Formatted: Superscript
Formatted: Superscript

694 neural networks and results are displayed in a $Fe_2O_3^T$+MgO vs K_2O+Na_2O diagram.
695 Colour coding corresponds to alteration types, as interpreted by Relvas et al. (2006).
696
697 **Figure 10.** Absolute mass changes, in grams, calculated for altered rocks from various
698 northern American porphyry deposits (samples from Dilles, 2012). The calculation was
699 performed using neural networks and results are displayed in a Na_2O vs K_2O diagram.
700 Colour coding corresponds to alteration types, as interpreted by Dilles (2012) using
701 detailed petrographic observations.
702
703
704 **Table captions**
705 **Table 1.** Optimal number of hidden neurons contained in the hidden layer of neural
706 networks
707
708 **Table 2.** Median errors (ME), rank-based correlation coefficients (r), standard deviation
709 of errors (SD) and normalized mean square errors (NMSE) for values predicted from the
710 *test set*. Median errors and correlations are calculated for the whole *test set* (n=1 452), as
711 well as for the following subsets: 1) subalkaline rocks (n=833); and 2) alkaline rocks
712 (n=619). Alkalinity is determined using the TAS diagram (Le-Bas et al., 1986).
713
714 **Table 3.** Rank-based correlation coefficients (r) for values predicted from the whole *test*
715 *set*, for tests using un-transformed data (see **Table 2**) and log-transformed data.
716

Formatted: Superscript

Formatted: Font: Bold

Table ~~3~~4. Effects that ±33% gains and losses of various immobile elements have on the predicted values calculated for selected samples. Modelled changes of elements' original values are in bold italics. Important prediction changes (>10%) are in bold.

Computer code

The method presented in this contribution is provided as a computer code (C++ language).

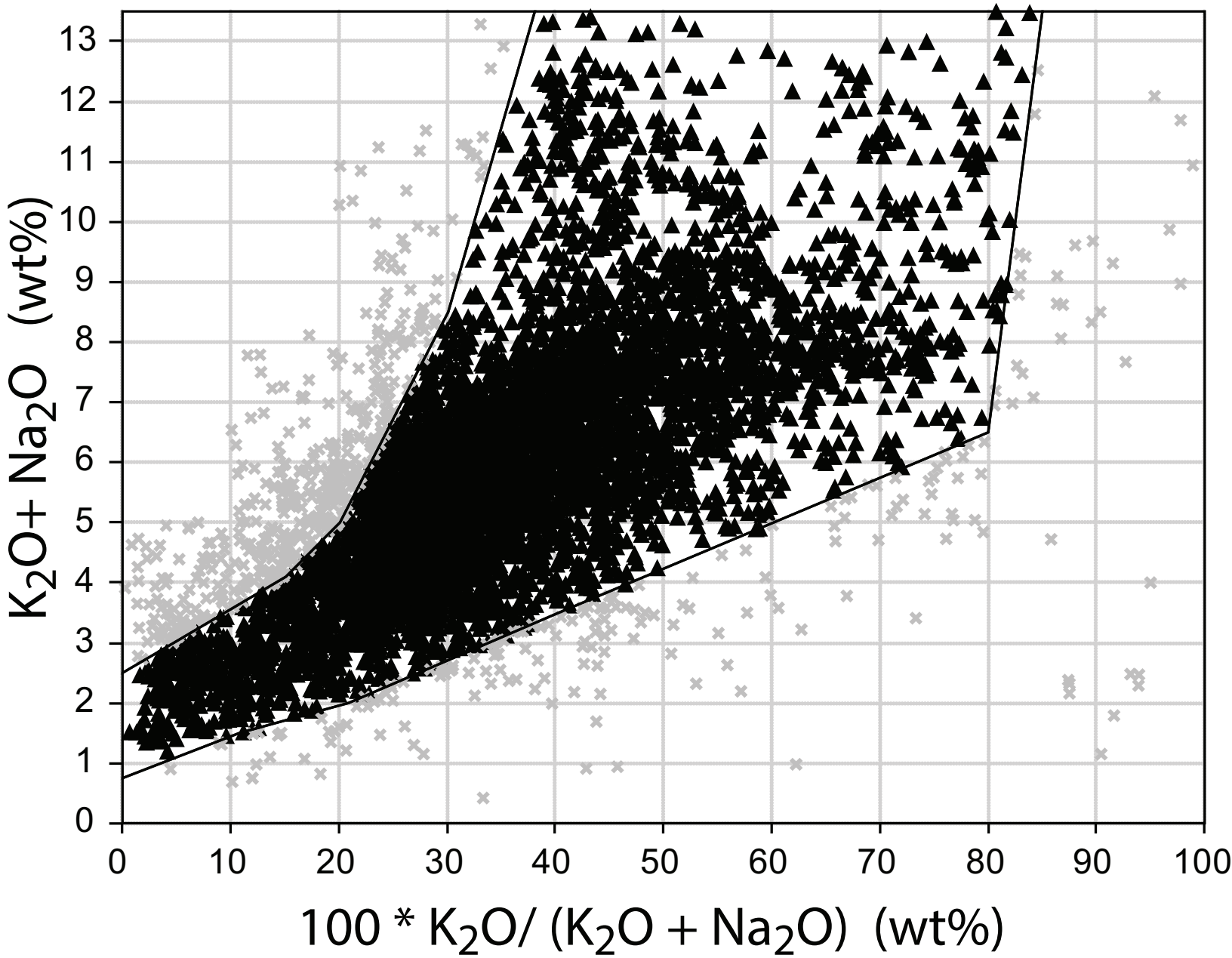
13
12
11
10
9
8
7
6
5
4
3
2
1
0
$K_2O+ Na_2O$ (wt%)
0
10
20
30
40
50
60
70
80
90
100
100 * K_2O/ (K_2O + Na_2O) (wt%)

Fig. 1

Figure(s)

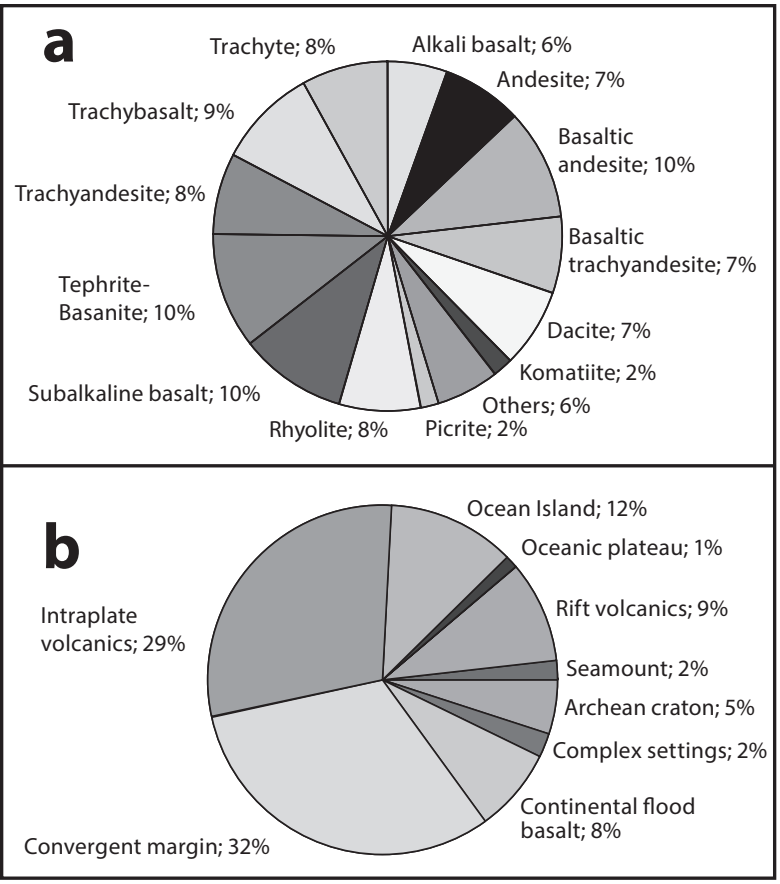

Fig.2 (black and white)

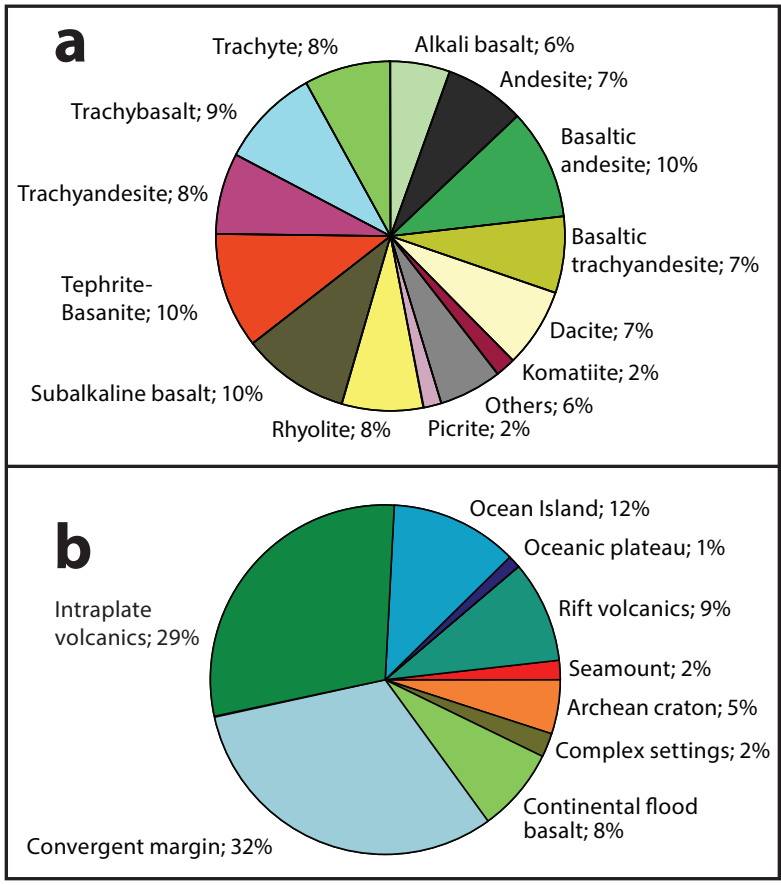

Fig. 2 (color, for online publication only)

Figure(s)

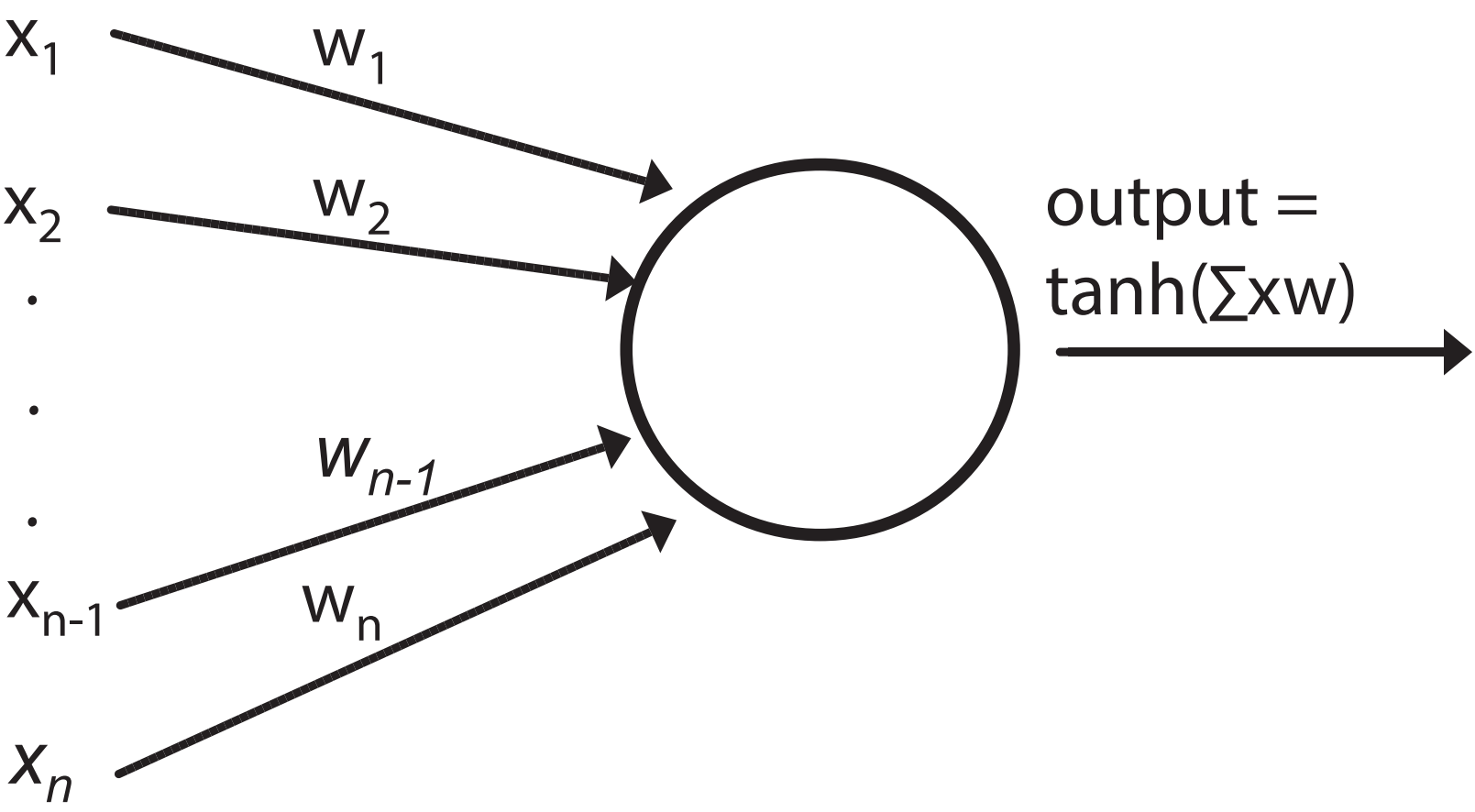

Fig. 3

Figure(s)

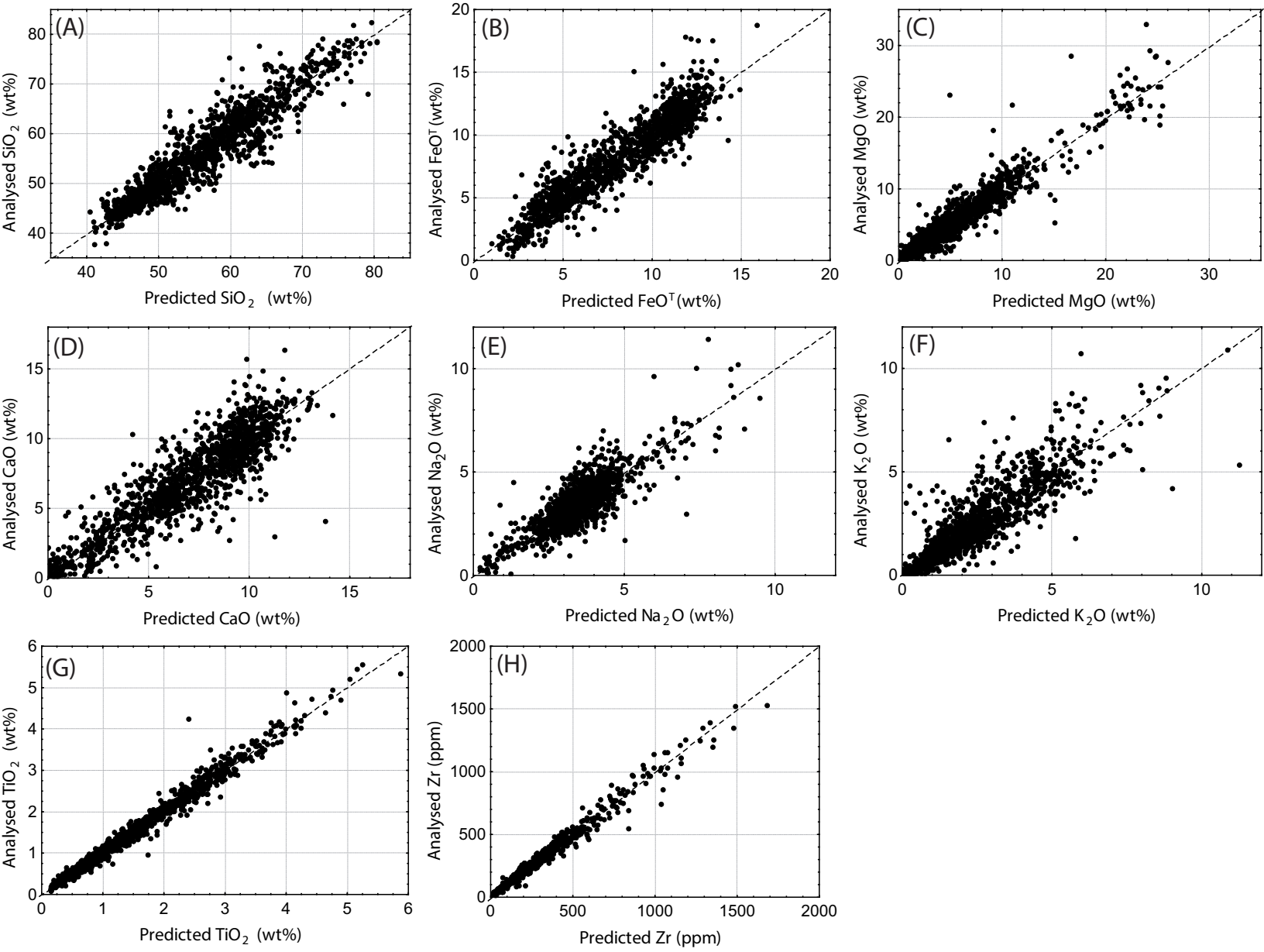

Fig.4

Figure(s)

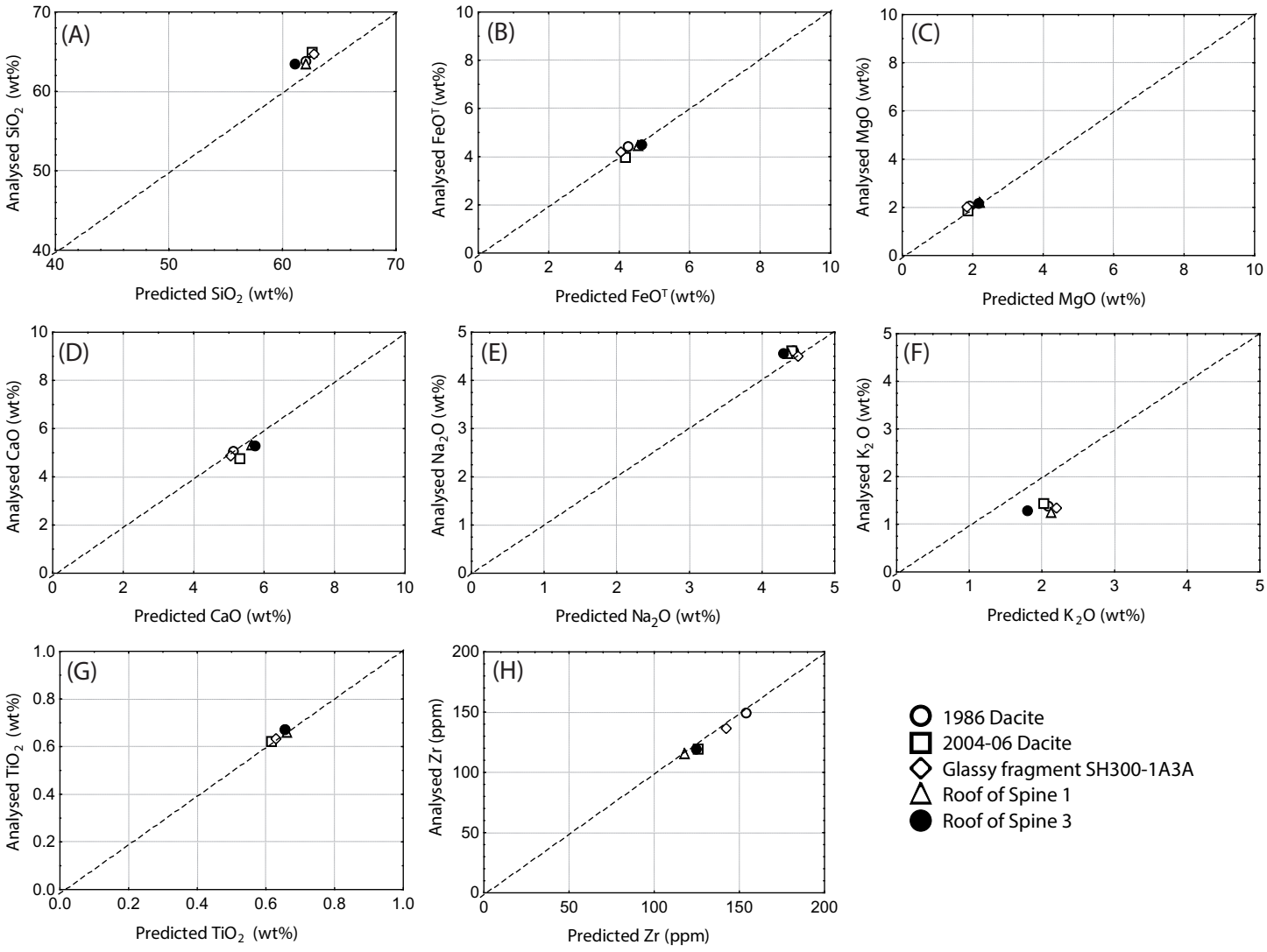

Fig.5

Figure(s)

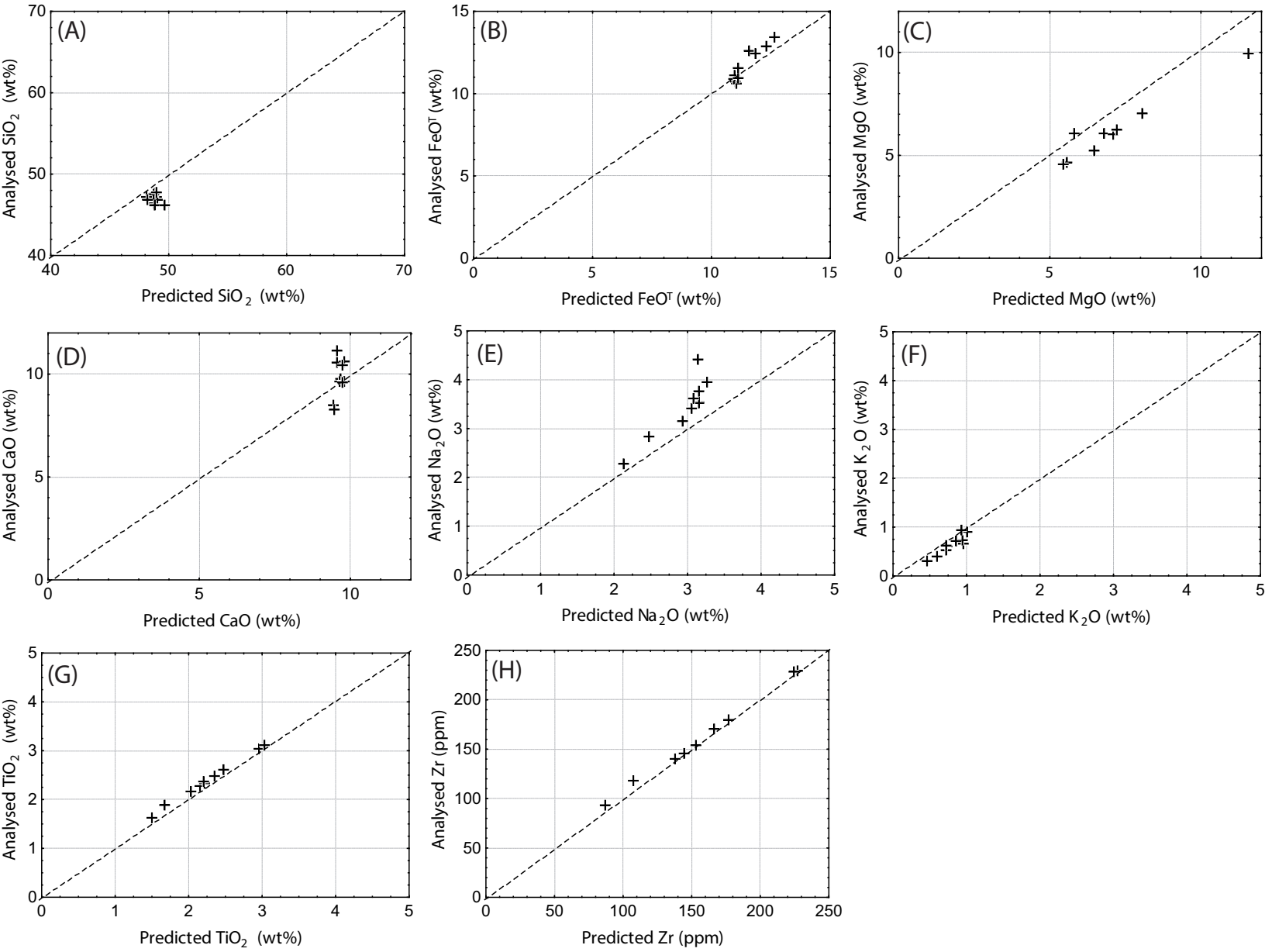

Fig.6

Figure(s)

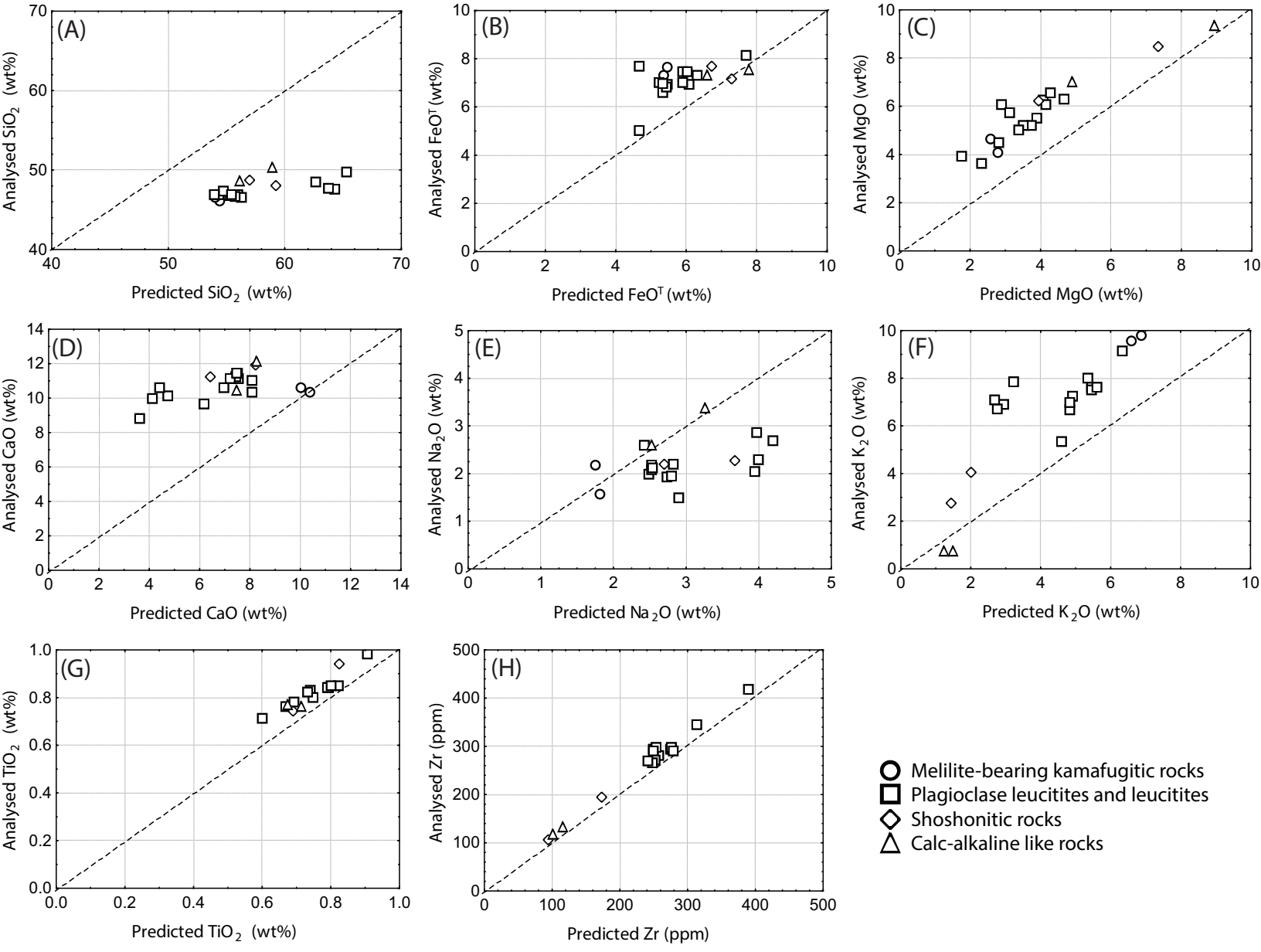

Fig.7

Figure(s)

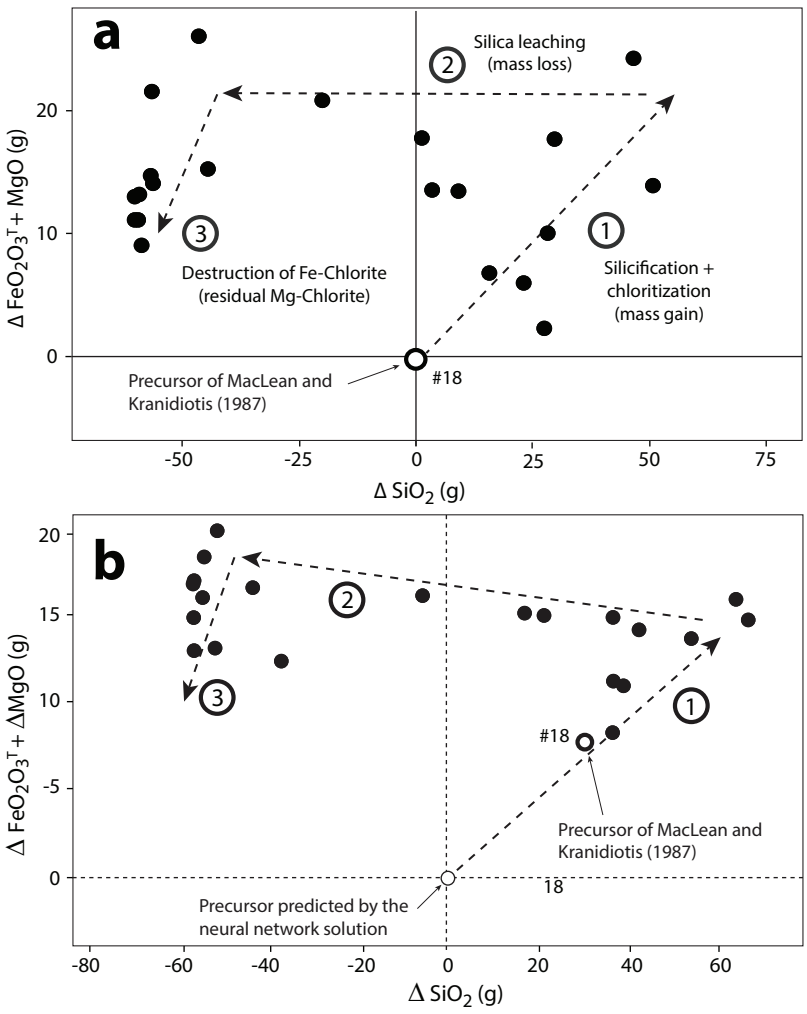

Fig. 8 (black and white)

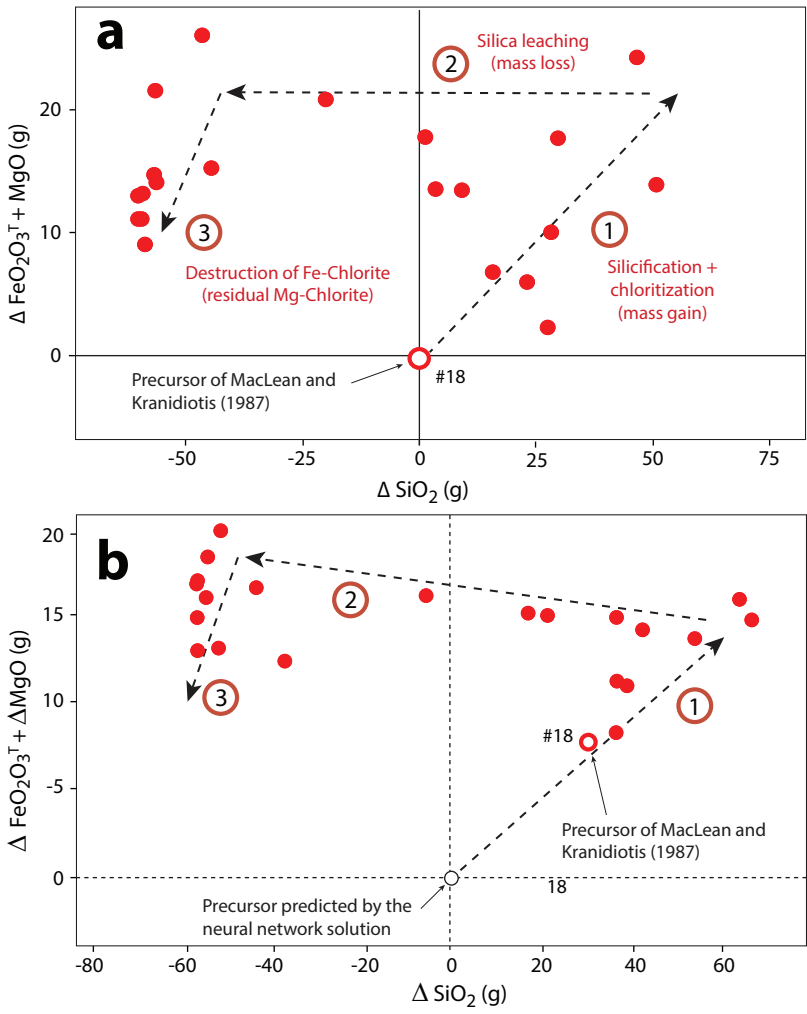

Fig. 8 (color, for online publication only)

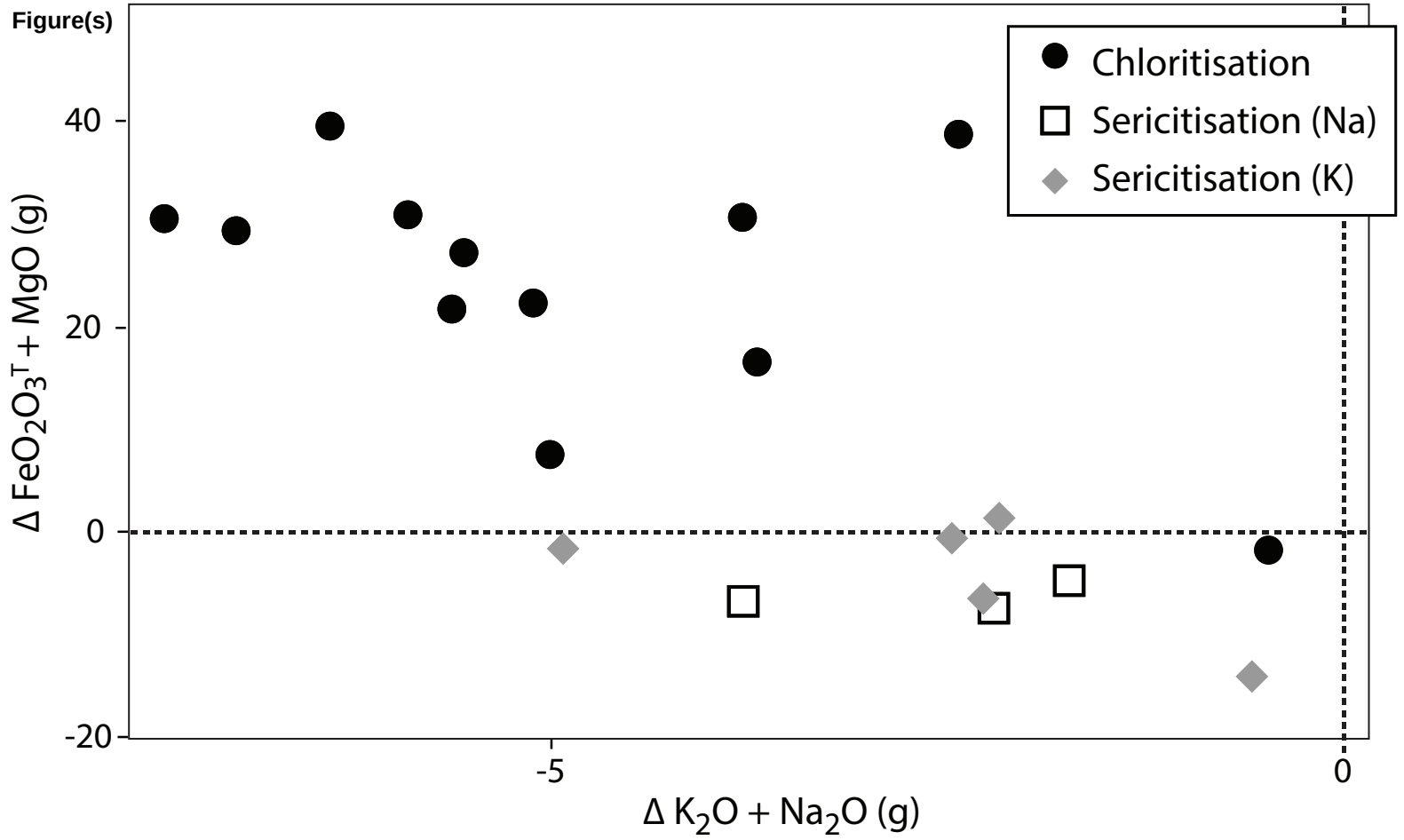

Fig. 9 (black and white)

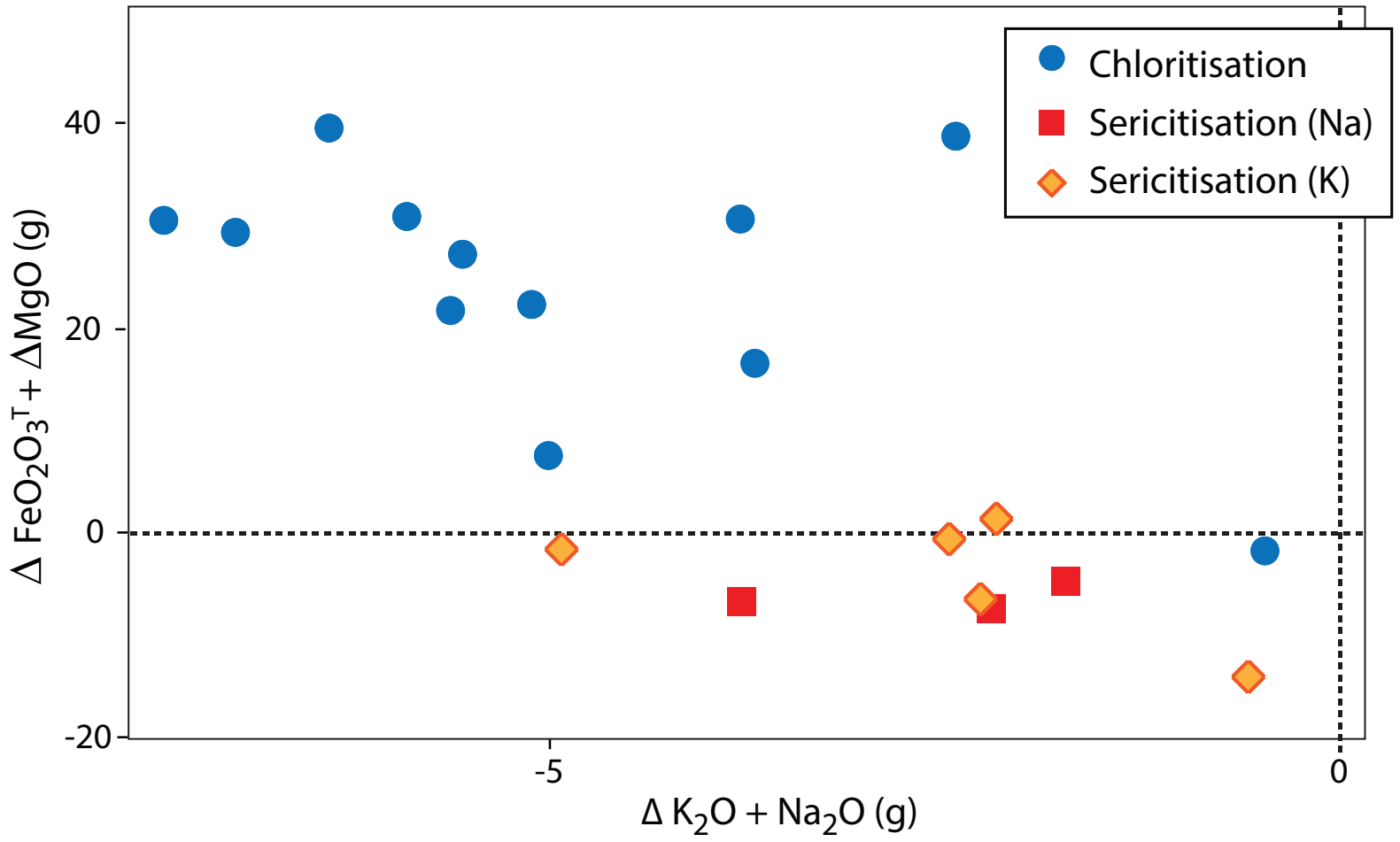

Fig. 9 (color, for online publication only)

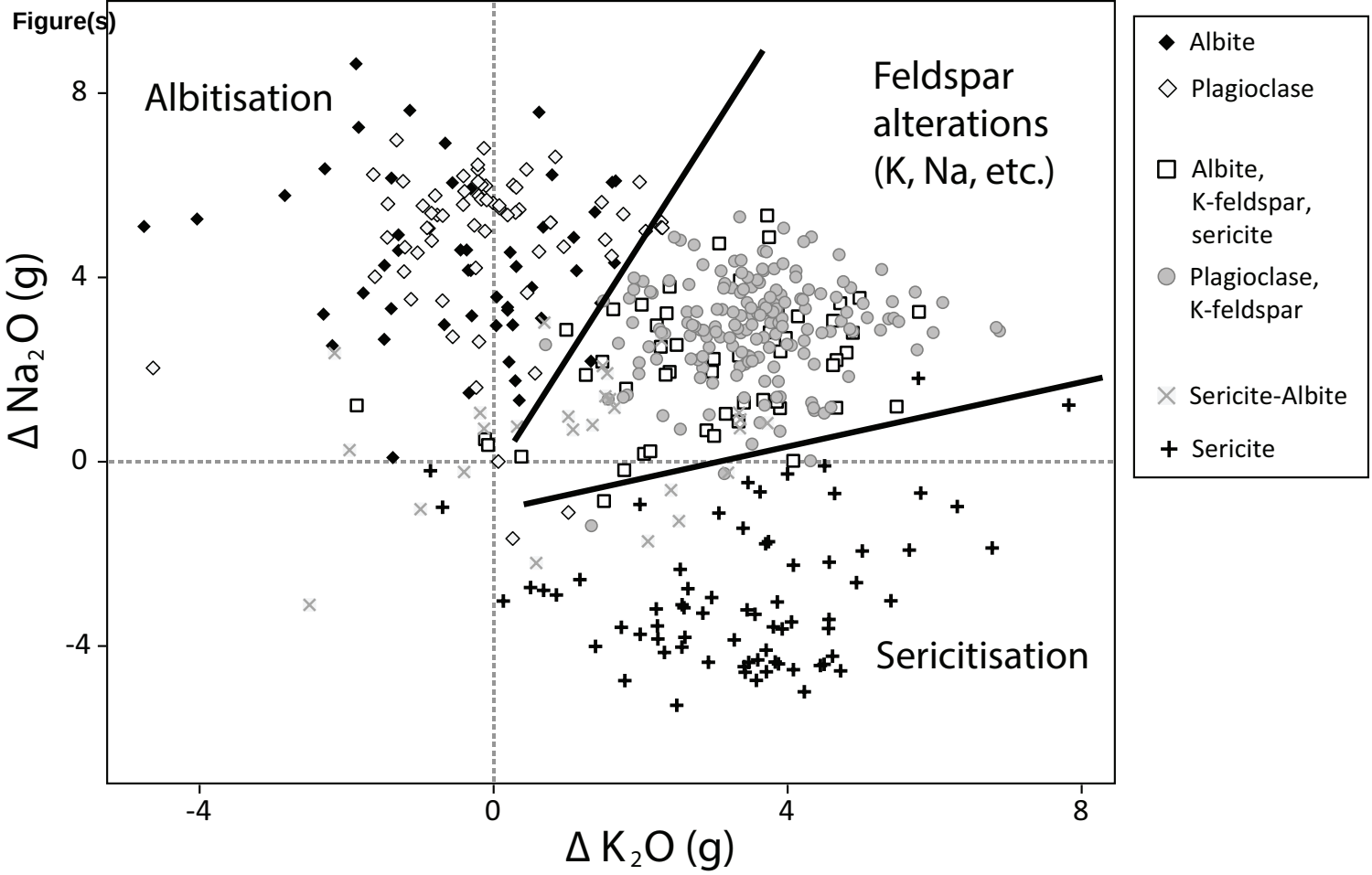

Fig. 10 (black and white)

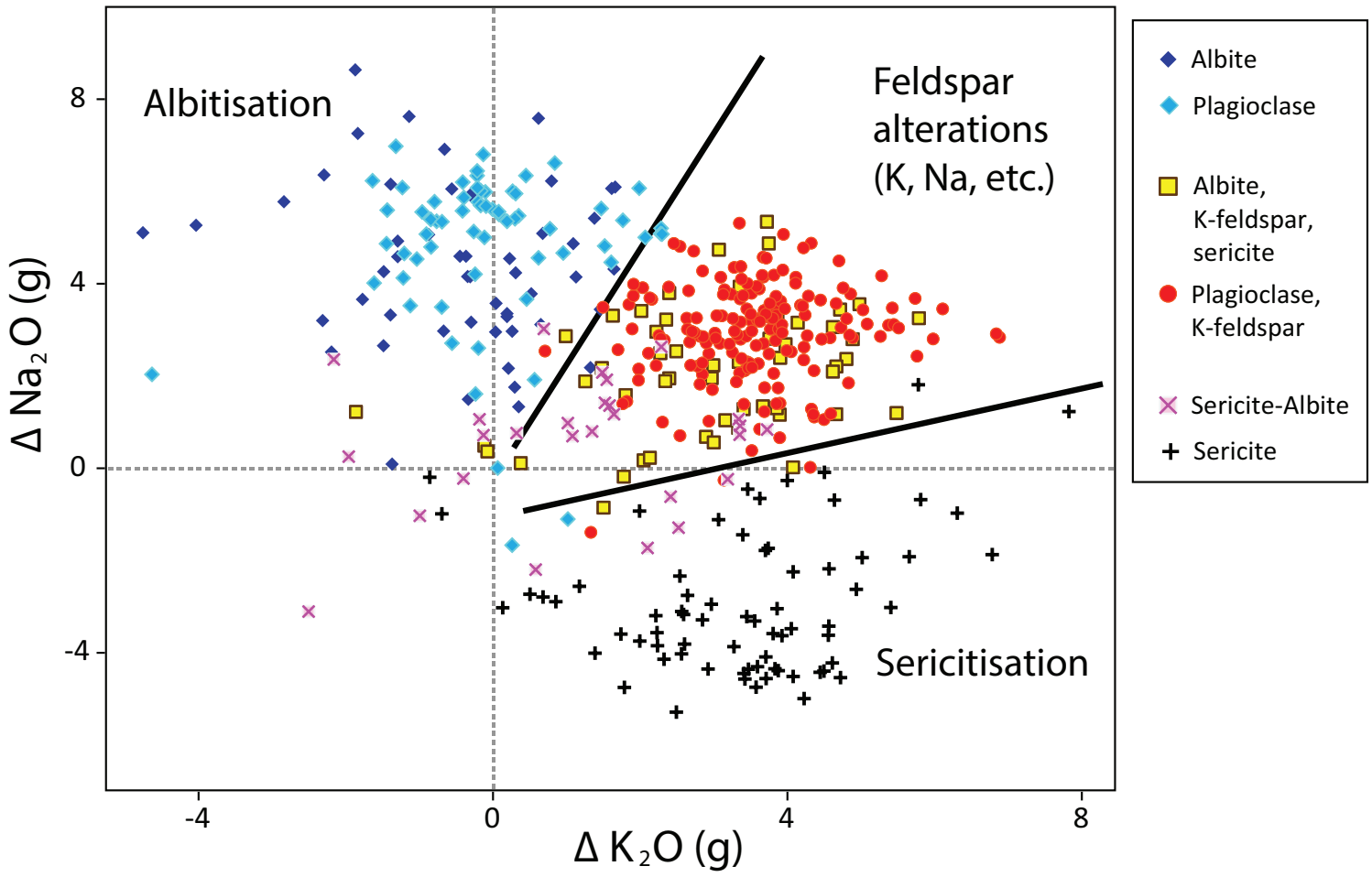

Fig. 10 (color, for online publication only)

Table 1
Optimal number of hidden neurons contained in the hidden layer of neural networks

Component	Optimal number of neurons
SiO_2	6
FeO^T	6
MgO	4
CaO	4
Na_2O	4
K_2O	4
TiO_2	2
Zr	6

Table 2

Median errors (ME), rank-based correlation coefficients (r), standard deviation of errors (SD) and normalized mean square errors (NMSE) for values predicted from the *test set*. Median errors and correlations are calculated for the whole *test set* (n=1 452), as well as for the following subsets: 1) subalkaline rocks (n=833); and 2) alkaline rocks (n=619). Alkalinity is determined using the TAS diagram (Le Bas et al., 1986).

	Global				Alkaline rocks				Subalkaline rocks			
Component	ME	r	NMSE	SD	ME	r	NMSE	SD	ME	r	NMSE	SD
SiO_2	0.095	0.935	0.139	3.091	0.793	0.918	0.163	2.817	-0.415	0.931	0.158	3.092
FeO^T	0.056	0.935	0.150	1.224	-0.173	0.913	0.149	1.157	0.224	0.937	0.167	1.258
MgO	0.015	0.946	0.103	1.472	-0.025	0.945	0.128	1.207	0.036	0.925	0.096	1.644
CaO	-0.021	0.878	0.231	1.474	-0.147	0.803	0.248	1.544	0.041	0.891	0.234	1.413
Na_2O	0.007	0.767	0.378	0.674	-0.064	0.726	0.415	0.725	0.057	0.798	0.375	0.598
K_2O	0.109	0.891	0.269	0.838	-0.044	0.838	0.337	0.996	0.213	0.921	0.220	0.651
TiO_2	0.000	0.992	0.018	0.129	-0.027	0.992	0.021	0.148	0.013	0.986	0.030	0.109
Zr	2.271	0.989	0.022	30.22	-0.460	0.989	0.026	35.09	3.438	0.984	0.022	24.92

Table 3

Rank-based correlation coefficients (r) for values predicted from the whole *test set*, for tests using un-transformed data (see **Table 2**) and log-transformed data.

Un-transformed		Centred log-transformed	
Component	r	Component	r
SiO_2	0.935	$Ln(SiO_2/Al_2O_3)$	0.791
FeO^T	0.935	$Ln(FeO^T/Al_2O_3)$	0.941
MgO	0.946	$Ln(MgO/Al_2O_3)$	0.951
CaO	0.878	$Ln(CaO/Al_2O_3)$	0.914
Na_2O	0.767	$Ln(Na_2O/Al_2O_3)$	0.756
K_2O	0.891	$Ln(K_2O/Al_2O_3)$	0.880

Table ~~3~~4

Effects that ±33% gains and losses of various immobile elements has on the predicted values calculated for selected samples. Modelled changes of elements' original values are in bold italics. Important prediction changes (>10%) are in bold.

	Original values (weight% for Al_2O_3 and TiO_2 and ppm for others)							Predicted values (weight% or ppm)								
Sample	Al_2O_3	TiO_2	Zr	Cr	Y	Nb	Th	SiO_2	FeOT	MgO	CaO	Na_2O	K_2O	TiO2	Zr	Note
S MB01-60A	14.37	0.61	364	2	43	18	17	69.8	4.20	0.64	2.25	4.18	4.37	0.67	370	Initial value
	14.37	0.61	364	2	43	***12***	17	68.6	4.10	0.68	2.26	4.42	4.24	0.67	373	-33% Nb
	14.37	0.61	364	2	43	***24***	17	69.7	4.31	0.61	2.24	4.16	4.64	0.67	374	+33% Nb
s TS/90/01	16.47	0.91	192	5	13	17	8.1	58.9	4.87	1.70	5.04	4.51	3.18	1.02	208	Initial value
	16.47	0.91	***126***	5	13	17	8.1	56.0	**5.72**	2.32	**6.22**	4.09	2.99	1.05	**142**	-33% Zr
	16.47	0.91	***255***	5	13	17	8.1	59.9	**4.35**	1.47	4.75	4.47	3.54	0.99	**260**	+33% Zr
s SA60	12.58	2.77	220	404	24	34	3.8	47.3	11.81	10.1	9.59	2.96	1.18	2.88	231	Initial value
	12.58	2.77	220	404	***16***	34	3.8	47.9	11.34	10.5	9.04	3.57	**1.94**	2.78	231	-33% Y
	12.58	2.77	220	404	***32***	34	3.8	46.8	12.03	9.64	10.28	2.68	0.98	2.89	231	+33% Y
s S10	16.54	0.68	163	76	13	14	10	62.4	4.55	2.63	4.47	4.05	2.61	0.72	169	Initial value
	16.54	0.68	163	***50***	13	14	10	62.4	4.52	2.38	4.47	4.13	2.68	0.72	171	-33% Cr
	16.54	0.68	163	***101***	13	14	10	62.4	4.58	2.85	4.47	3.99	2.55	0.71	167	+33% Cr
s OC93	15.72	1.07	244	154	23.4	13.3	11	58.5	6.66	4.69	6.70	3.56	2.81	1.08	249	Initial value
	15.72	***0.71***	244	154	23.4	13.3	11	63.0	**5.17**	**3.60**	**5.40**	3.72	3.17	**0.72**	244	-33% TiO_2
	15.72	***1.42***	244	154	23.4	13.3	11	56.1	**7.85**	5.29	**7.50**	3.41	2.51	**1.39**	249	+33% TiO_2

Click here to download Supplementary Materials: CAO.nsw

Click here to download Supplementary Materials: Main.obj

Click here to download Supplementary Materials: MassBalance.bsc

Click here to download Supplementary Materials: MassBalance.exe

Click here to download Supplementary Materials: MassBalance.ilk

Click here to download Supplementary Materials: MassBalance.pch

Click here to download Supplementary Materials: MassBalance.pdb

Click here to download Supplementary Materials: vc60.idb

Click here to download Supplementary Materials: vc60.pdb

Click here to download Supplementary Materials: FEOT.nsw

georoc.txt

Click here to download Supplementary Materials: Georoc.txt

Click here to download Supplementary Materials: Georoc.xls

Click here to download Supplementary Materials: K2O.nsw

Click here to download Supplementary Materials: Main.cpp

Click here to download Supplementary Materials: Main.h

Click here to download Supplementary Materials: MassBalance.dsp

Click here to download Supplementary Materials: MassBalance.dsw

Click here to download Supplementary Materials: MassBalance.ncb

Click here to download Supplementary Materials: MassBalance.opt

Click here to download Supplementary Materials: MassBalance.plg

Click here to download Supplementary Materials: MGO.nsw

Click here to download Supplementary Materials: NA2O.nsw

outputlimits.txt

Click here to download Supplementary Materials: outputLimits.txt

outputsgeoroc.txt

Click here to download Supplementary Materials: outputsGeoroc.txt

ratiolimits.txt

Click here to download Supplementary Materials: ratioLimits.txt

readme.txt

Click here to download Supplementary Materials: ReadMe.txt

Click here to download Supplementary Materials: SIO2.nsw

Click here to download Supplementary Materials: TIO2.nsw

Click here to download Supplementary Materials: ZR.nsw