APPLICATION OF AUTOMATIC DATA PROCESSING TO PROBLEMS OF EXPLORATION GEOCHEMISTRY

E.R. BOSMAN, D. ECKHART and K. KUBIK[1])

INTRODUCTION

Automatic data processing is applied with greatest success in those cases where large sets of data have to be handled. This is frequently the case in geochemical exploration. At the I.T.C., therefore, close corporation exists between the Departments of Mining and of Mathematics, where projects of geochemical exploration are executed with the aid of electronic computers, specifically:

1) After the field measurements are coded for the electronic computer, maps may be prepared showing the locations of the sample points together with the measured metal concentrations. These maps may be drawn at various scales, so that they can be superimposed on geographical or geological maps.
2) Analytical expressions are computed for a trend surface through the data points, and the trend surface is visualized via its contour lines, which are drawn employing an automatic drawing table.
3) A map of the residuals, by which the trend surface fails to pass through the data points (residual map), is drawn along with the contour map.
4) The dependence of the measured element concentrations and of the residuals on geographical and geological factors is analysed by statistical techniques such as correlation computation, linear, multi-linear and polynomial regression analysis and others.

[1]) ITC, Delft, Netherlands.

A short outline of the techniques employed in the computing and drawing of the trend surface and the residual map is presented below.

THE COMPUTATION OF THE TREND SURFACE

A trend surface is a surface u(x,y), which expresses the "large" scale dominant aspects of the element concentration, Z, in the area under consideration.

By removing the trend or regional effect from the observed data Z, small relatively obscure features are highlighted. Such features are frequently useful for locating conditions of economic or scientific importance. The concept of trend is difficult to define. "Large" is, in this connection, a relative term, and what is "large" for one purpose may be "small" for another. We resolved this ambiguity by predefining the average quantity of residuals (Z-u) which should remain after the elimination of the trend.

Various possibilities exist with regard to the analytic function u(x,y). Power series and trigonometric series are often used for this purpose. The selected function must be flexible enough to follow the trend over extended distances, but smooth enough to avoid large local undulations. Polynomials and trigonometric functions however, do not entirely satisfy the first condition.

Consequently, our computations employ piecewise polynomial functions for the trend surface. A piecewise function is that function u(x,y), which is continuous and possesses continuous derivatives in the whole area under consideration, and which equals polynomials of low degree in small subregions of the

area. The trend surface is thus defined by separate polynomials for different subregions, which join smoothly with the polynomials of the adjacent subregions. For the present computations piecewise polynomials of degree 3 are used. These functions are very flexible and can describe the trend over extended areas.

The trend surface should model and simulate the physical situation and the process, by which in the large the trace metal was deposited in the region. A very simple model for the trend surface u(x,y) is, that u changes as little as possible along the coordinate directions x and y:

$$\iint_{\text{region}} \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \, d\,x\,dy \rightarrow \text{minimum.}$$

This condition ensures a smooth trend surface. For a different history of the process another appropriate model should be chosen.

The condition of the mathematical model is now combined with the requirement, that the trend surface passes close to the data points, by formulating the condition:

$$P * \sum_{\text{points}} (u(x,y) - Z)^2 + \iint_{\text{region}} \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \, d\,x\,dy \rightarrow \text{minimum.}$$

The first term in this expression symbolizes the sum of squares of the residuals in all the data points and is a measure for the goodness of fit of the trend surface. The symbol P is a weighting factor for this first term. From this expression the parameters of the piecewise polynomial function are computed; thereby all the separate polynomials in the individual subregions are determined simultaneously. No condition is imposed from the computational aspect on the location of the data points.

The magnitude of the weight factor is chosen empirically in such a way, that the average magnitude of the residuals (u-Z) is equal to its pregiven value. If consequently after one run of the computation the values (u-Z) are still too large, the factor P is changed and the computation is repeated with this new value for P. In this manner trend surfaces are computed, which follow the data points with a pregiven magnitude of residuals.

The selection of a realistic model for the trend surface is very important to get a reliable result, especially with few data points or with a sparce and irregular arrangement of measurements. In these cases the superiority of this approach of trend surface computation to the classical polynomial fitting becomes very evident. This approach enables for a given set of data points the trend surface to be computed more accurately, or, inversely, to compute trend surfaces with less data points.

After the surface u(x,y) has been expressed numerically, it may also be represented in a graphical form so that it can be more easily visualized. A contour map of the surface is therefore prepared with help of an automatic drawing table; thereby either the contour lines are drawn themselves or areas of equal concentration are shaded in different tones of grey.

If required, there may also be computed the values u at arbitrary points of the area, for instance at those points regularly arranged in a rectangular grid covering the area.

THE DRAWING OF THE RESIDUAL MAP

After the computation of the trend surface the residual map can be prepared, which graphically shows the magnitudes of the residuals (u-Z) at the locations of the data points. Contour maps of the residuals are prepared, if desired, by the same technique as outlined for the trend surface.

FINAL REMARKS

A number of projects of the Mining Department of I.T.C., have been processed with the methods of automatic data processing outlined above. The introduction of those methods reduced significantly the manual work and moreover eliminated the subjective judgement of the engineer who constructs the trend surface map.

Discussion

H.J. de Wijs asked if the investigations into the use of piecewise polynomial functions for the sake of trend surface analysis should not have been preceded

by statistical studies on frequency distributions.

D. Eckhardt agreed that knowledge of the elementary curves was a prerequisite and stated that this principle had indeed been followed. However, he preferred to allow S. Dijkstra to provide further particulars.

S. Dijkstra then explained that the overall frequency distribution of all measured Zn values in streamsediments of the N.E. Ardennes was multimodal and appeared to include values belonging to different populations. A subsequent inspection of the raw data maps disclosed a spatial relationship between the values of the different populations and the geographical positions of sections having different lithostratigraphical settings. The boundaries of the sections coincided with sharp discontinuities in the geographical distribution of hxZn values. In the presence of these discontinuities the use of polynomial functions for trend surface analysis did not appear to be recommendable. This was one of the reasons behind the decision to use piecewise polynomial functions instead.

O. Koefoed queried if residual maps should not contain residuals only.

D. Eckhardt replied that should such maps be requested, they could be constructed without difficulty by adding extra conditions to the computer programme.

The chairman wondered how far one could go with cross-correlation analysis, and whether this had been explored by the author.

D. Eckhardt indicated that the mathematical department of ITC, together with the department of mining exploration, had performed some promising experiments, using limited numbers of pre-selected factors. He pointed out that, while in theory one could examine any number of factors, in actuality the amount of programming involved sets certain practical limitations.