

Another look at the constant sum problem in geochemistry

H. R. ROLLINSON

Department of Geology, University of Zimbabwe P.O. Box MP 167, Mount Pleasant, Harare, Zimbabwe

Abstract

Compositional data—that is data where concentrations are expressed as proportions of a whole, such as percentages or parts per million—have a number of peculiar mathematical properties which make standard statistical tests unworkable. In particular correlation analysis can produce geologically meaningless results. Aitchison (1986) proposed a log-ratio transformation of compositional data which allows inter-element relationships to be investigated. This method was applied to two sets of geochemical data—basalts from Kilauea Iki lava lake and granitic gneisses from the Limpopo Belt—and geologically 'sensible' results were obtained. Geochemists are encouraged to adopt the Aitchison method of data analysis in preference to the traditional but invalid approach which uses compositional data.

KEYWORDS: correlation analysis, constant sum problem, geochemistry.

Introduction

THE constant sum problem in geochemistry has been regarded by many as one of geochemistry's intractable problems. Meisch (1969) argued that 'the constant sum problem is one of the most important and most difficult encountered in the analysis and interpretation of geochemical data'. Ten years later the problem was summarised by Butler (1979) as follows: 'At this time it does not seem possible to satisfactorily separate the effect of percentage formation from fractionation, assimilation or whatever petrologic model the investigator may have in mind'. Recently, however, Aitchison (1986) has proposed a new method of expressing geochemical data which, it is claimed, is free from the problems of percentage data. Woronow (1987) in reviewing Aitchison's book challenges geologists to take up this tool in their analysis of geochemical data. As yet only a few limited applications have been made of Aitchison's method and the purpose of this short paper is to apply this new approach to two data sets typical of those used in igneous petrology.

The constant sum problem

Geochemists are used to expressing the major element compositions of rocks and minerals as

percentages, so that the sum of the major elements will always be about 100%. This standard form of data presentation is potentially unsound for three reasons:

1. It introduces a negative bias into correlations (Chayes, 1960).
2. It forces a correlation between components of the data set (Meisch, 1969).
3. Subcompositions such as 'A', 'F' and 'M' of the AFM diagram, do not reflect the variations present in the 'parent' data set (Aitchison, 1986). Clearly therefore, correlation, a statistical technique frequently used in geochemistry to examine the relationships between oxides and elements, is fraught with pitfalls when applied to percentage data. For whilst it may be possible to demonstrate that clear correlations exist, their meaning is in doubt if they can arise from the nature of the data itself rather than as a result of a geological process (see Skala, 1979; Butler, 1986).

This problem is familiar to geochemists although few have taken it seriously, chiefly because there has been no alternative. Recently, however, Aitchison (1981, 1982, 1984, 1986) has proposed a new approach to the statistical analysis of percentage data which has wide applications to all forms of percentage data including major element geochemical data. The purpose of this short paper, therefore, is to outline Aitchison's method, to apply it to a 'typical' set of major

element data and evaluate the usefulness of the method.

Aitchison's method

The main point of Aitchison's work is that 'the study of compositions is essentially concerned with the *relative* magnitudes of the ingredients rather than their absolute values; (this) leads naturally to a conclusion that we should think in terms of ratios' (Aitchison, 1986, p. 65). Thus when formulating questions about associations between variables in a geochemical data set, our thinking should be based upon ratios rather than upon percentages as has traditionally been the case. Aitchison goes on to point out that handling the variances and covariances of ratios is difficult and that a mathematically simpler approach is to take logarithms of these values. Thus he proposes that compositional data should be expressed as the covariances of (natural) log-ratios of the variables rather than the raw percentages.

Aitchison (1986) proposes three ways of handling log-ratio data:

1. As a *variation matrix* in which the log-ratio variances are plotted for every variable ratioed to every other variable. This matrix provides a measure of the relative variation of every pair of variables and can be used in a descriptive sense to identify relationships within the data array and in a comparative mode between data arrays.
2. To ratio every variable against a common divisor. The covariances of these log-ratios are presented as a *log-ratio covariance matrix*. The choice of variable as the divisor is immaterial because it is the structure of the matrix which is of importance rather than the individual values of the covariances.
3. To ratio every variable against its geometric mean. This gives rise to a *centred log-ratio covariance matrix*. This has the conceptual advantage for the geochemist that no one element is singled out as the divisor (see Butler and Woronow, 1986).

Data analysis

In this study two major element data sets are analysed using Aitchison's methodology. The first, a suite of basaltic volcanic rocks from Hawaii was chosen because of its apparent geological simplicity, that is the chemical variability appears to be the product of only one process. The second suite, metamorphosed tonalitic and trondhjemitic gneisses are thought to be more complex and their chemistry a function of a number of different processes.

Following the method of Aitchison (1986) the chemical analyses (Tables 1 and 2) were recalculated as log ratios and the variation matrix, the covariance matrix and the centred covariance matrix for the two data sets are presented in Tables 3 and 4. The covariance matrix was calculated using TiO_2 as the common divisor, although Aitchison (1986) makes it plain that any variable from the data set will do. The slight differences between the variances between TiO_2 and the other oxides calculated in the variation matrix (Tables 3a and 4a) and the variances in the log-ratio covariance matrix (Tables 3b and 4b) are a consequence of rounding errors from two separate calculation methods. The variation matrix was calculated in a spread-sheet and the covariance matrices were calculated using the statistical package SPSS-x.

The results of the data analysis will be considered in two different ways, in keeping with two different styles of data analysis. Firstly, the results are inspected 'to see what is there', without any particular hypothesis in mind. Secondly, they are examined in the light of a particular hypothesis in order to confirm or negate that hypothesis (Love and Woronow, 1988).

Example 1. Basalts from Kilauea Iki lava lake, Hawaii

The basalts of the Kilauea Iki lava lake, Hawaii were chosen for this study because there is good field and petrographic evidence to indicate that their chemical variability is due to the fractionation of magnesian olivine (Fo_{85}) from a single batch of magma (Richter and Moore, 1966). They contain olivine phenocrysts in a groundmass of fine grained clinopyroxene, plagioclase, titaniferous magnetite and glass with minor apatite, orthopyroxene and cristobalite. The conventional interpretation of these data using a correlation matrix (Table 1a and b) indicates strong associations between almost all element pairs except for those involving Mn. MnO concentrations are almost constant.

(i) *Data inspection.* In Table 3 the variation and covariance matrices are calculated after the method of Aitchison (1986). In the variation matrix (Table 3a) the highest values are for elements paired with MgO. In particular the pairs Mg-K, Mg-Ti, Mg-P, Mg-Na, Mg-Ca and Mg-Al show the greatest relative variation. The sum of the variance of all elements with MgO accounts for 65% of the total variance. The log-ratio covariance matrix (Table 3b) shows that the element pairs Mg-Si, Mg-Fe and Mg-Mn show the highest covariances. In this case the covar-

Table 1. Chemical analyses of rocks from Kilauea Iki lava lake, Hawaii (from Richter and Moore, 1966)

(a) PERCENTAGE DATA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
SiO ₂	48.29	48.83	45.61	45.50	49.27	46.53	48.12	47.93	46.96	49.16	48.41	47.90	48.45	48.98	48.74	49.61	49.20
TiO ₂	2.33	2.47	1.70	1.54	3.30	1.99	2.34	2.32	2.01	2.73	2.47	2.24	2.35	2.48	2.44	3.03	2.50
Al ₂ O ₃	11.48	12.38	8.33	8.17	12.10	9.49	11.43	11.18	9.90	12.54	11.80	11.17	11.64	12.05	11.60	12.91	12.32
Fe ₂ O ₃	1.59	2.15	2.12	1.60	1.77	2.16	2.26	2.46	2.13	1.83	2.81	2.41	1.04	1.39	1.38	1.60	1.26
FeO	10.03	9.41	10.02	10.44	9.89	9.79	9.46	9.36	9.72	10.02	8.91	9.36	10.37	10.17	10.18	9.68	10.13
MnO	0.18	0.17	0.17	0.17	0.17	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.17	0.18
MgO	13.58	11.08	23.06	23.87	10.46	19.28	13.65	14.33	18.31	10.05	12.52	14.64	13.23	11.18	12.35	8.84	10.51
CaO	9.85	10.64	6.98	6.79	9.65	8.18	9.87	9.64	8.58	10.55	10.18	9.58	10.13	10.83	10.45	10.96	11.05
Na ₂ O	1.90	2.02	1.33	1.28	2.25	1.54	1.89	1.86	1.58	2.09	1.93	1.82	1.89	1.73	1.67	2.24	2.02
K ₂ O	0.44	0.47	0.32	0.31	0.65	0.38	0.46	0.45	0.37	0.56	0.48	0.41	0.45	0.80	0.79	0.55	0.48
P ₂ O ₅	0.23	0.24	0.16	0.15	0.30	0.18	0.22	0.21	0.19	0.26	0.23	0.21	0.23	0.24	0.23	0.27	0.23

(b) CRUDE CORRELATION MATRIX

	SiO ₂	TiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MnO	MgO	CaO	Na ₂ O	K ₂ O	P ₂ O ₅
SiO ₂	1.000	0.907	0.990	-0.857	0.173	-0.997	0.962	0.906	0.705	0.933
TiO ₂		1.000	0.875	-0.703	-0.064	-0.900	0.770	0.931	0.649	0.985
Al ₂ O ₃			1.000	-0.858	0.202	-0.996	0.974	0.917	0.633	0.905
Fe ₂ O ₃				1.000	-0.124	0.848	-0.881	-0.762	-0.489	-0.717
MnO					1.000	-0.193	0.328	0.004	0.152	-0.029
MgO						1.000	0.969	-0.915	-0.679	-0.924
CaO							1.000	0.814	0.661	0.807
Na ₂ O								1.000	0.415	0.928
K ₂ O									1.000	0.678
P ₂ O ₅										1.000

iance between Mg and Mn, not seen elsewhere is a consequence of ratioing to TiO₂, since the variability in TiO₂ imparts a variance to the ratio MnO/TiO₂. The elements Na, K and P record very small covariances with all elements. In the centred log-ratio covariance matrix (Table 3c) the element pairs Mg-K, Mg-Ti, Mg-P, Mg-Na and Mg-Ca show the highest negative covariances and Mg-Fe a high positive covariance. In general the elements Si, Al, Fe, Mn and Ca show very low covariances.

These observations may be interpreted as follows. Firstly, the variation matrix (Table 3a) highlights the fact that the greatest relative variation exists between MgO (indicative of olivine as the fractionating phase) and the elements excluded from olivine and concentrated in the melt—K, Ti, P, Na, Ca and Al. Further, the magnitude of the variance of the log-ratios is related to the extent to which the element is partitioned between olivine and the melt. In other words the most incompatible elements have the highest log-ratio variances.

Secondly, the log-ratio covariance matrix shows that the greatest covariances are between the prime constituents of olivine (Mg-Fe-Si and Mn), the fractionating phase. Na, K and P which show very small covariances are elements excluded from olivine and concentrated in the

melt. In this particular case the results of the log-ratio covariance matrix should be interpreted with the caution because some ambiguity of interpretation can arise from an inappropriate choice of divisor for the log ratios.

Thirdly, in the centred log-ratio covariance matrix the negative covariances Mg-K, Mg-Ti, Mg-P, Mg-Fe, Mg-Na and Mg-Ca show the highest values, emphasising the strong antipathy between Mg (in olivine) and the elements K, Ti, P, Na and Ca which are concentrated in the melt. The positive covariance between Fe and Mg is reflected in the strong association between Mg and Fe in the mineral olivine. In contrast the elements with very small covariances Si, Al, Fe Mn and Ca show no strong association with most other elements.

(ii) Hypothesis testing. Confirmation of the olivine fractionation model comes from applying the end-member unmixing model for compositional data developed by Woronow (1991). This type of modelling is regarded as complementary to rather than replacing the log-ratio correlation analysis thus far described. Calculations based on the method of Woronow (1991) show that the rock compositions at Kilauea Iki lava lake are not controlled by pure Mg-Fe olivine but that additional ions present in the olivine are also important. These results are figured in Woronow

Table 2. Tonalitic and trondhjemitic gneisses from the northern marginal zone of the Limpopo Belt, Zimbabwe

(a) PERCENTAGE DATA

Rock No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SiO ₂	61.50	62.58	63.62	66.67	67.63	68.00	68.55	70.05	71.47	72.81	73.68	74.25	75.62	77.88	78.44
TiO ₂	0.61	0.56	0.61	0.72	0.47	0.78	0.47	0.32	0.25	0.39	0.33	0.18	0.13	0.19	0.34
Al ₂ O ₃	15.88	18.10	16.87	15.41	15.47	11.70	16.20	15.93	14.82	14.15	14.33	13.96	13.32	11.04	11.04
Fe ₂ O ₃	7.96	5.34	5.22	5.61	4.44	7.50	3.73	2.88	3.04	3.16	2.28	1.53	1.95	3.08	3.14
MnO	0.15	0.09	0.08	0.09	0.08	0.10	0.03	0.03	0.04	0.06	0.02	0.03	0.05	0.04	0.05
MgO	3.60	1.71	1.82	1.32	1.60	1.94	1.13	0.82	0.63	0.68	0.83	0.32	0.20	0.01	0.01
CaO	4.96	4.38	4.24	4.78	4.38	5.33	4.24	3.11	3.22	3.41	2.87	2.17	1.39	1.36	1.04
Na ₂ O	4.42	6.01	4.94	4.03	4.22	3.43	4.47	5.06	5.16	3.90	4.57	4.15	4.23	4.78	4.71
K ₂ O	0.82	1.30	1.70	1.05	1.02	0.54	1.31	1.95	1.04	1.75	2.83	2.76	2.98	1.23	1.75
P ₂ O ₅	0.28	0.20	0.17	0.17	0.09	0.16	0.13	0.12	0.04	0.10	0.07	0.05	0.04	0.01	0.07

(b) CRUDE CORRELATION MATRIX

	SiO ₂	TiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MnO	MgO	CaO	Na ₂ O	K ₂ O	P ₂ O ₅	
SiO ₂	1.000	-0.834	-0.787	-0.776	-0.725	-0.896	-0.893	-0.184	0.519	-0.892	SiO ₂
TiO ₂		1.000	0.432	0.876	0.720	0.796	0.942	-0.205	-0.646	0.813	TiO ₂
Al ₂ O ₃			1.000	0.238	0.243	0.522	0.586	0.478	-0.088	0.579	Al ₂ O ₃
Fe ₂ O ₃				1.000	0.911	0.881	0.806	-0.139	-0.776	0.854	Fe ₂ O ₃
MnO					1.000	0.857	0.661	-0.142	-0.618	0.818	MnO
MgO						1.000	0.838	-0.044	-0.553	0.910	MgO
CaO							1.000	-0.142	-0.648	0.794	CaO
Na ₂ O								1.000	0.029	0.058	Na ₂ O
K ₂ O									1.000	-0.492	K ₂ O
P ₂ O ₅										1.000	P ₂ O ₅

(in press). Support for this observation comes from mineral analyses (Helz, 1987) which show that the olivine contains 0.15–0.35 wt.% MnO and 0.15–0.30 wt.% CaO and suggest that a slightly impure olivine can account for the observed compositional range.

Example 2. Tonalitic and trondhjemitic gneisses from the Limpopo Belt

The analyses presented in Table 2a are of comagmatic, granulite facies tonalites and trondhjemites from the north marginal zone of the Limpopo belt, Zimbabwe. The most common mineral assemblage is quartz, plagioclase (An_{20–30}), orthopyroxene [Mg/(Mg + Fe) = 0.44] and biotite but iron–titanium oxides are present in some rocks as are garnet, retrogressive hornblende and potassium feldspar. The interpretation of these data is more difficult than for the previous example for there is no physical model for the chemical variability of these rocks, although Kramers and Ridley (1989) have proposed that they are magmatic granulites and so the observed mineralogy may be primary.

The conventional interpretation of these data based on the crude correlation matrix for the percentage data (Table 2b) indicates that the strongest associations of elements are the pairs Si–Ti, Si–Mg, Fe–Ti, Ca–Ti, P–Ti, Mn–Fe, Mg–

Fe, Ca–Fe, Mn–Mg, and Mg–Ca. The presence of the association of the elements Si–Mg–Fe–Mn–Ca–Ti might suggest the importance of a ferromagnesian silicate such as hornblende. Table 4 shows the variation and covariance matrices for the log ratios.

(i) *Data inspection.* Following the method of Aitchison (1986), however, the variation matrix in Table 4a shows that the element pairs with the greatest variation are Mg–K, Mg–Si, Mg–Al and Mg–Na. The smallest log-ratio variances are between the elements Si–Al–Na and Fe–Ti–Ca–Mn. The log-ratio covariance matrix (Table 4b) shows the greatest positive covariances between the pairs Si–K, Na–K, K–Al, Si–Na and Ca–P and negative covariances between Mg–K, Mg–Si and Mg–Na. The lowest covariances are associated with Ca and P. The centred log-ratio covariance matrix (Table 4c) shows the greatest positive covariance between Ca and P and negative covariances between the pairs Mg–K, Mg–Na and Mg–Si. Elemental associations with Mn, Fe, Ca and Ti show the lowest covariances.

These data suggest that there are two main associations amongst the major elements—a magnesian phase and a Si–Al–Na–K-bearing phase, most likely alkali feldspar. The association between Ca and P may also suggest the importance of the mineral apatite in the petrogenesis of the suite.

Table 3. Variation and covariance matrices of data from Hawaii (Table 1a)

(a) VARIATION MATRIX

	SiO ₂	TiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MnO	MgO	CaO	Na ₂ O	K ₂ O	P ₂ O ₅	Total Variance
SiO ₂	0	0.025	0.012	0.002	0.001	0.093	0.014	0.018	0.058	0.022	
TiO ₂		0	0.006	0.038	0.033	0.205	0.009	0.004	0.029	0.001	
Al ₂ O ₃			0	0.022	0.017	0.167	0.001	0.003	0.035	0.004	
Fe ₂ O ₃				0	0.001	0.07	0.025	0.029	0.074	0.035	
MnO					0	0.08	0.018	0.025	0.067	0.03	
MgO						0	0.173	0.183	0.259	0.199	
CaO							0	0.007	0.033	0.007	
Na ₂ O								0	0.047	0.004	
K ₂ O									0	0.028	
P ₂ O ₅										0	2.213

(b) LOG-RATIO COVARIANCE MATRIX

	SiO ₂ /TiO ₂	Al ₂ O ₃ /TiO ₂	Fe ₂ O ₃ /TiO ₂	MnO/TiO ₂	MgO/TiO ₂	CaO/TiO ₂	Na ₂ O/TiO ₂	K ₂ O/TiO ₂	P ₂ O ₅ /TiO ₂
SiO ₂ /TiO ₂	0.026	0.010	0.032	0.030	0.073	0.010	0.006	-0.003	0.002
Al ₂ O ₃ /TiO ₂		0.007	0.012	0.012	0.024	0.008	0.003	0.000	0.001
Fe ₂ O ₃ /TiO ₂			0.040	0.037	0.092	0.012	0.006	-0.004	0.002
MnO/TiO ₂				0.035	0.085	0.013	0.006	-0.003	0.002
MgO/TiO ₂					0.218	0.022	0.014	-0.013	0.004
CaO/TiO ₂						0.010	0.004	0.003	0.002
Na ₂ O/TiO ₂							0.004	-0.008	0.001
K ₂ O/TiO ₂								0.031	0.001
P ₂ O ₅ /TiO ₂									0.001

(c) CENTRED LOG-RATIO COVARIANCE MATRIX

	SiO ₂ /g	TiO ₂ /g	Al ₂ O ₃ /g	Fe ₂ O ₃ /g	MnO/g	MgO/g	CaO/g	Na ₂ O/g	K ₂ O/g	P ₂ O ₅ /g
SiO ₂ /g	0.002	-0.005	-0.002	0.004	0.003	0.013	-0.003	-0.003	-0.008	-0.005
TiO ₂ /g		0.014	0.006	-0.009	-0.008	-0.030	0.005	0.010	0.013	0.012
Al ₂ O ₃ /g			0.005	-0.005	-0.004	-0.018	-0.006	0.006	0.006	0.006
Fe ₂ O ₃ /g				0.008	0.006	0.024	-0.006	-0.006	-0.014	-0.009
MnO/g					0.005	0.019	-0.004	-0.005	-0.011	-0.007
MgO/g						0.083	-0.020	-0.023	-0.040	-0.029
CaO/g							0.007	0.005	0.008	0.005
Na ₂ O/g								0.010	0.002	0.009
K ₂ O/g									0.043	0.013
P ₂ O ₅ /g										0.012

There is however some ambiguity in these data, for inspection of Table 2a shows that samples 14 and 15 are slightly anomalous inasmuch as they have very low concentrations of MgO. If the data are reanalysed with these two samples omitted a rather different result (not shown) is obtained. The variation matrix shows that the greatest variability is between the element pairs K-Mg, K-Mn, K-Fe, K-Ti and K-P. The log-ratio covariance matrix shows high positive covariances between the elements K-Si-Na-Al. The centred log-ratio covariance matrix shows a high positive covariance between K-Si and K-Na and high negative covariance between K-Mg, K-P and K-Fe.

The positive covariances in the two covariance matrices support the association of the elements

Si-Al-Na-K indicated in the previous analysis and suggest the importance of alkali feldspar as a variable controlling the observed rock chemistry. The log-ratio variances and the negative covariances indicate antipathy between a potassic phase and Fe, Mg, Mn, Ti and P suggesting that the minerals apatite, a ferromagnesian mineral +/- iron-titanium oxides are also important.

(ii) Hypothesis testing. No firm conclusions were reached using the end-member unmixing model of Woronow (1991) although there is some evidence for potassium feldspar and albite control on the observed chemical variations. Support for the involvement of feldspar in the petrogenesis comes, however, from trace element studies. Several samples have rare earth element patterns which show a negative Eu anomaly. In addition

Table 4. Variation and covariance matrices for Limpopo belt gneisses calculated after methods of Aitchison (1986)

(a) VARIATION MATRIX

	SiO ₂	TiO ₂	Al ₂ O ₃	Fe ₂ O ₃	MnO	MgO	CaO	Na ₂ O	K ₂ O	P ₂ O ₅	Total Variance
SiO ₂	0	0.4100	0.0400	0.2700	0.3600	3.1500	0.3200	0.0300	0.1900	0.7600	
TiO ₂		0	0.2600	0.1000	0.2600	1.6900	0.0400	0.3700	0.8900	0.2700	
Al ₂ O ₃			0	0.2000	0.2900	2.5800	0.1800	0.0200	0.2300	0.5300	
Fe ₂ O ₃				0	0.0900	2.3700	0.1400	0.2400	0.7900	0.3700	
MnO					0	2.5100	0.2900	0.3400	0.8500	0.4600	
MgO						0	1.5900	3.0100	3.6600	1.4400	
CaO							0	0.2900	0.7400	0.2900	
Na ₂ O								0	0.2100	0.6900	
K ₂ O									0	1.1300	
P ₂ O ₅										0	34.94

(b) LOG-RATIO COVARIANCE MATRIX

	SiO ₂ /TiO ₂	Al ₂ O ₃ /TiO ₂	Fe ₂ O ₃ /TiO ₂	MnO/TiO ₂	MgO/TiO ₂	CaO/TiO ₂	Na ₂ O/TiO ₂	K ₂ O/TiO ₂	P ₂ O ₅ /TiO ₂
SiO ₂ /TiO ₂	0.437	0.336	0.129	0.164	-0.566	0.066	0.405	0.596	-0.049
Al ₂ O ₃ /TiO ₂		0.280	0.088	0.121	-0.335	0.066	0.327	0.492	-0.003
Fe ₂ O ₃ /TiO ₂			0.108	0.143	-0.307	-0.002	0.126	0.109	-0.004
MnO/TiO ₂				0.276	-0.299	0.005	0.154	0.157	0.032
MgO/TiO ₂					1.814	0.073	-0.508	-0.577	0.278
CaO/TiO ₂						0.041	0.065	0.103	0.006
Na ₂ O/TiO ₂							0.399	0.566	-0.030
K ₂ O/TiO ₂								0.955	0.015
P ₂ O ₅ /TiO ₂									0.283

(c) CENTRED LOG RATIO COVARIANCE MATRIX

	SiO ₂ /g	TiO ₂ /g	Al ₂ O ₃ /g	Fe ₂ O ₃ /g	MnO/g	MgO/g	CaO/g	Na ₂ O/g	K ₂ O/g	P ₂ O ₅ /g
SiO ₂ /g	0.219	-0.067	0.132	0.023	0.022	-0.590	-0.042	0.186	0.287	-0.213
TiO ₂ /g		0.085	-0.053	0.046	0.009	0.128	0.043	-0.065	-0.157	0.049
Al ₂ O ₃ /g			0.090	-0.003	-0.006	-0.345	-0.028	0.124	0.197	-0.133
Fe ₂ O ₃ /g				0.116	0.114	-0.219	0.001	0.022	-0.087	-0.013
MnO/g					0.211	-0.249	-0.029	0.013	-0.074	0.026
MgO/g						1.982	0.157	-0.529	-0.690	0.523
CaO/g							0.042	-0.042	-0.096	0.027
Na ₂ O/g								0.182	0.257	-0.198
K ₂ O/g									0.555	0.300
P ₂ O ₅ /g										0.683

these samples have negative Sr anomalies on 'spider diagrams'. Both these features are indicative of crystal-liquid equilibria involving feldspar either in the source or in a melt.

Discussion and conclusions

Many authors have argued that the traditional approach to compositional data in geochemistry is in error and that many deductions drawn simply from the inspection of correlation diagrams such as Harker diagrams or from correlation matrices are invalid.

An alternative approach, that of Aitchison, has been tested not as a comparison with the trad-

tional approach to see if it gives a 'better' answer, for if the traditional approach is in error it cannot be used as a standard for comparison. Rather we have used one example—Kilauea Iki—where the hypothesis of olivine fractionation has been advanced on geological grounds. The results of this new method confirm this hypothesis, thus lending weight to the applicability of this new method. The second example—Limpopo belt gneisses—leads to less certain conclusions, indicating that in this case the major element chemistry on its own, is unable to constrain a model for their origin.

The Aitchison method, at first sight clumsy and unnecessarily complex, is a necessary approach to

compositional data. However, if the data are manipulated by computer the computation need not be tedious or complex. Rather, it is incumbent upon the geochemical community to adopt this method in order to disentangle fact from fiction in their data. From the results of this empirical study the following observations are offered in the interpretation of log-ratio matrices.

1. The high values in the variation matrix identify the element pairs which show the greatest variability. In igneous rocks this may be between a crystallising mineral and the melt or between two or more crystallising minerals.

2. High positive values in the two covariance matrices indicate strong associations between the elements and this is interpreted to mean that they coexist in the same mineral.

3. Large negative values in the two covariance matrices tend to confirm the variability indicated in the variation matrix.

4. Total variance (Tables 3a and 4a) may turn out to be an important indicator of the processes operating in a suite of rocks. The total variance of the tonalitic and trondhjemitic gneisses (intuitively thought to be more complex than the basalts) is an order of magnitude greater than the total variance of the basalts. It should also be noted, however, that when the two samples were removed from the gneiss data set the total variance was more than halved.

5. The difference in interpretation of the two versions of the gneiss data set highlights the fact that this methodology is very sensitive to non-normal, small data sets.

Acknowledgements

I thank Gordon Lampitt, Alex Woronow and my late colleague Thorley Sweetman for helpful comments on an earlier draft of this manuscript.

References

Aitchison, J. (1981) A new approach to null correlations of proportions. *J. Math. Geol.*, **13**, 175–89.

— (1982) The statistical analysis of compositional data (with discussion). *J. Roy. Stat. Soc.*, **44**, 139–77.

— (1984) The statistical analysis of geochemical compositions. *J. Math. Geol.*, **16**, 531–64.

— (1986) *The statistical analysis of composition data*. Methuen, New York, 416 pp.

Butler, J. C. (1979) Numerical consequences of changing the units in which chemical analyses of igneous rocks are analysed. *Lithos*, **12**, 33–9

— (1986) The role of spurious correlation in the development of a komatiite alteration model. *J. Geophys. Res.* **91**, E275–80.

— and Woronow, A. (1986) Discrimination among tectonic settings using trace element abundances of basalts. *Ibid.* **91**, B10289–300.

Chayes, F. (1960) On correlation between variables with a constant sum. *Ibid.* **65**, 4185–93.

Helz R. T. (1987) Diverse olivine types in lava in the 1959 eruption of Kilauea volcano and their bearing on eruption dynamics. In *Volcanism in Hawaii* (Decker R. W., Wright T. L. and Stauffer, P. H., eds) U.S. Geol. Surv. Prof. Pap. 1350, 691–722.

Kramers, J. D. and Ridley, J. R. (1989) Can Archaean granulites be direct crystallisation products from a sialic magma layer? *Geology*, **17**, 442–5.

Love K. M. and Woronow, A. (1988) Bediasite source materials: A solution to an endmember mixing problem exploiting closed data. *Geochim. Cosmochim. Acta*, **52**, 2085–90.

Meisch, A. T. (1969) The constant sum problem in geochemistry. In *Computer Applications in the Earth Sciences* (Merriam, D. F., ed.), 161–76.

Richter, D. H. and Moore, J. G. (1966) Petrology of the Kilauea Iki Lava Lake Hawaii U.S.G.S. *Prof. Pap.*, 537-B, 26 pp.

Skala, W. (1979) Some effects of the constant sum problem in geochemistry. *Chem. Geol.*, **27**, 1–9.

Woronow, A. (1987) The statistical analysis of compositional data by John Aitchison. Book review. *Math. Geol.*, **19**, 579–81.

— (1991) End-member unmixing of compositional data. *Geochim. Cosmochim. Acta.*, **55**, 2351–53.

— (in press) Analysis of mixtures. *J. Geol.*

[Manuscript received 25 June 1991;
revised 10 February 1992]