CRITICAL COMMENT

Comment on “Rare earth elements in stream waters from the Rokko granite area, Japan: Effect of weathering degree of watershed rocks” by Nakajima and Terakado, Geochemical Journal, 37, 181–198, 2003

T. P. HUTCHINSON *

Department of Psychology, Macquarie University, Sydney, N.S.W. 2109, Australia

(Received August 21, 2003; Accepted December 25, 2003)

Nakajima and Terakado (2003) measured rare earth elements in stream waters from eleven locations in the Rokko granite area. The present paper examines their data in the light of a model in which each sample arises from mixing two sources, followed by concentration or dilution. The concentrations of rare earth elements in the specimens will then follow particular equations, and a straight line will be seen if the data are plotted in a certain way. It is concluded that data from some, perhaps all, of the locations is compatible with this model.

Keywords: rare earth element, REE pattern, stream water, mixing model, water-rock interaction

INTRODUCTION

It is common for researchers to report the concentrations of each of the rare earth elements (REEs) in each of several samples. For example, Nakajima and Terakado (2003, table 2) measured REEs in samples of stream waters from eleven locations on the steep slopes of the Rokko Mountain granite area, near Kobe, Japan. Nakajima and Terakado made the point that this is a simple geologic system—the streams are short, without anthropogenic pollution or biological activity to distort the results, and thus the chemical compositions of the waters reflect the interaction between water and granite materials. A table of eleven locations by ten REEs is 110 numbers that we need to understand—is there any way of making this process easier? One purpose of this paper is to show that there is. A second purpose is to present results suggesting that at least five of the samples are ordered in a sequence, and to invite Nakajima and Terakado to comment on whether this has substantive geologic meaning.

MODEL

One thing that would be a great help is existence of a model for how the concentrations arise. A model will give direction to the process of data analysis—without a model, there is no standard against which to compare the data.

A possible model is that of mixing plus concentration or dilution (Monecke et al., 2000). Suppose there are two sources of REEs, that are mixed in different proportions for the different samples, and also there is concentration or dilution, that again is different for different samples. This model may be written as follows. Let $C_{ij}$ be the concentration of rare earth element $i$ in sample $j$. Let $S_i$ and $T_i$ be the concentrations in the two sources. Let $p_j$ be the proportion the first source contributes to sample $j$, and $1 - p_j$ be the proportion the second source contributes to this sample. After mixing, the concentrations are $p_j S_i + (1 - p_j) T_i$. A process of concentration or dilution means that all the concentrations are multiplied by a common factor; let this be $k$. Therefore, $C_{ij} = k_j [p_j S_i + (1 - p_j) T_i]$. (Notice that $S$ and $T$ depend upon the element but not upon the sample, and $p$ and $k$ depend upon the sample but not upon the element.)

METHOD OF PLOTTING

If we have data for three samples only, there is a way of plotting this type of data that demonstrates the success or failure of the model (Hutchinson, 2002). Calculate the ratios $C_{11}/C_{12}$ and $C_{22}/C_{12}$ for each of the rare earth elements. Plot $C_{11}/C_{12}$ against $C_{22}/C_{12}$. It is easy to demonstrate algebraically that the result is a straight line if the model of mixing plus concentration/dilution is valid. A special case is that if the degree of concentration or dilution is the same for all three samples, then the straight line goes through the point $(1, 1)$. 
A straight line arises whichever of the three samples is used as denominator in calculating the two ratios. In presenting a scatterplot, it is probably best to choose the one showing a line with negative slope—a positive slope may arise artificially from having the same divisor in both the plotted quantities. A negative slope occurs when the sample used as the denominator is the one for which \( p \) is intermediate between the \( p \)'s for the other two samples.

In the paper by Monecke et al. (2000), there were only three samples (fluorites), and thus the proposed method of graphical presentation was easily used (Hutchinson, 2002). Is it practicable to adapt the method for use when there are many sets of concentrations of REEs in the dataset? Using the eleven samples of Nakajima and Terakado (2003), I will show that it is.

**DATA PROCESSING**

*Which elements?*

It is sometimes found that La, Ce, and Eu behave differently from the other REEs—La has the lowest atomic number and is consequently sometimes unusual, Ce because of the stability of the divalent ion, and Eu because of the stability of the tetravalent ion. Should these elements be excluded from the data analysis, or not? In the analysis below, all three were excluded, since an extra data point that obeys a relationship has only a little positive impact, whereas an extra data point that destroys a relationship has such a big negative impact.

**RESULTS**

By way of summary, in some of the 495 scatterplots, a negatively-sloping straight line was clear, in many a positively-sloping straight line was clear (but, as already
mentioned, this could arise from the two quantities having a divisor in common), while in others there was no discernible relationship between the two ratios plotted. Figure 1 shows a negatively-sloping straight line, in Fig. 2 it is difficult to see any relationship, and Figs. 3 and 4 are two further examples.

In the detailed account below, the eleven samples will be referred to as 1, 2, 3, 4, 5, 5’, 6, 7, 8, 9, and 10—see the paper by Nakajima and Terakado for the key to the names and locations of these places.

Convincing negatively-sloped straight lines for five samples

From examining the scatterplots that involve samples 3, 4, 5’, 6, and 8, it can be concluded that for these, (a) the data are consistent with the model, and (b) in respect of $p$, the samples are in this order: 3 6 5’ 8 4. That is, for these five samples, a plot of $R_{(u,w)}$ against $R_{(v,w)}$ is (approximately) a straight line, whatever samples $u, v, w$ are, with a negative slope whenever $w$ is in between $u$ and $v$ in the ordered list given. For example, $R_{(3,5')}i$ against $R_{(4,5')}i$ is a negatively sloping straight line, see Fig. 1.

For four samples, $k$ is the same

For four of the samples—3, 6, 5’, and 8—the straight lines pass close to the point (1, 1), and thus the $k$’s for these samples are approximately the same.

Where do the other samples fit into the sequence?

It is less clear that the scatterplots involving the other samples are straight lines. Three examples are given in Figs. 2–4.

One possible interpretation is that the whole dataset is consistent with the model, with the samples being in approximately the order 3 6 5’ (10 8) (2 7 9 4 1) 5 in respect of $p$; here, brackets have been used to mean that the samples within them have approximately the same value of $p$. That is, a plot of $R_{(u,w)}$ against $R_{(v,w)}$ is (approximately) a straight line, whatever samples $u, v, w$ are, and this has a negative slope if $w$ is in between $u$ and $v$ in the ordered list given. On this interpretation, the reason that some pairs of ratios (e.g., $R_{(10,7)i}$ and $R_{(9,7)i}$, see Fig. 2) show poor straight lines is that the samples are close together in the ordered list given, that is, their $p$’s are similar.

Another possible interpretation is that the model is not valid for some or all of these other samples. Against this, and in favour of the idea that most or all of the samples fall into one ordered list, is the fact that many of the extra pairs of ratios now brought into consideration do show quite good straight lines. For example, in Fig. 3, the appearance of a negative relationship suggests that sample 10 falls somewhere between the two extremes of 3 and 4. And in Fig. 4, the negative relationship suggests that sample 8 falls between 5 and 6, and therefore 5 is to the right of 8 in the list.

Elements La, Ce, and Eu

La, Ce, and Eu were excluded because the special features of their chemistry may lead to different behaviour. Do they in fact differ from the general pattern established by the other elements? If these elements are added to the scatterplots that involve samples 3, 6, 5’, 8, and 4, Eu is seen to follow roughly the same relationship as the other elements, but La and Ce are obvious outliers on some scatterplots.
Trends with atomic number

Some ratios consistently increase or consistently decrease as atomic number increases, from La to Lu. For example, $R_{3,5'}$ increases. This is outside the scope of the model.

DISCUSSION

Statistical testing?

Faced with some relationships that appear clear (e.g., Fig. 1) and others that are doubtful (e.g., Fig. 2), it is tempting to seek a statistical test that will tell us whether a relationship exists or not. However, in my view, this is not appropriate. One reason is that testing of a hypothesis is only valid when it is specified before analysing the data. If the hypothesis was suggested by the data, it is not surprising if it is found that the data supports it! That is, the analysis described is in essence one of description and exploration of data, not one of hypothesis testing, and it would not be valid to attempt to add a hypothesis test to this.

Another reason is the absence of an appropriate measure of variability. This leads me to suggest a change in the process of collecting data. What is needed is two or more replicates that can be viewed as a random sample from all those that might have been collected at that location. (This implies that the researcher would be equally happy to have any one of them as representative of that particular location.) The present common practice of collecting and analysing only one example may lead us to think that analytic error is the only source of variability, which surely cannot be right.

The case against these results

The possibility should be mentioned that the straight lines found are an overinterpretation of the data. The clearest relationships (such as Fig. 1) appear when two of the REEs are excluded and ratios calculated from a set of only five of the samples. Possibly La, Ce, and Eu are not exceptions, they should not be excluded, the proper view of the scatterplots is that no relationships exist, and the conclusion should be that the model is incorrect.

Interpretation

Having proposed an order of the samples in respect of the respective proportions of two sources, it may be asked whether this makes any sense in terms of the geology. From the order, can Nakajima and Terakado suggest what the two sources are? Judging by what they say in their paper, four possibilities are granite, clay, minor minerals having high REE concentrations, and sea spray. And, is it reasonable to interpret the concentration/dilution process as dilution by rainwater?

I hope that Nakajima and Terakado can comment on the order, on whether the fundamental idea of mixing plus concentration or dilution is credible for all or some of the samples, and on why La and Ce depart from the general pattern. In view of the possibility that I have overinterpreted the data, geological plausibility would greatly increase the value of the results. If Nakajima and Terakado are sceptical about my suggestions, at least I have described a method of analysis that may be useful with other datasets.

Without reference to the geology, can any interpretation be given based merely on the numbers? If one source has a much higher concentration of an element than the other source does, and the k’s for the different samples are the same, the order of the samples in respect of the concentration of that element will be the same as the order of the proportions of that source. It has been proposed that four of the samples are in the following order in respect of their p’s, 3 6 5’ 8, and that they have similar k’s. Referring back to the data in Nakajima and Terakado, it is found that this is also the order, from high to low, of the samples in respect of their concentrations of all of the REEs (recall that La, Ce, and Eu have been excluded), and also of Na, F, and NO3. The major difference occurs for the heavy REEs: the concentrations of Er, Yb, and Lu are respectively 8, 13, and 17 times higher in sample 3 than in sample 8. Thus it seems that one of the sources has much higher concentrations of heavy REEs than the other, and that the samples are ordered 3 6 5’ 8 (from high to low) in respect of the proportion this source contributes to the samples.

REFERENCES

