Optimum Spacing for Soil Sample Traverses

By W. A. HODGSON,* Ph.D., D.I.C., A.M.I.M.M.

SYNOPSIS

In order to be able to map the variation in the concentration of elements in soils, it is necessary to sample at intervals that will give adequate coverage of the area and will allow determination of the rate of fluctuation of concentration with distance in any environment. This rate of fluctuation with distance can be referred to as the continuity of concentration.

The continuity of concentration in soils varies according to the chemical characteristics of the element being considered, the underlying rock type or orebody and in response to the dispersion processes that operated during soil formation. Within any environment there is a maximum distance between sampling points within which adjacent sample values are related. This distance, the critical sampling interval, can be determined by means of a correlogram.

Where the sampling grid has been chosen to provide an adequate coverage over unmineralized rock, adjacent samples being related, anomalies may be missed if the element continuity in the anomaly is less than in the background. Results show that a knowledge of the critical sampling interval of the environment being examined may be used to select a more efficient sampling pattern for a geochemical survey.

INTRODUCTION

An orientation survey is normally carried out prior to a geochemical soil survey in order to determine the conditions of sampling that will demonstrate most clearly the variation in concentration of certain elements within the soils. The indicator elements considered are those which relate to the type of mineralization being sought. The optimum depth of sampling and the size fraction of the soil taken for analysis may be selected following a statistical examination of the variation of element concentration with respect to these parameters, but the choice of sample spacing is commonly arbitrary.

Soil samples are usually taken on a rectangular grid, with the sampling lines directed across the regional strike. It may be argued that the spacing between lines, or between sample points, should be sufficiently close as to guarantee that anomalies exceeding a certain size will not be missed. However, it is not enough to consider the significance of anomalies only in terms of size, for the dispersion from an economic mineral deposit of considerable volume may have little surface expression. This may be due to its shape or to the thickness and characteristics of the soil cover. It follows that minor anomalies may be just as significant as larger ones, and there is a tendency to reduce the sample spacing in order to guarantee that the anomalies with little surface expression are located.

Alternatively, where a geochemical anomaly has been discovered, it is often obvious that it would have been located by a much more widely-spaced sampling grid which, while being equally effective, would be less expensive.

The purpose of this paper, therefore, is to demonstrate the procedure for selecting the optimum spacing, neither too large nor too small, of the sampling grid, prior to conducting a geochemical survey. The results have been obtained from data that have been made available from a mineral exploration programme.

APPROACH

Geochemical soil surveys are carried out to identify areas characterized by high concentrations of specific elements, which may indicate a mineral deposit at depth. The soil samples should be taken with two different purposes. During the orientation survey, random samples from soil over unmineralized rock are analyzed to define the statistics of the background population. From these statistics a threshold concentration is established on the basis of probability, with the intention that sample values above this limit will be considered as anomalous. Other samples are collected on a systematic grid to guarantee effective coverage of the area. The values from the sampling grid are then examined in terms of the threshold concentration established for the unmineralized rock.

The concentration of an element in soils is a function of the primary dispersion of that element within the underlying rock, together with the effects of the secondary dispersion factors that were operative during soil formation. Soil samples are taken from a single body, the soil, in which the element concentration represents a continuum. Therefore, the values obtained from closely-spaced samples must be related. Matheron has used the term 'regionalized variable' to describe a parameter having values related in this way (Blais, *et al*, 1968).

If soil samples are taken on a closely-spaced grid, successive samples along a line will provide values that are related in a manner analogous to a time series (Agterberg, 1965). A plan of the sampling grid will show a clustering of related sample values around the extremes of the background population. As the grid spacing is enlarged, the relation between adjacent sample values is decreased. For each geological environment there is a distance between sampling points beyond which sample values can be considered to be unrelated or random. This distance is the critical sampling interval (Agterberg, 1965) and can be determined by means of a correlogram. Samples to be used in establishing the statistics of the background population should be collected at a spacing greater than the critical sampling interval, whereas samples intended to provide coverage of an area should be collected at a spacing less than this interval.

Where a sample line traverses the soil above an orebody, the values for an indicator element from the soil samples along that line can be represented by the model in Fig. 1. Both the background and the anomalous populations in the model are characterized by their own specific mean and standard deviations. The differences between the two means (h in Fig. 1) may vary in different situations. In the model the range and the standard deviation of the anomalous population are smaller than those of the background popu-

*Institute of Mining Research, University of Rhodesia.

lation. This is not always so, for the concentration of an element within an anomalous population in soils could range up to ore grade. A further characteristic of the model is the difference in the rate of fluctuation of concentration with distance between sampling points. This characteristic can be observed in Fig. 1, as the more frequent fluctuation of values in the anomaly than in the background, for an equal sample spacing.

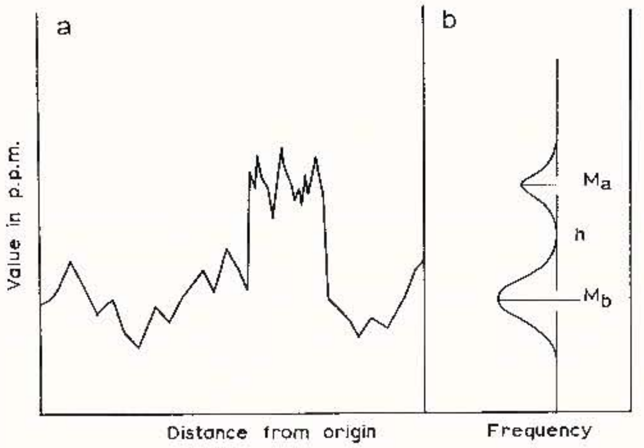

Fig. 1 (a). Diagramatic representation of the variation of element concentration along a soil sample line crossing an anomaly.

(b). Frequency distribution diagram of the element concentration in soil samples along the same line. The mean values of the anomalous and background populations are marked M_a and M_b, respectively, and h refers to the difference between the mean values for the two populations.

Three possibilities should be examined with respect to the model and to the difference in continuity of the element concentration in the anomaly and in the background population. These are:

(i) the continuity in the anomaly is the same as in the background,

(ii) the continuity is greater in the background than in the anomaly, and

(iii) the continuity is less in the background than in the anomaly.

The first two of these cases are illustrated diagramatically in Fig. 2. This diagram combines two concepts. The frequency distribution of the background and anomalous populations is shown in the lower part of the diagram while the spread of data from adjacent sample points is plotted with respect to the X and Y axes. The parallel lines in Fig. 2 represent probability limits to the spread of data from adjacent pairs of sample values, X and Y, separated by a constant sampling interval. The distance between the limits is related to the degree of continuity within the environment and to the degree of correlation between adjacent sample values. Although each diagram shows the spread of data in both the background and the anomaly, all sample pairs belong to either one or other of the populations.

The first example (Fig. 2a) is the model usually adopted in geochemical exploration. The continuity of the element concentration in the background and in the anomaly is not considered to be significantly different, $r_b = r_a$. The populations can be separated by examining the probability of values in the anomaly occurring as part of the background population. If the sample spacing is considerably smaller than the critical sampling interval, a plan of the data may place undue emphasis on the extremes of the background population, because of the regional clustering of related high values. Conversely, if the samples are spaced so widely that adjacent values are unrelated, these anomalous values will represent only a small portion of the total data and may be treated as extremes of the background population.

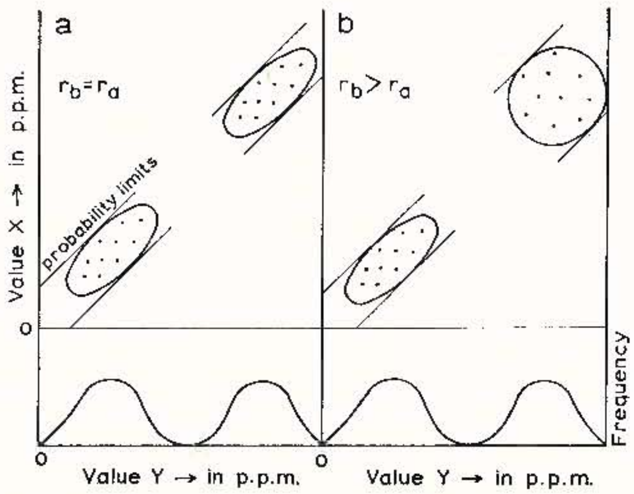

Fig. 2. Scattergrams showing the correlation of values from adjacent pairs of soil samples in the background and an anomalous population, for a given sample spacing. In (a) the correlation and element continuity between adjacent sample values are the same for the two populations, $r_a = r_b$. In (b) the correlation and element continuity are greater in the background than in the anomaly, $r_b > r_a$. A smaller sampling distance would be required in the anomaly to give the same correlation between adjacent values as in the background.

In the second example (Fig. 2b) the spread of data in the background is less than in the anomaly. The closer probability limits in the range of the background indicate a greater degree of correlation and a greater degree of continuity between adjacent sample values in the background, $r_b > r_a$. If the spacing for a sampling grid is based on a background that has good element continuity, then it is possible for an anomaly with lower element continuity to be small enough to be missed by the sampling grid, or at best to be recognized in a single sample value. Unless the grid spacing that is used is less than the critical sampling interval of the anomaly, such isolated values may well be considered as extreme members of the background population. It should be noted that the difference in spread of the data in the anomaly for Figs. 2a and 2b does not alter the form of the frequency distribution.

In the third case mentioned earlier the continuity of sample values is greater in the anomaly than in the background. Provided the sample spacing is less than the critical sampling interval in either population, any anomalous values will be recognized.

Clearly, the sampling interval for a regional survey should be based upon the critical sampling interval of the environment with the lowest continuity of element concentration, whether this occurs in the background or in the anomaly. Under these conditions the fluctuation in the environment with the lowest element concentration continuity can be noted. The smallest extent of any anomaly outlined is related to the critical sampling interval in the anomalous population, and the chance of missing a target of this size is related to the probability level upon which the critical sampling interval has been determined.

The continuity of element concentration values, or the correlation between adjacent sample values, can be studied by means of a correlogram (Agterberg, 1965). The correlogram in Fig. 3 illustrates the difference in element continuity within the background and anomalous populations of the model. In the correlogram the correlation coefficient r is calculated from the equation $r = (\Sigma xy)/[(\Sigma x^2)(\Sigma y^2)]^{\frac{1}{2}}$ for the values of a number, n, of randomly-selected pairs of samples, X and Y, separated by a given sample spacing that is plotted on the abscissa. In the equation, x and y represent the differences between each sample value and the population mean.

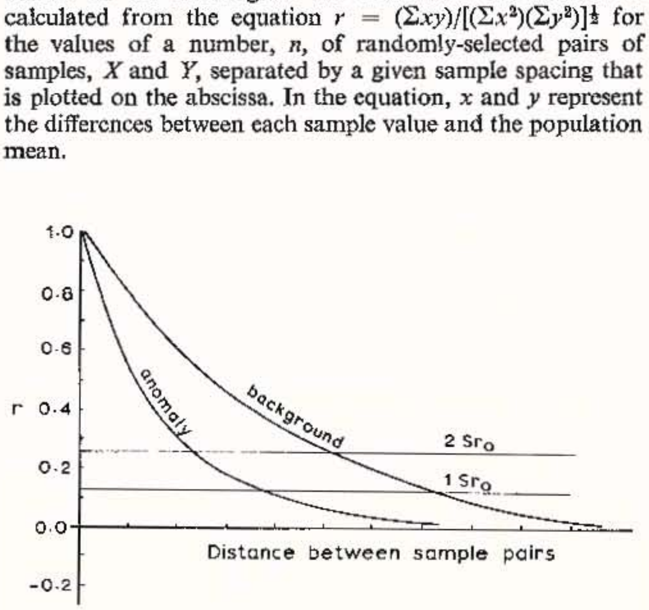

Fig. 3. Correlogram showing the change in value of the correlation coefficient r with increasing distance between adjacent sample points. The curves demonstrate the difference in element continuity for the background and the anomalous populations of the model shown in Figs. 1 and 2a. The horizontal lines represent probability limits for a correlation coefficient of zero from unrelated sample values.

Repeated values of r are determined for other sets of random pairs but at a different sampling interval. As the distance between the sample pairs is increased, the value of the correlation coefficient approaches zero asymptotically. This indicates the increasing probability of a random relationship between the values from sample pairs. The critical sampling interval is that distance between sample points at which the relationship between adjacent sample values can be considered to be random.

For a correlation coefficient of zero, indicating a random relationship, the standard error of the correlation coefficient is given by $S_o = (1-r^2)/(n-1)^{\frac{1}{2}}$ .Thus when the value of r falls within three standard errors of zero, the relation between sample values is small and may result from a random relationship. Values of r greater than three standard deviations from zero would result from random samples only three times in a thousand. The serial correlation coefficient (Agterberg, 1965) has not been used because the data are distributed in two dimensions and comparison is being made between random sample pairs rather than within a single sample line.

RESULTS

The results presented below were obtained from a geochemical survey and the continuity of element concentration has been examined using a program written for the ICL 1901A computer at the University of Rhodesia. The program is available from the Institute of Mining Research at the University. The flow sequence for the computer is as follows:

(i) Select an area characterized by a uniform geological and pedological environment in which the analytical data represent a single population.

(ii) Consider an initial spacing between a sample pair equal to that provided by the data.

(iii) Select 100 sample pairs at random and determine a correlation coefficient.

(iv) Increase the spacing between a sample pair and return to (iii).

(v) Repeat (iv) until the correlation coefficient is reduced to a value selected on the basis of probability and controlled by the standard error of the correlation coefficient for random samples.

(vi) Repeat the sequence (i) to (v) for a different lithology or for an anomaly.

Some of the results obtained, using 100 sample pairs, are shown in Fig. 4, and data for the means and standard deviations of the populations used, are listed in Table I. The correlation coefficient r is related to the sampling interval in units of 100 ft, and the standard error for a correlation coefficient of zero, with 100 samples, is taken as 0.1. There is a 99.7 per cent probability that a correlation coefficient, r, determined from random samples will fall within three standard errors of zero, that is, r will be less than or equal to $\pm$ 0.3. Conversely, there is a very low probability, 0.15 per cent, that a value of r greater than + 0.3 will be obtained from random samples. Therefore, if the value of r exceeds + 0.3 the sample pair can be considered to be related.

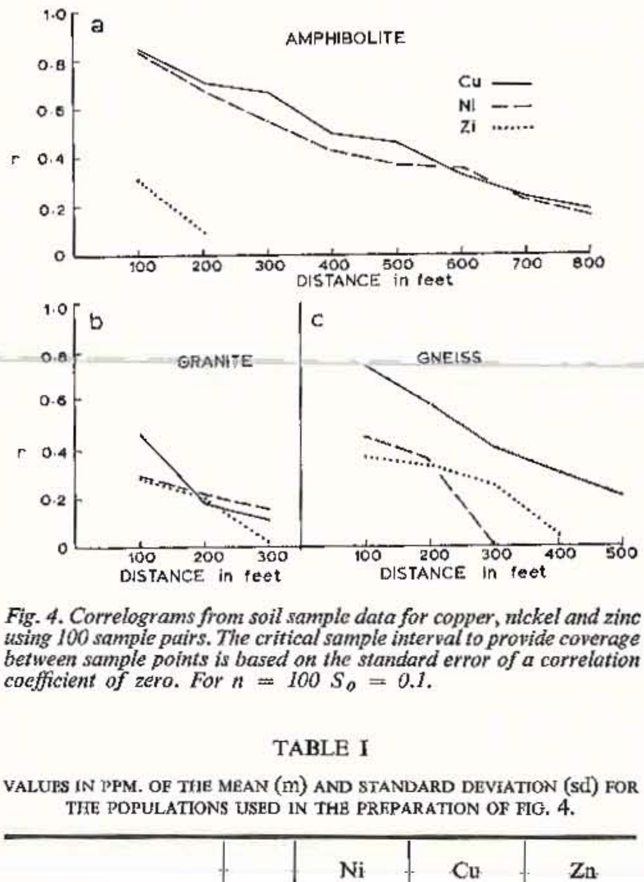

Fig. 4. Correlograms from soil sample data for copper, nickel and zinc using 100 sample pairs. The critical sample interval to provide coverage between sample points is based on the standard error of a correlation coefficient of zero. For n = 100 S_o = 0.1.

TABLE I

VALUES IN PPM. OF THE MEAN (m) AND STANDARD DEVIATION (sd) FOR THE POPULATIONS USED IN THE PREPARATION OF FIG. 4.

		Ni	Cu	Zn
Amphibolite . . .	m	374	253	42
	sd	317	220	10
Granite	m	23	30	32
	sd	10	23	13
Gneiss	m	48	40	39
	sd	19	17	10

The critical sampling interval, to the nearest 100 ft, is that which gives a value for r closest to, but greater than $+ 0.3$. Over amphibolite this interval is 600 ft for both copper and nickel and 100 ft for zinc. At this sampling interval the implication of the 0.15 per cent probability of getting a higher correlation coefficient from random samples is a 0.15 per cent probability of missing an anomaly which must be smaller in size than the critical sampling interval determined.

In granite, the critical sampling interval for copper is 100 ft, whereas samples for both nickel and zinc taken at this distance can be considered to be random. If a probability level of 95.5 per cent is accepted, based on two standard errors of a correlation coefficient of zero, then the critical sampling interval would be 200 ft for copper and zinc. However, there would be a 2.25 per cent probability that a correlation coefficient greater than 0.2 could be obtained where there was no correlation between adjacent samples. This is equivalent to a 2.25 per cent probability of missing an anomaly smaller than 200 ft.

Over gneiss, the critical sampling interval is 400 ft for copper and 200 ft for nickel and zinc. More accurate figures for the critical sampling interval could be obtained if a larger number of sample pairs and a smaller initial sampling interval were used.

DISCUSSION

The results show clearly the value of determining the critical sampling interval for each element during an orientation survey.

In the examples considered all the sample pairs were taken from within sampling lines directed across the regional strike. Since the dispersion pattern may be controlled by the geological structure and other features, it would be necessary to determine the critical sampling interval along the strike before establishing the spacing between lines.

It may not be convenient to determine the critical sampling interval during all orientation surveys. It would be useful, therefore, to accumulate data for the continuity of indicator elements in soils over different rock types under various conditions of chemical dispersion. In particular this applies to the element continuity in anomalous areas, where the amount of available data is usually small.

Sampling based on the concept of element continuity and the critical sampling interval takes into account the effects of secondary dispersion on different elements. From the data analyzed it appears that the sampling interval for zinc should be less than that for copper in the same geological environment.

The data in Table I show that some of the populations considered, particularly the values of copper and nickel in amphibolite, are not distributed normally but are probably lognormal. For these populations the correlation coefficients were recalculated after taking logarithms of the values to base ten. This had little effect on the value of the correlation coefficients and consequently the critical sampling interval remained unchanged.

A further related possibility is that the element continuity may vary within the range of a single population, for example, values varying more rapidly with distance at the upper extreme of the population. This effect could be influenced by the method of chemical analysis. The investigation of these questions remains to be carried out.

CONCLUSION

The results show that the use of the critical sampling interval to select the optimum sample spacing in geochemical surveys is justified. The concept of element continuity can be used to increase substantially the effectiveness of this technique in mineral exploration.

ACKNOWLEDGEMENTS

The writer acknowledges the assistance of the Anglo American Corporation of South Africa (Rhodesia) Limited who made available the data from some areas of geochemical exploration. Mr P. F. Ridler and Mrs A. McCabe, of the University of Rhodesia, assisted with the computer program. Thanks are also due to Mr P. Clarke, of the University of Rhodesia, for discussion concerning the statistical approach, and to Dr K. A. Viewing, Director of the Institute of Mining Research, for his support of this work and for his review of the manuscript.

REFERENCES

Blais, R. A. and Carlier, P. A. (1968). Applications of geostatistics in ore evaluation in 'Ore reserve estimation and grade control' *Canad. Inst. Min. Metall.* Special vol. 9, pp. 41-68.

Agterberg, F. P. (1965). The technique of serial correlation. *J. Geol.* vol. 73, pp. 142-154.