

# Studies in Stoichiometry with Compositional Data

E. C. Grunsky<sup>1</sup>, B.A. Kjarsgaard<sup>1</sup>, J.J. Egozcue<sup>2</sup>,  
V. Pawlowsky-Glahn<sup>3</sup> and S. Thio-Henestrosa<sup>3</sup>

<sup>1</sup> Geological Survey of Canada, Ottawa, Ontario, Canada. egrunsky@nrcan.gc.ca

<sup>2</sup> Departament Matemàtica Aplicada III, Universitat Politècnica de Catalunya, Barcelona, Spain

<sup>3</sup> Departament Informàtica i Matemàtica Aplicada, Universitat de Girona, Girona, Spain

## Abstract

Geochemical data that is derived from the whole or partial analysis of various geologic materials represent a composition of mineralogies or solute species. Minerals are composed of structured relationships between cations and anions which, through atomic and molecular forces, keep the elements bound in specific configurations. The chemical compositions of minerals have specific relationships that are governed by these molecular controls. In the case of olivine, there is a well-defined relationship between Mn-Fe-Mg with Si. Balances between the principal elements defining olivine composition and other significant constituents in the composition (Al, Ti) have been defined, resulting in a near-linear relationship between the logarithmic relative proportion of Si versus (MgMnFe) and Mg versus (MnFe), which is typically described but poorly illustrated in the simplex.

The present contribution corresponds to ongoing research, which attempts to relate stoichiometry and geochemical data using compositional geometry. We describe here the approach by which stoichiometric relationships based on mineralogical constraints can be accounted for in the space of simplicial coordinates using olivines as an example. Further examples for other mineral types (plagioclases and more complex minerals such as clays) are needed. Issues that remain to be dealt with include the reduction of a bulk chemical composition of a rock comprised of several minerals from which appropriate balances can be used to describe the composition in a realistic mineralogical framework. The overall objective of our research is to answer the question: In the cases where the mineralogy is unknown, are there suitable proxies that can be substituted?

**Kew words:** Aitchison geometry, balances, mineral composition, oxides.

# 1 Introduction

Geochemical analysis of whole rock, weathered rock, soils, glacial tills, lake/stream sediments and chemical precipitates reflect the compositions of minerals. Minerals are composed of molecules of atoms that are in the form of regular structures (e.g. chain silicates, ortho-silicates, sheet silicates, etc.) (Deer, Howie and Zussman, 1963). The ordered structure of minerals requires that the atomic constituents arrange themselves in specific configurations, which are controlled by thermodynamic conditions. The molecular structure of minerals can usually be defined in terms of cation/anion relationships. Anions, for example, represent SiO<sub>2</sub> tetrahedra in many silicates, while the cations represent elements which are sited within these anion structures and balance the charges between the two. The combination of molecules that make up a specific mineral is termed stoichiometry. The physical principles governing the ordered assembly of atomic constituents in minerals is explained by Crystal Field Theory (Zoltai and Stout, 1984, Albède, 1995). Stoichiometry can be defined as the combination of elements that coincide with molecular structures such that the ratios of these elements are constant.

When a rock is geochemically analyzed, the resulting reported constituents (SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, FeO, Fe<sub>2</sub>O<sub>3</sub>, MgO, CaO, Na<sub>2</sub>O, K<sub>2</sub>O, TiO<sub>2</sub>, MnO, P<sub>2</sub>O<sub>5</sub> plus many other elements) represent a bulk composition and not the compositions of the minerals that make up the material. Thus, a chemical composition of rock does not represent the proper distribution/allocation of cations and anions. Therefore, an investigation of the “natural constraints” that mineralogy imposes on the chemical composition of a rock is warranted.

This study examines a relatively simple mineral, olivine. Given the stoichiometric equations of olivines in Appendix 3 of the book “An Introduction to the Rock-Forming Minerals” from W.A. Deer, R.A. Howie and J. Zussman (1992, 2<sup>nd</sup> edition) we decided to work with only six parts: SiO<sub>2</sub>, TiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, FeO, MnO and MgO. Small amounts of Mn, Ni, and Ca can substitute for Mg and Fe in the olivine structure (Mg, Fe)<sub>2</sub>[SiO<sub>4</sub>], whereas Ti and Al substitute for Si. In this study we have restricted our investigation to the principal constituents of olivine, namely Mg, Fe, Mn, Si, Al and Ti.

A set of 359 olivine analyses from Hawaiian volcanic rocks were obtained from a database of olivines taken from: <http://georoc.mpch-mainz.gwdg.de/georoc/Entry.html>.

## 2 Evaluating the olivine dataset

### 2.1 Exploratory analysis

The table of compositional descriptive statistics (Table 1), the variance of the centred logratio transformed data (clr-variance, Table 2), and the variation array (Table 3) already reflect interesting information concerning the behaviour of the oxides to be expected. This information is complemented by the biplot presented in Figure 1.

**Table 1.** Compositional descriptive statistics

	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MgO	MnO
Centre	39.8552	0.0129	0.0165	15.5804	44.3286	0.2064
Min	36.6167	0.0100	0.0099	9.4561	31.3499	0.0711
Max	41.2794	0.0612	0.1218	31.6414	49.9744	0.4080
Q25	39.0457	0.0101	0.0101	12.5204	41.7215	0.1620
Median	39.6452	0.0102	0.0102	15.0828	44.9826	0.2016
Q75	40.2268	0.0103	0.0301	18.8795	47.0487	0.2654

Table 1 shows that TiO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub>, followed by MnO, are proportionally very small compared to the other elements, the largest part corresponding in mean to MgO, as shown by the centre. It also shows that TiO<sub>2</sub> has a very short range of variability between the minimum (0.0100) and the third quantile (0.0103) compared to the maximum (0.0612), and analogously for Al<sub>2</sub>O<sub>3</sub> between the minimum (0.0099) and the median (0.0102). This behaviour might be related to the role played by these elements in the structure of olivines, as shall be seen later.

Table 2 reflects the variance of the clr transformed data (Aitchison, 1982). The sum of these variances is 0,7370, known as the total variance. As can be seen, the clr-component corresponding to  $\text{Al}_2\text{O}_3$  explains about 50% of the variability in the data, whereas all the other components are more or less the same size.

**Table 2.** clr-variance

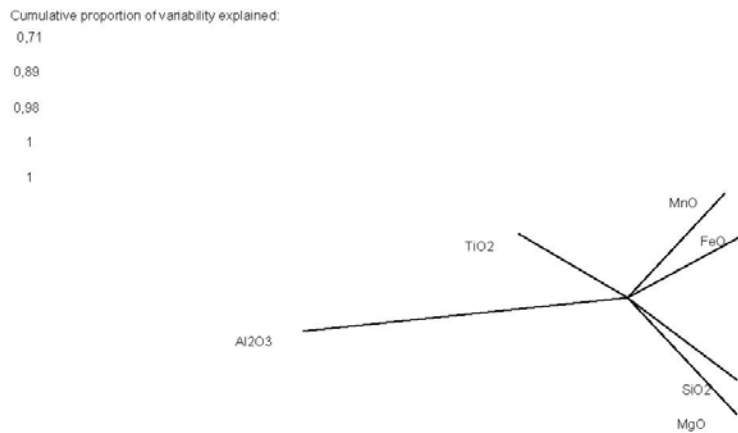
$\text{clr}(\text{SiO}_2)$	$\text{clr}(\text{TiO}_2)$	$\text{clr}(\text{Al}_2\text{O}_3)$	$\text{clr}(\text{FeO})$	$\text{clr}(\text{MgO})$	$\text{clr}(\text{MnO})$
0.0601	0.0970	0.3483	0.0707	0.0835	0.0775

The variation array (Aitchison, 1982) in Table 3 indicates that the variability is related mainly to  $\text{Al}_2\text{O}_3$ . In fact, except for the logratio variance  $\ln(\text{TiO}_2/\text{Al}_2\text{O}_3)$ , which is equal to 0.2796, all the other variances involving  $\text{Al}_2\text{O}_3$  have a value above 0.6.

**Table 3.** Variation Array (upper triangle: logratio variances, lower triangle: logratio means)

	$\text{SiO}_2$	$\text{TiO}_2$	$\text{Al}_2\text{O}_3$	$\text{FeO}$	$\text{MgO}$	$\text{MnO}$
$\text{SiO}_2$		0.2622	0.6164	0.0846	0.0050	0.1291
$\text{TiO}_2$	8.0374		0.2796	0.2506	0.2926	0.2337
$\text{Al}_2\text{O}_3$	7.7903	-0.2471		0.6648	0.6357	0.6301
$\text{FeO}$	0.9392	-7.0981	-6.8511		0.1283	0.0327
$\text{MgO}$	-0.1064	-8.1438	-7.8967	-1.0456		0.1764
$\text{MnO}$	5.2633	-2.7741	-2.5270	4.3241	5.3697	

This pattern of variability is graphically represented in the biplot in Figure 1. The first thing to be noticed is the high proportion of variability explained, namely 89%. Recall that the biplot represents the plane of the first two principal components (Aitchison and Greenacre, 2002). This means that the overall pattern of the data is essentially 2-D, or at most 3-D, taking into account that 98% of the variability is explained by the first three components.



**Figure 1.** Biplot of the Hawaii olivines.

The biplot also shows the large variability related to  $\text{Al}_2\text{O}_3$  and the nearly aligned pattern of the vertices corresponding to  $\text{SiO}_2$ ,  $\text{MgO}$ ,  $\text{FeO}$  and  $\text{MnO}$ .

## 2.2 Isometric logratio transformation of the data

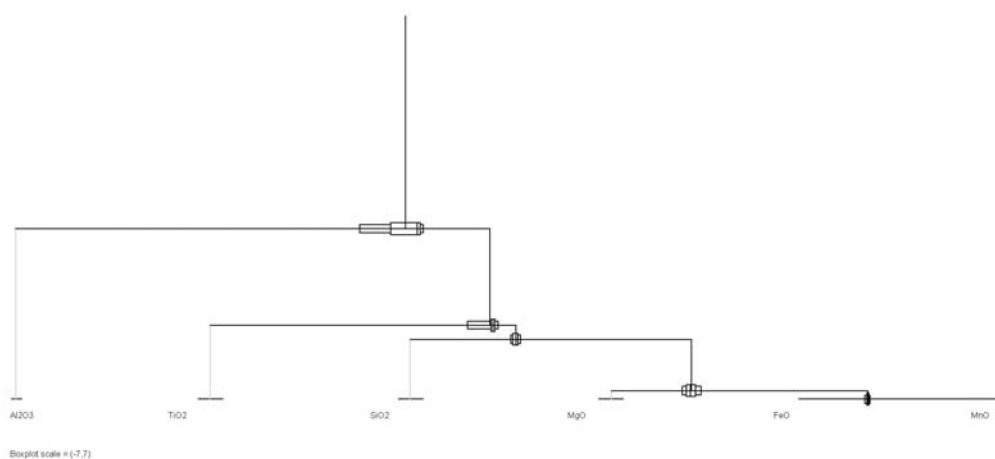
A balance analysis was carried out based on the stoichiometric equation of olivines,  $(\text{Mg}, \text{Mn}, \text{Fe})_2[\text{SiO}_4]$ , and on the results of the previous exploratory analysis. A balance analysis consists of an isometric logratio transformation in order to obtain coordinates in the real space (Egozcue and Pawlowsky-Glahn, 2005; Egozcue and others, 2003). Table 4 represents the sequential binary partition (SBP) which defines the balances chosen. Given their larger variability, first  $\text{Al}_2\text{O}_3$  and then  $\text{TiO}_2$  were separated, so that the basic relationships between  $\{\text{SiO}_2, \text{MgO}, \text{MnO}, \text{FeO}\}$ , the main elements defining the olivines, would be reflected in the last three balances. The balance-dendrogram presented in Figure 2 reflects the variability explained by each balance, as well as the summary statistics given in Table 5 in the form of box-plots.

The correlation between the balances is given in Table 6 and is represented in form of a scatterplot matrix in Figure 3.

**Table 4.** Sequential binary partition corresponding to the dendrogram shown in Figure 2.

Balance	SiO <sub>2</sub>	TiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	FeO	MgO	MnO
1	-1	-1	1	-1	-1	-1
2	-1	1	0	-1	-1	-1
3	1	0	0	-1	-1	-1
4	0	0	0	-1	1	-1
5	0	0	0	1	0	-1

As can be observed in Figure 2 and in Table 5, most of the variability is due to the balance Al<sub>2</sub>O<sub>3</sub> vs. all the others, and to the balance TiO<sub>2</sub> vs. {SiO<sub>2</sub>, MgO, MnO, FeO}. Thus, the first two balances describe exclusivity of both Ti and Al relative to the other elements. The smallest amount of variability is explained, according to this dendrogram, by the last balance, relating FeO and MnO.



**Figure 2.** Balance Dendrogram of 6 elements in olivine compositions. Boxplot scale(-7,7).

The previous statements are numerically reflected in Table 5. In fact, The largest variance is associated with balance B1 (0.4168), followed by balance B2 (0.1795), while the smallest corresponds to balance B5 (0.0163).

**Table 5:** Summary statistics of balances.

Balance	B1	B2	B3	B4	B5
Mean	-4.5311	-5.8257	1.7598	2.6190	3.0576
Std.Dev.	0.6456	0.4237	0.1627	0.3096	0.1278
Variance	0.4168	0.1795	0.0265	0.0958	0.0163
Min	-5.1409	-6.2534	1.3882	1.8044	2.4930
Max	-2.8337	-4.4921	2.1368	3.2349	3.6123

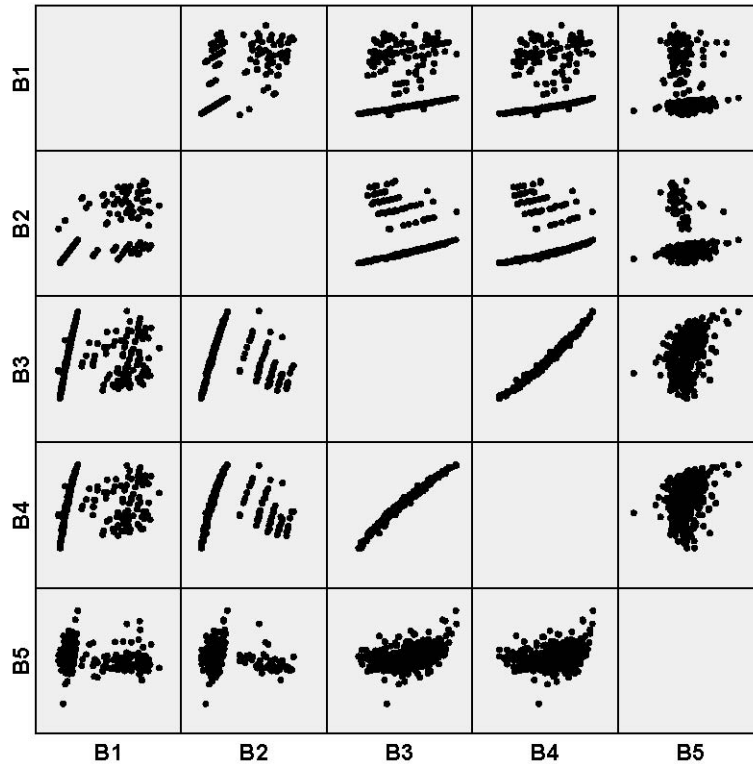
The relationship between Mg, Mn and Fe (balances B4 and B5) are consistent with the fact that these three elements share the same sites in the crystal structure. Aluminium is known to substitute for Si in the silica tetrahedral of olivine. Titanium also substitutes for Si, but less commonly. Alternatively, both Ti and Al can be present as defects in the crystal lattice structure in very small amounts. The fact that they play no structural role is reflected in the large amount of variability explained by the first two balances in which they participate.

Neither the balance dendrogram, nor Table 5, reflect the relationship between balances, which should potentially reveal the structure inherent in the data. This relationship appears in the correlation matrix presented in Table 6 and the scatterplot matrix in Figure 3.

**Table 6:** Correlation matrix of balances

	B1	B2	B3	B4	B5
B1	1	0.6808	0.0446	0.0472	-0.1942
B2	0.6808	1	-0.1288	-0.1277	-0.2336
B3	0.0446	-0.1288	1	0.9923	0.4192
B4	0.0472	-0.1277	0.9923	1	0.3497
B5	-0.1942	-0.2336	0.4192	0.3497	1

In Figure 3 balances B1 and B2 show distinctive features that likely describe atomic substitution for Si by Ti and Al. These patterns are shown as straight lines. The presence of Ti and Al that occur as defects in the olivine crystal structure are likely those observations that do not follow the straight line trends. Alternatively, these patterns may be due to analytical differences from different laboratories. Data from laboratories with detailed quality control parameters could resolve such issues. Unfortunately such information is not available in the database from which the olivine analyses were derived. However, looking at the correlation matrix in Table 6 and the scattergrams in Figure 3, the strong relationship between balances B3 and B4 is striking. The correlation coefficient is 0.9923 and the scattergram shows a near-linear trend. This near linear trend clearly defines the stoichiometric relationship between Si with Mg, Fe and Mn (B3) and Mg with Fe and Mn (B4). Balance B5 (Fe - Mn) shows a dispersed linear trend with balances B3 and B4 and a bimodal association with balances B1 and B2.



**Figure 3.** Scatterplot of balances for the Hawaiian olivines.

## 2.2 Analysis of the relationship between balances B3 and B4

The equations of balances B3 and B4 are, for the Silicate Framework ( $\text{SiO}_4$  tetrahedron):

$$B3 = \sqrt{\frac{3}{4}} \ln \frac{\text{Si}}{(\text{MgMnFe})^{1/3}} ,$$

and for the Divalent Cation Site (M1):

$$B4 = \sqrt{\frac{2}{3}} \ln \frac{Mg}{(MnFe)^{1/2}} .$$

To better illustrate the pattern in the sample, we have computed the principal components for B3 and B4. The first component PC1 explains  $\approx 99.6\%$  and PC2 the remaining  $\approx 0.4\%$ . After some calculations we obtain the equivalent expressions

$$PC1 \approx 0.140 \ln SiO_2 + 0.205 \ln MgO - 0.210 \ln MnO - 0.136 \ln FeO ,$$

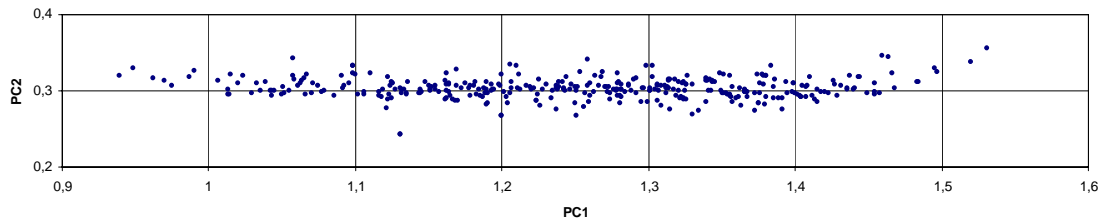
$$PC2 \approx 0.001 \ln SiO_2 - 0.017 \ln MgO - 0.074 \ln MnO + 0.090 \ln FeO .$$

Multiplying and dividing by convenient constants so that the coefficients become approximately integers we can approximate these equations by

$$PC1 \approx 1/7 \ln (SiO_2/FeO) + 1/5 \ln (MgO/ MnO);$$

$$PC2 \approx 1/10 \ln (FeO /MnO) \approx \text{constant}.$$

It is the second equation which describes the equilibrium we see in the scattergram between B3 and B4 and, given that the mean of  $\ln (FeO /MnO) \approx 4.3241$  (see Table 3), we obtain that the equilibrium in this data set is described by the ratio of FeO to MnO, which is approximately 75.5.



**Figure 4.** Scatterplot of the principal components between B3 and B4.

Despite the large amount of variability explained by the first principal component, this is though a rough approximation because, as can be seen in the scatterplot of PC1 versus PC2 (Figure 4), the relationship is slightly curved.

### 3 Discussion

Based on the known controls on the distribution of Si, Mg, Fe and Mn in the crystal structure of olivine, balances can be constructed that describe a well defined near-linear trend in the compositional range of olivines for Hawaiian volcanic rocks. Elements which are not controlled by the stoichiometric process, as reflected in balances B1 and B2, show minimal correlation with the B3 and B4 balances. The B5 balance shows a dispersed association of Mn with Fe relative to the stoichiometric equations of B3 and B4.

The slightly curved pattern of the scatterplot of the two principal components between B3 and B4 calls for an alternative model which better fits the data, which could be related to the crystallization process of olivines. Crystallization in general consists of two complementary processes: elimination of elements from the magma and accumulation of crystallized material. The first can be described by a linear process in the simplex, as it seems reasonable to think that elimination will be proportional to availability. The second has been called complementary process (Pawlawsky-Glahn and others, 2007) and further studies are required to adjust a proper model to the present data set.

The patterns revealed in this initial study are not surprising. The issue is how can the inclusion of stoichiometric equations be used in compositional data analysis? When geochemical patterns are observed in bulk geochemical analyses, can these be interpreted in terms of mineral compositions or mixtures of different minerals? Further work is required to tackle these issues.

### Acknowledgements

This research has been supported by the Spanish Ministry of Education and Science under project Ingenio Mathematica (i-MATH) No. CSD2006-00032 (Consolider – Ingenio 2010) and project MTM2006-03040.

## References

- Aitchison, J. 1982. The statistical analysis of compositional data (with discussion). *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 44(2), 139–177.
- Aitchison, J. & Greenacre, M. 2002. Biplots for compositional data. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 51(4), 375–392.
- Albarède, F. (1995). *Introduction to Geochemical Modeling*, Cambridge University Press, 543 p.
- Deer, W.A., Howie, R.A., and Zussman, J. (1966). *An Introduction to the Rock-forming Minerals*, Longman, London, 528 p.
- Egozcue, J.J. and Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology*, 37(7), 795–828.
- Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G. & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3), 279–300.
- Pawlowsky-Glahn, V., Egozcue, J.J., Tolosana-Delgado, R. (2007). Lecture notes on compositional data analysis. <http://hdl.handle.net/10256/297>, Univ. Girona. 95 p.
- Zoltai, T. and Stout, J.H. (1985). *Mineralogy, Concepts and Principles*, Bugess Publishing Company, Minneapolis, Minnesota, 505 p.