

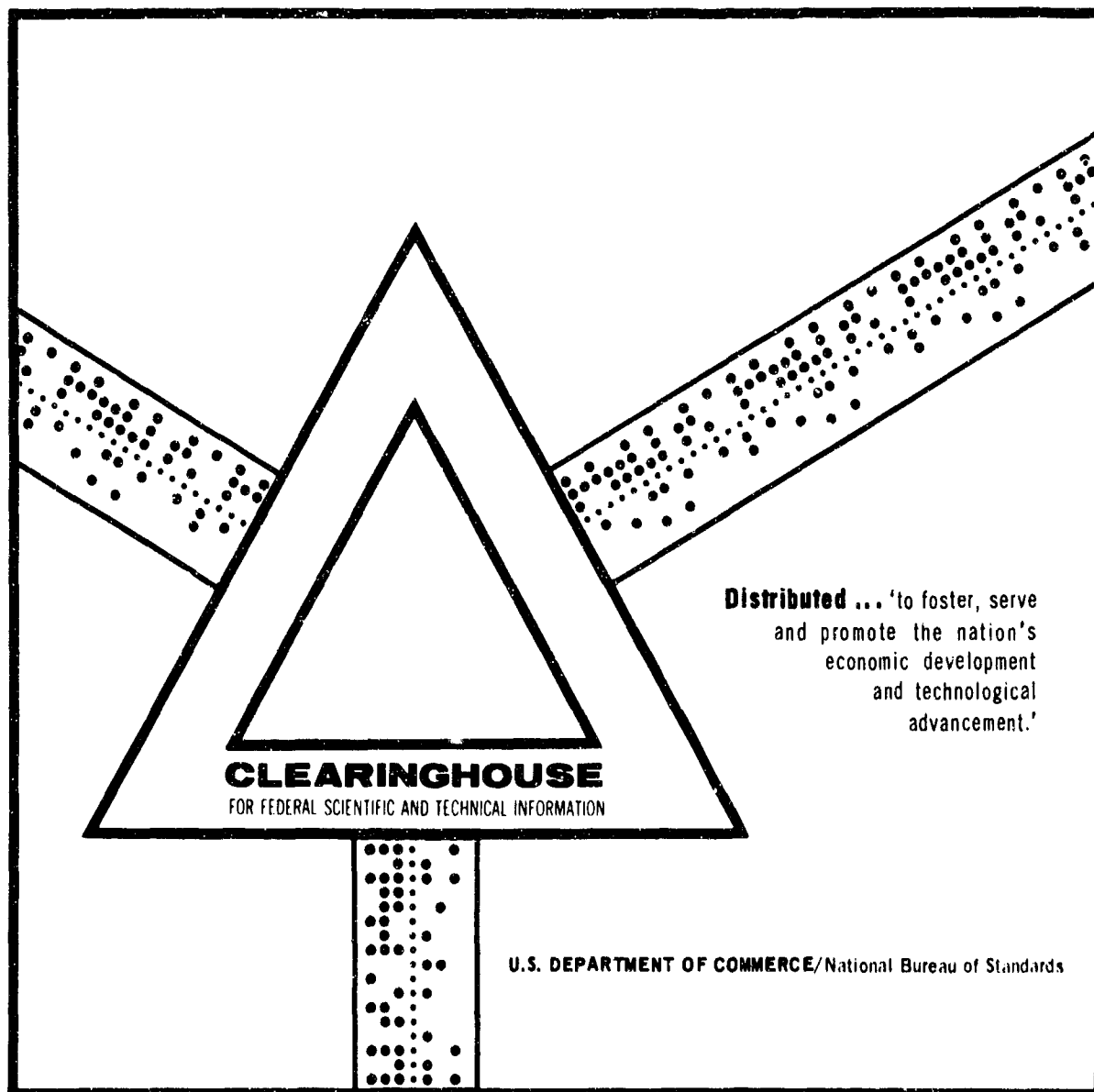
AD 699 163

TREND SURFACE ANALYSIS AND SPATIAL CORRELATION

Geoffrey S. Watson

Johns Hopkins University
Baltimore, Maryland

September 1969



This document has been approved for public release and sale.

AD699163

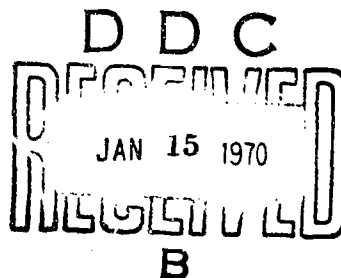
TREND SURFACE ANALYSIS AND SPATIAL CORRELATION[†]

by

Geoffrey S. Watson*

Technical Report No. 124

September, 1969



[†]To be read to "Quantitative Geology Symposium" at the Atlantic City Meeting of the American Geological Society, November, 1969.

*Research supported by the Office of Naval Research under contract NONR 4010(09) awarded to the Department of Statistics, The Johns Hopkins University. This paper in whole or in part may be reproduced for any purpose of the United States Government.

DEPARTMENT OF STATISTICS
THE JOHNS HOPKINS UNIVERSITY
BALTIMORE, MARYLAND

Reproduced by the
CLEARINGHOUSE
for Federal Agencies
under contract with the
Department of Commerce, Washington, D.C.

TREND SURFACE ANALYSIS AND SPATIAL CORRELATION

by

Geoffrey S. Watson*

1. Introduction.

Mapping has always been a favorite tool of geologists so that, with the increased accessibility of computers, it is scarcely surprising that geologists should now be very interested in studying spatial distributions mathematically. Their aim is to use some mathematical model to smooth or contour a net of data points, or to interpolate between them or to calculate some average over the area. (Two recent general references with substantial bibliographies are [1], [2].)

There are a number of possible methods, each with advocates, and some controversy has arisen. In the U.S., the most popular methods have assumed that

$$\text{value at data point} = \text{value of deterministic function} + \text{random error} \quad (1)$$

and consequent application of the least squares techniques. In South Africa and France [3], the preference seems to be for

$$\text{value at data point} = \text{value at a point on a random function} \quad (2)$$

and application of moving averages. In actual fact, these two models are not really distinct. It seems useful, without attempting to say what should be used, to try to explain the theoretical background of these methods, because this is not available in an elementary exposition. At the outset it should be said that this area requires considerable statistical research; it has so far been mainly the preserve of very theoretical workers. What little there is in the literature by way of applications has been written

* Research supported by the Office of Naval Research under contract NONR 4010(09) awarded to the Department of Statistics, The Johns Hopkins University. This paper in whole or in part may be reproduced for any purpose of the United States Government.

by geophysicists perhaps the earth sciences as a whole will help to create a new class of statistical methods (see some of the references in [1]).

2. Conceptual Problems.

Let us restrict discussion to the two dimensional problem. Thus suppose we have a quantity y measured at points (x_1, x_2) in the plane, distributed over some region R . Geometrically, we have points (with coordinates in three dimensions) (x_1, x_2, y) . Through this cloud of points we desire to put a smooth surface. It is hoped that this surface will show how y would vary with position (x_1, x_2) if there were no local or small scale variation the effect of large scale disturbances is supposed to be incorporated into the smooth surface, which may be called to trend surface.

Thus

$$\text{observable surface} = \text{trend surface} + \text{variation surface} \quad (3)$$

The separation into two parts is made on the vague basis of "large" and "small" scale variation. Our interest may be in either part. The trend surface seems appropriate for interpolation and averages, the second for the detection of anomalies, e.g., discrete ore bodies. It is clear that, unless we can introduce more prior knowledge of the particular phenomenon under study, or have a bright idea about the mechanism after a first look at the data, any analysis must be quite empirical. It follows that one cannot say any method is "best".

In section 3, we will discuss some specific mechanisms which might account realistically for some geological distributions. Here we will continue with a more general analysis of (3). If (3) is compared with (1), we notice that the two parts are assumed to be of a different nature. The trend is assumed to be a deterministic function and the variation to be random or stochastic. If (3) is compared with (2), there is no dissection but "random" appears again.

We therefore discuss first how randomness might enter. In the first place, y will usually be measured with error, i.e., if the site is revisited, sampled and the sample measured, a different value would be obtained. This is well understood and so can be ignored for the moment. Suppose now that our interest is only in the region R that was sampled at a finite number of points - we will later consider the case where region R itself has been chosen because it is typical of many similar area, i.e., our interest is in the population from which R was somehow chosen. There are infinitely many hypothetical surfaces passing near a finite number of points. Some will be more credible geologically than others. It is a standard tactic in science to deal statistically with an infinite hypothetical populations. Explicitly, we must define the class of admissible surfaces and suppose our particular case is a randomly chosen member of the class. Apart then from measurement errors, it is in this way that "randomness" enters when attention is confined to one area.

The population of admissible surfaces may be defined in several ways, corresponding to (1) to (2). We will return to this in a moment.

If the region R itself has been chosen as typical of many other regions, even complete knowledge of the surface over R would be insufficient for conclusions about the other regions. It is necessary to define again a class of admissible surfaces and to regard the surface over R as a randomly drawn member of the class if we are to use statistical arguments.

To define a class of admissible surfaces, we are guided by scientific plausibility and mathematical and computational convenience. The older method is that of (1) and the commonest choices for the deterministic function

are (i) an algebraic polynomial in x_1 and x_2 (ii) a trigonometric polynomial, i.e., the first few terms of a two dimensional Fourier series. It is usually implicit in papers on this basis that the choice of function, the highest degree terms used, etc., should make the "random error" have a zero mean, be uncorrelated and have constant variance. For example, the most thorough study in the algebraic polynomial case - Fraser Grant [4] - makes this quite explicit. His methods may be repeated using trigonometric polynomials, or any other set of functions. In detail, they depend strongly on having observations on a regular grid,

However, the model (1) method does not have this requirement and so has appealed to geologists - but the "error" should be studied if the model is to be checked at all with reality.

The "errors" referred to in the last paragraph should be thought of as the heights of a random surface $e(x_1, x_2)$ above the observing points - because we might have used other observing points. If $P'=(x_1', x_2')$, $P''=(x_1'', x_2'')$ are any two observing points, the above assumptions read

$$\begin{aligned} E(e(x_1', x_2')) &= 0, & E(e^2(x_1', x_2')) &= \sigma^2, \\ E(e(x_1', x_2') e(x_1'', x_2'')) &= 0. \end{aligned} \tag{4}$$

Clearly the last assumption is unreasonable if P' and P'' are close together and reasonable if they are far apart - for this is the way we expect spatial correlation to behave. We are supposing that $e(x_1, x_2)$ describes small scale disturbances only. It would be preferable if we could do an analysis which replaced the last two assumptions of (4) by

$$E(e(P') e(P'')) = c(P', P'') \tag{5}$$

where the covariance function $c(P', P'')$ decreased as P'' moved away from P' .

If we assume that

$$c(P', P'') = c_1(P' - P'') \quad (6)$$

i.e., that $c(P', P'')$ depends only on the vector $(x_1'' - x_1', x_2'' - x_2')$ the error function is said to be a homogeneous random function. If we assume more, namely that

$$c(P', P'') = c_2(r) \quad (7)$$

where

$$r = \sqrt{(x_1' - x_1'')^2 + (x_2' - x_2'')^2}$$

the error function is an isotropic random function, because the covariance now depends on the length but not the direction of the vector separating P' and P'' . In neither case does the average depend on the absolute position of P' and P'' . These are the generalizations to two dimensions of the notion of stationarity, now common in time series analysis - which is the one dimensional equivalent of our problem.

The spectral methods of time series analysis are now familiar, particularly to geologists who have followed computer applications of the Kansas group. The weakness of the model (1) work has come from ignoring the two dimensional time series aspect of the error model implied in model (1).

Having come this far, one sees that it may be unnatural to separate the observable surface into a deterministic and a random function. For the small and large scale disturbance notions find a natural definition in spectral terms - high and low frequencies. And there is no more an unnatural

division between the two. It may be better to define the class of admissible functions in terms of averages over the class, to assume homogeneity or isotropy and to base the analysis on this assumption. We cannot, in this short paper, explain how the analysis would then go, except to note that moving averages appear naturally. It is not too different from the trigonometric equivalent of Fraser Grant's method. Thus it is very awkward without regularly spaced data. Also the general weak assumptions made imply that one should have a considerable number of data points. This is one form of model (2). We hope to deal with methods based on this model in detail at a later date. The reader will find a list of relevant references at the end of the paper [6].

Matheron [5] advocates another version of model (2). He does not assume that the averages are independent of the absolute position but that the averages of changes are. Explicitly, he considers random surfaces $f(x_1, x_2)$ such that, for all $(x_1, x_2), (h_1, h_2)$

$$\frac{1}{2} E\{[f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2)]^2\} = \gamma(h)$$

where

(8)

$$h = \sqrt{h_1^2 + h_2^2}$$

This is not the same as saying that

$$s(x_1, x_2) = f(x_1, x_2) - f(0, 0) \quad (9)$$

is an isotropic random function because (8) does not allow the calculation of $c(P', P'')$ when P' and P'' are distinct. The extensive literature [5]

generated by Matheron and his co-workers, particularly Serra has only recently come to our attention. At the time of writing it is not clear how assumption (8), even with a particular choice for the function $\gamma(h)$, is made to support a method of analysis. Matheron however endorses the use of moving average methods which he attributes to Krige. In fact he uses "Kriging" to describe the act of taking moving averages. Matheron claims his methods do not require regularly spaced data points.

To conclude this section, we return to measurement errors. If they are present, one must add to the random surfaces discussed above a random surface which is isotropic and in which the spatial correlation falls off very rapidly, i.e., the equivalent to the white-noise of time series analysis.

3. Some mechanistic models.

There are two relatively simple ways of building probabilistic mechanisms that may have application to trend surface work, i.e., generate classes of random functions of geological pertinence. The key paper, with a large bibliography, is Whittle [6]. The tract by Matérn [7] is also a basic course.

An example of the first method would be the following construction. Suppose ore bodies are initially distributed randomly in the plane with different sizes and that, after a given time, their "influence" is felt at neighboring points. For example, their contents may diffuse outwards from their initial positions which we might approximate by points. If the medium is homogenous, $c(w,r)$ could stand for the concentration at a point, a vector \underline{r} away from a body of mass w . If masses w_i fell at positions \underline{R}_i then a measurement at \underline{x} , $y(\underline{x})$, free of measurement error, will yield

$$y(\underline{x}) = \sum_i c(w_i, \underline{x} - \underline{R}_i) \quad (10)$$

since for the i th mass, $\underline{r} = \underline{x} - \underline{R}_i$. The sum in (10) is over all the masses. If the masses are distributed over the whole plane by a Poisson process so that the mean number in area A is λA and if the probability of a mass lying in $(w, w, + dw)$ is $f(w) dw$ independently of its position, the stochastic properties of (10) may be worked out easily. In particular, the process is homogeneous. If $f(w, \underline{r})$ depends only on the length of \underline{r} , it is also isotropic. If it is relevant, then interesting problems arise because now the data should enable us to estimate λ and $f(w)$.

This model could also be used for linear reefs - one would then distribute lines rather than points. Another generalization would arise if the Poisson distribution is inadequate and has to be replaced.

The statistical treatment of data on these models is in its infancy. Models of type (10) are called moving average representations.

Another class of models is suggested by classical mathematical physics in which the mathematical description usually has the form

$$\left. \begin{aligned} Ly(\underline{x}) &= 0, & \underline{x} \text{ inside } \underline{R}, \\ y(\underline{x}) &= e(\underline{x}), & \underline{x} \text{ on } \underline{S} \end{aligned} \right\} \quad (11)$$

where \underline{S} is the boundary of \underline{R} and where L is a linear operator. For example, if $y =$ temperature at equilibrium, $L = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$. In usual practical cases, the temperature distribution over the boundary determines the interior temperature distribution. Thus

$$y(\underline{x}) = \int_S b(\underline{x}, \underline{x}') e(\underline{x}') d\underline{x}' \quad (12)$$

where $b(x, x')$ is related to the Green's function of the problem and is determined uniquely by L and B . Thus if $a(x')$ is a random function on B , $y(x)$ is a random function in the interior. It will be seen that (12) is again a moving average representation like (10). It is merely obtained in a different way. Fluid flow is described similarly so it seems this approach should be relevant for some problems.

A much more difficult situation may be relevant to many geological problems. Above we have been talking about boundary value problems in an isotropic medium but with random boundary values. The converse situation -- a medium whose properties vary randomly and simple boundary conditions -- is much harder to formulate and deal with.

b. References.

- [1] Harbaugh, J. W. and Merriam, D. F. (1968). Computer applications in stratigraphic analysis, John Wiley and Sons, New York.
- [2] Merriam, D. F. and Cooke, Man C. (1967). Computer applications in the Earth Sciences: Colloquium on trend analysis, Computer Contribution 12, State Geological Survey, The University of Kansas, Lawrence, Kansas.
- [3] Matheron, G. (1967). Kriging or polynomial interpolation procedures, The Canadian Mining and Metallurgical Bulletin, LXX, pp. 240-244.
- [4] Grant, F. (1957). A problem in the analysis of geophysical data Geophysics, XXII-2, pp. 309-314.
- [5] Matheron, G. (1965). Les variables régionalisées et leur estimation, 306 pages; Masson, Paris.
- Matheron, G. (1967). Eléments pour une théorie des milieux poreux, 107 pages; Masson, Paris.
- Matheron, G. (1968). Osnovy Prikladnoi Geostatistiki, 408 pages; Editions Mir, Moscou.
- Matheron, G. (1966). Structure et composition des perméabilités, Revue de l'Institut Français du Pétrole XXI-4, 564-582.
- Matheron, G. (1966). Genèse et signification énergétique de la loi de Darcy, Revue de l'Institut Français du Pétrole, XXI-11, 1697-1706.
- Matheron, G. (1967). Composition des perméabilités en milieu poreux hétérogène: Méthode de Schwydler et règles de pondération; Revue de l'Institut Français du Pétrole, XXII-3 443-466.
- Matheron, G. (1968). Composition des perméabilités en milieu poreux hétérogène: Critique de la règle de pondération géométrique, Revue de l'Institut Français du Pétrole, XXIII-2, 201-218.
- Matheron, G. (1966). Comparaison entre les échantillonnages à poids constant et à effectifs constants, Revue de l'Industrie Minière, VIII, 609-621.
- Serra, J. (1967). Echantillonnage et estimation locale des phénomènes miniers de transition (Thèse), 670 pages, Nancy.
- Serra, J. (1968). Les structures gigognes: Morphologie mathématique et interprétation métallogénique, Mineral. Deposita, 3, 135-154.
- Serra, J. (1968). Morphologie mathématique et genèse des concrétions carbonatées des minerais de fer de Lorraine, Sedimentology, 10, 183-208.

4. References (Continued).

- Serra, J. (1967). But et réalisation de l'analyseur de textures, Revue de l'Industrie Minérale, 49 (9), 14 pages.
- Hass, A., Matheron, G., et Serra, J. (1967). Morphologie mathématique et granulométrie en place; Annales des Mines XI, 735-753, et XII, 768-782.
- [6] Whittle, P. (1963). Stochastic processes in several dimensions, Bulletin of the 34th Session of I.S.I., pp. 974-985.
- [7] Matérn, B. (1960). Spatial variation, Medd. Skogsforskn Inst., 49, No. 5, 144 pages.

Appendix.

Some Comments on
"Les variables regionalisees et leur estimation"
Manson et Cie, 1965, by G. Matheron.

There have been many publications by Matheron and his colleagues. The central reference is the above book of 305 pages of mathematics. This book is not easy reading and has no point of contact with the usual literature - in fact, it has no references at all, except to text books on stochastic processes. It seems worthwhile therefore to attempt a survey of its contents. Since it develops a great deal of formalism, we try only to describe in our own words and notation the essential problem Matheron tackles and his methods - also we give references to key sources in the English statistical literature. This may understate the range and depth of the book. If so, we hope it provoked someone else to translate his methods at greater length.

Let $f(\underline{x})$ be a function of $\underline{x} = (x_1, x_2, \dots, x_n)$, a point in n dimensional Euclidean space. We would like to estimate

$$Q = \int f(\underline{x}) d\underline{x}, \quad (1)$$

where $d\underline{x} = dx_1 \dots dx_n$, given some information on $f(\underline{x})$. To fix ideas, $f(\underline{x})$ might be the density of a metal at a point \underline{x} in $n = 3$ space. Then Q is the total quantity of metal; the integral makes sense because $f(\underline{x})$ is usually zero outside some region R . One other example is important. Suppose $k_R(\underline{x})$ is zero for \underline{x} outside R and unity for \underline{x} in R . Then

$$V = \int k_R(\underline{x}) d\underline{x} \quad (2)$$

is the volume of the region R . Again, given information on the values of

$k_R(x)$, we might wish to estimate V .

To evaluate (1) when $f(x) = 0$ outside a region R , and when the functional form of $f(x)$ is given in a matter of integration - numerical integration if the form of $f(x)$ is complicated. One divides R into M subregions R_i with volumes v_i , evaluates $f(x)$ at a typical point x_i in each R_i and forms

$$B = \sum_{i=1}^M f(x_i) v_i \quad (4)$$

Faulser methods approximate f locally by simpler functions, but the final formula will have a form like (3). The error made is simply Q-B, and books on numerical integration suggest ways of estimating this using one's knowledge of the formula for $f(x)$. This is all classical mathematics. If nothing is known about $f(x)$ except its values at x_1, \dots, x_M , nothing can be said about the error Q-B.

Two examples show an interesting result (Aitken (1949), Yates (1948), Kendall (1942)). In one dimension the use of points equally spaced h units apart reveals the following facts:

$$\sum_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(n+nh)^2} h \text{ estimates } \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(x+0)^2}}{\sqrt{2\pi}} \cdot x = 1 \quad (4)$$

very well even for h as large as two σ while

$$\sum_{-\infty}^{\infty} \frac{h}{\pi(1-(n+nh)^2)} \text{ estimates } \int_{-\infty}^{\infty} \frac{dx}{\pi(1+(x+0)^2)} = 1 \quad (5)$$

much more poorly for similar values of h . As Kendall showed, the reason for the difference in the behavior is that the characteristic function (Fourier transform), $e^{-|t|}$ of $(1+x^2)^{-1}$ decreases at a slower rate than the characteristic function, $e^{-t^2/2}$, of $e^{-x^2/2}$ as $|t| \rightarrow \infty$.

Survey sampling is now an essential part of statistical theory and practice. In particular, the methods of stratified and systematic sampling have been widely discussed. These methods are usually set up for a finite population of N objects or units u_1, u_2, \dots, u_N . Some character of unit u_j is measured by y_j . The usual estimate $\frac{1}{N} \sum_{j=1}^N y_j$ from a sample of the units. If the units can be grouped together into strata so that the within strata variation is less than the between strata variation, selecting some of the units at random from each stratum (stratified sampling) has merit. Often the objects have a natural order (say that of their subscripts) and it is easier to take every k th unit. If the first unit is drawn at random from u_1, u_2, \dots, u_k , we have a systematic sampling procedure. A comparison of these methods and the relation of the latter to harmonic analysis is known - see Cochran (1966), Hansen (1967).

Borish (1966) reexamined the problems and methods of Kendall by choosing θ at random (see (6)) in the interval $(0, 1)$, i.e., he considered interval integration by systematic sampling. Thus, generally in one dimension, $Q = \int_{-\infty}^{\infty} f(x) dx$ is approximated by

$$S(\theta) = h \sum_{-\infty}^{\infty} f(\theta + nh), \quad (6)$$

a random variable since θ is a random variable. Immediately

$$\begin{aligned} E(S(\theta)) &= \int_0^1 \frac{d\theta}{h} h \sum_{-\infty}^{\infty} f(\theta + nh), \\ &= \sum_{-\infty}^{\infty} \int_{nh}^{(n+1)h} f(x) dx, \\ &= \int_{-\infty}^{\infty} f(x) dx, \end{aligned} \quad (7)$$

so that $B(u)$ is an unbiased estimator of Q . Also

$$\text{var}[B(u)] = E(B^2(u)) - [E(B(u))]^2$$

where, after several manipulations,

$$E[B^2(u)] = h^{-2} \int_{-\infty}^{\infty} f(x) f(x+uh) dx, \quad (8)$$

and

$$\text{var}[B(u)] = \sum_m \left[E\{f^{(m)}\} \right]^2, \quad (9)$$

where the sum in (9) is from $-\infty$ to ∞ excluding $m = 0$, and

$$\phi(t) = \int_{-\infty}^{\infty} e^{ixt} f(x) dx \quad (10)$$

is the Fourier transform of $f(x)$. The error in $B(u)$ will be smaller on the average the smaller (9) is. For fixed h , this means $|\phi(t)|$ tending to zero faster as $|t| \rightarrow \infty$. The behavior of (4) and (5) can now be predicted. If θ is not taken as a random variable, Kendall showed how to get similar results by noticing that $B(\theta)$ and the error $Q-B(\theta)$ are periodic functions of θ , the period being h , and so representable as Fourier Series

Finally, the variance of $B(b)$ would also be known if we knew

$$\mu(y) = \int_{-\infty}^{\infty} f(x) f(x+y) dx \quad (11)$$

for every y - because then we could evaluate (8). Matheron calls (11) the covariogram of $f(x)$. He has given the n -dimensional version of the above results the only changes required are notational. He also defines the semi-varlogram

$$\gamma(y) = \frac{1}{2} \int_{-\infty}^{\infty} [f(x+v) - f(x)]^2 dx \quad (12)$$

Thus

$$\gamma(y) = \frac{1}{2} \int_{-\infty}^{\infty} (f^2(x+y) - 2f(x+y)f(x) + f^2(x)) dx, \quad (13)$$

$$\mu(0) = \mu(y) \quad .$$

However there are functions for which (11) is infinite but for which (12) is not - this is the reason why Matheron prefers (12). We will return later to these questions and to the key question of choosing a form for $g(y)$ or $\gamma(y)$ in practice.

Sampling methods for solving mathematical problems (we have just seen an example) are now called Monte Carlo methods and are very well known - see, e.g., the book byammersley and Handscombe (1964). They all depend on statistical theory. The above method is by no means the only, or the best, way of evaluating integrals by sampling.

Certain geometrical problems are closely related to the above. For example Kendall (1948) considered estimating the area of a region R by counting the number of lattice points which fall in it when the lattice is positioned at random. The theory is as above using the function $k_R(x)$ defined for equation (2). Such problems are of great mathematical interest and can be traced through books and papers on Integral Geometry or Geometrical Probability - see e.g., Kendall and Moran (1963), Moran (1969). Matheron calls

$$K(h) = \int k_R(x) k_R(x+h) dx \quad (14)$$

the geometric covariogram.

So far we have here regarded $f(x)$ as an ordinary function - the only stochastic element has been the sampling process. But workers in many fields including geology have considered the case where $f(x)$ is a random function with certain properties - the reason why and the basic references have been given in the body of the paper.

In this case even a complete knowledge of $f(x)$ so that (1) may be evaluated is not enough. The value of Q obtained is just one possible value out of an ensemble whose average, $E(Q)$, we seek. But

$$\begin{aligned} E(Q) &= \int E(f(x)) dx, \\ &= \int m(x) dx, \end{aligned} \quad (15)$$

where

$$E(f(x)) = m(x) \quad . \quad (16)$$

Since $\text{var}(Q) = E(Q^2) - E(Q)^2$, we need

$$E(Q^2) = E \iint f(x) f(x') dx dx' \quad .$$

But setting $x' = x + y$, we have

$$E(Q^2) = E \int \int (f(x) f(x+y)) dx dy, \quad (17)$$

and using (11),

$$E(Q^2) = \int E g(y) dy. \quad (18)$$

This should be compared with (8),

$$E(S^2(\theta)) = h \int_{-\infty}^{\infty} g(nh) . \quad (19)$$

The relationship is closer than one might have thought. In this approach we usually ask whether our understanding of the problem suggests a form to assume for, not $E(\rho(y))$ but,

$$\text{cov}(f(x), f(x+y)) = E(f(x) f(x+y)), \quad (20)$$

called the covariance function. Equivalently one might try to specify a new version of (12)

$$E(\gamma(y)) = \frac{1}{2} E [f(x+y) - f(x)]^2 \quad (21)$$

which, if known, implies the value of (20) when that exists.

To estimate $E(Q)$, given only $f(x)$ at x_1, \dots, x_n we might consider the estimator

$$P = \lambda_1 f(x_1) + \dots + \lambda_n f(x_n) \quad (22)$$

- AB -

where the λ_i 's are to be chosen to minimize

$$E(R - Q)^2 = \text{mean square error of } R \quad (23)$$

But

$$\begin{aligned} E(R - Q)^2 &= \iint (f(x) f(x')) dx dx' \\ &= 2 \sum \lambda_i \int E(f(x_i) f(x)) dx \\ &\quad + \sum \lambda_i \lambda_j E(f(x_i) f(x_j)) \end{aligned} \quad (24)$$

Hence

$$\begin{aligned} \frac{\partial E(R-Q)^2}{\partial \lambda_i} &= -2 \int E(f(x_i) f(x)) dx \\ &\quad + 2 \sum \lambda_j E(f(x_i) f(x_j)) \\ &= 0 \quad (i = 1, \dots, n) \end{aligned} \quad (25)$$

is a set of "normal" equations for the $\lambda_1, \dots, \lambda_n$. They may be solved if the covariance function (20) is known. Matheron calls this method Kriging.

There is some literature on how best to choose the points x_i in V , where the random function is observed so that

$$\bar{f} = \frac{1}{N} \sum_1^N f(x_i)$$

"best" approximates

$$\frac{1}{V} \int_V f(x) dx$$

Hajek (1959) has a result justifying equal spacing when V is an interval. Talbot, Hajek and Fabrzvekl (1960) have studied the planar case.

So far we have considered only the estimation of the integral (1) - this seems to be the only problem attacked in the book although the estimation of

$$I = \int p(x) f(x) dx \quad (26)$$

where $p(x)$ is known, is considered. If $p(x) = \delta(x-x_0)$, the Dirac delta function, $I = f(x_0)$. Thus in particular (26) raises the problem of estimating $f(\cdot)$ at unobserved points. This problem is considered in our first reference to Matheron's works. His approach seems to be the following. Suppose $f(\cdot)$ is observed at x_1, \dots, x_n and required at x_0 . Write

$$\hat{f}(x_0) = c_1 f(x_1) + \dots + c_n f(x_n) \quad (27)$$

and choose c_1, \dots, c_n so that

$$E(\hat{f}(x_0) - f(x_0))^2 = \text{minimum} \quad (28)$$

This least squares problem may be solved if one knows the covariance function (20). The coefficients are clearly dependent on x_0 , so that (27) is a kind of moving average. Matheron also calls this kriging.

All the above theory can be rewritten with $f(\cdot)$ vector valued and Matheron does this.

Thus all the above theory is quite simple and not novel. Its application is seen always to involve the choice of a function - the variogram, semi-variogram or covariance. It is difficult to see just what Matheron would recommend, and why! Should anyone be willing to accept the challenge to put

Matheron's work into the English literature, this is the point that needs most attention. It is clear that he favors functions which are isotropic. Even given this restriction, the estimation of the above functions from data is non-trivial and this does not seem to be attempted.

References.

- Yates, F. (1948). Philos. Trans. A 241 345-77.
- Aitkin, A. C. (1939). Statistical Mathematics Oliver and Boyd, Edinburgh.
- Kendall, D. G. (1942). Quart. J. Math. 13 172-84.
- Kendall, D. G. (1948). Quart. J. Math. 19 1-26.
- Moran, P. A. P. (1950). Proc. Camb. Phil. Soc. 46 111-115.
- Cochran, W. G. (1946). Ann. Math. Statist. 17 164-77.
- Hannan, E. J. (1962). Biometrika, 19, 281-3.
- Fisher, R. A. (1921) Philos. Trans. Roy. Soc. A 222 309.
- Kendall, M. G. and Stuart, A. (1955). Advanced Theory of Statistics, Vol. 1 Griffin and Co., London.
- Kendall, M. and Moran, P. A. P. (1963). Geometrical Probability, Griffin Monograph No. 10, London.
- Moran, P. A. P. (1969). Adv. App. Prob., Vol. 1, No. 1, 73-89.
- Dalenius, T., Hajek, J., Zubrzycki (1960). Proc. Fourth Berkeley Symposium. 125-50.

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Department of Statistics The Johns Hopkins University Baltimore, Maryland 21218		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE TREND SURFACE ANALYSIS AND SPATIAL CORRELATION			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report - September 1969			
5. AUTHOR(S) (First name, middle initial, last name) Watson, Geoffrey S.			
6. REPORT DATE September 11, 1969		7a. TOTAL NO. OF PAGES 10	7b. NO. OF REFS 7
8a. CONTRACT OR GRANT NO. NONR 4010(09)		9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 124	
b. PROJECT NO. NR 042-232		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. DISTRIBUTION STATEMENT			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Office of Naval Research Logistics and Mathematical Statistics Branch Washington, D. C. 20360	
13. ABSTRACT The methods for analysing data in one dimension (usually time) are highly developed. However, in several dimensions, most applied workers only use one class of methods -- that in which the data is assumed to be the sum of a deterministic trend and an uncorrelated random error. The more general model with a spatially correlated error has been neglected in most fields -- oceanography is a conspicuous exception. This neglect is partly due to the lack of expositions for the applied workers. However, the manner in which much geological data is now collected does make this application difficult. The aim of this paper is to explain the relevance of the possible models and methods and to indicate their data requirements. An appendix has been added on Matheron's work.			

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
trend surface analysis						
spatial correlation						
uncorrelated random error						