ORKUSTOFNUN
NATIONAL ENERGY AUTHORITY
GEOTHERMAL DIVISION

Jan A. Czubek

INSTITUTE OF NUCLEAR PHYSICS
ul. Radzikowskiego 152
31 - 342 Krakow, Poland.

INTRODUCTION TO GEOSTATISTICS

Lecture notes

Lectures presented by J. A. Czubek in March 1981 at Orkustofnun, Iceland, as a part of the International Atomic Energy Agency Experts Mission, Project ICE/8/003

OS81012/JHD08
Reykjavík, July 1981

ORKUSTOFNUN
NATIONAL ENERGY AUTHORITY
GEOTHERMAL DIVISION
GRENSÁSVEGUR 9,
108 REYKJAVÍK ICELAND

G V Johnsen

Jan A. Czubek

INSTITUTE OF NUCLEAR PHYSICS
ul. Radzikowskiego 152
31 - 342 Krakow, Poland.

INTRODUCTION TO GEOSTATISTICS

Lecture notes

Lectures presented by J. A. Czubek in March 1981 at Orkustofnun, Iceland, as a part of the International Atomic Energy Agency Experts Mission, Project ICE/8/003

OS81012/JHD08
Reykjavík, July 1981

FOREWORD

During three weeks in March 1981 I presented lectures on geostatistics to the staff of the National Energy Authority (Orkustofnun) in Reykjavík. Being in Iceland as an expert of the International Atomic Energy Agency in Vienna, these lectures have been a part of my duty here. The idea, however, to have lectures just on geostatistics as a part of the knowledge needed to geophysicists for analysis of field data is due to Dr. Valgardur Stefánsson of Orkustofnun, whose enthusiasm and preserving will to reach the goal pushed me to write the notes for my lectures. They have been written daily, as the sequence of the lectures was, and almost instantly typewritten and figures drawn by the technical staff of the Orkustofnun are appearing in the form as you have in your hands. This part of the job was a big and harassing work which was perfectly done. I am therefore deeply indebted to Ms. Sigrídur Valdimarsdóttir and Ms. Audur Ágústsdóttir for their big effort to typewrite the whole text and draw the figures. This enterprise could be accomplished due to the organizing effort of Dr. Valgardur Stefánsson who is highly acknowledged.

My hope is that this introduction to geostatistics will permit to somebody from the Orkustofnun people to have an easier access to the more advanced geostatistical literature which finally will result in a more sophisticated applications of this technique.

J.A. Czubek

Reykjavík, April 1, 1981.

TABLE OF CONTENTS

LECTURE 1

J.A. Czubek

LECTURE 1 Reykjavík, March 3, 1981

INTRODUCTION. WHAT MEANS GEOSTATISTICS ?

"Geostatistics" was introduced by G. MATHERON in France in the fifties of this century to describe some special approach to the statistical treatment of the geological data.

As geological data one can mention: porosity, bulk density, ore grade, permeability, bed thickness etc. All these phenomena can be characterized by the spatial distribution (one-, two-, or three dimensional) of measurable quantities. We are going to call these quantities the "regionalized variables". We denote any regionalized variable by z known at the point x , thus: z(x). Usually we know the values of z at several points x_i , thus we have a set of values $\{z(x_i)\}$

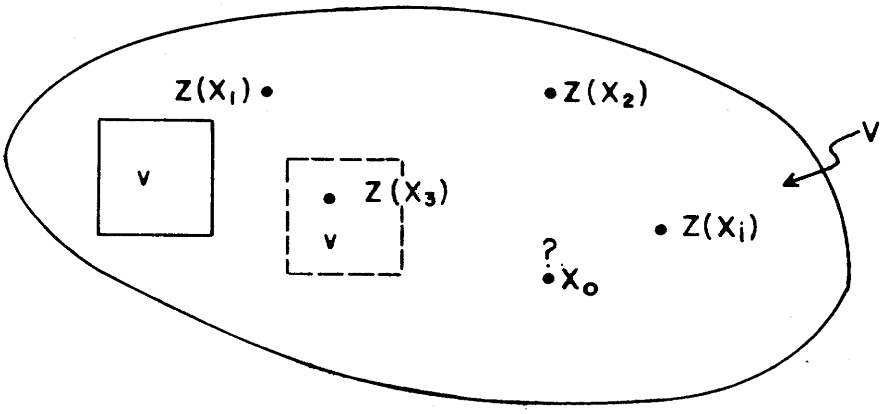


Fig. 1

connected with a given geological formation having the total volume V. We can treat the set $\{z(x_i)\}$ as a particular realization of some random variable Z(x):

$$Z(x)=\{Z(x_i),\ \forall x_i \in V\} \qquad (1)$$

which is nothing else than to state that Z(x) is a stochastic (in general three-dimensional) process.

Now, what are the problems of geostatistics? They are several. For example:

1. Can we predict the value Z at point x_o :

$$E\ [Z(x_o)]\quad ?$$

and what is the variance D of such estimation ?

2. Can we predict the average value of Z inside some panel (block) v, does not matter if some known, point like, value $Z(x_i)$ is or is not inside this volume, i.e.

$$E\ [Z_v(x)]\ ?\ ,$$

where x is a coordinate of a center of the block v. What is the variance D of such estimation ?

3. Can we determine the k variable probability distribution function

$$F_{x_1\ ,\ x_2 \ldots\ x_k}\ (z_1\ ,\ z_2\ , \ldots z_k) = \mathrm{Prob}\{Z(x_i) < z_1, \ldots\ , Z(x_k) < z_k\} \quad (2)$$

and to find its moments ?

These three problems presented above using the statistical language can now be translated into the geological language. For example, in problem No 1 we want to know an average porosity (permeability, density etc.) at a new drilled well (for a given formation) at the point x_o when this parameter is still known at other wells situated at the points $x_1\ , x_2\ \ldots\ x_i$. Doing the same operation but for the point x_o "walking around" the whole area we are arriving at the notion of the interpolation method for drawing the maps of isolines, which together with the second question for this problem will give the second map. This one of the precision with which the map of isolines is determined. The same procedure can be used also to treat different geophysical fields like the gravimetric or magnetic data, or meteorological data etc. Thus, it is, simply speaking, some interpolation procedure, based on the geostatistical principles. Just to give some more comprehensive examples:

Let us suppose some underground body B with linear form

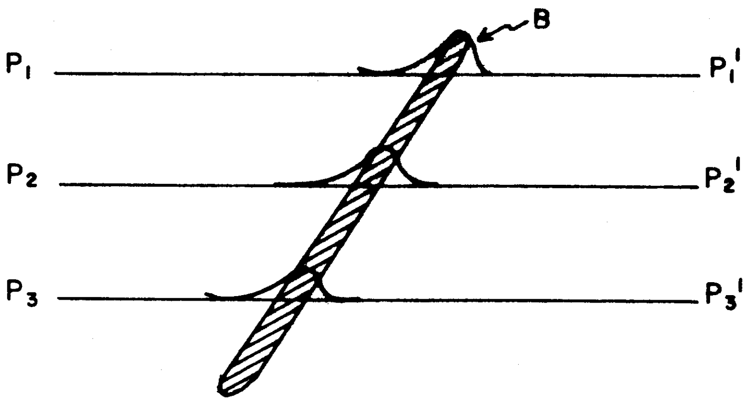


Fig. 2

We are flying along the profiles $P_1 P_1'$, $P_2 P_2'$ etc obtaining at some regions the anomalies above this unknown body, i.e. the net result of measurements is now:

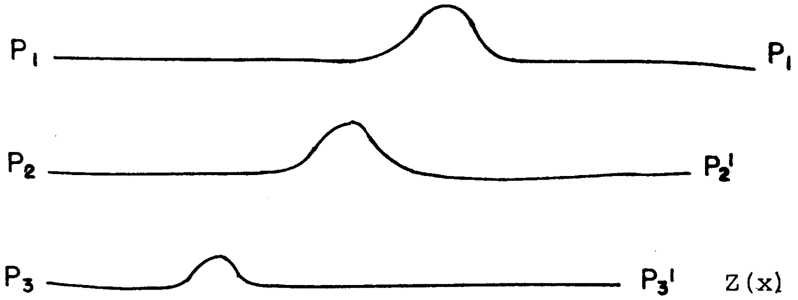


Fig. 3

Z(x) profiles

When the usual polynomial method of interpolation is now used (together with the least square method) to get the map of isolines, the result is:

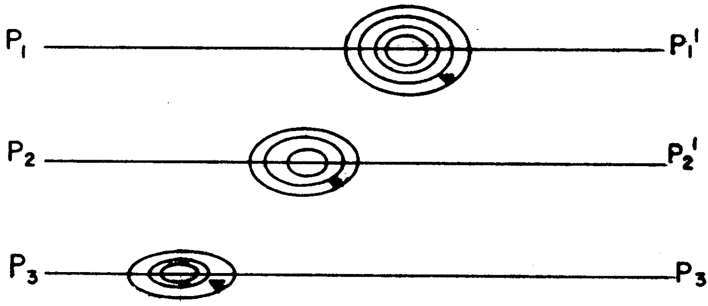


Fig. 4

Z(x) map

which gives three-nested structure. When the geostatistical method of interpolation is used (so called punctual Kriging), the resulting

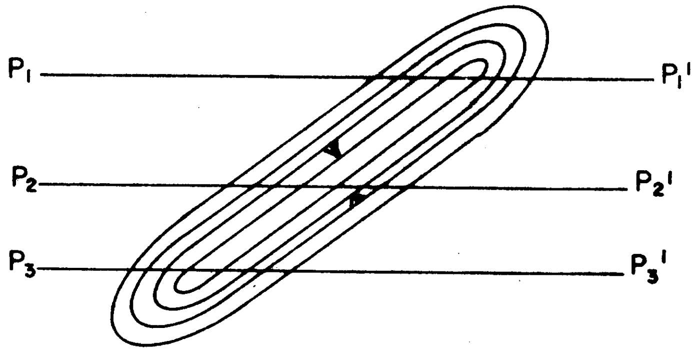


Fig. 5

Z(x) map

map of the parameter will be as shown in Fig. 5. The one-standard deviation map for the isoline map in Fig. 5 will be of the form shown in Fig. 6.

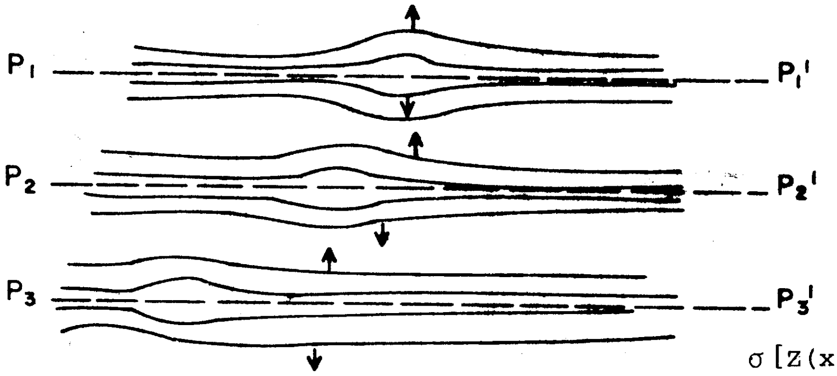


Fig. 6

$\sigma [Z(x)]$ map

A typical example for the problem No 2 can be as follows:

During the mining operation by blocks we want to know what will be the daily production (in tons of metal for example), and what is the precision of this estimate.

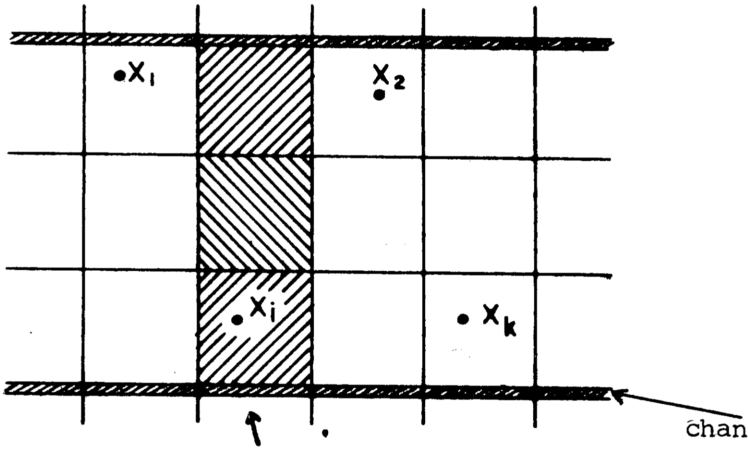
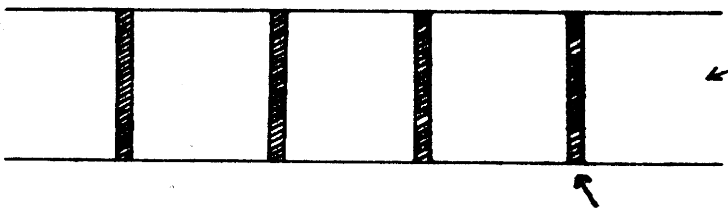


Fig. 7

blocks planned for exploitation

channel sample along a drift

The ore body is sampled either by drill-holes at the points x_1 , x_2 , x_i , x_k etc. or it can be channel samples along drifts, or vertical channel samples in a drift, as below

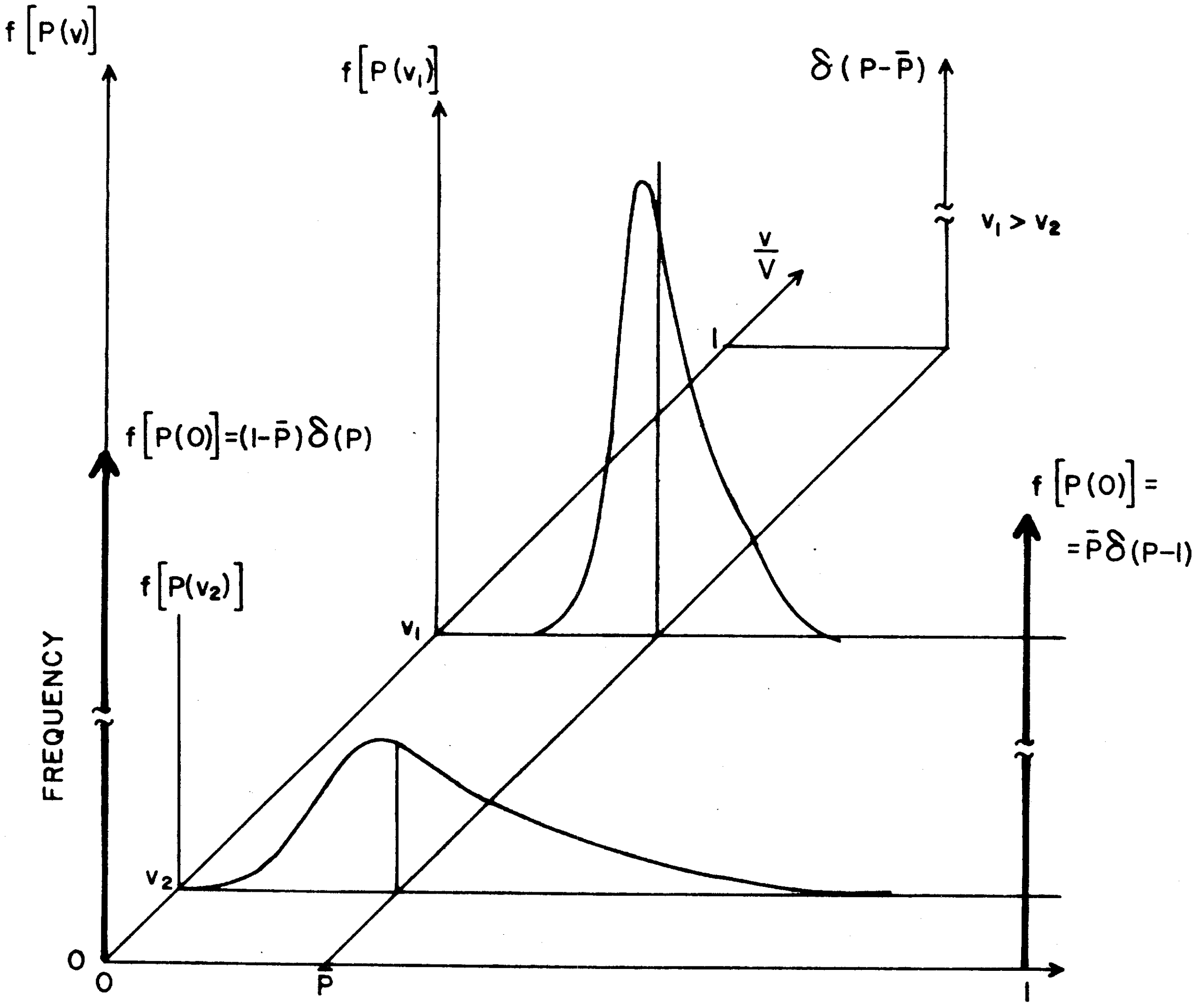


drift wall

Fig. 8

vertical channel sample on the drift wall

Another example for this type of problem is:
we know the ore grade (elemental content, porosity etc.) measured on chips taken from percussion borehole along its axis. Each sample contains a given mass (let say m_1) of the rock material. For the N_1 samples taken along the borehole axis we know the average value and the variance of the parameter being investigated. Now we use some geophysical method in this borehole to measure the same geological parameter. Will the result be the same as previously? Certainly not, because now the "geophysical sample" is much larger and contains more rock material (let say $m_2 > m_1$). Thus, how to "translate" the picture obtained at the laboratory on the borehole samples onto the picture observed in situ by the borehole methods?
To depict better this problem Fig. 9 was prepared.



ore grade (geological parameter of interest)

Fig. 9

This is a three-dimensional plot of the frequencies f of the geological parameter P measured on samples having the volume v (we call that procedure the regularization of the random variable P on the volume v). V is the total volume of a given geological formation. Let the parameter P being investigated be the grade of mineral (or porosity). We take first the samples having each the volume v = o (this is a theoretical example). Thus, for each particular sample we are either in the barren grain or in the mineral grain, and the frequency distribution will be the two Dirac delta functions:

$$f[P(o)] = (1-\bar{P})\delta(P)+\bar{P}\ \delta(P-1) \tag{3}$$

and because by definition it should be

$$\int_{-\infty}^{+\infty} f\ (P)\,dp = 1 \tag{4}$$

we really have

$$\int_{-\infty}^{+\infty} f[P(o)]dP = (1-\bar{P}) \int_{-\infty}^{+\infty} \delta\ (P)\,dP + \bar{P} \int_{-\infty}^{+\infty} \delta\ (P-1)\,dP = 1-\bar{P}+\bar{P} = 1 \tag{5}$$

Here $\bar{P}$ is the expected value of the parameter P in the whole deposit:

$$\bar{P} = E[P] = \int_{-\infty}^{+\infty} P\cdot f(P)\,dP \tag{6}$$

Now we start to sample the deposit taking the samples of volume $v_2>o$ each (for example 20 cm^2). We obtain in this case the frequency distribution

$$f[P(v_2)] \tag{7}$$

which is again normalized to 1 and has the same first moment $\bar{P}$. When the $\bar{P}$ value is small in comparison with 1 (but $\bar{P} > o$ by definition), the distribution (7) is usually asymmetrical. But what about the dispersion D ?

The dispersion D of the statistical distribution f(P) is, simply speaking, a second central moment, e.g.

$$D(P) = \int_{-\infty}^{+\infty} (P-\bar{P})^2\ f(P)\,dP \tag{8}$$

For the distribution given by E_2. (3) one has:

$$D[P(o)] = \int_{-\infty}^{+\infty} (P-\bar{P})^2 \cdot f[P(o)]dP =$$

$$= \int_{-\infty}^{+\infty} (P^2-2\,P\bar{P}+\bar{P}^2)\cdot[(1-\bar{P})\delta(P)+\bar{P}\;\delta(P-1)]dP =$$

$$=(1-\bar{P}) \int_{-\infty}^{+\infty} P^2\delta(P)dP + \bar{P} \int_{\infty}^{\infty} P^2\delta(P-1)dP-$$

$$- 2\bar{P}(1-\bar{P}) \int_{-\infty}^{+\infty} P\delta(P)dP - 2\bar{P}^2 \int_{-\infty}^{+\infty} P\cdot\delta(1-P)dP +$$

$$+ \bar{P}^2(1-\bar{P}) \int_{-\infty}^{+\infty} \delta(P)dP + \bar{P}^3 \int_{-\infty}^{+\infty} \delta(P-1)dP =$$

$$= 0 + \bar{P} - 0 - 2\bar{P}^2 + \bar{P}^2\;(1-\bar{P}) + \bar{P}^3$$

$$= \bar{P} - 2\bar{P}^2 + \bar{P}^2 - \bar{P}^3 + \bar{P}^3 = \bar{P} - \bar{P}^2 = D_o\;(P) \tag{9}$$

or

$$D_o(P) = \bar{P}\;(1-\bar{P}) \tag{10}$$

Now for the distribution $f[P(v_2)]$ let us assume that it follows well the beta distribution

$$\beta[P/p,q] \qquad 0<P<1\;,\quad p>0\;,\; q>0$$

defined as

$$\beta[P/p,q] = \frac{\Gamma\;(p+q)}{\Gamma(p)\;\Gamma\;(q)}\;P^{p-1}\;(1-P)^{q-1} \tag{11}$$

For this distribution the n-th moment is simply:

$$\int_o^1 P^n\;\beta[P/p,q]dP = \frac{\Gamma\;(p+n)}{\Gamma(p)} \cdot \frac{\Gamma(p+q)}{\Gamma(p+q+n)} \tag{12}$$

Thus the expected value is

$$\bar{P} = \frac{p}{p+q} \tag{13}$$

$$D(P) = \frac{p \cdot q}{(p+q)^2 (p+q+1)} \tag{14}$$

For $p>1$ and $q>1$ there is a unique mode at the point

$$P = \frac{p-1}{p+q-2} \tag{14}$$

(which is different from $\bar{P}$!). A lot of geological data fit well the beta distribution when the p and q parameters are well adjusted.

Now, if

$$f\ [P(v_2)] = \beta\ [P/p,q] \tag{15}$$

according to Eq. (14) and (13) we have

$$D\ [P(v_2)] = D_{v_2}\ (P) = \frac{p}{p+q}\ \frac{p+2-p}{(p+q)\ (p+q+1)} =$$

$$= \frac{\bar{P}\ (1-\bar{P})}{p+q+1} \tag{16}$$

and because, by definition

$$p+q+1>1 \tag{17}$$

by a simple comparison with Eq. (10) we have

$$D[P(v_2)] < D_o(P) \tag{18}$$

Thus, the bigger the volume v of the sample the smaller the variance of the samples.

Increasing further the volume v of the samples from v_2 to v_1 ($v_1 > v_2$), in the case when the statistical distribution fits again the beta distribution (which is not obvious at all !!!) it appears as an increase in the values at the p and q parameters (for the value $\bar{P}$ given in Eq. (13) remaining constant), which is equivalent to the diminishing again of the D $[P(v_1)]$ value in comparison with the D $[P(v_2)]$ value, thus

$$D\ [P(v_2)] > D\ [P(v_1)]$$

if (19)

$$v_2 < v_1$$

Finally we can take the whole deposit as a sample. In that case the volume of the sample is $v = V$, we have one unique value of $P = \bar{P}$ and the statistical distribution is

$$f\ [P(V)] = \delta\ (P-\bar{P}) \tag{20}$$

with

$$E(P) = \bar{P} = \int_{-\infty}^{+\infty} P\delta(P-\bar{P})\,dP = \bar{P} \tag{21}$$

and with the variance

$$D(P) = \int_{-\infty}^{+\infty} (P-\bar{P})^2\delta(P-\bar{P})\,dP = \int_{-\infty}^{+\infty} P^2\delta(P-\bar{P})\,dP - 2\bar{P}\int_{-\infty}^{+\infty} P\delta(P-\bar{P})\,dP +$$

$$+ \bar{P}^2\int_{-\infty}^{+\infty} \delta(P-\bar{P})\,dP = \bar{P}^2 - 2\bar{P}^2 + \bar{P}^2 = o \tag{22}$$

which is quite trivial because the variance of a constant $\bar{P}$ should be equal to zero !

The last case ($v = V$) is still not as theoretical as the first one (for $v=o$). We are really doing this when the whole deposit is still exploited - we know its total volume V and the total amount of ore extracted, thus, we know $\bar{P}$. But, of course, in this case the knowledge of these parameters does not help very much, <u>because the deposit is still exploited</u> !

By discussing the cases depicted in Fig. 9 we have obtained a firm feeling that the bigger the volume of the geological samples the smaller the dispersion of the geological parameter measured on them.
But how to translate them from one v value to another one? Is it possible to translate the dispersions only,or may be the whole frequency distributions $f[P(v)]$? The answer to these questions is just given by the solutions of the problems No 2 and 3 of geostatistics presented above.

The translation of the dispersions only is an easier task than the translation of the whole frequency distributions. This last case is solved by the so called non-linear methods of geostatistics. For this case an example below is given:

The point values (v very small) came from a porphyry copper deposit and were log-normally distributed with a mean of 1.19% and a variance of 0.20 $(\%)^2$. It was 67 200 points on 50 ft x 50 ft x 12 ft grid over a field measuring 2000 ft x 2000 ft x 504 ft. The frequency distribution is given in Fig. 10.

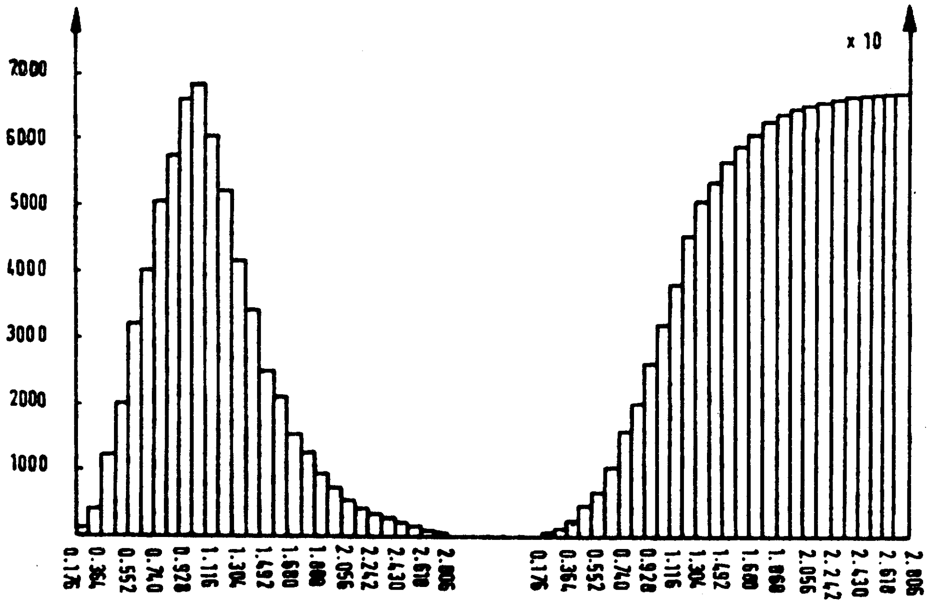


Fig. 10

RELATIVE CUMULATIVE
Frequency distribution of point samples

Next, using the so-called conditional simulation [DOWD and DAVID 1975] the new frequency distribution has been calculated, In this case for 2000 blocks each having the volume 100 ft x 100 ft x 100 ft. The result is given in Fig. 11 . The expected value for the blocks is now again 1.19% but the variance is 0.08 $(\%)^2$.

Knowing the frequency distribution of big blocks a lot of useful economical data for the mining operations can be drawn.

I have presented during this lecture some basic problems which can be solved by the geostatistical methods without, however, any description of

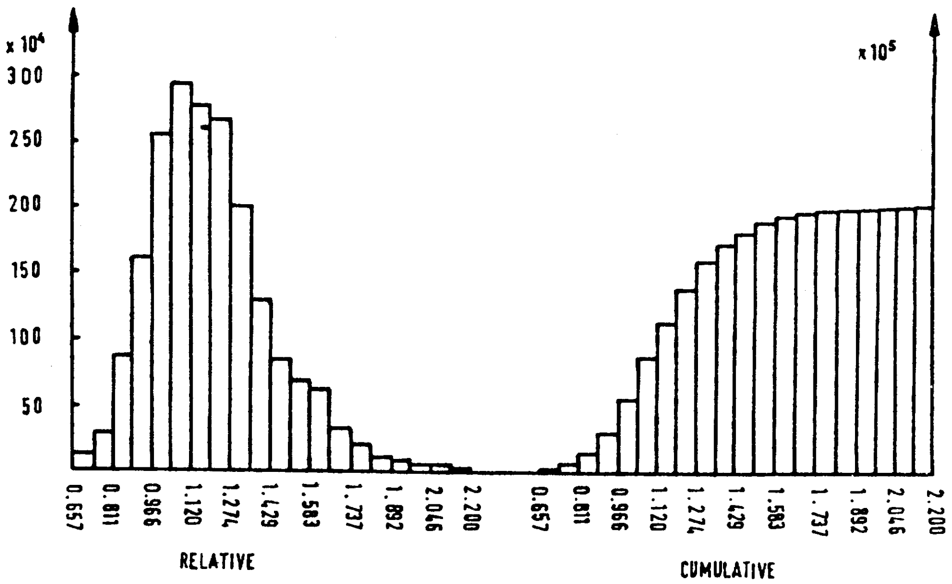


Fig. 11

Frequency distribution of kriged block values.

the geostatistical methods themselves. This will be the subject of my following lectures. What I want to emphasize now is, that the application of geostatistics is not restricted to the geological data. These methods are now used to treat geophysical data, in geomorphology, hydrology, meteorology and other branches of natural sciences.

References

1. A.G. Journel and Ch. J. Huijbregts: Mining Geostatistics. Academic Press, London, 1978.

2. M. David: Geostatistical Ore Reserve Estimation. Elsevier Sci. Publ. Co., Amsterdam, 1977.

3. M. Guarascio, M. David, C. Huijbregts (editors). Advanced Geostatistics in the Mining Industry. NATO Advanced Study Inst. Series. D. Reidel Publ. Co., Dordrecht-Holland, 1976.

4. P.A. Dowd, M. David: Planning from Estimates: Sensitivity of Mine Production Schedules to Estimation Methods. pp. 163-183 in [3].

5. H. Cramer: Mathematical Methods of Statistics. Princeton Univ. Press, Princeton, 1951.

LECTURE 2

J.A. Czubek

LECTURE 2 Reykjavík March 4, 1981

PROPERTIES OF REGIONALIZED VARIABLE

Our random variable Z(x) is called regionalized because it is defined inside the field V ($x \varepsilon V$) and moreover the Z value is always defined on some support v - we call it in that case Z_v (x), and when v=o (point like sampling) - we use simply Z (x) notation. We have seen in Fig. 9 (p. 13) that this variable Z has some features of randomity and at the same time, when v is big, some features which can be called deterministic. We have still mentioned the joined probability function for this variable (Eq. 2).

MOMENTS

We can speak about the moments of the random variable. If the distribution function $F_{x_1\ x_2 \cdot\cdot x_k}(z_1\ ,\ z_2 \ldots z_k)$ of Z(x) has an expectation (and we shall suppose that it has), then this expectation is, in a general case, a function of x, and is written as

$$E\{Z(x)\} = m(x) \tag{23}$$

Note, please, the meaning of Eq. 23 . For the sake of simplicity we can reduce the space V to the one dimensional line x. We can have, generally speaking, an infinite number of realizations of the random variable Z(x):

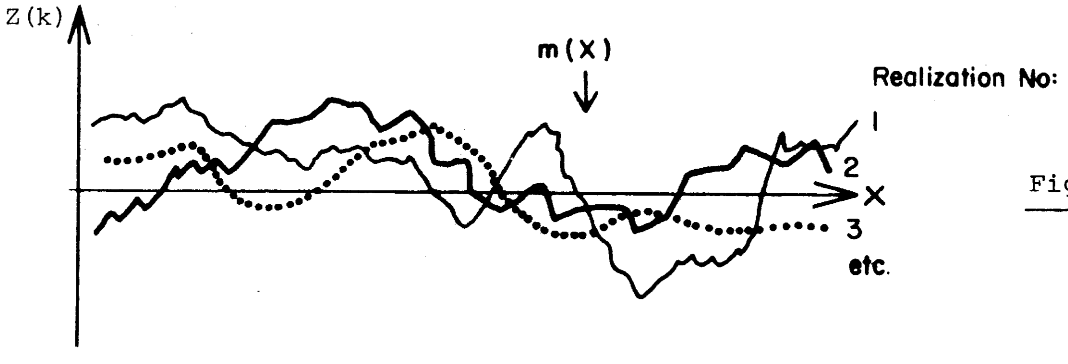


Fig. 12

Doing many realizations of Z(x) we calculate the value E {Z(x)} for a given x according to the definition of the expected value. But in geology we have only <u>ONE</u> realization of Z (x) ! And we are not able to repeat it! (except in the case of the simulation technique).

Another moments which are of interest here are the second order moments. In geostatistics we are using three kinds of the second order moments:

(a) The variance $D[Z(x)]$: this is the so called "a priori" variance of $Z(x)$, and if it exists, it is defined as:

$$D[Z(x)]= E\ \{[Z(x)-m(x)]^2\} \tag{24}$$

thus, the variance D is again, in the general case, a function of x.

(b) The covariance $C(x_1, x_2)$ - when the two random variables $Z(x_1)$ and $Z(x_2)$ have the variances at the points x_1 and x_2, they also have a covariance, which is defined as:

$$C(x_1, x_2)=E\{[Z\ (x_1)-m\ (x_1)]\cdot[Z\ (x_2)-m(x_2)]\} \tag{25}$$

(c) The variogram $2\cdot\gamma(k_1, k_2)$ which is defined as the variance of the increment $Z(x_1)-Z(x_2 = \varepsilon(x_1, x_2)$:

$$2\ \gamma(x_1, x_2)= D\ [Z(x_1)-Z(x_2)] =$$

$$= E\ \{[Z^\circ x_1)-Z(x_2)]^2\} \tag{26}$$

The function $\gamma\ (x_1, x_2)$ is called the semi-variogram. Just this semi-variogram (called sometimes for brevity the variogram) is the most important relation in geostatistics.

The higher order moments (especially the fourth order ones) are sometimes used in geostatistics for some very specialized purposes.

The definitions given in Eqs. (23) to (26) are valid for any stochastic process $Z(x)$ and these moments are functions of either the position x or of the two positions x_1 and x_2. To be able to perform some calculations, however, on the geological data, which are a unique realization of a given process, we have to introduce some restrictions on the character of the random process $Z(x)$.

STATIONARITY

For the special case, when

$$E[Z(x)]=m(x)=m \qquad Vx \tag{27}$$

i.e. the expected value is independent of the position x and the second order moments depend only upon the distance h between x_1 and x_2

$$h = |x_1 - x_2| \tag{28}$$

the process is said to be weakly stationary or stationary in a wide sense. When all possible moments and joint moments (not only the second order ones) are x invariant, such process is called strongly stationary. In geostatistics we need only the weak stationarity. Thus, except of definition given by Eq. 27 we have, for the weak stationarity:

$$C(h)=E\{Z(x+h)\cdot Z(x)\}-m^2 \qquad Vx \tag{29}$$

and immediately:

$$D[Z(x)]=E\{[Z(x)-m]^2\} = C(o), \qquad Vx \tag{30}$$

and

$$\gamma(h)=\frac{1}{2}\, E\{[Z(x+h)-Z(x)]^2\}=C(o)-C(h) \qquad Vx \tag{31}$$

As one can see from relation given by Eq. 31 , under the hypothesis of a weak stationarity the covariance and the variogram are equivalent each other.

INTRINSIC HYPOTHESIS

As a matter of fact, in geostatistics we need even less strict assumptions about stationarity than these presented above. Looking at the relation (31) we can see, that in some cases when the C(h) function is not finite, the difference C(o)-C(h), i.e. the variogram γ(h) can be finite. And this is just the assumption which we need and it is called the intrinsic hypothesis. Thus:

a random function Z(x) is said to be intrinsic when:

a) its mathematical expectation exists and does not depend upon the supposed point x:

$$E[Z(x)] = m \ , \qquad \forall x; \qquad (27)$$

and

b) for all vectors h the increment $[Z(x+h)-Z(x)]$ has af finite variance which does not depend on x:

$$D\{Z(x+h)-Z(x)\}=E\{[Z(x+h)-Z(x)]^2\}= 2\ \gamma(h) \qquad (32)$$

Note, please, that the second order (weak) stationarity implies the intrinsic hypothesis but the inverse is not true. The intrinsic hypothesis can also exist when the random process is not even weak stationary.

All geostatistical methods are based on the intrinsic hypothesis and this is the main difference between the geostatistics and the usual stochastic process approach.

In practice we cannot extend the vector h towards infinity because any geological body has a finite dimension (remember that Z(x) and Z(x+h) have to be taken from the same geological formation!). Let this limit value of h be b (here b can be considered as a dimension of the ore body, scale of experiment, etc.), thus

$$|h| \leq b$$

In this case we say that we need the quasi-stationarity or the quasi-intrinsic hypothesis.

PROPERTIES OF COVARIANCE AND VARIOGRAM

These properties appear directly from the stochastic process theory.

Let Z(x) be a stationary random function with the expected value m and covariance C(h) (or semi-variogram $\gamma(h)$). Let us take a linear combination of Z(x):

$$Y = \sum_{i=1}^{n} \lambda_i \cdot Z(x_i) \qquad (33)$$

for any weights λ_i .

This linear combination is also a random variable and its variance should never be negative, i.e.

$$D(Y) \geq 0$$

which explicitly is written as:

$$D(Y) = \sum_i \sum_j \lambda_i \lambda_j \; C(x_i - x_j) \geq 0 \qquad (34)$$

Thus the covariance function C(h) must be such that it furnishes always a positive variance (or zero). By definition the function C(h) is said to be "positive definite" (thus not any function can be used as a covariance function!).

Taking into account Eq. (31) the last expression (36) could be written as:

$$D(Y) = C(0) \cdot \sum_i \lambda_i \cdot \sum_i \lambda_j - \sum_i \sum_j \lambda_i \lambda_j \; \gamma(x_i - x_j) \qquad (35)$$

In the case, when C(0) does not exist, and only the intrinsic hypothesis is assumed, and the variance of Y is defined, it follows from (35) that

$$\sum_{i=1} \lambda_i = 0 \qquad (36)$$

and

$$D(Y) = - \sum_i \sum_j \lambda_i \lambda_j \; \gamma \; (x_i - x_j) \qquad (37)$$

In the last expression when $D(Y) \geq 0$ and with the condition (36), the function $\gamma(h)$ is said to be a "conditional positive definite" function. Note, please, that when $C(0)$ does not exist, it means that the variance of the random variable $Z(x)$ does not exist either!

PROPERTIES OF THE COVARIANCE

$C(0) = D[Z(x)] \geq 0$ - a priori variance cannot be negative (38)

$C(h) = C(-h)$ - the covariance is an even function (39)

$|C(h)| \lesseqgtr C(0)$ Schwarz's inequality (40)

When the covariance function $C(h)$ has a high value we say that the two random variables $Z(x)$ and $Z(x+h)$ with distance h from each other are well correlated, and when $C(h) \to 0$ we say, that there is no correlation. Usual case is that for a big h values $C(h)$ tends to zero.

$$C(h) \to 0 \text{ , when } h \to \infty \tag{41}$$

The variogram $\gamma(h)$, when for $h \to 0$ is $\gamma(h) \to 0$, we say that the random variable $Z(x)$ is well correlated. Usually when $h \geq a$ where a is a certain distance the variogram $\gamma(h)$ attains some sill (its maximum value), which according to Eqs. (30) and (31) corresponds to the variance of the random variable. Finally the usual behaviour of the $C(h)$ and $\gamma(h)$ functions is (Fig. 13)

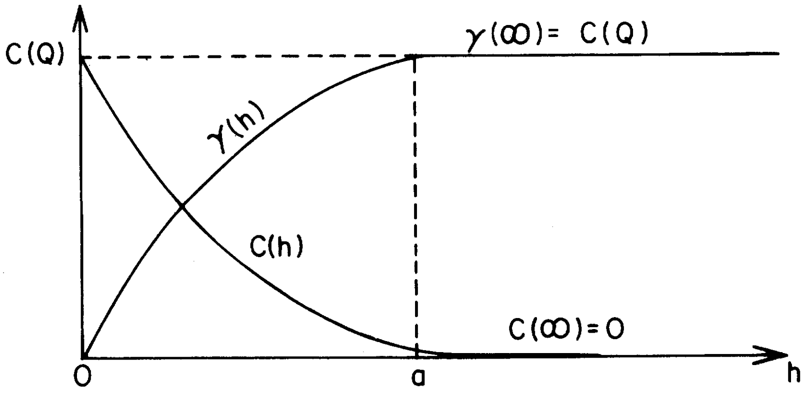


Fig. 13

The value $h=a$ in Fig. 13 is called the zone of influence or the range of variogram. Outside the zone of influence the values of the random variables $Z(x)$ and $Z,(x+h)$, $h>a$, are independent of each other.

PROPERTIES OF THE VARIOGRAM

From the above discussion we have:

$$\begin{aligned} \gamma(0) &= 0 \\ \gamma(h) &= \gamma(-h) \geq 0 \\ \gamma(\infty) &= D[Z(x)] = C(0) \end{aligned} \tag{42}$$

Anisotropy: It is quite obvious that the behaviour of the $\gamma(h)$ function can be different when the vector h has different directions. In this case we are calling that fenomenon the anisotropy. When any geological formation has the same variogram $\gamma(h)$ independent of the direction of the h vector, such formation is called isotropic. An example of the anisotropy is given in Fig. 14 for the ore body of the lens type

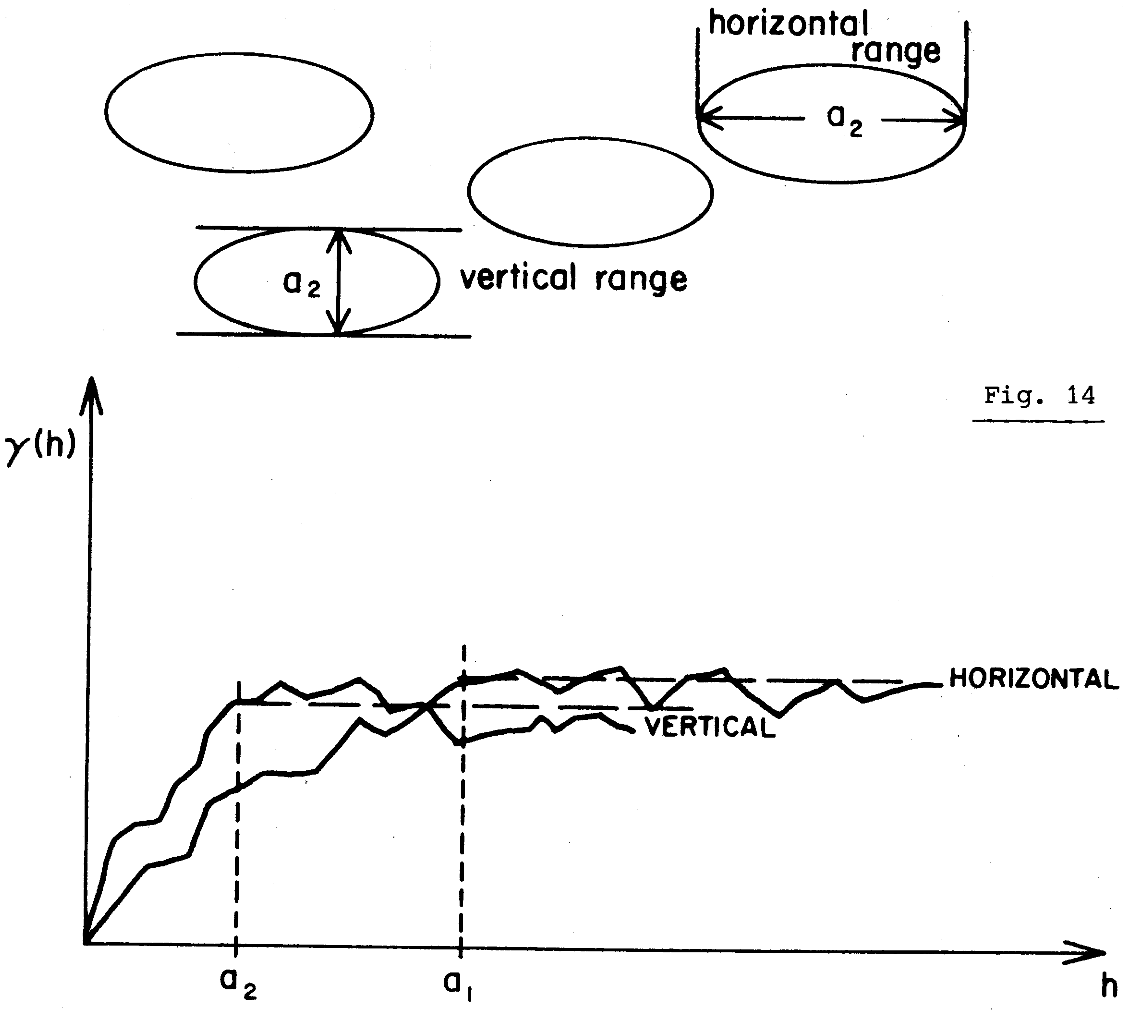
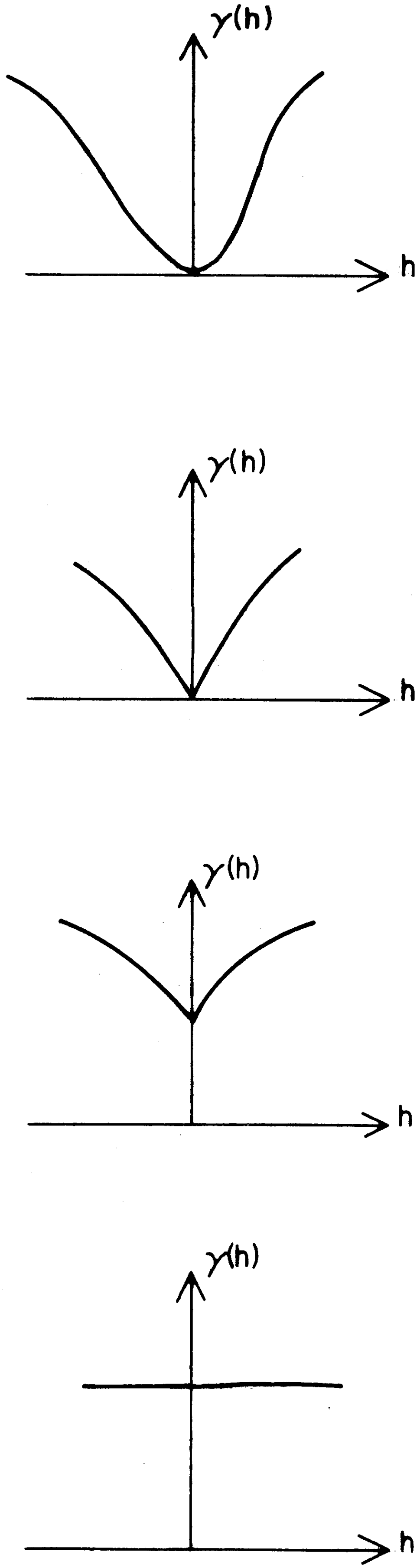


Fig. 14

BEHAVIOUR OF THE VARIOGRAM NEAR THE ORIGIN



Parabolic $\gamma(h) \sim A|h|^2$ for $h \to 0$

$\gamma(h)$ is twice differentiable -
in highly regular spatial variability

Linear $\gamma(h) \sim A|h|$ for $h \to 0$

Discontinuity at the the origin -
nugget effect

$\gamma(0) \neq 0$ although by difinition, from the theory should be $\gamma(0)=0$. It is the effect of the so called white noise due to the micro-variability

Pure nugget effect $\gamma(h) = c_0$ -

- there is no space correlation -
- the values of $Z(x)$ function are completely independent of each other.

Fig. 15

BEHAVIOUR OF THE VARIOGRAM AT THE INFINITY

As it follows from Eq (37) the function

$$- \gamma(h)$$

is a conditional positive definite function, and it can be shown that the variogram increases always more slowly at infinity than does $|h|^2$:

$$\lim_{|h|\to\infty} \frac{\gamma(h)}{|h|^2} = 0 \qquad (43)$$

If one observes for some geological formation that for large h its variogram increases more rapidly than h^2 it means, that the random function Z(x)is no more stationary, even intrinsic, that is,one has in this case some drift (trend):

$$E\ \{Z(x)\} = m\ (x) \qquad (44)$$

which depends upon x .

As an example

$$\bar{}Z(x) = a_o + a_1 x + Z^*(x) \qquad (45)$$

where $a_o + a_1 x$ is a linear drift and $Z^*(x)$ is a random variable. Calculating according to Eq. (31) the $\gamma(h)$ value for Z (x) and bearing in mind that

$$E[Z^*(x+h)-Z^*(x)] = 0 \qquad (46)$$

(because $Z^*(x)$ is a random function !), one arrives to the expression:

$$2\ \gamma(h) = E\{[Z(x+h)-Z(x)]^2\} =$$

$$= a_1^2 h^2 + 2\gamma^*(h) \qquad (47)$$

where

$$2\ \gamma^*(h) = E\ \{[Z^*(x+h)-Z^*(x)]^2\} \qquad (48)$$

When the drift is not so strong, for example, it has a form

$$m(x) = a_o + a_1 \sqrt{X} \qquad (49)$$

one has to calculate according to the general formula:

$$2\gamma(h) = E\ \{[m(x+h)-m(x)]^2\} + 2\gamma^*(h) \tag{50}$$

HOW TO CALCULATE THE VARIOGRAM ?

To be more instructive we now turn out the problem to be of only one-dimension, say X. In this case, when the drift is absent, one has, according to Eq. (31):

$$\gamma(h)=\frac{1}{2}\ E\ \{[Z(x+h)-Z\ (x)]^2\} =$$

$$= \frac{1}{2}\ \lim_{X\to\infty}\ \frac{1}{X}\int_0^X [Z(x+h)-Z(x)]^2\, dx \tag{51}$$

This is the so called theoretical variogram which is computed over an infinite number of pairs in a supposedly infinite field. We of course never obtain it in practice. Thus two other variograms will appear:

Instead of calculating the integral in (51) over an infinite field we do it in the finite field L over which samples are spread. Thus:

$$\gamma^1(h) = \frac{1}{2(L-h)}\int_0^{L-h} [Z(x+h)-Z\ (x)]^2\ dx \tag{52}$$

The variogram $\gamma^1(h)$ is called the local variogram and again it contains an infinite number of pairs of the Z values. We will never obtain it either. What is really computed is the so called experimental variogram $\gamma^*(h)$ which can be considered as an estimate of the $\gamma(h)$ or $\gamma^1(h)$ variograms:

$$\gamma^*(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)} [Z(x_i+h)-Z(x_i)]^2 \tag{53}$$

where N(h) is the number of pairs $[Z(x_i+h)-Z(x_i)]$ available for a distance h equal to a multiple of the sampling interval.

It was proved by MATHERON (1965) [6] that the expected value of $\gamma^1(h)$ and $\gamma^*(h)$ is always $\gamma(h)$:

$$E[\gamma^*(h)] = E[\gamma^1(h)] - \gamma(h) \tag{54}$$

However, one should take into account the statistical discrepancy which may occur in a given particular case. According to MATHERON one has:

FLUCTUATION - it is the difference between the local $\gamma^1(h)$ and the theoretical $\gamma(h)$ variogram, and

ESTIMATION ERROR - it is the difference between the local variogram $\gamma^1(h)$ and the experimental variogram $\gamma^*(h)$

One has for the fluctuation variance:

$$E\{[\gamma(h)-\gamma^1(h)]^2\} = \begin{cases} A\cdot(\frac{h}{L})\cdot[\gamma(h)]^2 & \text{for } h < \frac{L}{3} \\ \text{prohibitive} & \text{for } h > \frac{L}{2} \end{cases} \tag{55}$$

where A is a constant ($A=\frac{4}{3}$ for linear model $\gamma(h) = |h|$)

and for the estimation error the variance is:

$$E\{[\gamma^1(h)-\gamma^*(h)]^2\} \simeq 4\ \gamma(h)\cdot \frac{D(Z)}{N^1} \tag{56}$$

where N^1 is the number of the experimental data pairs and D(Z) is the variance of the point like support variable Z .

The practical rule here is that when one wants to get some idea what is the behaviour of the theoretical variogram $\gamma(h)$ having the experimental data on $\gamma^*(h)$, one can do this on basis of

---

6. G. MATHERON: Les variables régionalisées et leur estimation. Editions MASSON, Paris, 1965.

the behaviour of the experimental variogram γ^* (h) within the range $0 < h < L/3$ only ! And again the number of pairs should be as big as possible (N^1 big !).

The practical rule of calculation of the $\gamma^*(h)$ values according to the formula (53) is following:

a) when the sampling is performed at constant intervals:

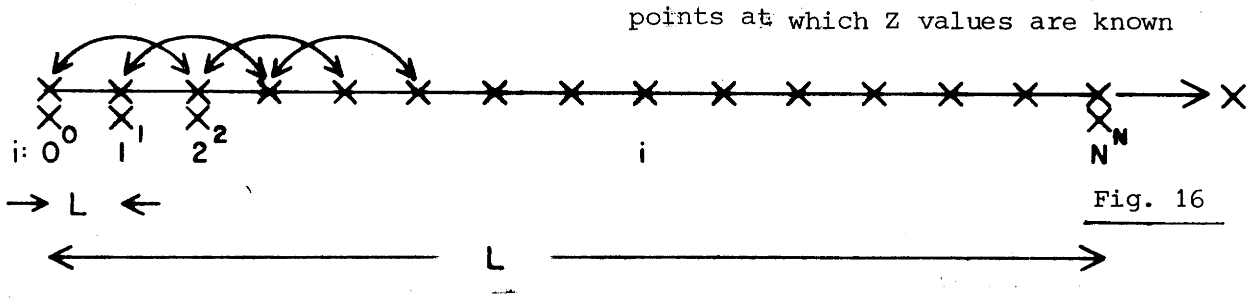


Fig. 16

$L = N \cdot 1$ 1 - sampling interval

$h = n \cdot 1$ Number of samples: N+1

$$\gamma^*(h) = \gamma^*(n) = \frac{1}{2(N+1-n)} \sum_{i=o}^{N-n} [Z(i+n)-Z(i)]^2$$

$$x_i = i \cdot 1$$

and in Fig. 16 and example is shown for n = 2.

Repeating the calculations for the consecutive values

$n = 1, 2, 3 \ldots N$

one obtains a plot presented in Fig. 17

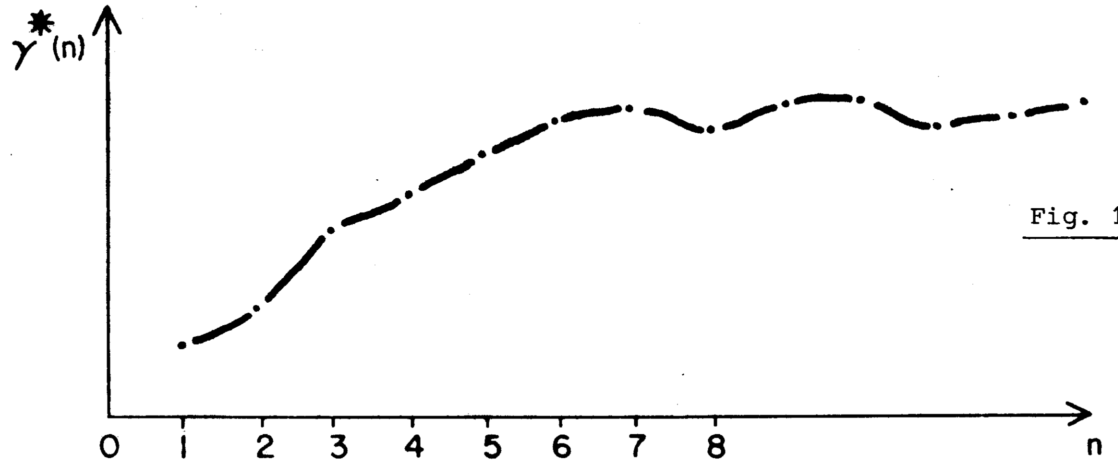


Fig. 17

b) when the sampling is performed more or less randomly inside some space. Let us take the two-dimensional space - cf - Fig. 18

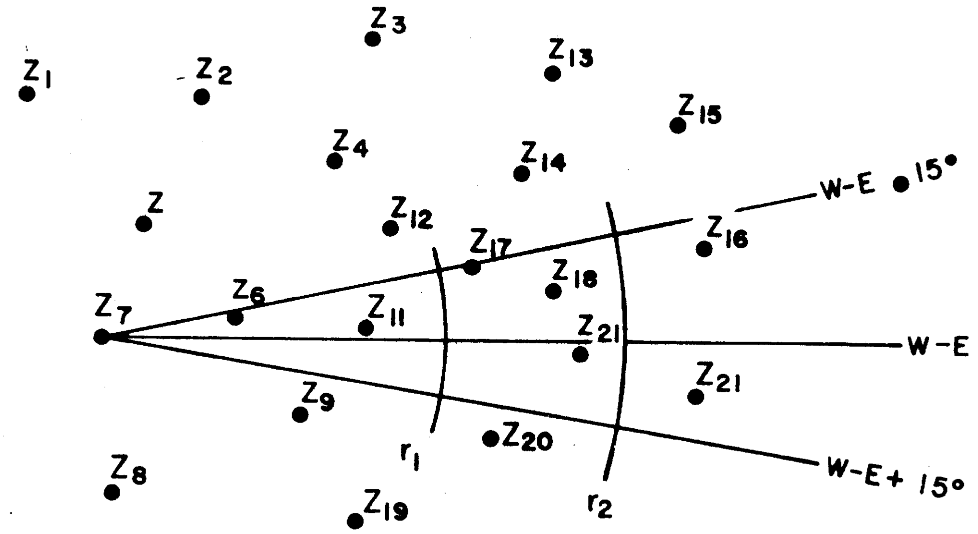


Fig. 18

In this case the experimental variogram is calculated for a given interval of the azimuth and a given interval of distance. One situates in this case the origin of the polar coordinates at a given point x_i - for example $x_i = x_7$. One takes for example the direction W - E ± 15° and the interval of distances between r_1 and r_2 , i.e. $h_1 = \frac{r_2 - r_1}{2}$. Thus, for the point x_7 one has in this case 3 pairs of data:

$z_7 - z_{17}$

$z_7 - z_{18}$

$z_7 - z_{13}$

which enter as the three pairs for the formula to calculate $\gamma(h_1)$. Next one moves the origin of the polar coordinates to the next point, for example z_8 and one repeats the same operations for the same range of angles and distances obtaining by this manner the further pairs of data in the formula for the $\gamma(h_1)$ calculation.

LECTURE 3

J.A. Czubek

LECTURE 3 Reykjavík March 8, 1981

MORE ABOUT VARIOGRAM

When the sampling interval is constant (which is, for example the case of any well logging operation) the variogram can be very easily calculated using a pocket programmable calculator. Such program for the TI-59 is given in the following pages. This program calculates the $\gamma^*(h)$ values for $h = i\cdot l.$; $i = 1,2, \ldots 10$, the average $\bar{z}$ value according to the formula:

$$E\{Z(x)\} \simeq \bar{z} = \frac{1}{N}\sum_{k=1}^{N} z_K \tag{58}$$

and the dispersion variance $s^2(z)$:

$$D\{Z(x)\} \simeq s^2(Z) = \frac{1}{N-1}\sum_{k=1}^{N}(z_k-\bar{z})^2 =$$

$$= \frac{1}{N-1}\left(\sum_{k=1}^{N} z_k^2 - N\cdot\bar{z}^2\right) \tag{53}$$

For variogram calculated according to the formula

$$\gamma^*(h) = \gamma^*(i\cdot l) = \frac{1}{2N^1}\sum_{k=1}^{N^1}[z(k+i)-z(k)]^2 , \tag{60}$$

here the notation is used:

$$z(x_k) \equiv z_k \equiv z(k) , \tag{61}$$

The number of pairs N^1 in Eq. (60) is equal to:

$$N^1 = N-10 \tag{62}$$

where N is the number of data in Eqs. (58) and (59).
This program, VARIOGRAM 10, is presented in pages 38-41 of this lecture. The N^1 value is kept constant for all $h = i\cdot l$ values which results in the fact that not all data are always used for the calculation of the $\gamma^*(i\cdot l)$ data - this is just the boundary effect due to the fact that the total number of data is limited at the beginning (there is no data for $k < 1$ and at the end -

there is no data for $k > N^1$). The organizing diagram showing which data are used only once or even not at all is given in Fig. 19.

DATA USED FOR CALCULATION OF $\gamma^*(i \cdot l)$

VALUES IN THE PROGRAM VARIOGRAM 10

k i: ↓	1	2	3	4	5	6	7	8	9	10
1	x	x	x	x	x	x	x	x	x	x
2	0	x	DATA USED ONLY ONCE							x
3	0	0	x			(x)				x
4	0		0	x						x
5	0			0	x					x
6	0				0	x				x
7	0					0	x			x
8	0	DATA USED TWICE(0)					0	x		x
9	0							0	x	x
10	0								0	x
11	0									0
.	0									0
.	0									0
.	0									0
.	0									0
.	0									0
$k=N^1$	0	0	0	0	0	0	0	0	0	0
k+1	x	x	x	x	x	x	x	x	x	x
k+2		x	DATA USED ONLY ONCE							x
k+3			x					(x)		x
k+4				x						x
k+5					x					x
k+6		DATA NOT				x				x
k+7		USED					x			x
k+8								x		x
k+9									x	x
k+10=N										x

Fig. 19

Let us take for example $i=6$. For the k-values ranging from $k=1$ up to $k=6$ the data

$$z_1 \ , \ z_2 \ \dots \ z_6$$

are used only once in Eq. 60 and they act with the data

$$z_7 \ , \ z_8 \ \ldots \ z_{12}$$

respectively. Starting from k=7 the data z_K are used twice in Eq. 60 , because, just for k=7 they enter as:

$$(z_7 - z_1)^2 \quad \text{and} \quad (z_{13} - z_7)^2$$

This is the situation up to a certain $K=N^1$ where N^1 is the total number of pairs $(z_{k+6} - z_k)$ of data used in the calculation. Now, starting from $k = N^1 + 1$ up to $k = N^1 + 6$ the

$$z_{N^1+1} \ , \ z_{N^1+2} \ \ldots . \ z_{N^1+6}$$

data are used only once in the respective pairs:

$$(z_{N^1+1} - z_{N^1-5})^2_1, (z_{N^1+2} - z_{N^1-4})^2 \ \ldots, (z_{N^1+6} - z_{N^1})^2 \ .$$

The data for $k > N^1+6$ are not used at all for the $\gamma^*(6 \cdot 1)$ calculation.

This certain assymmetry in the data utilization can in some cases give discrepancies when the variogram is calculated down-wards or up-wards, because (slightly) different set of data is used. This is especially the case, when the process being investigated is not entirely stationary and when certain drift exists.

When the printer is not used while operating the program VARIOGRAM 10 , at the positions: 147, 166, 181, 190, 199, 208, 217, 226, 235, 244, 253 and 262 the instruction R/S should be given instead of PRT.. This program fails also when some z_k data are equal to zero, thus it always should be $z_k \neq 0$.

EXERCISE No 1

Calculate the experimental variograms using the program VARIOGRAM 10 for the porosity Ø and for the bulk density ρ data reported in Table 1 at the one meter intervals. Start k=1 for the depth 1708.45 m. The plot of these data is in Fig. 20. Use the last data N = k+10 at the depth 1847,45 m. Thus in this case N = 140 .

TITLE VARIOGRAM 10 PAGE 1 OF 4

PROGRAMMER CZUBEK DATE DEC. 11. 1980

TI Programmable
Program Record

Partitioning (Op 17) ___ Library Module ___ Printer YES Cards ___

PROGRAM DESCRIPTION

$\gamma(n) = \frac{1}{2N}\sum_{i=1}^{N}\left[z(i) - z(i+n)\right]^2$ $n = 1,2,3,4,5,6,7,8,9,10$

$N > n$, N NON LIMITED $\bar{z} = \frac{1}{N+10}\sum_{i=1}^{N+10} z_i$

$s^2(z) = \frac{1}{N+9}\sum_{i=1}^{N+10} z_i^2 - \frac{N+10}{N+9}\bar{z}^2$

USER INSTRUCTIONS

STEP	PROCEDURE	ENTER	PRESS			DISPLAY
1	INSERT PROGRAM		RST	CMS	R/S	
2		z_1			R/S	0
3		z_2			R/S	0
- -	etc					
11		z_{10}			R/S	0
12		z_{11}			R/S	$2N\gamma(10)$
13		z_{12}			R/S	$2N\gamma(10)$
N+11		z_{N+10}			R/S A	$2N\gamma(10)$
N+12		PRINT				$\bar{z}$ $s^2(z)$ $\gamma(1)$ $\gamma(2)$ - - - $\gamma(10)$
		z_{N+11} etc			R/S	$2N\gamma(10)$

USER DEFINED KEYS		DATA REGISTERS (INV ■)			
Ⓐ		0	used	10	used
B		1	used	11	$2N\gamma(1)$
C		2	used	12	$2N(\gamma(2)$
D		3	used	13	$2N\gamma(3)$
E		4	used	14	$2N\gamma(4)$
A'		5	used	15	$2N\gamma(5)$
B'		6	used	16	$2N\gamma(6)$
C'		7	used	17	$2N\gamma(7)$
D'		8	used	18	$2N\gamma(8)$
E'		9	used	19	$2N\gamma(9)$

LABELS (Op 08)

INV ___	lnx ___	CE ___	CLR ___	x⇄t ___	x² ___
√x ___	1/x ___	STO ___	RCL ___	SUM ___	yˣ ___
EE ___	(___	) ___	÷ ___	GTO ___	× ___
SBR ___	− ___	RST ___	+ ___	R/S ___	. ___
+/− ___	= ___	CLR ___	INV ___	log ___	CP ___
tan ___	Pgm ___	P→R ___	sin ___	cos ___	CMs ___
Exc ___	Prd ___	|x| ___	Eng ___	Fix ___	Int ___
Deg ___	Pause ___	x=t ___	Nop ___	Op ___	Rad ___
Lbl ___	x≥t ___	Σ+ ___	x̄ ___	Grad ___	St flg ___
If flg ___	D.MS ___	π ___	List ___	Write ___	Dsz ___
Adv ___	Prt ___				

FLAGS 0	1	2	3	4	5	6	7	8	9

TITLE VARIOGRAM 10 PAGE 2 OF 4

TI Programmable
Program Record

PROGRAMMER ______ DATE ______

Partitioning (Op 17) ______ Library Module ______ Printer YES Cards ______

PROGRAM DESCRIPTION

WHEN Ⓐ IS PRESSED APPEARS PRINT: $\bar{z}$, $s^2(z)$, $\gamma(1)$, $\gamma(2)$ etc $\gamma(10)$
ONE CAN PUT FURTHER DATA (N+1, N+2 etc) PRESSING R/S
AND AGAIN Ⓐ TO HAVE A NEW SET $\bar{z}$, $s^2(z)$ AND $\gamma(i)$
RESULTS OBTAINED FOR THE NEW N VALUE

USER INSTRUCTIONS

STEP	PROCEDURE	ENTER	PRESS	DISPLAY

USER DEFINED KEYS	DATA REGISTERS (INV ■)		
A	20 $2N\,\gamma(10)$	30 $\gamma(10)$	
B	21 $\gamma(1)$	31 $\sum z_i$	
C	22 $\gamma(2)$	32 $\sum z_i^2$	
D	23 $\gamma(3)$	33 $N+10$	
E	24 $\gamma(4)$	34 $\bar{z}$	
A'	25 $\gamma(5)$	35 $s^2(z)$	
B'	26 $\gamma(6)$	36	
C'	27 $\gamma(7)$	37	
D'	28 $\gamma(8)$	38	
E'	29 $\gamma(9)$	39	

LABELS (Op 08)

INV ___	lnx ___	CE ___	CLR ___	x⇌t ___	x² ___
√x ___	1/x ___	STO ___	RCL ___	SUM ___	yˣ ___
EE ___	(___	) ___	÷ ___	GTO ___	× ___
SBR ___	− ___	RST ___	+ ___	R/S ___	. ___
+/− ___	= ___	CLR ___	INV ___	log ___	CP ___
tan ___	Pgm ___	P→R ___	sin ___	cos ___	CMs ___
Exc ___	Prd ___	|x| ___	Eng ___	Fix ___	Int ___
Deg ___	Pause ___	x=t ___	Nop ___	Op ___	Rad ___
Lbl ___	x≥t ___	Σ+ ___	x̄ ___	Grad ___	St flg ___
If flg ___	D.MS ___	π ___	Ist ___	Write ___	Dsz ___
Adv ___	Prt ___				

FLAGS									
0	1	2	3	4	5	6	7	8	9

USER DEFINED KEYS	DATA REGISTERS (INV ■)		
A	40	50 N	
B	41	51 $(2N)^{-1}$	
C	42	52	
D	43	3	
E	44	4	
A'	45	5	
B'	46	6	
C'	47	7	
D'	48	8	
E'	49	9	

LABELS (Op 08)

INV ___	lnx ___	CE ___	CLR ___	x⇌t ___	x² ___
√x ___	1/x ___	STO ___	RCL ___	SUM ___	yˣ ___
EE ___	(___	) ___	÷ ___	GTO ___	× ___
SBR ___	− ___	RST ___	+ ___	R/S ___	. ___
+/− ___	= ___	CLR ___	INV ___	log ___	CP ___
tan ___	Pgm ___	P→R ___	sin ___	cos ___	CMs ___
Exc ___	Prd ___	|x| ___	Eng ___	Fix ___	Int ___
Deg ___	Pause ___	x=t ___	Nop ___	Op ___	Rad ___
Lbl ___	x≥t ___	Σ+ ___	x̄ ___	Grad ___	St flg ___
If flg ___	D.MS ___	π ___	Ist ___	Write ___	Dsz ___
Adv ___	Prt ___				

FLAGS									
0	1	2	3	4	5	6	7	8	9

TLE *VARIOGRAM 10* PAGE *3* OF *4*

ROGRAMMER DATE

TI Programmable Coding Form

LOC	CODE	KEY	COMMENTS
00	91	R/S	
01	44	SUM	
02	31	31	$\Sigma\ z_i$
03	33	X²	
04	44	SUM	
05	32	32	$\Sigma\ z_i^2$
06	34	√X	
07	48	EXC	
08	00	00	z (i+10)
09	48	EXC	
10	01	01	z (i+9)
11	48	EXC	
12	02	02	z (i+8)
13	48	EXC	
14	03	03	z (i+7)
15	48	EXC	
16	04	04	z (i+6)
17	48	EXC	
18	05	05	z (i+5)
19	48	EXC	
20	06	06	z (i+4)
21	48	EXC	
22	07	07	z (i+3)
23	48	EXC	
24	08	08	z (i+2)
25	48	EXC	
26	09	09	z (i+1)
27	48	EXC	
28	10	10	z (i)
29	00	0	
30	32	X:T	
31	43	RCL	
32	10	10	
33	67	EQ	
34	00	00	
35	00	00	
36	01	1	
37	44	SUM	
38	50	50	$\Sigma\ i=N$
39	43	RCL	
40	10	10	
41	75	-	
42	43	RCL	
43	09	09	
44	95	=	
45	33	X²	
46	44	SUM	
47	11	11	2N γ(1)
48	43	RCL	
49	10	10	
50	75	-	
51	43	RCL	
52	08	08	
53	95	=	
54	33	X²	
055	44	SUM	
056	12	12	2N γ(2)
057	43	RCL	
058	10	10	
059	75	-	
060	43	RCL	
061	07	07	
062	95	=	
063	33	X²	
064	44	SUM	
065	13	13	2N γ(3)
066	43	RCL	
067	10	10	
068	75	-	
069	43	RCL	
070	06	06	
071	95	=	
072	33	X²	
073	44	SUM	
074	14	14	2N γ(4)
075	43	RCL	
076	10	10	
077	75	-	
078	43	RCL	
079	05	05	
080	95	=	
081	33	X²	
082	44	SUM	
083	15	15	2N γ(5)
084	43	RCL	
085	10	10	
086	75	-	
087	43	RCL	
088	04	04	
089	95	=	
090	33	X²	
091	44	SUM	
092	16	16	2N γ(6)
093	43	RCL	
094	10	10	
095	75	-	
096	43	RCL	
097	03	03	
098	95	=	
099	33	X²	
100	44	SUM	
101	17	17	2N γ(7)
102	43	RCL	
103	10	10	
104	75	-	
105	43	RCL	
106	02	02	
107	95	=	
108	33	X²	
109	44	SUM	
110	18	18	2N γ(8)
111	43	RCL	
112	10	10	
113	75	-	
114	43	RCL	
115	01	01	
116	95	=	
117	33	X²	
118	44	SUM	
119	19	19	2N γ(9)
120	43	RCL	
121	10	10	
122	75	-	
123	43	RCL	
124	00	00	
125	95	=	
126	33	X²	
127	44	SUM	
128	20	20	2N γ(10)
129	81	RST	
130	76	LBL	
131	11	A	
132	43	RCL	
133	50	50	
134	85	+	
135	01	1	
136	00	0	
137	95	=	
138	42	STO	
139	33	33	N+10
140	35	1/X	
141	65	×	
142	43	RCL	
143	31	31	
144	95	=	
145	42	STO	
146	34	34	$\bar{z}$
147	99	PRT	
148	33	X²	
149	65	×	
150	43	RCL	
151	33	33	
152	75	-	
153	43	RCL	
154	32	32	
155	95	=	
156	94	+/-	
157	55	÷	
158	53	(	
159	43	RCL	

MERGED CODES

62 Pgm Ind
63 Exc Ind
64 Prd Ind
72 STO Ind
73 RCL Ind
74 SUM Ind
83 GTO Ind
84 Op Ind
92 INV SBR

TEXAS INSTRUMENTS
INCORPORATED

TITLE VARIOGRAM 10 PAGE 4 OF 4

PROGRAMMER ____ DATE ____

TI Programmable Coding Form

LOC	CODE	KEY	COMMENTS
160	33	33	
161	75	-	
162	01	1	
163	95	=	
164	42	STO	
165	35	35	$\sigma^2(z)$
166	99	PRT	
167	43	RCL	
168	50	50	
169	65	×	
170	02	2	
171	95	=	
172	35	1/X	
173	42	STO	
174	51	51	$(2N)^{-1}$
175	65	×	
176	43	RCL	
177	11	11	
178	95	=	
179	42	STO	
180	21	21	$\gamma(1)$
181	99	PRT	
182	43	RCL	
183	51	51	
184	65	×	
185	43	RCL	
186	12	12	
187	95	=	
188	42	STO	
189	22	22	$\gamma(2)$
190	99	PRT	
191	43	RCL	
192	51	51	
193	65	×	
194	43	RCL	
195	13	13	
196	95	=	
197	42	STO	
198	23	23	$\gamma(3)$
199	99	PRT	
200	43	RCL	
201	51	51	
202	65	×	
203	43	RCL	
204	14	14	
205	95	=	
206	42	STO	
207	24	24	$\gamma(4)$
208	99	PRT	
209	43	RCL	
210	51	51	
211	65	×	
212	43	RCL	
213	15	15	
214	95	=	
215	42	STO	
216	25	25	$\gamma(5)$
217	99	PRT	
218	43	RCL	
219	51	51	
220	65	×	
221	43	RCL	
222	16	16	
223	95	=	
224	42	STO	
225	26	26	$\gamma(6)$
226	99	PRT	
227	43	RCL	
228	51	51	
229	65	×	
230	43	RCL	
231	17	17	
232	95	=	
233	42	STO	
234	27	27	$\gamma(7)$
235	99	PRT	
236	43	RCL	
237	51	51	
238	65	×	
239	43	RCL	
240	18	18	
241	95	=	
242	42	STO	
243	28	28	$\gamma(8)$
244	99	PRT	
245	43	RCL	
246	51	51	
247	65	×	
248	43	RCL	
249	19	19	
250	95	=	
251	42	STO	
252	29	29	$\gamma(9)$
253	99	PRT	
254	43	RCL	
255	51	51	
256	65	×	
257	43	RCL	
258	20	20	
259	95	=	
260	42	STO	
261	30	30	$\gamma(10)$
262	99	PRT	
263	92	RTN	
264	81	RST	
265	00	0	
266	00	0	
267	00	0	
268	00	0	

MERGED CODES

62 Pgm Ind	72 STO Ind	83 GTO Ind
63 Exc Ind	73 RCL Ind	84 Op Ind
64 Prd Ind	74 SUM Ind	92 INV SBR

TEXAS INSTRUMENTS INCORPORATED

TI-24181

TABLE 1

DEPTH (m)	RESISTIVITY (Ωm)	POROSITY (ALP)	DENSITY (CGS)
1705.45	94.70	0.521010	1.3630
1706.45	113.60	0.496950	2.0530
1707.45	654.10	0.187980	2.8590
1708.45	994.40	0.065260	2.7380
1709.45	964.10	0.080340	2.7890
1710.45	1004.20	0.078920	2.6870
1711.45	1163.80	0.073900	2.6900
1712.45	1431.30	0.064160	2.7330
1713.45	1413.20	0.065340	2.7270
1714.45	1333.70	0.070200	2.7700
1715.45	1175.40	0.073160	2.7800
1716.45	1025.20	0.078470	2.7580
1717.45	864.20	0.093590	2.8100
1718.45	638.10	0.125800	2.8850
1719.45	408.10	0.150760	2.8410
1720.45	302.40	0.193980	3.1010
1721.45	273.10	0.073700	3.1320
1722.45	1609.10	0.036480	3.0610
1723.45	2620.50	0.042670	3.1080
1724.45	2397.30	0.037270	3.0220
1725.45	1772.90	0.030930	3.0440
1726.45	1733.10	0.031310	3.0480
1727.45	2476.10	0.031830	3.0120
1728.45	2862.80	0.034470	2.9920
1729.45	2819.20	0.033990	3.0310
1730.45	2583.70	0.033020	2.9970
1731.45	2225.10	0.026830	3.0300
1732.45	1807.80	0.024650	3.0160
1733.45	1537.50	0.066290	2.7800
1734.45	913.40	0.147170	2.7470
1735.45	314.50	0.171620	2.7820
1736.45	258.90	0.167240	2.8600
1737.45	251.40	0.115180	3.0730
1738.45	369.10	0.071110	3.0930
1739.45	1082.10	0.052650	3.0720
1740.45	1468.70	0.047260	3.1730
1741.45	1397.50	0.038120	3.1750
1742.45	1486.00	0.039900	3.1940
1743.45	1345.60	0.035370	3.1420
1744.45	1446.70	0.068960	2.9650
1745.45	868.00	0.126700	2.9330
1746.45	550.30	0.125830	2.9150
1747.45	602.10	0.083350	2.9900
1748.45	965.70	0.073230	2.9730
1749.45	980.60	0.080530	3.0570
1750.45	926.60	0.064950	3.1020
1751.45	1107.20	0.077530	2.8340
1752.45	509.30	0.172040	2.8450
1753.45	371.50	0.129840	2.8760
1754.45	511.00	0.084360	2.9860
1755.45	696.10	0.085670	2.9600
1756.45	582.50	0.071830	2.8860
1757.45	829.40	0.077230	2.8220
1758.45	524.40	0.090110	3.1410
1759.45	849.00	0.050630	3.0880
1760.45	1270.70	0.030400	3.2130
1761.45	1465.50	0.024510	3.1630
1762.45	1297.10	0.031690	3.2220

1763.45	1182.60	0.023120	3.1820
1764.45	1255.20	0.024720	3.1850
1765.45	1430.50	0.023720	3.1580
1766.45	1478.80	0.025900	3.0370
1767.45	1597.80	0.020520	2.8910
1768.45	1865.60	0.013280	3.1690
1769.45	1669.70	0.019300	3.0750
1770.45	1214.60	0.040020	2.9060
1771.45	690.60	0.078030	2.8340
1772.45	398.90	0.091870	2.8180
1773.45	554.60	0.101760	3.0140
1774.45	593.10	0.075790	3.0910
1775.45	945.60	0.031480	2.9690
1776.45	1250.70	0.022900	3.0190
1777.45	1105.30	0.041980	2.9500
1778.45	882.80	0.067810	3.0180
1779.45	702.30	0.058750	2.9390
1780.45	689.00	0.079380	2.9940
1781.45	608.60	0.090750	2.9510
1782.45	593.60	0.081610	2.8660
1783.45	681.20	0.091250	2.9720
1784.45	692.70	0.103980	2.9430
1785.45	674.20	0.116430	2.9380
1786.45	653.00	0.109930	3.0370
1787.45	811.70	0.062260	3.1360
1788.45	1248.80	0.029430	3.1280
1789.45	1409.00	0.038540	3.0390
1790.45	1337.80	0.059520	3.1800
1791.45	1078.50	0.038600	3.0970
1792.45	931.90	0.044780	3.0290
1793.45	1013.60	0.045740	3.1220
1794.45	965.20	0.037140	3.1560
1795.45	967.90	0.036010	3.0690
1796.45	1187.00	0.036120	3.0880
1797.45	1206.00	0.037670	3.0530
1798.45	1167.50	0.044110	3.1470
1799.45	1104.20	0.041540	3.1940
1800.45	1057.80	0.036800	3.2920
1801.45	1020.60	0.033470	3.2550
1802.45	1048.70	0.029250	3.1300
1803.45	1123.40	0.027470	3.2010
1804.45	1253.90	0.026950	3.1470
1805.45	1397.00	0.025320	3.1320
1806.45	1469.00	0.023380	3.1480
1807.45	1537.40	0.023640	3.2650
1808.45	1737.00	0.024930	3.1680
1809.45	1923.00	0.026130	3.1510
1810.45	1962.60	0.026750	3.0940
1811.45	1923.70	0.027180	3.1750
1812.45	1865.00	0.032980	3.2180
1813.45	1711.90	0.037840	3.2640
1814.45	1393.30	0.034920	3.1390
1815.45	975.20	0.039610	3.2450
1816.45	1030.40	0.040060	3.2190
1817.45	1003.00	0.035170	3.2660
1818.45	878.40	0.032310	3.1550

1819.45	1170.10	0.022770	3.1380
1820.45	1332.70	0.031200	3.0250
1821.45	1591.10	0.047760	3.2180
1822.45	1662.80	0.026650	3.2560
1823.45	1064.80	0.029900	3.1120
1824.45	1308.50	0.060480	3.0520
1825.45	1449.80	0.093760	3.1470
1826.45	1257.10	0.053010	3.1280
1827.45	839.90	0.019560	3.1440
1828.45	1313.40	0.019930	3.1040
1829.45	2085.50	0.022410	3.1160
1830.45	2187.80	0.021850	3.1870
1831.45	2140.70	0.020810	3.0870
1832.45	1967.30	0.021580	3.2410
1833.45	2108.30	0.025590	3.1500
1834.45	2386.20	0.026570	3.1690
1835.45	1964.20	0.036420	3.1210
1836.45	1724.90	0.049040	3.0800
1837.45	1744.70	0.040710	3.0990
1838.45	1619.30	0.040030	3.0990
1839.45	1947.50	0.040860	3.1700
1840.45	1677.50	0.033030	3.1130
1841.45	1681.80	0.038920	3.1130
1842.45	1767.70	0.038860	3.1580
1843.45	1610.20	0.054540	3.0380
1844.45	1684.80	0.060100	3.1690
1845.45	1349.90	0.052900	3.1610
1846.45	1398.00	0.054980	3.1290
1847.45	2016.20	0.039800	3.0960
1848.45	2231.70	0.035710	2.7810

The results of calculation are given in Table 2 and the plots of these variograms are in Fig. 21. The variogram of porosity is very regular and for $h \geq 6$ m the values $\gamma^*(h \geq 6 \text{ m})$ approach the $s^2(\emptyset)$ value very well. This is not the case, however, for the bulk density variogram. The values $\gamma^*(h)$ for this variogram attain certain sill, which is far, however, from the $s^2(\rho)$ value. Moreover, the shape of the density variogram is not so regular as it was for the porosity variogram. There can be two reasons for such behaviour of the density variogram:

1). the maximum distance $h = 10$ m is too small to reach the variogram saturation level. In this case it means that the range of correlation for density is much higher than the one for the porosity, where it was $a = \text{range} = \sim 6$ m.

2). The density data are not representative for the stationary or intrinsic hypothesis process.

To check these two possibilities do Exercise No 2.

EXERCISE No 2

Using the same data as in Exercise No 1 and the program VARIOGRAM 10 calculate the density variogram, taking now $l = 4$ m. Do this variogram downwards and upwards.

Using the same set of data given in Table 1 we take now each fourth data as an input data for the VARIOGRAM 10. We run this program four times take each time as $k = 1$ the first, second, third and the fourth data, i.e.

	Run No 1	Run No 2	Run No 3	Run No 4
k				
1	2.738	2.789	2.687	2.690
2	2.733	2.727	2.777	2.780
3	2.758	2.810	2.885	2.841
etc.	etc.	etc.	etc.	etc.

Thus, we have four sets of the $\gamma^*(i)$ values: $\gamma_1^*(i)$, $\gamma_2^*(i)$, $\gamma_3^*(i)$ and $\gamma_4^*(i)$. The final result is an average value for a given i index:

$$\gamma^*(i) = \frac{1}{4}\{\gamma_1^*(i) + \gamma_2^*(i) + \gamma_3^*(i) + \gamma_4^*(i)\} \tag{63}$$

Table 2

Results of the variogram calculation for the data in Table 2.

Depth interval: 1708.45 - 1847.45 m . Variogram calculated downwards.

$\bar{z}$		5.621 %	3.0478 g cm^{-3}
$s^2(z)$		12.8499 $\%^2$	$20.3323 \cdot 10^{-3}$ g^2 cm^{-6}
N^1		140	140
N^1		130	130
h	$\gamma^*(i)$	$\%^2$	
(m)			$x10^{-3}$ g^2 cm^{-6}
1	$\gamma^*(1)$	2.8537	4.2830
2	$\gamma^*(2)$	6.9701	7.0337
3	$\gamma^*(3)$	9.7328	8.9541
4	$\gamma^*(4)$	11.4762	9.8031
5	$\gamma^*(5)$	12.2710	10.7828
6	$\gamma^*(6)$	12.3711	12.1516
7	$\gamma^*(7)$	12.5417	14.0251
8	$\gamma^*(8)$	12.7281	14.2962
9	$\gamma^*(9)$	12.6241	14.5804
10	$\gamma^*(10)$	12.3829	13.8520

Remark The programs in the FORTRAN language for computers are given in [1], [2] and [7]

7. Y. C. KIM, H. P. KNUDSEN : GEOSTATISTICAL ORE RESERVE ESTIMATION FOR A ROLL-FRONT TYPE URANIUM DEPOSIT. (PRACTITIONER's GUIDE). GJBX-3 (77) - GRAND JUNCTION DOE REPORT, JANUARY 1977.

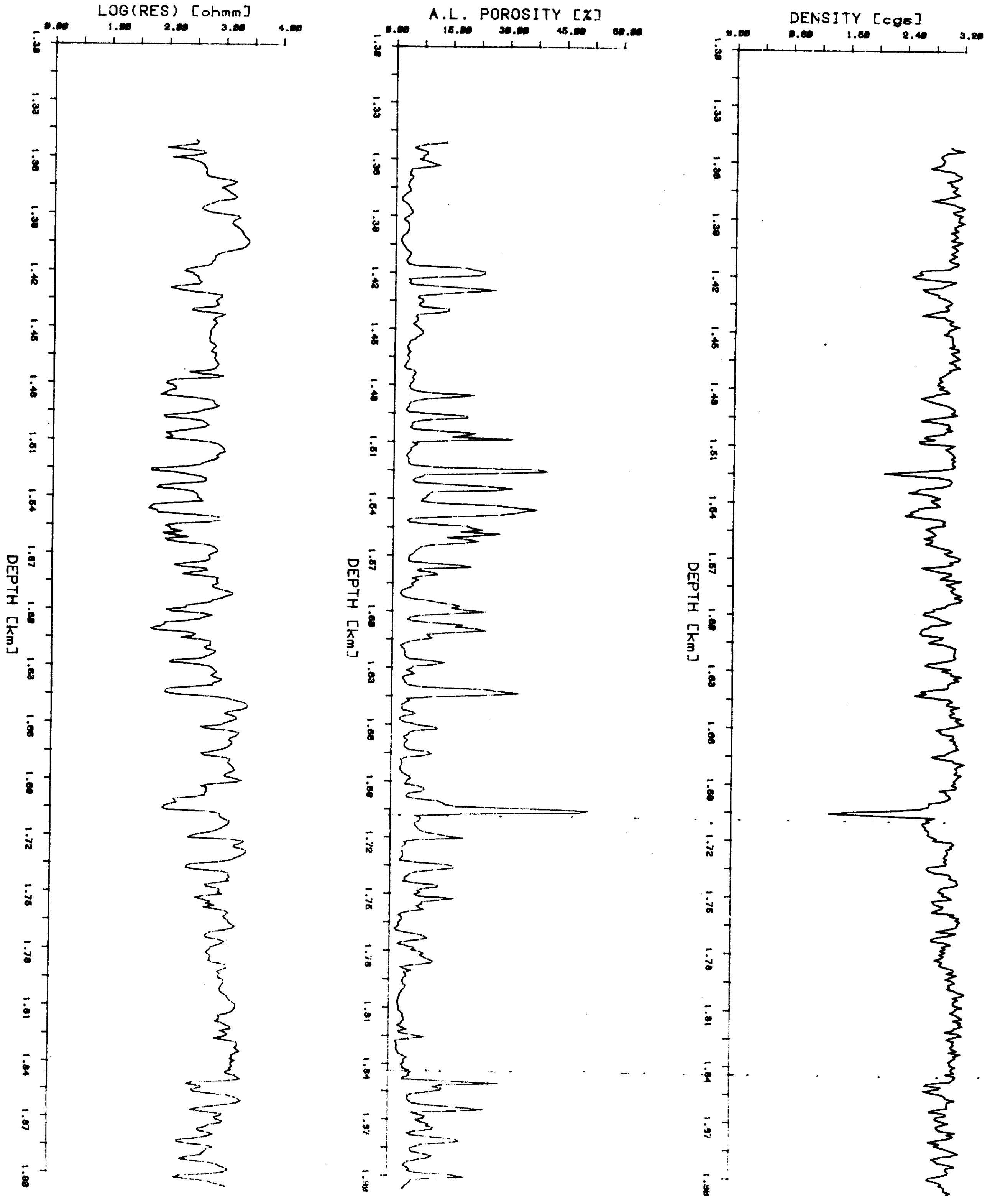

LOG(RES) [ohmm]
0.00
1.00
2.00
3.00
4.00
DEPTH [km]
A.L. POROSITY [%]
0.00
15.00
30.00
45.00
60.00
DEPTH [km]
DENSITY [cgs]
0.00
0.80
1.60
2.40
3.20
DEPTH [km]
1.30
1.33
1.36
1.39
1.42
1.45
1.48
1.51
1.54
1.57
1.60
1.63
1.66
1.69
1.72
1.75
1.78
1.81
1.84
1.87
1.90

Fig. 20

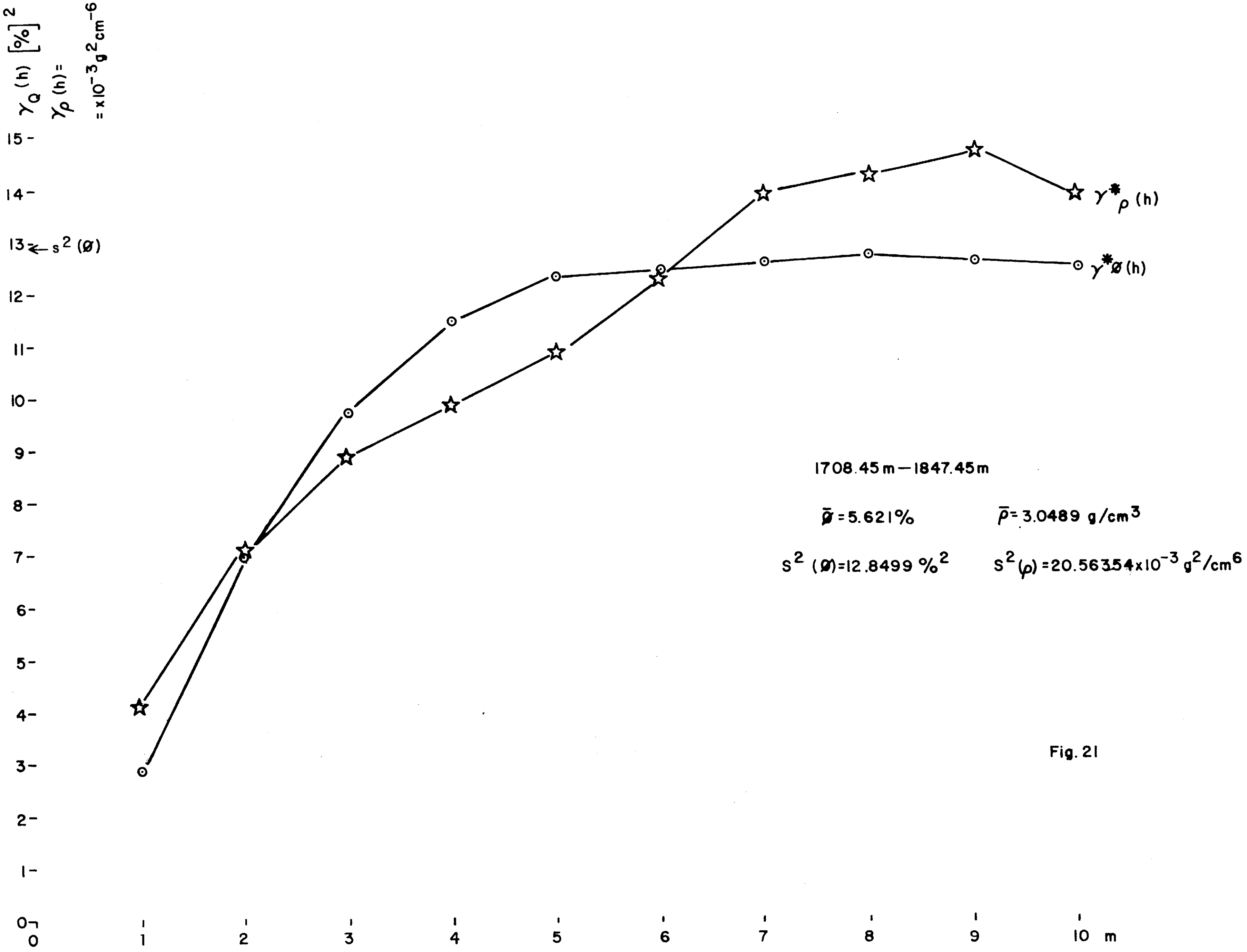

$\gamma_{\phi}(h)$ $[\%]^2$
$\gamma_{\rho}(h) =$
$= \times 10^{-3} g^2 cm^{-6}$
15
14
13 ← $s^2(\phi)$
12
11
10
9
8
7
6
5
4
3
2
1
0
0
1
2
3
4
5
6
7
8
9
10 m
$\gamma^*_{\rho}(h)$
$\gamma^*_{\phi}(h)$
1708.45 m — 1847.45 m
$\bar{\phi} = 5.621\%$
$\bar{\rho} = 3.0489$ g/cm3
$s^2(\phi) = 12.8499\ \%^2$
$s^2(\rho) = 20.56354 \times 10^{-3}\ g^2/cm^6$

Fig. 21

Repeat those calculations for the upward direction.

The results of calculation are given in Tables 3 and 4 and they are plotted in Fig. 22, where also for 1 = 2 m the values of the average variogram have been plotted.

Table 3

Results of the variogram of the bulk density calculation for the data in Table 1.

Depth interval 1708.45 - 1847.45 m . Downward variogram

$\bar{\rho}$	g cm^{-3}	3.0422	3.0504	3.0594	3.0395	3.0479
$s^2(\rho)$	g^2 cm^{-6}	$22.2294\cdot10^{-3}$	20.8362	18.9823	20.8345	20.7206
N		35	35	35	35	140
N^1		25	25	25	25	100
h	$\gamma(i)$	$\gamma_1^*(i)$	$\gamma_2^*(i)$	$\gamma_3^*(i)$	$\gamma_4^*(i)$	$\gamma^*(i)$
(m)	i		$x10^{-3}$ g^2 cm^{-6}			
4	1	13.8833	11.7825	11.5882	10.1077	11.8404
8	2	17.4517	17.6495	17.2920	17.0817	17.3687
12	3	20.8106	18.1939	14.8486	18.7237	18.1442
16	4	20.0587	17.4335	13.9190	12.1748	15.8965
20	5	14.4296	12.6248	12.9968	13.3225	13.3434
24	6	19.6047	14.3967	14.7707	17.0242	16.4491
28	7	18.8380	18.3681	18.3275	22.6167	19.5336
32	8	21.5341	19.1273	15.3088	19.3046	18.8187
36	9	14.3060	17.6192	10.4331	16.8680	14.8066
40	10	12.9682	13.1157	16.6064	16.8174	14.8769

Table 4

Results of the variogram calculations for the bulk densities for data in Table 1.

Depth inverval 1708.45 - 1847.45 m
Upward variogram.

N^1		25	25	25	25	100
h	$\gamma(i)$	$\gamma_1^*(i)$	$\gamma_2^*(i)$	$\gamma_3^*(i)$	$\gamma_4^*(i)$	$\gamma^*(i)$
(m)	i		x 10^{-3} g^2	cm^{-6}		
4	1	9.9013	5.5263	5.8101	5.6832	6.7302
8	2	14.7326	9.0277	7.9113	10.7307	10.6005
12	3	16.3118	10.4313	7.1634	10.5101	11.1041
16	4	14.8986	14.7162	10.2120	9.1780	12.2512
20	5	12.0684	10.7972	9.8776	10.5290	10.8180
24	6	15.0489	9.6220	10.3734	13.3975	12.1104
28	7	14.2367	13.1690	11.3309	17.4886	14.0560
32	8	16.7817	15.5397	9.7575	15.1249	14.3009
36	9	13.2754	17.2065	9.1836	15.1387	13.7010
40	10	12.9682	13.1157	16.5069	16.7685	14.8398

Remark: Note please a significant difference between the variograms calculated downward and upward. It means that the bulk density data set contains some trend (drift).

Discussion of results

Porosity ϕ has a very regular variogram (cf. Fig. 21) which we shall call the spherical variogram (will be explained in the following lectures). The range of correlation for porosity is of the order of 5 or 6 m . It means that samples taken at closer distances influence each other.

The variogram of bulk density shows some nested structure (at h=20 m) and depends strongly upon the direction (downwards or upwards) at which it is calculated. The experimental variance $s^2(\rho)$ is always much higher than the values reached by the variogram at any point. These facts give the reason to the hypothesis that on a long range the bulk density data contain some drift. On the other hand, looking on the beginning of the variogram $\gamma_\rho^*(h)$ only, one can see that up to the range of about 10 meters the data are intercorrelated. In this case we can say that the bulk density data follow the intrinsic hypothesis locally (within the range of ~ 10 m).

There is another remark too. If one assumes the constant matrix density ρ_M , the bulk density is given as:

$$\rho = (1-\phi)\ \rho_M + \phi \cdot \rho_w = \rho_M - \phi \cdot (\rho_M - \rho_w) \qquad (64)$$

where ρ_w is the water density. Because ρ and ϕ are depth dependent, we can write

$$\rho(x) = \rho_M - \phi(x) \cdot (\rho_M - \rho_w) \qquad (64a)$$

and inserting this expression into the formula (51) for the variogram we have:

$$\gamma_\rho(h) = (\rho_M - \rho_w)^2 \cdot \gamma_\phi(h) \qquad (65)$$

which means that the variogram of bulk density

$$\gamma_\rho(h)$$

should follow up to the constant multiplication factor $(\rho_M - \rho_w)^2$ the variogram $\gamma_\phi(h)$ of the porosity. If it is not the case it can mean that either the rock matrix density is not constant or the bulk density or the porosity data or even both, are wrong. In this particular

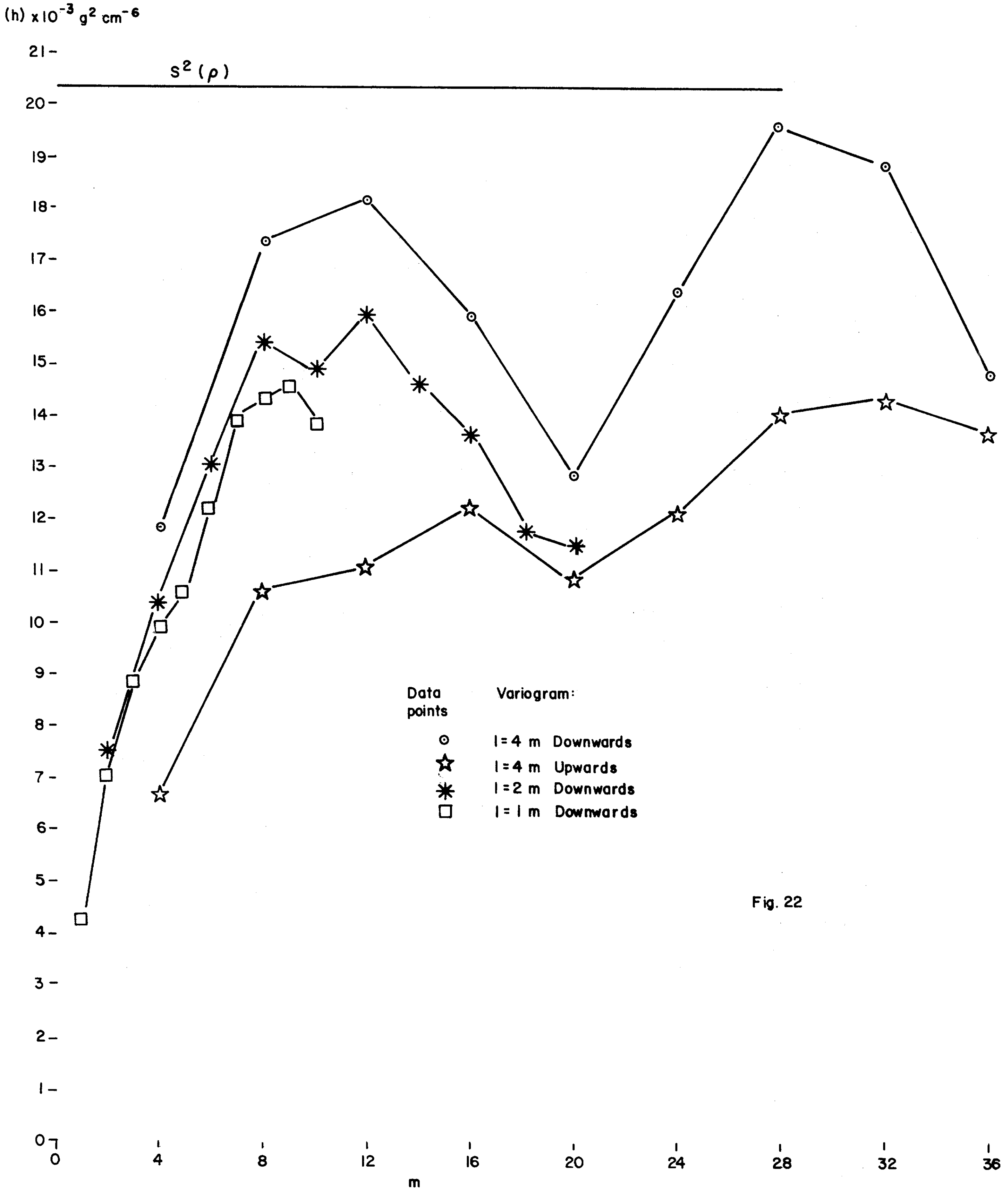

(h) x 10^{-3} g^2 cm^{-6}
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
S^2 (ρ)
0
4
8
12
16
20
24
28
32
36
m
Data points
Variogram:
⊙ l = 4 m Downwards
☆ l = 4 m Upwards
✳ l = 2 m Downwards
□ l = 1 m Downwards

Fig. 22

case we can opt for the hypothesis that the bulk density data are probably not correct. Just to emphasize this hypothesis one can calculate from Eq. (65) the ρ_M value as being:

$$\rho_M = \sqrt{\frac{\gamma_\rho(h)}{\gamma_\phi(h)}} + \rho_w \qquad (66)$$

which for the data given in Table 2 gives the result which is plotted in Fig. 23 (for ρ_w = 1.0)

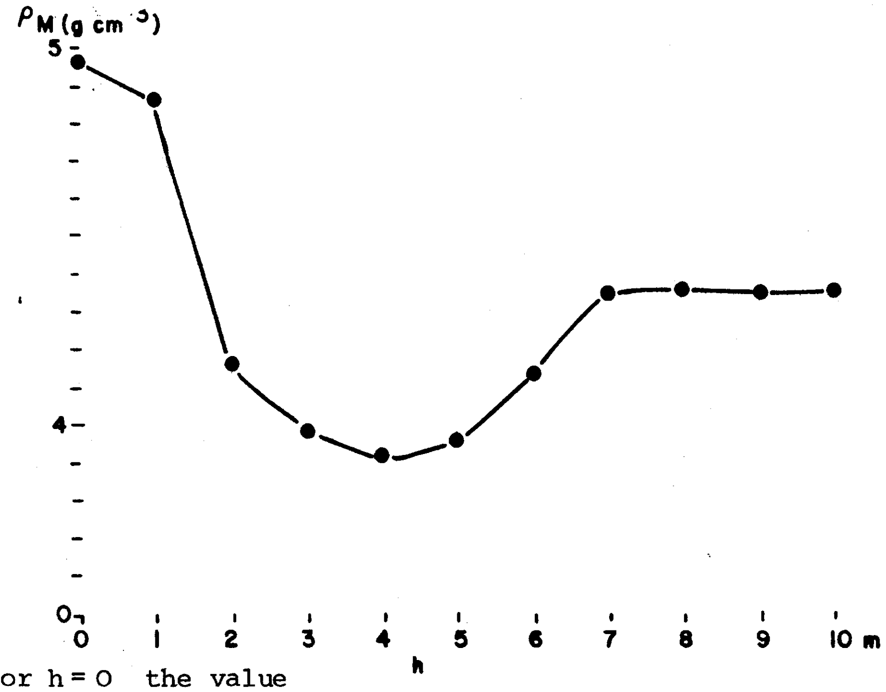


Fig. 23

For h = 0 the value

$$\rho_M = \sqrt{\frac{s^2(\rho)}{s^2(\phi)}} + \rho_w \qquad (67)$$

is plotted.

The discussion of the reliability of the plot given in Fig. 23 is out of the scope of this lecture and will be done elsewhere.

LECTURE 4

J.A. Czubek

LECTURE 4

Reykjavík, March 8, 1981

ESTIMATION PROBLEMS

STATEMENT OF THE PROBLEM:

There exist two blocks, say V and v (usually the letter V or v stands for a volume in the three-dimensional case, surface in the two-dimensional case and the line in the one-dimensional case). We can assume (but it is not necessary) that $v < V$. The volume v can be inside or outside of the volume V

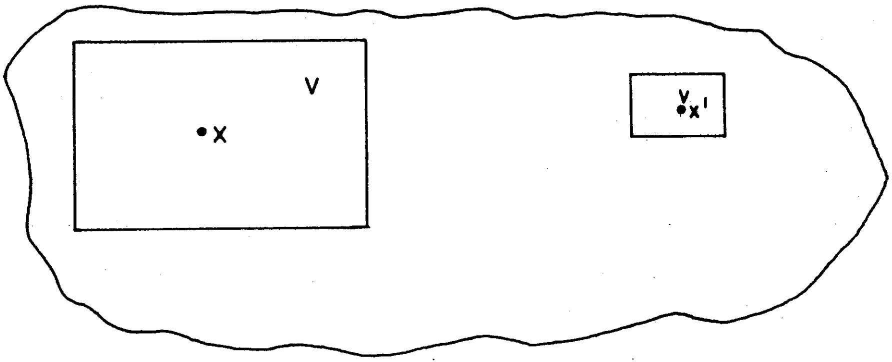


Fig. 24

Both volumes V and v are inside a given geological formation of interest which is a particular realization $z(x)$ of a certain random function $Z(x)$ having the intrinsic hypothesis (at least). The volume V is centered at the point x, whereas the volume v at the point x^1.

The expected values of the random function $Z(x)$ within the blocks V and v are

$$E\{Z_V(x)\} = \frac{1}{V} \int_{V(x)} Z(x)\,dx \tag{68}$$

and

$$E\{Z_V(x^1)\} = \frac{1}{v} \int_{v(x^1)} Z(x)\,dx \tag{69}$$

and because they are taken under the stationarity hypothesis (or intrisic) they are equal each other:

$$E\{Z_V(x)\} = E\{Z_V(x^1)\} = E\{Z(x)\} = m \tag{70}$$

For a given particular realization of the random variable $Z(x)$, however, we can at least know the values:

$$z_V(x) = \frac{1}{V} \int_{V(x)} z(x)\,dx \tag{71}$$

and

$$z_v(x^1) = \frac{1}{v} \int_{v(x^1)} z(x)\,dx \tag{72}$$

which are the experimentally known $z(x)$ values averaged over the volume V and v, respectively. Note, please, that dx in Eqs. (68) to (72) has a meaning of length, surface or volume (elementary) whether the problem is one-, two- or three-dimensional.

Now we attribute the value $z_v(x)$ to the volume $V(x)$, which is admissible in average because of the validity of equation (70). But for this particular realization of the random function what is the error in such estimation of the $z_V(x)$ value by the $z_v(x^1)$ value ? This error R will be, of course:

$$R(x) = z_V(x) - z_v(x^1) \tag{73}$$

but we are unable to calculate it unless the value $z_V(x)$ is known. When it is not the case, however, we can try to find the expected value of its square:

$$E\{R^2(x)\} = E\{[z_V(x)-z_v(x^1)]^2\} = \sigma_E^2[V(x),v(x^1)] \tag{74}$$

which we shall call the estimation variance of the volume $V(x)$ by the volume $v(x^1)$.

Developing Eq. (74) and inserting the meanings of the $z_V(x)$ and $z_v(x^1)$ values according to the equations (71) and (72) gives:

$$\sigma_E^2\,[V(x),\ v(x^1)] = E\{z_V^2(x)\} + E\{z_v^2(x^1)\} - 2E\{z_V(x)\cdot z_v(x^1)\}$$

$$= E\{\frac{1}{V} \cdot \int_{V(x)} z(x_1)\,dx_1 \cdot \frac{1}{V} \int_{V(x)} z(x_2)\,dx_2\} +$$

$$+ E\{\frac{1}{v} \int_{v(x^1)} z(x_1^1)\,dx_1^1 \cdot \frac{1}{v} \int_{v(x^1)} z(X_2^1)\,dx_2^1\} -$$

$$- 2\,E\{\frac{1}{V} \int_{V(x)} z(x_1)\,dx_1 \cdot \frac{1}{v} \int_{v(x^1)} z(x_2^1)\,dx_2^1\} \qquad (75)$$

Here, to emphasize that each integration of the variable $z(x)$ is independent each other we have introduced the integration variables x_1 , x_2 , x_1^1 and x_2^1 .

We can interchange the expectation and the integration signs (they are independent each other), which gives:

$$\sigma_E^2\,[V(x), v(x^1)] = \frac{1}{V^2} \int_{V(x)} dx_1 \int_{V(x)} dx_2\; E\{z(x_1) \cdot z(x_2)\} +$$

$$+ \frac{1}{v^2} \int_{v(x^1)} dx_1^1 \int_{v(x^1)} dx_2^1\; E\{ z(x_1^1) \cdot Z(x_2^1)\} =$$

$$- \frac{2}{V \cdot v} \int_{V(x)} dx_1 \int_{v(x^1)} dx_2^1\; E\{z(x_1) \cdot z(x_2^1)\} \qquad (76)$$

But what are the expected values of the product of random variable taken at two different points? According to equations (29) and (31) they are:

$$E\{ z(x_1) \cdot z(x_2)\} = C(x_1 - x_2) + m^2 =$$

$$= C(0) + m^2 - \gamma\,(x_1 - x_2) \qquad (77)$$

Insertion of eq. (77) in Eq. (76) gives (remember that $C(0)$ and m^2 are constants):

$$\sigma_E^2\,[V(x), v(x^1)] = \frac{2}{V \cdot v} \int_{V(x)} dx_1 \int_{v(x^1)} dx_2^1\; \gamma(x_1 - x_2^1) -$$

$$- \frac{1}{V^2} \int_{V(x)} dx_1 \int_{V(x)} dx_2\; \gamma(x_1 - x_2) - \frac{1}{v^2} \int_{v(x^1)} dx_1^1 \int_{v(x^1)} dx_2^1\; \gamma\,(x_1^1 - x_2^1) \qquad (78)$$

Thus, we have obtained an expression for the estimation variance of the block V by the block v where the average values of the variogram $\gamma(h)$ appear. The averaging of a variogram is done by this way that the two extremities of the distance h:

$$|h| = x_1-x_2 \tag{79}$$

are "walking" independent each other throughover the volume V or v depending upon which integral is involved in Eq.(78). As a matter of fact three operations of the double integration are involved here:

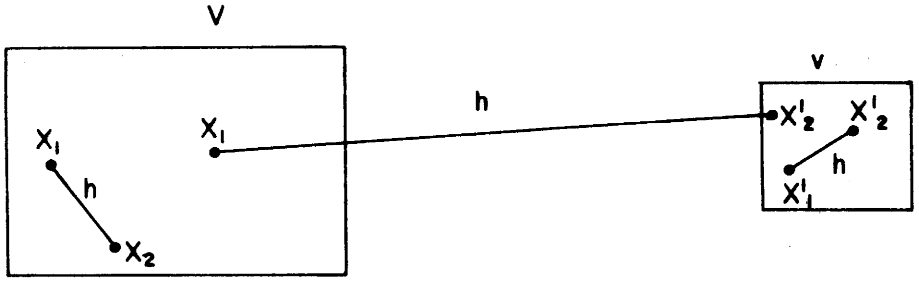


Fig. 25

and in the case of the three-dimensional space it means sextuple integrations which are involved in Eq.(78).

DISPERSION VARIANCE

Equation (78) is a most general formula giving the estimation variance and it will serve to calculate all other particular cases. One of them is the problem of the dispension variance:

We take again two volumes V and v assuming

$$v << V$$

and we put v inside V, that is:

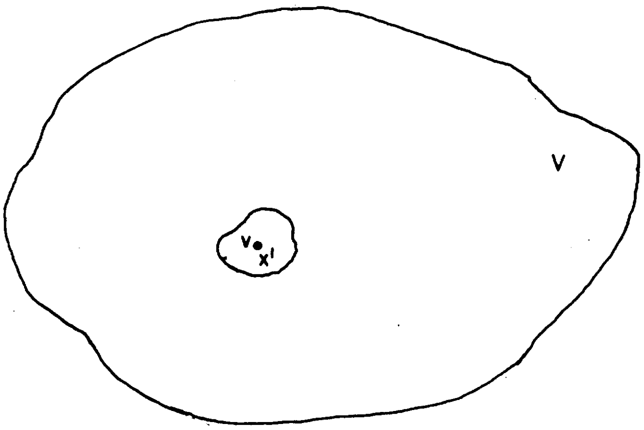


Fig. 26

When we want to express the z_V value using the $z_v(x^1)$ value, the variance of such estimation is just given by Eq. (78). But what will be if we move the volume v towards another point x^{11}, again and again? What will be the estimation variance of such procedure when the volume v will "walk" all possible positions inside V ? The physical meaning of such procedure is just equivalent to the question, what is the variance of the z_v value inside the field V ? Mathematically one has again to calculate the average value of

$$\sigma_E^2 [\, V(x),\ v(x^1)\,]$$

averaged over all positions of v inside V , that is:

$$D(v/V) = \frac{1}{V} \int_{V(x)} dx^1 \, \sigma_E^2 \, [V(x),\ v(x^1)\,] \qquad (80)$$

Here D(v/V) is called the dispersion variance of v inside V.

To calculate the integral in Eq. (80) let us rewrite Eq. (78) using the notation of "average" instead of integration:

$$\sigma_E^2 \, [V(x),\ v(x^1)\,] = 2\,\bar{\gamma}\,[V(x),\ v(x^1)\,] - \bar{\gamma}\,[V(x),\ V(x)\,] - $$

$$- \bar{\gamma}[v(x^1),\ v(x^1)\,] \qquad (81)$$

The functions inside the brackets [,] indicate on which volumes the variogram is averaged. As the random variable Z(x) is stationary inside the field V, the averaged variogram values $\bar{\gamma}$ in the last two terms do not depend upon the position x^1, thus:

$$\left.\begin{array}{ll} & \bar{\gamma}\,[V(x),\ V(x)\,] \equiv \bar{\gamma}\,(V,\ V) \\ \text{and} & \\ & \bar{\gamma}\,[v(x^1),\ v(x^1)\,] \equiv \bar{\gamma}\,(v,\ v) \end{array}\right\} \qquad (82)$$

and as a consequence, these two terms remain invariant when being taken on average over the field V(x), as it is needed by Eq. (80). The first term in (81) when inserted into (80) becomes:

$$\frac{1}{V} \int_{V(x)} dx^1 \cdot \bar{\gamma}\,[V(x),\ v(x^1)\,] =$$

$$= \frac{1}{V} \cdot \frac{1}{Vv} \int_{V(x)} dx^1 \int_{V(x)} dx_1 \int_{v(x^1)} dx_2^1 \, \gamma(x_1 - x_2^1) =$$

$$= \frac{1}{V} \int_{V(x)} dx_1 \frac{1}{V} \int_{V(x)} dx^1 \frac{1}{v} \int_{v(x^1)} dx_2^1 \, \gamma\, (x_1 - x_2^1) =$$

$$= \frac{1}{V} \int_{V(x)} dx_1 \cdot \frac{1}{V} \int_{V(x)} dx_2^1 \, \gamma(x_1 - x_2^1) = \bar{\gamma}\, (V,\ V) \qquad (83)$$

Finally equation (80) becomes:

$$D(v/V) = \bar{\gamma}(V,\ V) - \bar{\gamma}(v,\ v) =$$

$$= \frac{1}{V^2} \int_V dx_1 \int_V dx_2 \, \gamma(x_1 - x_2) - \frac{1}{v^2} \int_v dx_1 \int_v dx_2 \, \gamma(x_1 - x_2) \qquad (84)$$

Equation (84) has a very important sense. When v is attributed to the volume of the sample and V to the volume of the whole geological formation in question (deposit etc.), the first integral becomes a constant:

$$A = \frac{1}{V^2} \int_V dx_1 \int_V dx_2 \, \gamma(x_1 - x_2) \qquad (85)$$

Usually the knowledge of the constant A is not really needed in practice and even when the mathematical form of the function $\gamma(h)$ is known, the calculation of the integrals in (85) are very often troublesome because of the uncertainity in the volume V. This integral can be, however, obtained in another way:

When some formation is sampled by a set of N samples having each the volume v and the space distribution of samples is either completely random or strictly regular we can approach the D(v/V) value by a well known estimator:

$$D(v/V) \approx s_v^2(Z) = \frac{1}{N-1} \sum_{K=1}^{N} (z_k - \bar{z})^2 \qquad (86)$$

where z_k are the values of the random variable Z, measured on the samples having each the volume v. Thus, in this case

$$A = s_v^2(Z) + \frac{1}{v^2} \int_v dx_1 \int_v dx_2 \, \gamma(x_1 - x_2) \tag{87}$$

which demands now the knowledge of the $\gamma(h)$ function only, the volume v being known.

The main importance of Eq. (84) lies however in another point. Assume the whole geological field V is divided into N_1 blocks of volume V_1 each and next each block is again divided into N_2 small blocks having the volume V_2 each. Thus we have:

$$V = N_1 \cdot V_1 = N_1 \cdot N_2 \cdot V_2 \tag{88}$$

and we assume that both N_1 and N_2 are large.
According to Eq (84) we can write:

$$D(V_2/V) = \frac{1}{V^2} \int_V dx_1 \int_V dx_2 \, \gamma(x_1 - x_2) - \frac{1}{V_2} \int_{V_2} dx_2 \, \gamma(x_1 - x_2) =$$

$$= \frac{1}{V^2} \int_V dx_2 \, \gamma(x_1 - x_2) - \frac{1}{V_1^2} \int_{V_1} dx_1 \int_{V_1} dx_2 \, \gamma(x_1 - x_2) +$$

$$+ \frac{1}{V_1^2} \int_{V_1} dx_1 \int_{V_1} dx_2 \, \gamma(x_1 - x_2) - \frac{1}{V_2^2} \int_{V_2} dx_1 \int_{V_2} dx_2 \, \gamma(x_1 - x_2) =$$

$$= D(V_1/V) + D(V_2/V_1) \tag{89}$$

Thus the dispersion variance of the small samples within the whole field is the same of the variance of these samples within the blocks plus the variance of these blocks within the whole field. This property, first observed experimentally by D.G.
Krige has been explained mathematically in the way as above, by Matheron.

Another problem which can be solved using Eq. (84) is following:

Some geological formation has been sampled by a set of samples each having the volume v_1. One has the experimental dispersion variance

$$s^2_{v_1}(2)$$

How will the dispersion variance be if the same formation is sampled by another set of samples each having another volume v_2 ?

This new dispersion variance is

$$s^2_{v_2}(z) \approx D(v_2/V) \tag{90a}$$

and together with the first dispersion variance

$$s^2_{v_1}(z) \simeq D(v_1/V) \tag{90b}$$

by a single subtraction and taking into account Eq. (84) one obtains:

$$s^2_{v_2}(z) - s^2_{v_1}(z) =$$

$$= \frac{1}{v_1^2}\int_{v_1}\int_{v_1} \gamma(x_1-x_2)\,dx_1\,dx_2 - \frac{1}{v_2^2}\int_{v_2}\int_{v_2} \gamma(x_1-x_2)\,dx_1 dx_2 \tag{91}$$

Thus, knowing the variogram $\gamma(h)$, the volumes v_1 and v_2 and the value of $s^2_{v_1}(z)$ from the experiment the value of a new dispersion variance

$$s^2_{v_2}(z)$$

can be readily calculated from Eq. (91). This is just the case when one wants to compare logging data with core sample data obtained in the laboratory (cf. Fig. 9).

LECTURE 5

J.A. Czubek

LECTURE 5

Reykjavík, March 9, 1981

MODELS FOR VARIOGRAMS

Although the variograms are calculated using the experimental data (cf. Eq. 60), the knowledge of its theoretical form is of a great importance in all geostatistical practice and calculations.

Generally speaking we have a few theoretical models to present the semi-variogram $\gamma(h)$, the most important of them are:

LINEAR MODEL:

$$\gamma(h) = A|h| + C_o \tag{92}$$

SPHERICAL MODEL:

$$\gamma(h) = C \cdot \left[\frac{3}{2}\left|\frac{h}{a}\right| - \frac{1}{2}\left(\left|\frac{h}{a}\right|\right)^3\right] + C_o \qquad h < a$$

$$= C + C_o \qquad h \geq a \tag{93}$$

EXPONENTIAL MODEL:

$$\gamma(h) = C_1 \left[1 - \exp\left(- \left|\frac{h}{a_1}\right|\right)\right] + C_o \tag{94}$$

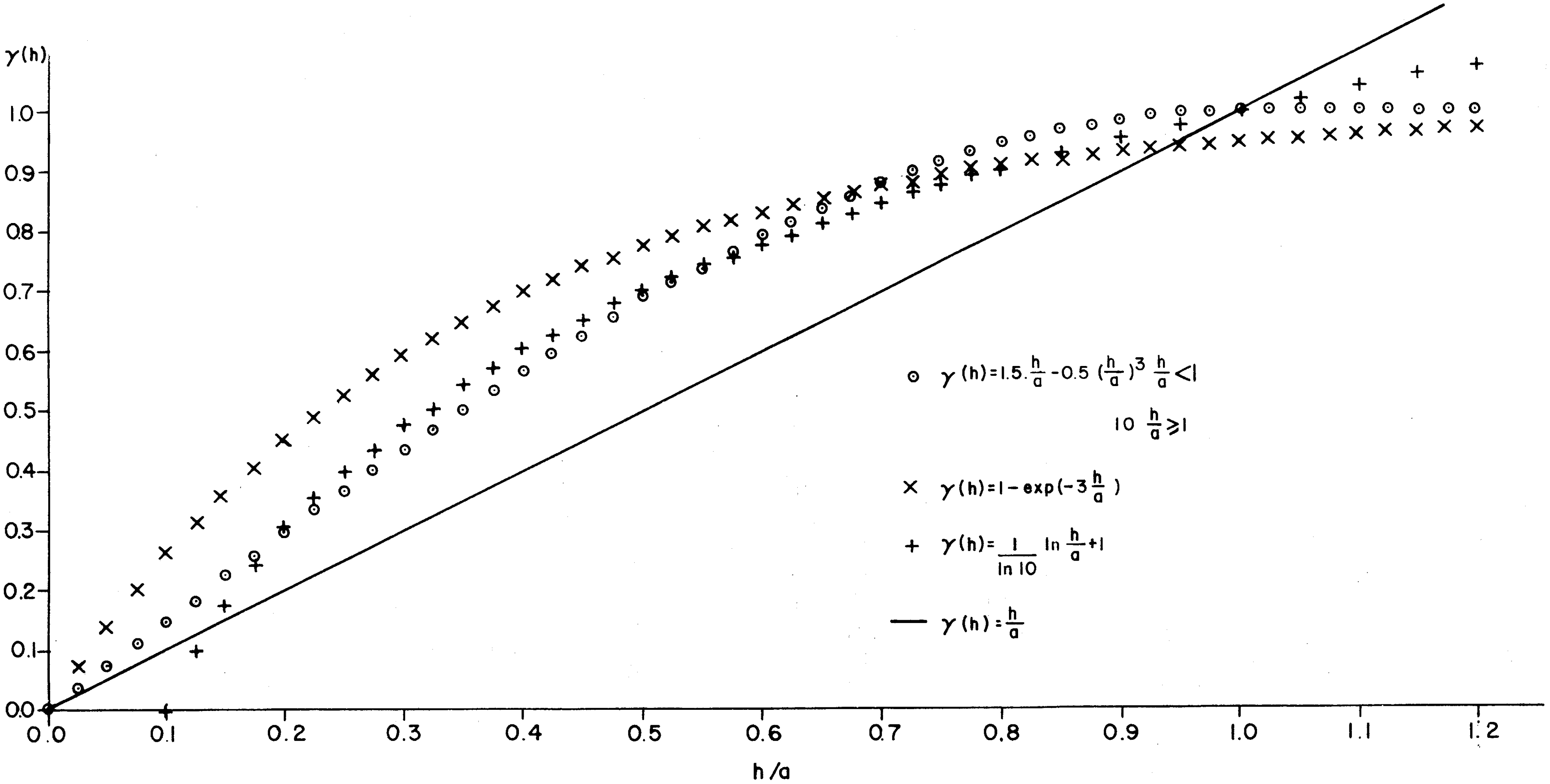
De Wijs MODEL:

$$\gamma(h) = a_2 \ln h + b_2 \tag{95}$$

In these formulas the C_o is the nugget effect value and C, C_1, a, a_1, a_2, b_2 are some constants characterizing the behaviour of variograms. The behaviour of these variograms is given in Fig. 27.

The linear and the de Wijs models have the variograms increasing to infinity for the infinite h values. Moreover the de Wijs model has a negative infinite value of $\gamma(h)$ when $h \to o$. This model cannot be used for point-like samples.

Fig.27. Behaviour of different models for variograms



In the spherical model the constant a is called the range of a variogram. For the lags h larger than the range a the variogram has a constant value C (the so-called sill) which means that the samples at distances $h > a$ are independent of each other. The name "spherical" is due to the fact that when one takes two spheres with radius R each and when the centers of spheres are at distance from each other the volume of the first sphere which is intersected by the other one (in relative units) is

$$\frac{V}{V_1} = 1 - \frac{3}{2}\,\frac{h}{2R} + \frac{1}{2}\left(\frac{h}{2R}\right)^3 = 1 - \frac{1}{C}\,\gamma(h) \qquad (96)$$

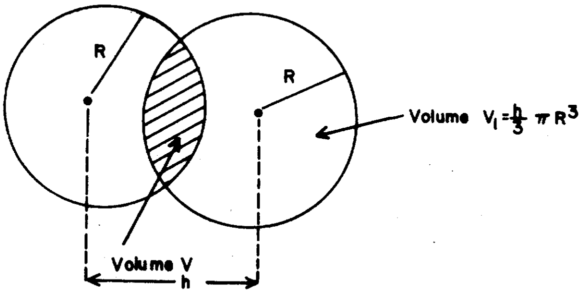


Fig. 28

In the exponential model the variogram reaches the value C_1 asymptotically, but the practical rule is that for the distance

$$h \simeq 3a_1 \qquad (97)$$

this limit value C_1 is practically reached. Thus, for the exponential model the distance $h = 3a_1$ is called the range.

The theoretical models of variogram presented above concern the point-like samples which never occur in geological practice, of course. The random variable $Z(x)$ is always regularized (i.e. averaged) over the volume v of the sample, that is

$$Z_v(x) = \frac{1}{v} \int_{v(x)} Z(y)\,dy \qquad (98)$$

where the volume v is centered at the point x. To calculate the experimental variogram we are using the $Z_v(x)$ data, thus the question about the regularized variogram $\gamma_v(h)$ arises and how it is related to the theoretical variogram $\gamma(h)$.

Starting from the definition of variogram one has:

$$2\gamma_v(h) = E\{[Z_v(x+h) - Z_v(x)]^2\} \qquad (99)$$

To calculate this expected value we can consider the expression (99) as the variance of the estimation of the mean grade $Z_v(x)$ by the mean grade $Z_v(x+h)$ separated by the vector h - cf. Eq. (79),(78) and Eq. (81). Treating the problem in this way we have:

$$2\gamma_v(h)=2\bar{\gamma}[v(x),\ v(x+h)]-\bar{\gamma}\ [v(x),\ v(x)]-\bar{\gamma}\ [v(x+h),\ v(x+h)] \qquad (100)$$

Since the point semi-variogram $\gamma(h)$ is stationary (i.e. it has in average for a fixed h value the same $\gamma(h)$ value wherever the vector h is situated), the last two terms in Eq. (100) are equal to each other and thus:

$$\gamma_v(h) = \bar{\gamma}\ [v(x),\ v(x+h)] - \bar{\gamma}\ (v,\ v) \qquad (101)$$

In this expression $\bar{\gamma}[v(x),\ v(x+h)]$ means an average value of the variogram calculated for the case when one extremity of the vector h is "walking" inside the volume sample v centered at x and the other end of this vector describes the same volume v but centered at the point x+h , as it is shown in Fig. 29

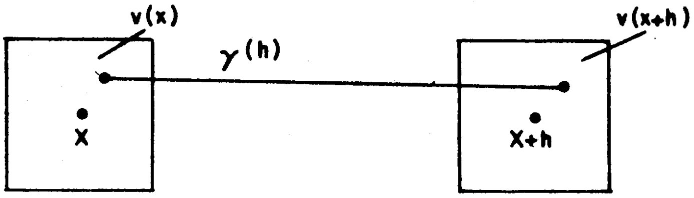


Fig. 29

The $\bar{\gamma}(v,\ v)$ is the same as in Eq. (82). When the dimension of v is small compared to the distance h between the samples, we can use an approximation:

$$\bar{\gamma}[v(x),\ v(x+h)] \simeq \gamma(h) \qquad (102)$$

and in this case the regularized variogram is

$$\gamma_v(h) \simeq \gamma\ (h) - \bar{\gamma}(v,\ v) \qquad (103)$$

Formula (103) can be applied for example, for the variogram of the core samples when the distance between the samples is much greater than the dimension of the sample itself.

Just in the case of borehole, when the core segments of the length 1 and the cross section s is small enough in comparison with 1 (i.e. $\sqrt{s} << 1$), the regularized random variable $Z(x)$ is:

$$Z_v(x) \simeq Z_l(k) = \frac{1}{l} \int_{l(x)} Z(y)\,dy \qquad (104)$$

and in this case the regularized variogram is

$$\gamma_l(h) = \bar{\gamma}\,[l(x),\ l(x+h)] - \bar{\gamma}\,(l, l) \qquad (105)$$

for the situation presented in Fig. (30)

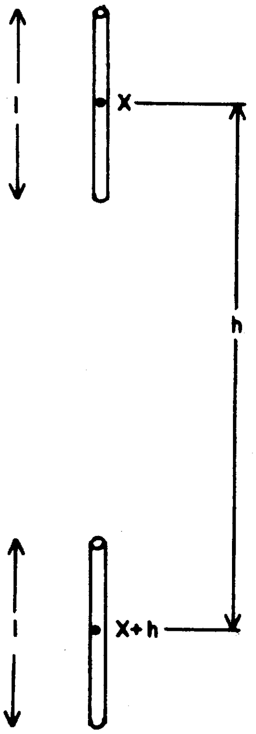


Fig. 30.

Now let us try to calculate the regularized $\gamma_l(h)$ variograms for the four models presented above.

The linear model

We take, for brevity, Eq. (92) in the form:

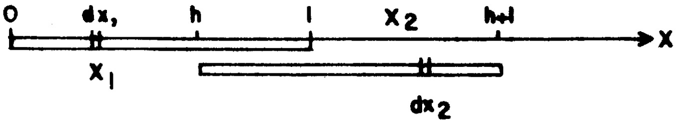
$$\gamma(h) = |h| \qquad (92.a)$$

Thus, according to Eq. (105) one has:

$$\bar{\gamma}\,[l(x),\ l(x+h)] = \frac{1}{l^2} \int_o^l dx_1 \int_h^{h+l} |\,x_1 - x_2\,|\, dx_2 \qquad (106)$$

which one has to solve for the two situations:

when the cores are overlying



$|h| < l$

Fig. 31

when the cores are apart of each other.

Fig. 31

$h > l$

Solving the integral in Eq. (106) for both cases one has:

$$\bar{\gamma}[l(x),\ l(x+h)] = \frac{1}{3}\ \frac{h^2}{l^2}\ (3\ l-h) + \frac{1}{3} \quad \text{for } |h| \leq l \tag{107}$$

and

$$\bar{\gamma}[l(x),\ l(x+h)] = h \qquad \text{for} \quad |h| \geq l \tag{108}$$

the special case of equation (107) for $h = o$ gives

$$\bar{\gamma}[l(x),\ l(x)] = \bar{\gamma}\ (l,\ l) = \frac{1}{3} \tag{109}$$

which is just the second term in Eq. (105). Thus, finally one has:

$$\gamma_1(h) = \begin{cases} \dfrac{h^2}{3\,l^2}\ (3\ l-h) & \text{for } |h| \leq l \\[2ex] h - \dfrac{1}{3} & \text{for } |h| \geq l \end{cases} \tag{110}$$

The regularized variogram $\gamma_1(h)$ is depicted in Fig. 32. Because for the experiment the region $|h| \leq l$ is practically unaccessible, the regularized linear variogram gives an apparent negative nugget effect for extrapolated value towards $h = 0$. This apparent nugget effect is

$$\gamma_1(0) = -\frac{1}{3} \tag{111}$$

Note that for $|h| \geq l$ the approximated formula (103) gives a strict equation.

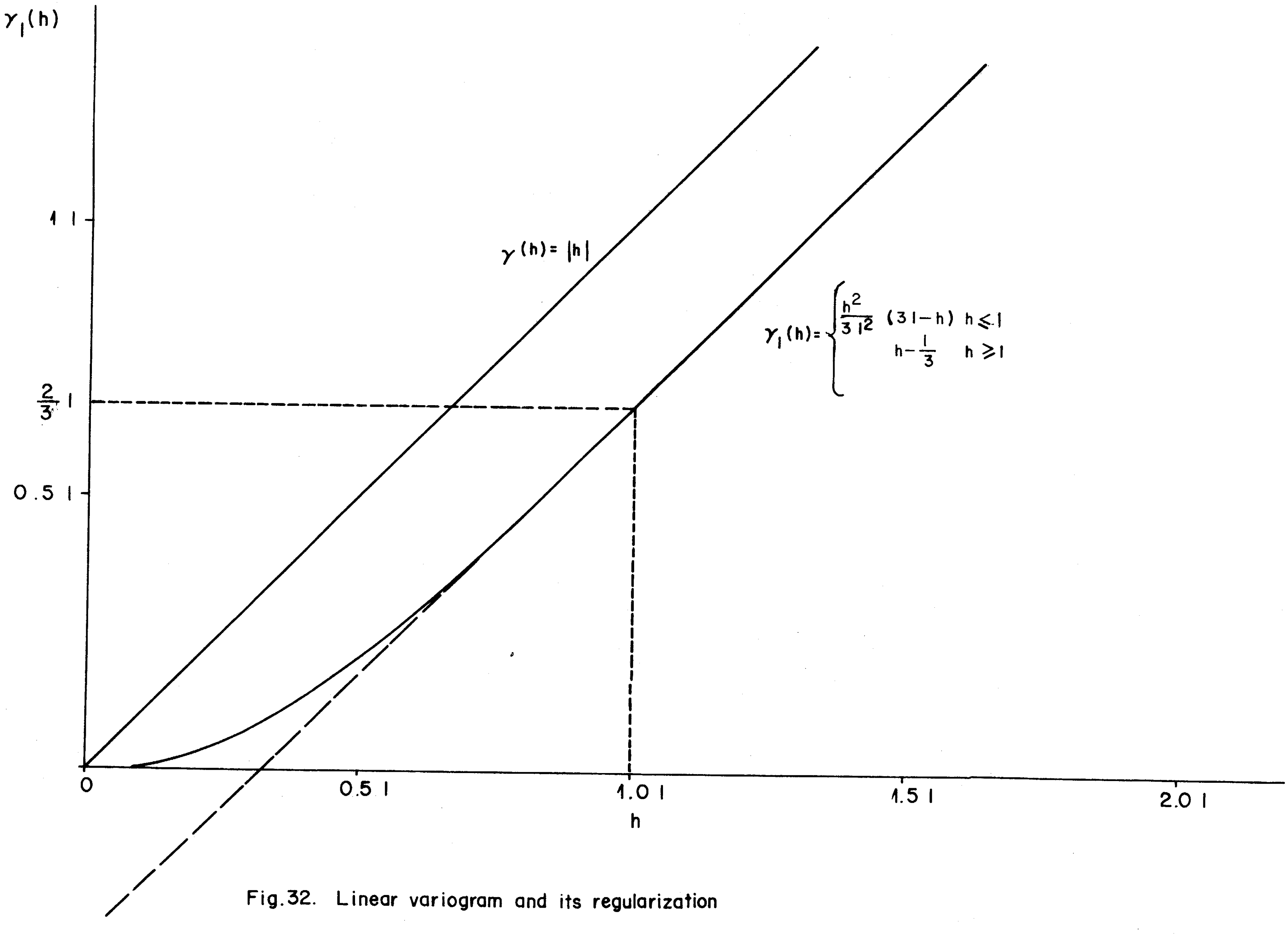


Fig.32. Linear variogram and its regularization

The De Wijs model

Repeating now the same kind of integration for the De Wijs model but written, for the sake of simplicity in the form

$$\gamma(h) = \ln |h| \tag{112}$$

one obtains for the regularized $\gamma_1(h)$ value:

for $h \leq 1$ $\frac{h}{1} = t$ $o \leq t \leq 1$

$$\gamma_1(h) = \gamma_1(t) = \frac{t^2}{2} \ln(\frac{1}{t^2} - 1) + \frac{1}{2} \ln (1-t^2) + t \ln \frac{1+t}{1-t} \tag{113}$$

here one has

$$\bar{\gamma}(1, 1) = \ln 1 - \frac{3}{2} \tag{114}$$

and for $h \geq 1$ $\frac{h}{1} = t'$ $t' \geq 1$

$$\gamma_1(h) = \gamma_1(t') = \frac{t'^2}{2} \ln (1- \frac{1}{t'^2}) + \frac{1}{2} \ln(t'^2-1) + t' \ln \frac{t'+1}{t'-1} \tag{115}$$

which for $t' >> 1$ becomes

$$\gamma_1(h) = \gamma_1(t') \simeq \ln t' + \frac{3}{2} \tag{116}$$

which again corresponds to the approximation given by Eq. (103).

and for $t = 1$ $(h = 1)$ both formulas (113) and (115) give

$$\gamma_1(1) = \ln 4 = 1.38629 \tag{117}$$

The plot of Eqs. (113), (115) and (116) is given in Fig. 33.

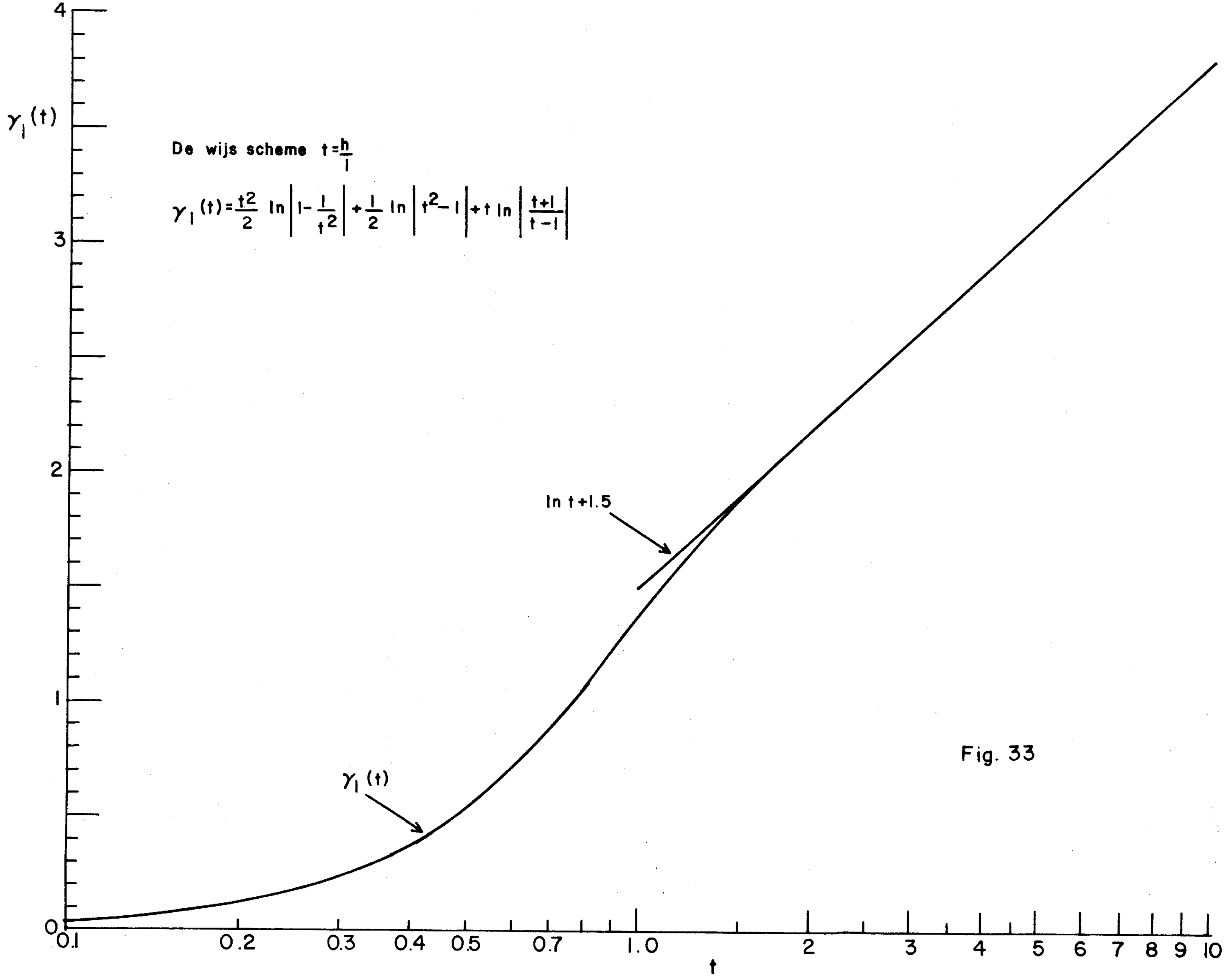

$\gamma_1(t)$
De wijs scheme $t=\frac{h}{l}$
$\gamma_1(t)=\frac{t^2}{2}\ln\left|1-\frac{1}{t^2}\right|+\frac{1}{2}\ln\left|t^2-1\right|+t\ln\left|\frac{t+1}{t-1}\right|$
4
3
2
1
0
ln t+1.5
$\gamma_1(t)$
0.1
0.2
0.3
0.4
0.5
0.7
1.0
2
3
4
5
6
7
8
9
10
t

Fig. 33

The spherical model

The regularization of the spherical model given by Eq. (93) along the core length l gives for this variogram a long and awkward expression. The characteristic points of the regularized variogram are:

for h=l

$$\gamma_l(l) = \frac{A}{l} + C\left(\frac{l}{a} - \frac{7}{10}\cdot\frac{l^3}{a^3}\right) \qquad (118)$$

for h = ∞ is a sill

$$\gamma_l(\infty) = \frac{A}{l} + C\left(1 - \frac{l}{2a} + \frac{l^3}{20a^3}\right) \qquad (119)$$

where A is connected with the nugget effect and it is a constant value.

The range of the regularized variogram is now

$$a + l$$

and knowing the values of the regularized variogram at points h=l and h→∞ we can draw from Eq. (119) and (118) the values of C and A once the range a is known. The plots of the regularized spherical variograms are given in Figs. 34, 35 and 36 according to [1] and [2].

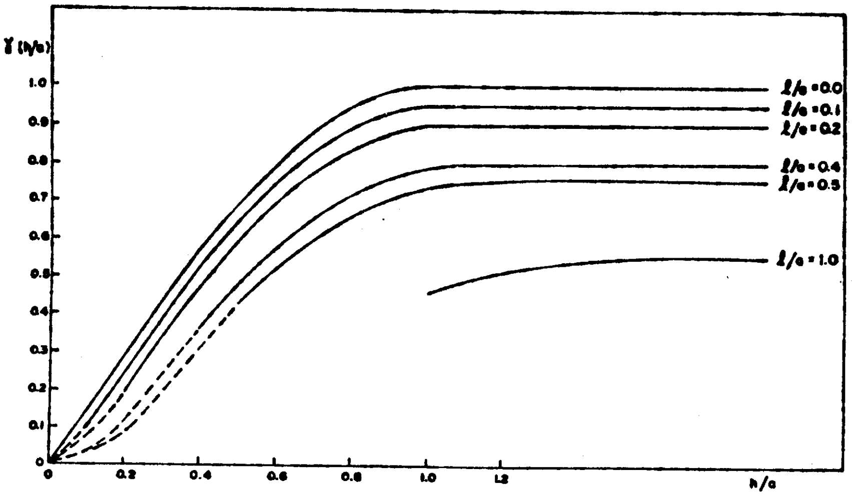


Fig. 34

Smoothing of a spherical variogram; variogram of samples of length *l* parallel to the direction of computation of the variogram (pieces of core in a D.D.H.).

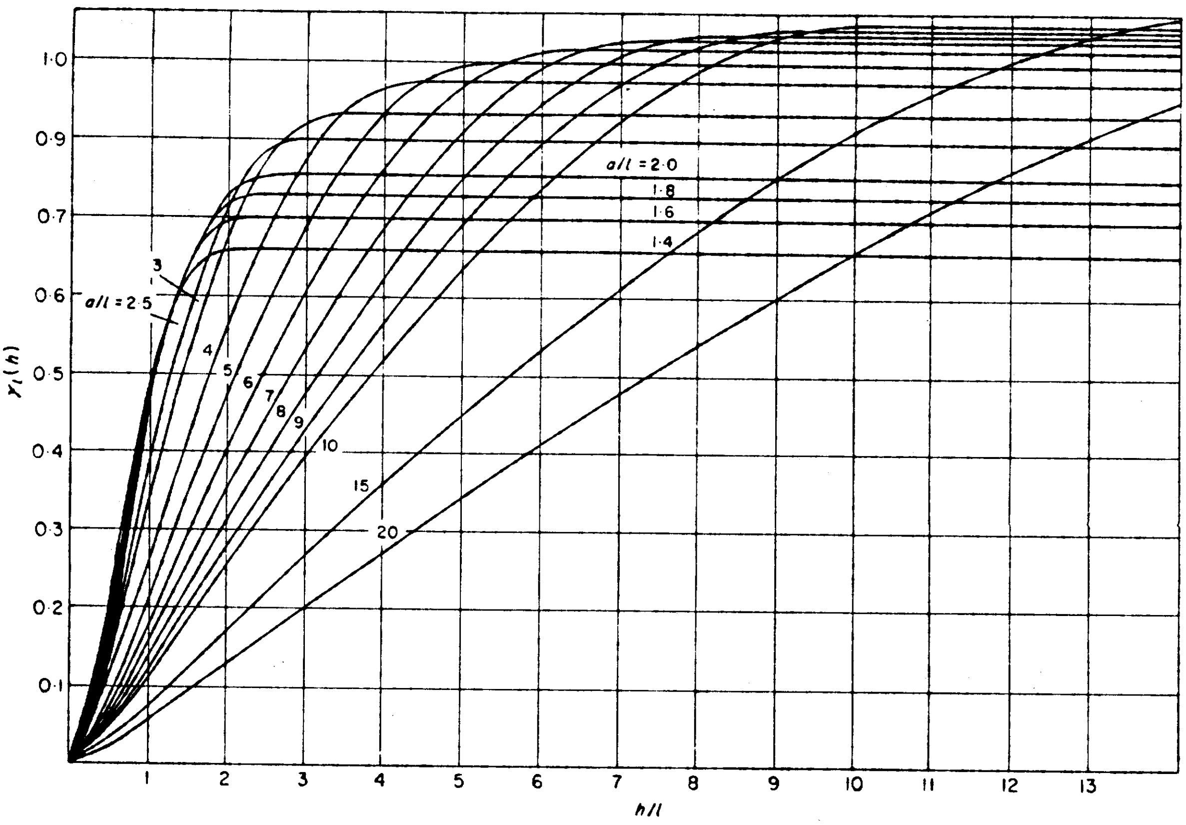


Fig. 35 **Spherical model. Regularized semi-variogram $\gamma_l(h)$.**

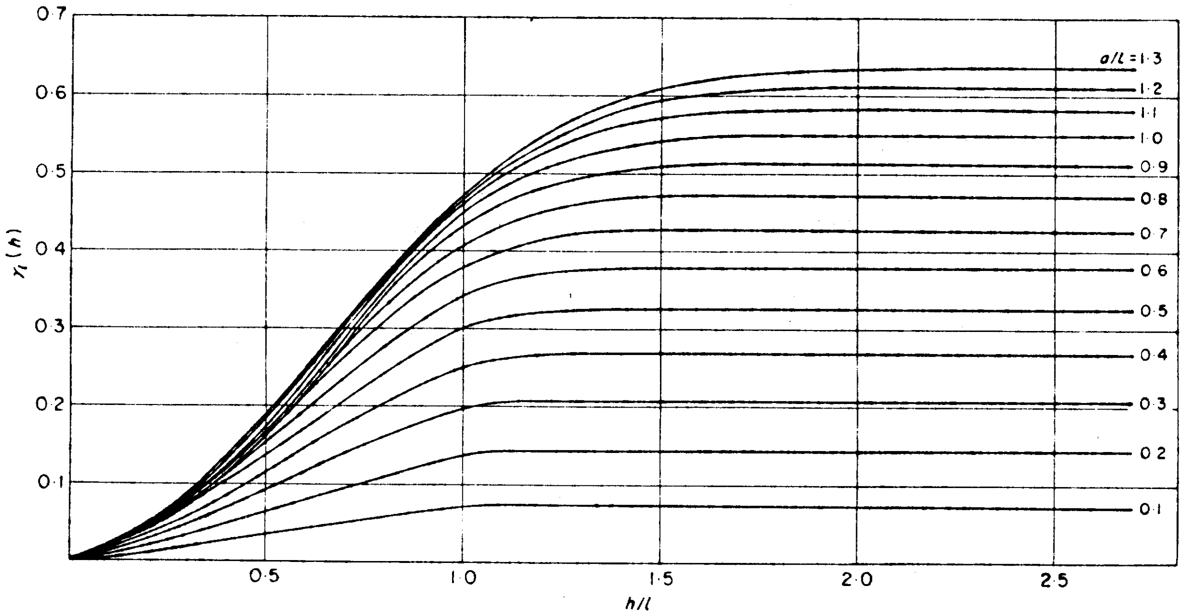


Fig. 36 **Spherical model. Regularized semi-variogram $\gamma_l(h)$.**

As an example let us take the variogram of porosity presented in Fig. 21 (lecture No 3). This is just the regularized variogram of the spherical type. The regularization is done by the fact that the porosity used for calculation of the experimental variogram $\gamma^*(h)$ corresponds to the average porosity for 1 m borehole section.

From Eqs.(118) and (119) we have:

$$C = \frac{\gamma_l(\infty) - \gamma_l(l)}{1 - \frac{3l}{2a} + \frac{3l^3}{4a^3}} \tag{120}$$

$$A = l\ [\gamma_l(l) - C(\frac{l}{a} - \frac{7}{10}\ \frac{l^3}{a^3})\ ] \tag{121}$$

The experimental range in Fig. 21 is 6 m thus for $l = 1$ m one has

$$a = 5$$

Next: $\gamma_l(\infty) = \sigma^2(\phi) = 12.85 \cdot \%^2$

$\gamma_l(l) = \gamma^*_\phi(l) = 2.85\ \%^2$

Thus from Eqs.(120) and (121) one has:

$$\begin{aligned} C &= 14.16\ \%^2 \\ A &= 0.096\ \text{m} \cdot \%^2 \end{aligned} \tag{122}$$

thus the nugget effect is negligible here.

The $\bar{\gamma}(l, l)$ function for the spherical variogram is:

$$\bar{\gamma}(l, l) = \frac{l}{2a} - \frac{l^3}{20a^3} \qquad (\text{for } l < a) \tag{119a}$$

The exponential model

We take to simplify the formulae the values $C_1 = 1$ and $C_o = 0$ in Eq. (94). Thus the regularized exponential variogram will be of the form:

$$\gamma_1(h) = \begin{cases} \frac{2a_1{}^2}{1^2}\left(e^{-1/a_1} - 1\right) + \frac{2a_1 h}{1^2} + \frac{a_1{}^2}{1^2}\, e^{-\frac{h}{a_1}} \cdot \left(2 - e^{-1/a_1}\right) - \\ \qquad\qquad - \frac{a^2}{1^2}\, e^{(h-1)/a_1} \quad \text{for } h \leq 1 \\ \frac{a_1^2}{1^2}\left[e^{-1/a_1} - e^{1/a_1} + \frac{2\,1}{a_1} + \left(e^{-1/a_1} + e^{1/a_1} - 2\right)\left(1 - e^{-h/a_1}\right)\right] \\ \qquad\qquad \text{for } h \geq 1 \end{cases}$$

We can see that for $h \geq 1$ the regularized exponential variogram is again exponential with some kind of "nugget effect"

$$C_{o1} = \frac{a_1^2}{1^2}\left[e^{-\frac{1}{a_1}} - e^{\frac{1}{a_1}} + \frac{2\,1}{a_1}\right] \tag{124}$$

and the new sill value:

$$C_{1,1} = C_1 \cdot \frac{a_1^2}{1^2}\left(e^{-\frac{1}{a_1}} + e^{\frac{1}{a_1}} - 2\right) \tag{125}$$

An approximation according to Eq. (103) gives for $h > 1$:

$$\gamma_1(h) = C_o + C_1\left(1 - e^{-h/a_1}\right) - \bar{\gamma}(1,1) \tag{R6}$$

where $$\bar{\gamma}(1,1) = 1 - \frac{a_1^2}{1^2}\left[2e^{-1/a_1} + \frac{2\,1}{a_1} - 2\right] \tag{127}$$

EXERCISE No 3

We can come back again to the example shown in EXERCISE No 1 (LECTURE No 3). From the experimental variogram of porosity we have just calculated the parameters of the underlying theoretical variogram of the spherical type, assuming the range

$$a = 5 \text{ m}$$

we have obtained

$$C = 14{,}16 \,\%^2$$

and negligible nugget effect.

Now, we can again take the porosity data shown in Table 1 (page 32, 33, 34) and calculate using the program VARIOGRAM 10 again the variogram but taking now as an input data the porosities averaged by the segments of 2 meters. Thus, instead of introducing into the program the set

$$\emptyset(x_1)\ ,\ \emptyset(x_2),\ \emptyset(x_3),\ \emptyset(x_4)\ ,\ \ldots\text{etc}\quad \emptyset(x_{139}),\ \emptyset(x_{140})$$

we introduce the sequence:

$$\emptyset_1 = \frac{1}{2}\,[\emptyset(x_1)+\emptyset(x_2)]\ ;\ \emptyset_2 = \frac{1}{2}\,[\emptyset(x_3) + \emptyset(x_4)]\ ;\ \ldots\ldots.$$

$$\ldots\ \emptyset_K = \frac{1}{2}\,[\emptyset(x_{2K-1}) + \emptyset(x_{2K})]\ldots;\ \emptyset_{70} = \frac{1}{2}\,[\emptyset(x_{139})+\emptyset(x_{140})]$$

The result of this calculation is given in Table 5 and is plotted in Fig. 37. Some nested structure at the $h = \sim 18$ m is visible on this variogram and it becomes more similar to that one of the bulk density given in Fig. 22.

Using the a and C values given above one can calculate from Eq. (119a) the

$$C\cdot\bar{\gamma}\,(1,1) \quad \text{and} \quad C\cdot\bar{\gamma}\,(2,2)$$

values and the variance of the 2 m long samples can be obtained from the relation (91) which in this case is written as:

$$D(2/V) - D(1/V) = C\cdot[\bar{\gamma}(1.1) - \bar{\gamma}(2.2)] \qquad (128)$$

Table 5

Calculation of the variogram of porosity for the 2 m borehole segments (after the data in Table 1).

Average porosity	5.612 %
Variance of porosity	11.480 $\%^2$
Number of data used for calculation of average	70
Number of pairs in variogram	60

h m	i	$\gamma^*(i)$	
2	1	$\gamma^*(1)$	5.471
4	2	$\gamma^*(2)$	10.549
6	3	$\gamma^*(3)$	11.690
8	4	$\gamma^*(4)$	11.779
10	5	$\gamma^*(5)$	11.740
12	6	$\gamma^*(6)$	11.647
14	7	$\gamma^*(7)$	10.470
16	8	$\gamma^*(8)$	8.976
18	9	$\gamma^*(9)$	8.694
20	10	$\gamma^*(10)$	10.647

We have (a = 5 m)

$\bar{\gamma}(1,1) = 0.0996$

$\bar{\gamma}(2,2) = 0.1968$

$C[\bar{\gamma}(.1) - \bar{\gamma}(2.2)] = -1.3763\ \%^2$

and finally, for

$D(1/V) = 12.8499\ \%^2$ one has

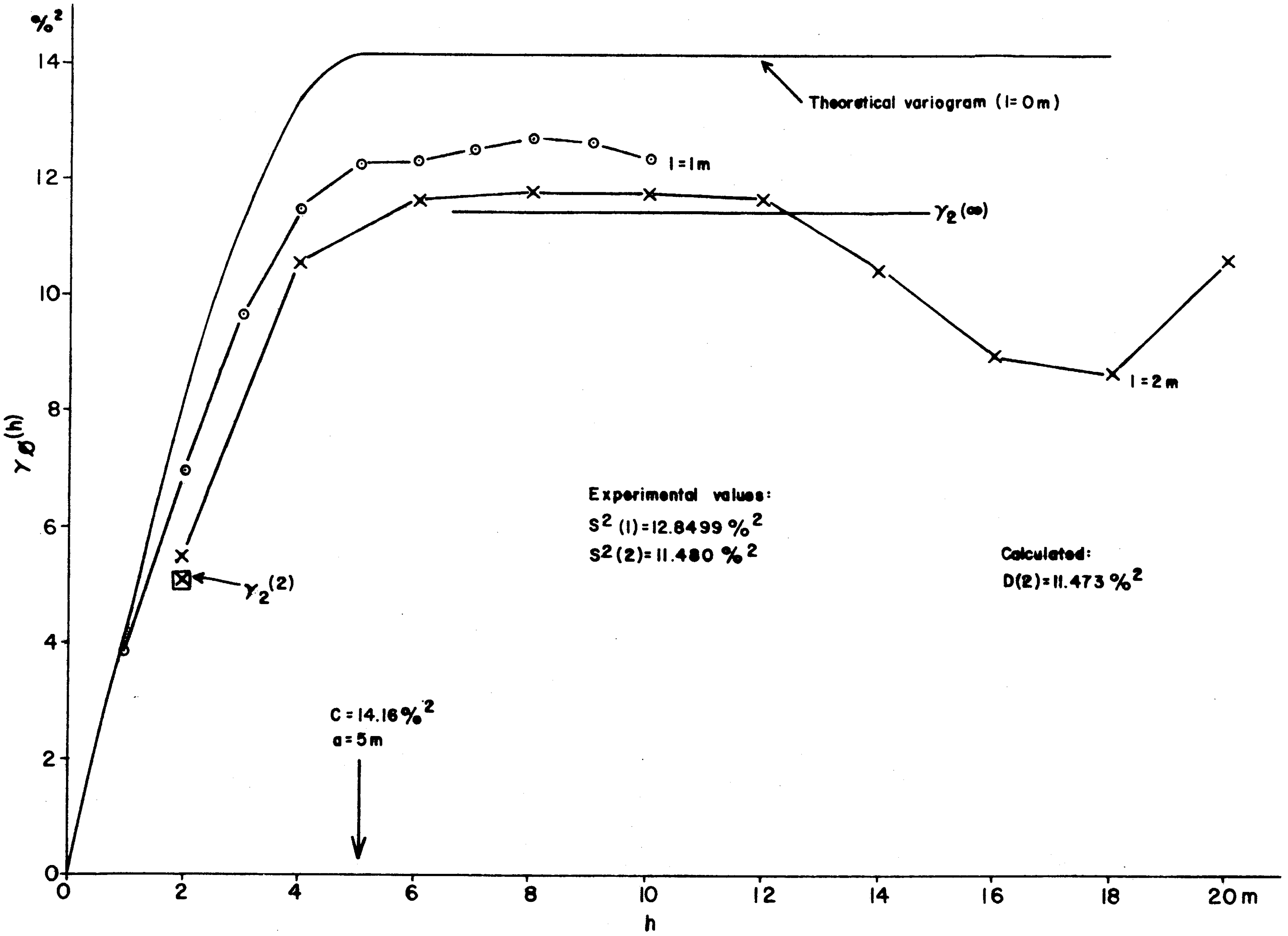


Fig. 37 Experimental proof of the variogram regularization by the core segments

from Eq. 128 :

$$D(2/V) = 11.474\ \%^2$$

whereas from the experiment (cf. Table 5) one obtained

$$s^2\ (\emptyset) = 11.480\ \%^2$$

for $l = 2$ m

Let us remark that if one takes $a = 4$ m or $a = 6$ m the result for the $D(2/V)$ calculation will be 11.512 $\%^2$ and 11.796 $\%^2$ respectively.

LECTURE 6

J.A. Czubek

LECTURE 6 Reykjavík, March 11, 1981

MORE ABOUT REGULARIZATION

Another example of regularization is the so called grading. Let for example, in the three-dimensional space the point-like variogram be

$$\gamma(h) = \gamma(\sqrt{x^2+y^2+z^2}) \tag{129}$$

where x, y, z are considered as the components of the vector h (i.e. $x = x_1-x_2$; $y = y_1-y_2$; $z = z_1-z_2$ x_1, y_1, z_1 and x_2, y_2, z_2 being a coordinates at the two points P_1 and P_2). The measurements are performed in the boreholes along some layer thickness l . Thus, in the vertical plan the situation is:

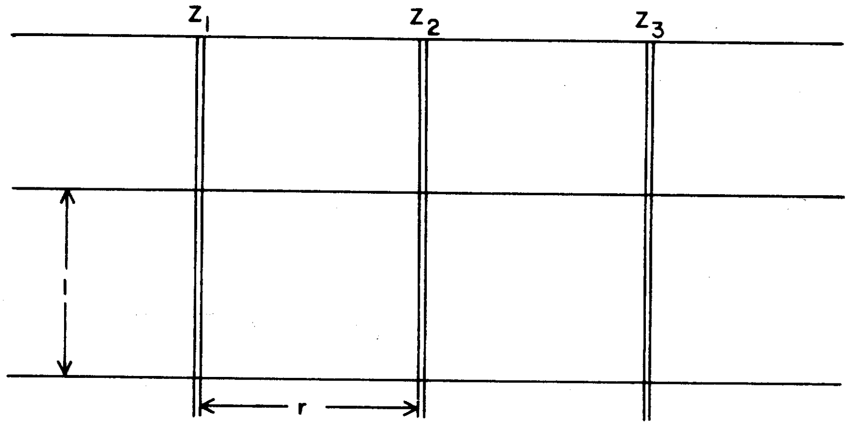


Fig. 38

We know the average values of the random function in each borehole. It is averaged along the layer thickness l , and if the distance between two boreholes is r

$$z_G = \frac{1}{l} \int_{l(x_i)} z(u)\,du \tag{130}$$

are the values at points x_i on the map

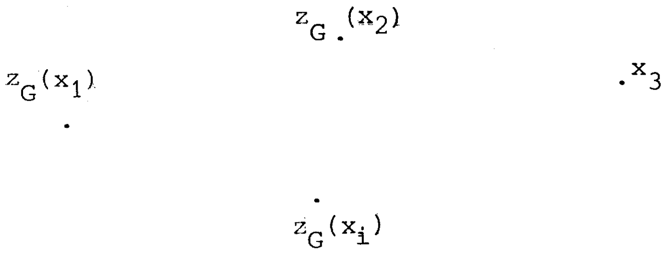


Fig. 39

This operation is called in geostatistics, "grading" (it comes from the French translation of expression "to be promoted to higher grade" as in the army, for example). When the experimental variogram is calculated for the graded values of Z , it is:

$$\gamma_G(r) = \frac{1}{2} E\{ [Z_G(x+r,) - Z_G(x)]^2\} =$$

$$= \bar{\gamma}[\, l(x),\ l(x+r)\,] - \bar{\gamma}(l, l) \qquad (131)$$

Here we have used the general relation (101) applied to this particular case. Thus, what one has to do is to calculate the average values:

$$\gamma\,[\, l(x),\ l(x+r)\,] = \frac{1}{l^2} \int_o^l du_1 \cdot \int_o^l du_2\ \gamma\,(\,\sqrt{r^2+(u_1-u_2)^2} \qquad (132)$$

and

$$\bar{\gamma}\,(l,\ l\,) = \underset{r=o}{\bar{\gamma}}\ [\,l(x),\ l(x+r)\,] \qquad (133)$$

where the variables u_1 and u_2 are as in Fig. 40

$h = \sqrt{r^2+(u_1-u_2)^2}$

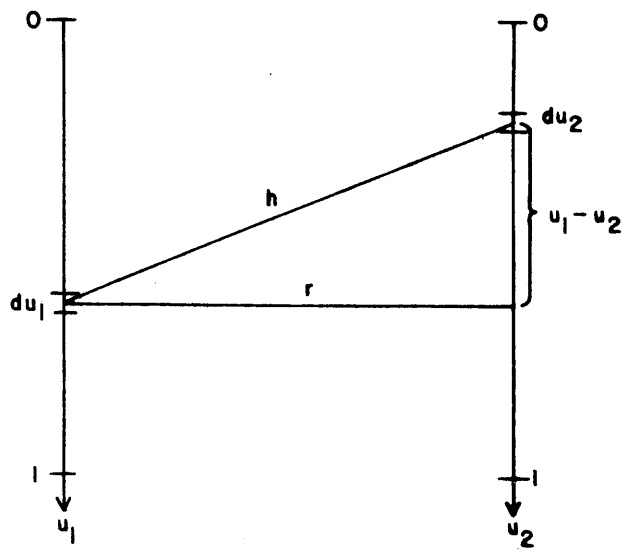


Fig. 40

For

$$|u_1-u_2| = y$$

the double integral in (132) can be rewritten in the form

$$\frac{1}{1^2}\int_0^1 du_1 \int_0^1 du_2\ \gamma\left(\sqrt{r^2+(u_1-u_2)^2}\right) = \frac{2}{1^2}\int_0^1 dy \int_0^1 du_2\ \gamma\left(\sqrt{r^2+y^2}\right) =$$

$$= \frac{2}{1^2}\int_0^1 (1-y)\gamma\left(\sqrt{r^2+y^2}\right) dy \qquad (135)$$

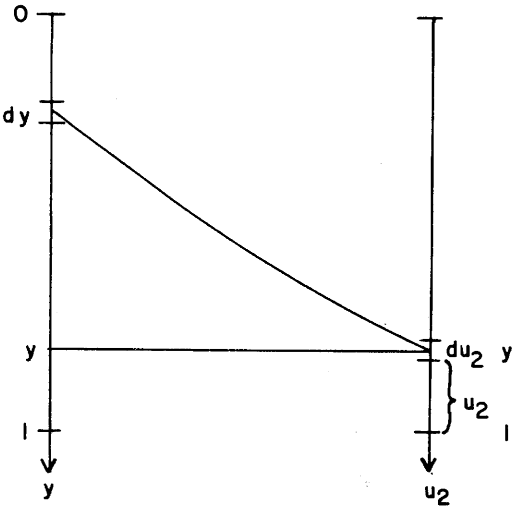


Fig. 41

The coefficient 2 in (135) appears because of the sign of the absolute value in (134) - simply speaking we change here the integration domain from u_1, u_2 onto y_1 u_2 , as shown in Fig. 42

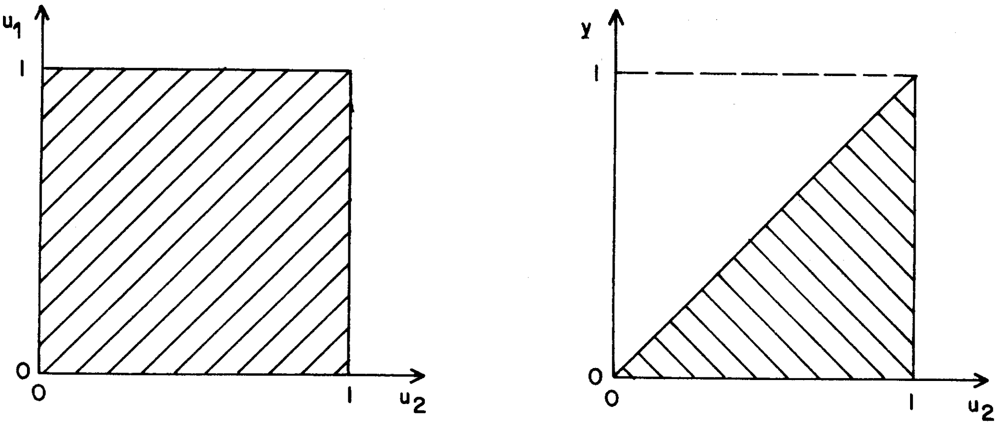


Fig. 42

Similarily for $\bar{\gamma}(1, 1)$ one has:

$$\bar{\gamma}(1,1)=\frac{2}{1^2}\int_0^1 (1-y)\gamma(y)\,dy \tag{136}$$

which corresponds just to the integral used in the regularization along the borehole cores.

The linear model

For the linear model

$$\gamma(h)=|h|$$

the exact expression for $\gamma_G(h)$ according to Eqs. (131), (132), (133) and (135) is:

$$\gamma_G(r)=\frac{1}{3}\sqrt{r^2+1^2}+\frac{r^2}{1}\ln\frac{1+\sqrt{r^2+1^2}}{r}+\frac{2}{3}\frac{r^2}{1^2}\left[r-\sqrt{r^2+1^2}\,\right]-\frac{1}{3} \tag{137}$$

with

$$\bar{\gamma}(1,1)=\frac{1}{3}$$

and in practice, when $r\geq 1$ one uses the standard approximation given by Eq.(103)

$$\gamma_G(r)\approx\gamma(r)-\bar{\gamma}(1,1)=r-\frac{1}{3} \tag{138}$$

Let us remark, that the graded variogram in the linear model for large r values is indistinguishable from the model $\gamma_1(h)$ regularized by cores - cf. Eq.(110).

The de Wijs scheme (logarithmic model)

Here the graded variogram for the model

$$\gamma(h)=\ln|h|$$

is

$$\gamma_G(r)=\frac{r}{L}\cdot 2\,\mathrm{arc\ ctg}\,\frac{r}{1}+\frac{1}{2}\ln\left(1+\frac{r^2}{1^2}\right)-\frac{1}{2}\frac{r^2}{1^2}\ln\left(1+\frac{1^2}{r^2}\right) \tag{139}$$

with again

$$\bar{\gamma}(1,1)=\ln 1-\frac{3}{2} \tag{140}$$

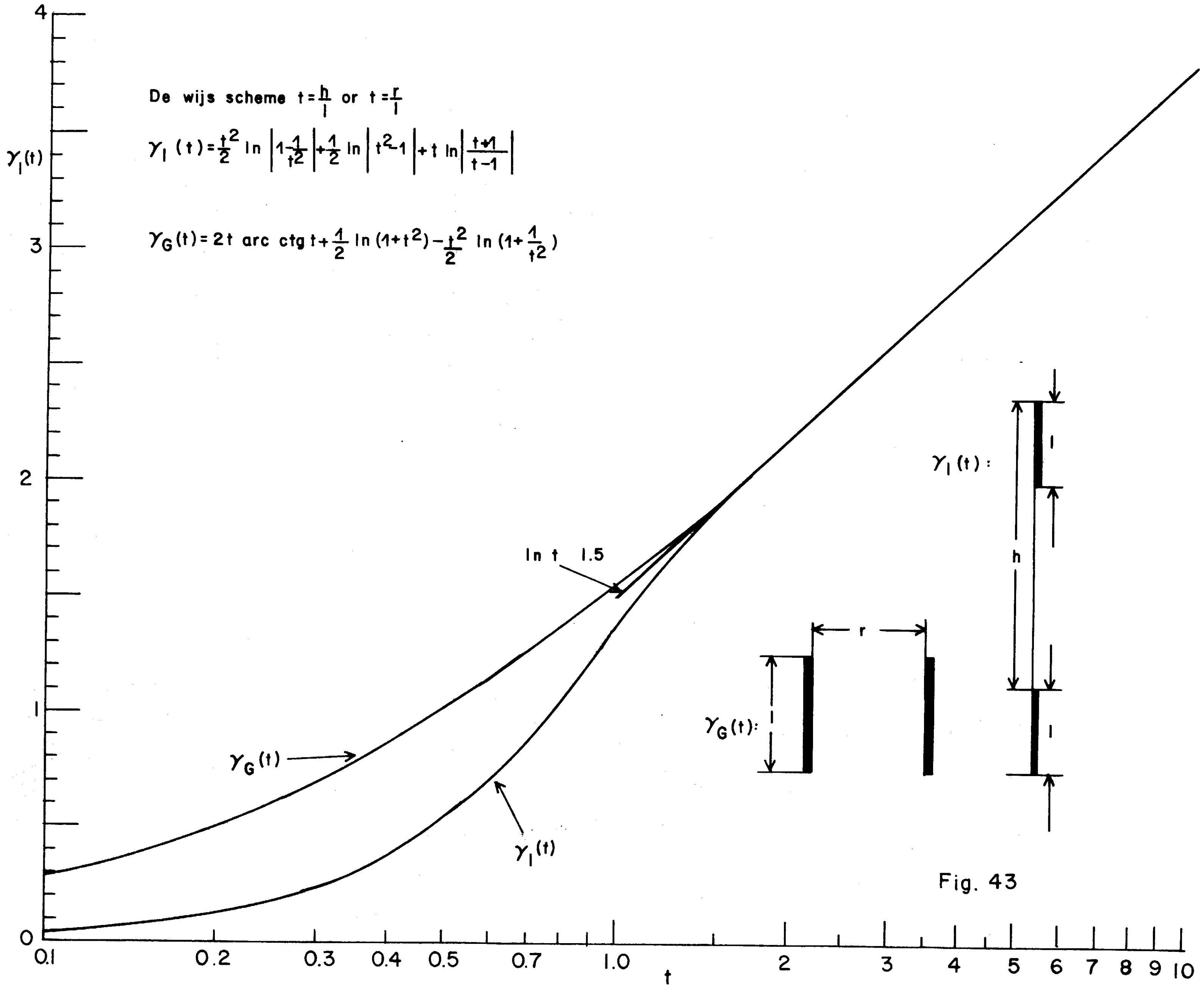

De wijs scheme $t=\frac{h}{l}$ or $t=\frac{r}{l}$
$\gamma_l\,(t)=\frac{t^2}{2}\ln\left|1-\frac{1}{t^2}\right|+\frac{1}{2}\ln\left|t^2-1\right|+t\ln\left|\frac{t+1}{t-1}\right|$
$\gamma_G(t)=2t\ \mathrm{arc\ ctg}\ t+\frac{1}{2}\ln(1+t^2)-\frac{t^2}{2}\ln(1+\frac{1}{t^2})$
$\gamma_l(t)$
4
3
2
1
0
0.1
0.2
0.3
0.4
0.5
0.7
1.0
t
2
3
4
5
6
7
8
9
10
ln t 1.5
$\gamma_G(t)$
$\gamma_l(t)$
$\gamma_G(t)$:
r
$\gamma_l(t)$:
h
l
l

Fig. 43

The variogram $\gamma_G(r)$ given by Eq. (139) is shown in Fig. 43 together with the variogram $\gamma_1(h)$ for the de Wijs scheme. One can see very easily that both of them approach the line

$$\ln \frac{r}{1} + \frac{3}{2}$$

for the $\frac{r}{1} > 2$, whereas for $\frac{r}{1} < 1$ they differ considerably.

Spherical and exponential models

For the spherical and exponential models the formulae for the $\gamma_G(r)$ variograms are long and not very convenient to handle. For the spherical model (with $C = 1$ and $C_0 = 0$) it is given in Fig. 44 (according to Journel and Huijbregts 1978). The sill value is here

$$\gamma_G(\infty) = 1 - \bar{\gamma}(1, 1) = 1 - \frac{1}{2a} - \frac{1^3}{20\cdot a^3} \qquad (141)$$

and for $r \geq a$ the range of the graded model is equal to the range a of the point model, and when $r >> 1$ one uses

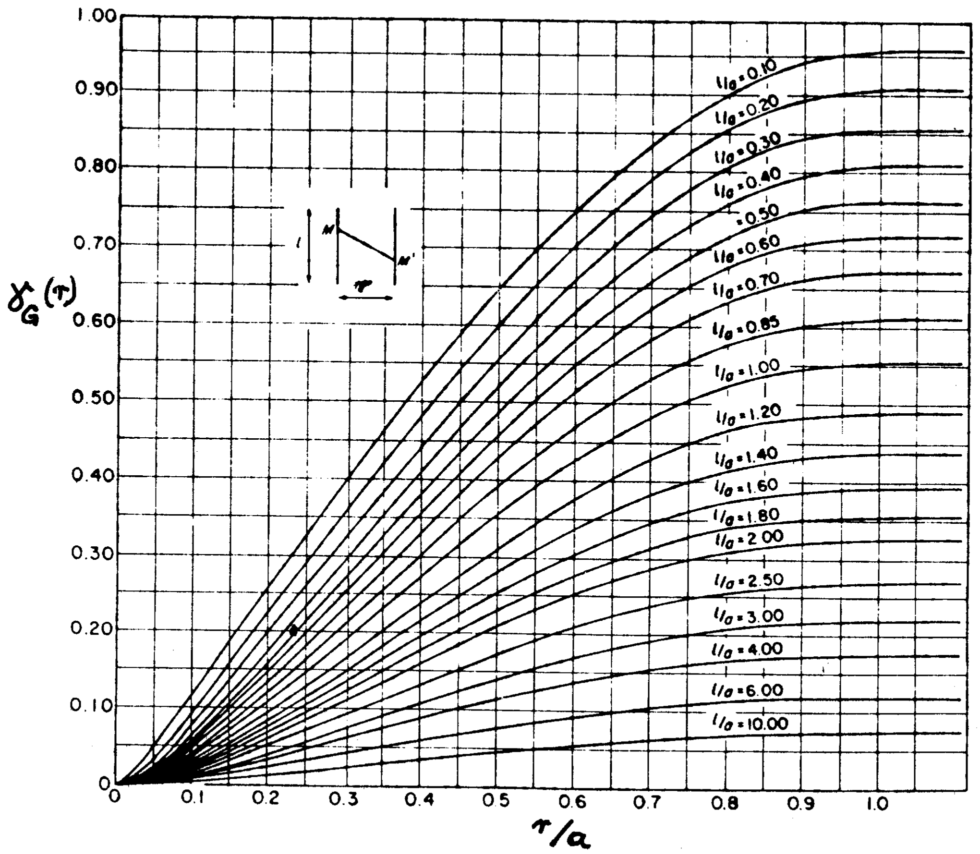


Fig. 44 . Spherical model. Grading over *l*: $\gamma_G(r)$

$$\gamma_G(r) \approx \gamma(r) - \bar{\gamma}(1, 1) \qquad (142)$$

For distances $r \leq a$ one has to use the exact formulae to calculate the $\gamma_G(r)$ values in order to be able to find out the parameters of the underlying point variogram. These formulae are given in the doctoral thesis of J. Serra in Fontainebleau (Center of Mathematical Geomorphology) and have never been published (but everybody can calculate them starting from equation 131 !). When $a >> 1$ (in practice, when $a > 31$), and $a >> r$, or when

$$a > 3\sqrt{r^2+1^2}$$

it means, when we are at the beginning of the underlying spherical variogram, we can neglect the contribution of the second term in it, i.e.

$$\gamma(h) = C \cdot [\frac{3}{2}\frac{h}{a} - \xcancel{\frac{1}{2}(\frac{h}{a})^3}]$$

and then this variogram becomes a linear variogram with the slope equal to

$$1.5\ C$$

In this case we are using Eq. (137) to fit the model.

A similar situation occurs for the graded variogram over the exponential model. For big r values in comparison with the 1 value one uses the approximation given by Eq. 142 , whereas when

$$a_1 > \sqrt{r^2+1^2}$$

(where a_1 is given in Eq. 94) we can at the origin (when $r \rightarrow 0$) treat this graded variogram as being from the linear model with the slope equal to

$$\frac{C_1}{a_1}$$

The plot of variograms graded over the exponential model are given in Fig. 45 (according to Journel and Huijbregts, 1978).

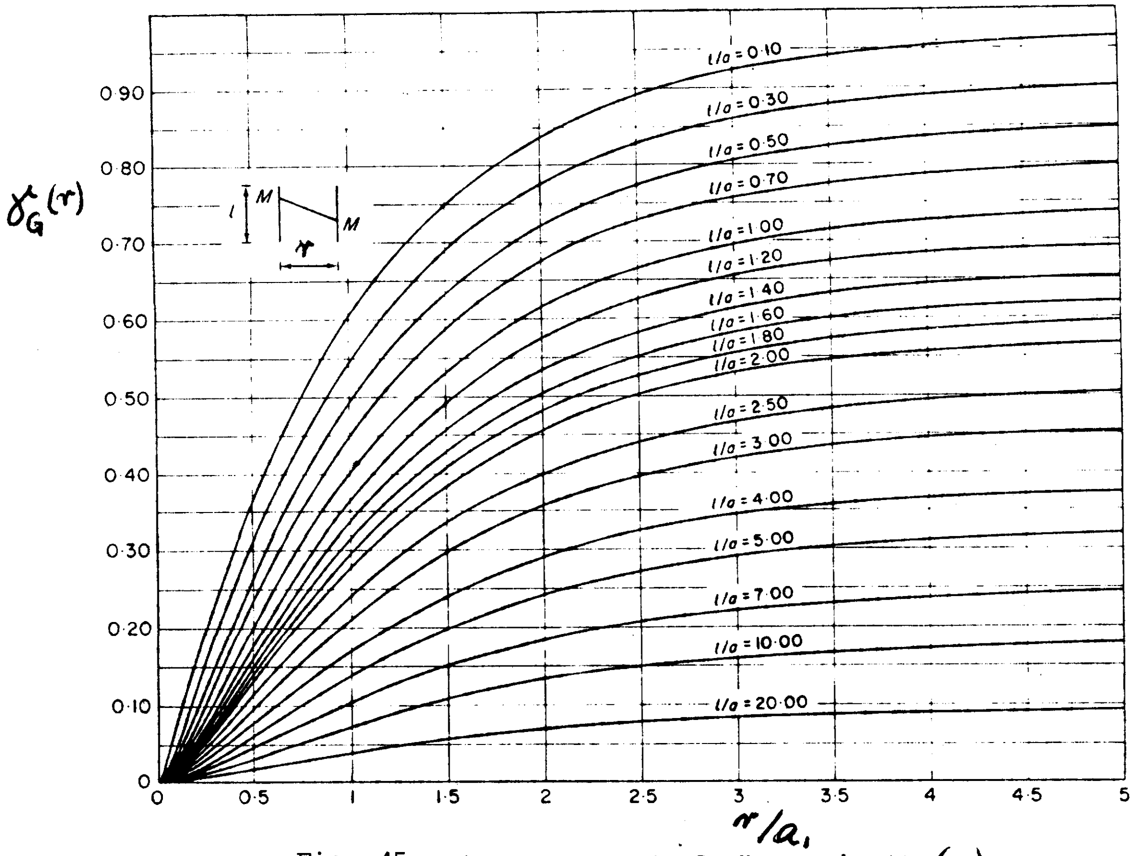


Fig. 45 **Exponential model. Grading over *l*:** $\gamma_G(r)$

In practice, any kind of regularization of the variogram requires more or less some particular discussion, but the general approach is always starting from Eq. 131. There is only a technical problem now to calculate the double or quadruple or even sextuple integrals involved. Very often one uses here a numerical quadratures to get them. How important this question is, we are going to see in the next subject treated in this lecture.

MORE ABOUT DISPERSION VARIANCE

LINEAR EQUIVALENTS

Let us go back to Eq. 84 which is a basic equation for the dispersion variance:

$$D(v/V) = \bar{\gamma}(V,V) - \bar{\gamma}(v,v) \tag{84}$$

As mentioned the $\bar{\gamma}(V,V)$ value is a constant, very often unknown, characteristic for a given geological formation and the size of the geological samples has no influence on it. The whole dispersion variance can be changed by the change of the $\bar{\gamma}(v,v)$ function only. This function, $\bar{\gamma}(v,v)$, however is the average value of a given variogram (point-like model), where the averaging procedure is performed over the whole volume v of the sample. It is quite evident, that different geometrical forms of v can have the same average value of $\bar{\gamma}(v,v)$! If it is so, this means that the samples of different sizes can have the same dispersion variance $D(v/V)$ inside a given geological field ! We can find an infinite number of such corresponding volumes for which the function

$$\bar{\gamma}(v,v)$$

is invariant. Among all these possibilities one is especially interesting, when the sample has the shape of a line segment of the length l^*. We call the value l^* a linear equivalent of a given volume v. How to find the linear equivalent ? Just from the equation:

$$\bar{\gamma}(v,v) = \bar{\gamma}(l^*,l^*) \tag{143}$$

which has to be solved for a given shape v of the sample and for a given underlying point-like variogram. More developed form of Eq. 143 is:

for the 3-dimensional samples (cf. Eq. 135)

$$\frac{1}{v^2} \underset{v}{\iiint} dv_1 \underset{v}{\iiint} dv_2\, \gamma\,(|r|) = \frac{1}{l^{*2}} \underset{l^*}{\int} dx_1 \underset{l^*}{\int} dx_2\, \gamma\,(|x_1 - x_2|) \tag{144}$$

where dv_1 and dv_2 are the two elementary volumes situated at the two extremities of the vector r inside the volume v as it is shown in Fig. 46

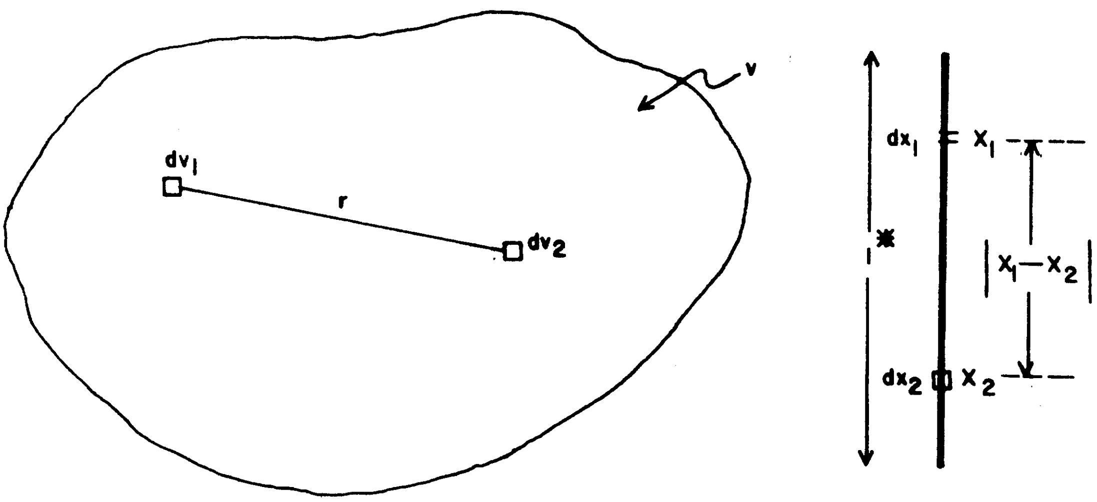


Fig. 46

In the two dimensional space the translation of Eq. 144 is immediate:

$$\frac{1}{S^2}\iint_S dS_1 \iint_S dS_2\ \gamma(|r| = \frac{1}{l^{*2}}\int_{l^*} dx_1 \int_{l^*} dx_2\ \gamma\,(|x_1-x_2|) \tag{145}$$

where dS_1 and dS_2 are again the two elementary volumes situated at the two extremities of the vector r.

It can be also some confluent cases, (but rather artificial, though not impossible), when the sample does not have plain surface or the line sample is not a straight line (for example a perimeter of circle or rectangular, etc.).

Now, to calculate the linear equivalents, we have to take into account the particular forms of the samples and the particular forms of the variograms. The easiest way, when the question is how to calculate the integrals involved in Eqs.(144) and (145), is to take the parallelepipeds and rectangles as the forms of the samples.

LECTURE 7

AUXILIARY FUNCTIONS

All calculations of variances and covariances in geostatistics are based on the calculation of some specially averaged values of the variogram $\gamma(h)$. We can distinguish few averaging procedures which are of special interest and once these cases are pre-calculated they are then very useful in the routine practice.

ONE-DIMENSIONAL CASE

For the segment $AB = L$ as in Fig. 47

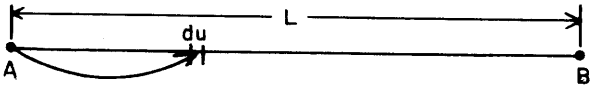


Fig. 47

one calculates the average value of $\gamma(h)$ when ore extremity of h is fixed at the one end (say A) and the other extremity describes the whole segment AB, thus:

$$\chi(L) = \bar{\gamma}\ (A\,;\,AB) = \frac{1}{L}\int_0^L \gamma(u)\,du \tag{146}$$

and it is a simple average of $\gamma(h)$ taken along the segment AB.

The auxiliary function $F(L)$ is defined as the mean value of $\gamma(h)$ when the two extremities of the vector h are "walking" independently of each other along the whole segment L, thus

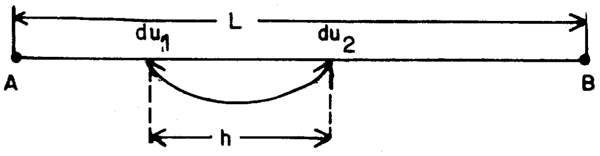


Fig. 48

$$F(1) = \bar{\gamma}\ (AB\,;\,AB) = \frac{1}{L^2}\int_0^L du_1 \int_0^L du_2\ \gamma(u_1 - u_2) =$$

$$= \frac{2}{L^2}\int_0^L (L-u)\ \gamma\ (u)\,du = \frac{2}{L^2}\int_0^L u\chi(u)\,du \tag{147}$$

We have met this function previously in Eq. (136).

FOR TWO DIMENSIONS

One considers the rectangle ABCD:

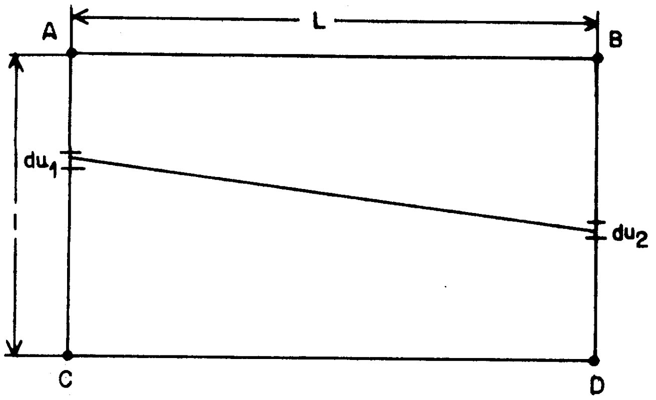


Fig. 49

and the two-variable function $\alpha(L ; l)$ is defined as: (cf. Fig. 49)

$$\alpha(L ; l) = \bar{\gamma}(AC ; BD) \tag{148}$$

i.e. the two extremities of the vector h are moving along the sides AC and BD of the rectangle ABCD, that is:

$$\alpha(L ; l) = \bar{\gamma}(AC ; BD) = \frac{1}{l^2} \int_0^l du_1 \int_0^l du_2 \, \gamma(\sqrt{L^2+(u_1-u_2)^2}) =$$

$$= \frac{2}{l^2} \int_0^l (l-u)\, \gamma(\sqrt{L^2+u^2})\, du = \bar{\gamma}[l(x) ; l(x+L)] \tag{149}$$

we have just met this function in Eqs. 132 and 134. Note, that the function $\alpha(L ; l)$ is not symmetrical in its variables:

$$\alpha(L ; l) \neq \alpha(l ; L) \tag{150}$$

The two-variable auxiliary function $\chi(L ; l)$ is defined as a mean value of $\gamma(h)$ when one extremity of the vector h is moving along the side l, whereas another one is walking along the whole rectangle ABCD as is shown in Fig. 50 . Thus one has:

$$\chi(L;l) = \bar{\gamma}(AC ; ABCD) = \frac{1}{L} \int_0^L \alpha(u,l)\, du \tag{151}$$

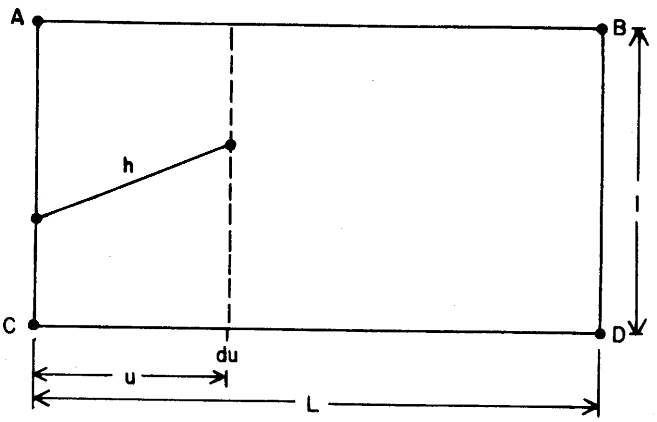


Fig. 50

which is a simple average of the function $\alpha(L;l)$ and in view of Eq. (149) is given as a triple integral of the variogram $\gamma(h)$. This function is again non-symmetrical in its variables:

$$\chi(L;l) \neq \chi(l;L) \qquad (152)$$

The two-variable auxiliary function $F(L;l)$ is defined as the mean value of $\gamma(h)$ when both extremities of h vector are describing, independently of each other, the entire rectangle ABCD, as in Fig. 51

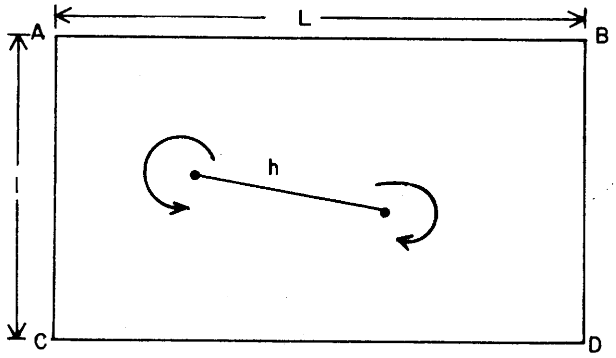


Fig. 51

Thus

$$F(L;l) = \bar{\gamma}(ABCD;ABCD) = \frac{1}{L^2 l^2} \iint_{Lxl} ds_1 \iint_{Lxl} ds_2\, \gamma\left(\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}\right) =$$

$$= \frac{2}{L^2}\int_0^L (L-u)\,\alpha(u;l)\,du = \frac{2}{L^2}\int_0^L u\chi(u;l)\,du \qquad (153)$$

Function $F(L;l)$ is symmetrical in its variables.

This function can be also given as:

$$F(L;l)=\frac{1}{L^2}\int_0^L du_1\int_0^L du_2\;\alpha(|u_1-u_2|;l) \qquad (153.a)$$

Finally the two-variable auxiliary function $H(L;l)$ is defined as a mean $\gamma(h)$ value when one extremity of the vector h is fixed at any of the corners of the rectangle, and the other extremity describes the entire rectangle ; as it is shown in Fig. 52 , thus

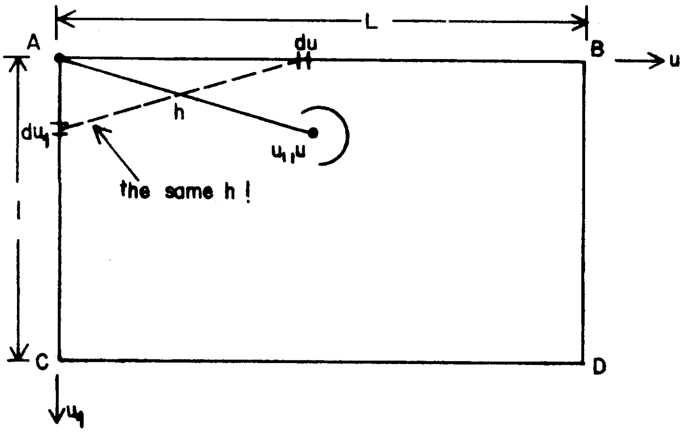


Fig. 52

$$H(L;l)=\bar{\gamma}\;(A;\;ABCD)=\frac{L}{L\cdot l}\int_0^L du\int_0^l du_1\;\gamma\,(\sqrt{u_1{}^2+u^2})=$$

$$=\bar{\gamma}\;(\,AC\;;\;AB) \qquad (154)$$

Thus, the function $H(L;l)$, symmetrical in its variable, is also an average value of the $\gamma(h)$ variogram when each extremity of the vector h is describing the adjacent sides of the rectangle.

All these auxiliary functions, being some average values of $\gamma(h)$ can be derived one from another, and the following relations (except those integral ones given above) hold:

$$\alpha(L;l)=\frac{1}{2}\,\frac{\partial^2}{\partial L^2}\,[L^2\cdot F(L;l)]=\frac{\partial}{\partial L}\,[L\cdot\chi(L;l)] \qquad (155)$$

$$\chi(L;l)=\frac{1}{2L}\,\frac{\partial}{\partial L}\,[L^2\cdot F(L;l)] \qquad (156)$$

$$H(L;l)= \frac{1}{2l}\ \frac{\partial}{\partial l}\ [l^2 : \chi(L;l)]= \frac{1}{4\,l\cdot L}\ \frac{\partial^2}{\partial l\partial L}\ [l^2\cdot L^2\cdot F(L;l)] \qquad (157)$$

Next, contracting one of the sides of the rectangle ABCD to zero,one can obtain from the two-variable auxiliary functions the one-variable auxiliary function

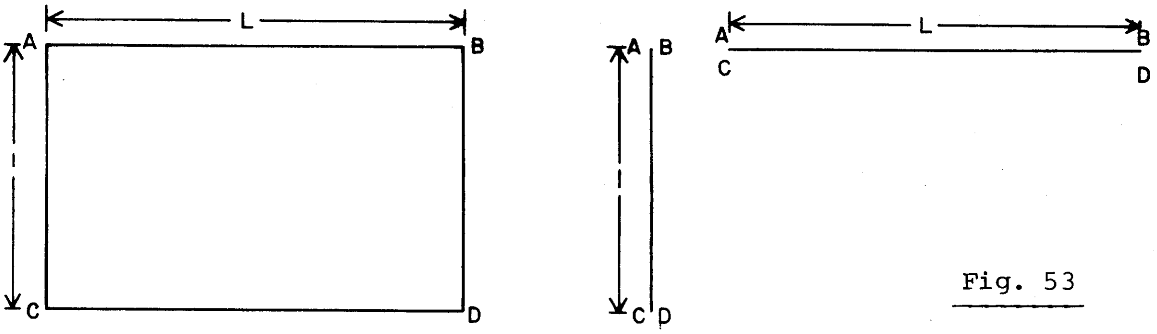


Fig. 53

which gives:

$$\left.\begin{aligned}
\alpha(L;0) &= \bar{\gamma}\ (A,B) = \gamma\ (L)\\
\alpha\ (0;l) &= \bar{\gamma}\ (AC,AC) = F(l)\\
\chi\ (L;0) &= \bar{\gamma}\ (A,AB) = \chi(L)\\
\chi\ (0;l) &= \bar{\gamma}\ (AC,AC) = F(l)\\
F(L;0) &= \bar{\gamma}\ (AB,AB) = F(L)\\
F(0;l) &= \bar{\gamma}\ (AC,AC) = F(l)\\
H(L;0) &= \bar{\gamma}\ (A,AB) = \chi(L)\\
H(0;l) &= \bar{\gamma}\ (A,AC) = \chi(l)
\end{aligned}\right\}\ (158)$$

The formulae (155) to (158) are very useful when one has to check the correctness of a given formula.

IN THREE DIMENSIONS

One has the three-variable auxiliary functions similar to the two-dimensional case. These functions are defined on the paralellepiped P = ABCDEFGH and just for sake of simplicity one takes a square base of this parallelepiped, thus its dimensions are $L X l^2$ as in Fig. 54.

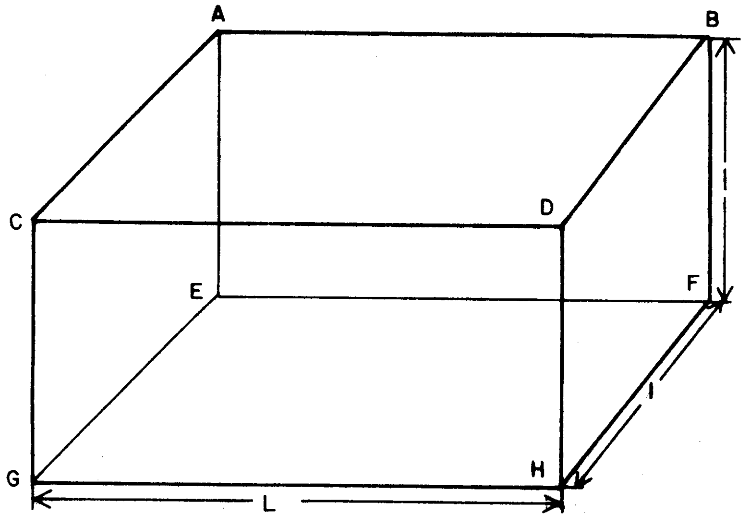


Fig. 54

The following auxiliary functions are defined:

1) the mean value of $\gamma(h)$ between the two square bases l^2 of P:

$$\alpha(L;l^2) = \bar{\gamma}\ (ACEG\ ;\ BDHF) \qquad (159.a)$$

2) the mean value of $\gamma(h)$ between one square face and the whole parallelepiped P:

$$\chi(L;l^2) = \bar{\gamma}\ (ACEG;P) \qquad (159.b)$$

3) the mean value of $\gamma(h)$ within P:

$$F(L;l^2) = \bar{\gamma}\ (P;P) \qquad (159.c)$$

4) the mean value of $\gamma(h)$ between two adjacent faces of P which is equal to the mean value between a side l and P:

$$H(L;l^2) = \bar{\gamma}\ (ACEG;ABEF) = \bar{\gamma}(AE;P) \qquad (159.d)$$

Now by contracting one of the sides of the parallelepiped P ($L \to 0$ or $l \to 0$) one obtains the two-variable auxiliary functions, or when two sides are simultaneously contracted to zero, the one variable auxiliary functions.

$$\alpha(L;l^2 \Rightarrow 0\ ) = \gamma\ (\ L) \qquad (160.e)$$

$$\alpha(0;l^2) - F(l;l) \qquad (160.b)$$

$$\left.\begin{array}{ll} \chi\ (L;l^2\rightarrow 0) = \chi(L) & (c)\\ \chi\ (0;l^2) = F(l;l) & (d)\\ F\ (L;l^2\rightarrow 0) = F(L) & (e)\\ F\ (0;l^2) = F(l;l) & (f)\\ H\ (L;l^2\rightarrow 0) = \chi(L) & (g)\\ H\ (0;l^2) = \chi\ (l,l) & (h) \end{array}\right\} \quad (160)$$

Now one has to discuss all these auxiliary functions for different models of the theoretical variograms $\gamma(h)$.

1. THE LINEAR MODEL

$$\gamma(h) = |h|$$

In one dimension:

$$\left.\begin{array}{l} \chi(L) = \frac{L}{2}\\ \\ F(L) = \frac{L}{3} \end{array}\right\} \quad (161)$$

In two dimensions:

$$t = \frac{l}{L}$$

$$\frac{1}{L}\ \alpha(L;l) = \frac{1}{3}\sqrt{1+t^2} + \frac{2}{3}\frac{1}{t^2}\left(1-\sqrt{1+t^2}\right) + \frac{1}{t}\ \ln\left(t+\sqrt{1+t^2}\right) \qquad (162.a)$$

$$\frac{1}{L}\ \chi(L;l) = \frac{1}{6}\cdot\frac{1}{t^2} + \sqrt{1+t^2}\ \left(\frac{1}{4}-\frac{1}{6}\cdot\frac{1}{t^2}\right) + \frac{1}{3}\cdot\frac{1}{t}\ \ln\left(t+\sqrt{1+t^2}\right) +$$

$$+ \frac{1}{12}\ t^2\ \ln\frac{1+\sqrt{1+t^2}}{t} \qquad (162.b)$$

$$\frac{1}{L}\ F(L;l) = \left(\frac{1}{5}-\frac{1}{15\cdot t^2}-\frac{t^2}{15}\right)\ \sqrt{1+t^2} + \frac{1}{15}\ \left(\frac{1}{t^2}+t^3\right) +$$

$$+ \frac{1}{6t}\ \ln\left(t+\sqrt{1+t^2}\right) + \frac{t^2}{6}\ \ln\frac{1+\sqrt{1+t^2}}{t} \qquad (162.c)$$

$$\frac{1}{L}\,H(L;l) = \frac{1}{3}\sqrt{1+t^2} + \frac{t^2}{6}\ln\frac{1+\sqrt{1+t^2}}{t} + \frac{1}{6t^2}\ln\left(t+\sqrt{1+t^2}\right) \qquad (162.d)$$

Note here the special case for $L = l$ $(t=1)$ which corresponds to the rectangle:

$$\left.\begin{aligned} \alpha\ (l;l) &= 1.0765\ l \\ \chi\ (l;l) &= 0.65176\ l \\ F\ (l;l) &= 0.5213\ l \\ H\ (l;l) &= 0.7652\ l \end{aligned}\right\} \qquad (163)$$

2. <u>THE LOGARITHMIC MODEL (DE WIJS SCHEME):</u>

$$\gamma(h) = \ln|h|$$

In one dimension:

$$\left.\begin{aligned} \chi(L) &= \ln L - 1 \\ F(L) &= \ln L - \frac{3}{2} \end{aligned}\right\} \qquad (164)$$

In two dimensions: (here again $t = \frac{l}{L} = t_g\theta$)

$$\alpha\ (L;l) = \ln l - \frac{3}{2} + \frac{2}{t}\,\mathrm{arc\,tg}t + \frac{1}{2}\ln\left(1+\frac{1}{t^2}\right) - \frac{1}{2t^2}\ln(1+t^2) \qquad (165.a)$$

$$\chi\ (L;l) = \ln l - \frac{11}{6} + \frac{1}{t}\,\mathrm{arc\ tg}t + \frac{1}{2}\ln\left(1+\frac{1}{t^2}\right) - \frac{1}{6t^2}\ln(1+t^2) +$$

$$+ \frac{t}{3}\left(\frac{\pi}{2} - \mathrm{arc\ tg}t\right) \qquad (165.b)$$

$$F(L;l) = \ln l - \frac{25}{12} + \frac{1}{2}\left(1 - \frac{t^2}{6}\right)\cdot\ln\left(1+\frac{1}{t^2}\right) - \frac{1}{12\cdot t^2}\ln(1+t^2) +$$

$$+ \frac{2}{3}\,t\left(\frac{\pi}{2} - \mathrm{arc\ tg}t\right) + \frac{2}{3}\,\frac{\mathrm{arc\ tg}t}{t} = \qquad (165.c)$$

$$= \ln l - \ln\sin\theta - \frac{25}{12} + \frac{1}{6}\,tg^2\theta\cdot\ln\sin\theta + \frac{1}{6}\,\frac{\ln\cos\theta}{tg^2\,\theta} + \frac{2}{3}\,\frac{\theta}{tg\,\theta} +$$

$$+ \frac{2}{3}\,tg\,\theta\cdot\left[\frac{\pi}{2} - \theta\right] \qquad (165.d)$$

For the values $t > 1$ or $t < 1$ Eqs. (165) develops into the rapidly convergent series:

For $t > 1$: $t = \frac{1}{L}$

$$\alpha\ (L;1) = \ln 1 - \frac{3}{2} + \frac{\pi}{t} + \frac{1}{t^2} \ln \frac{1}{t} - \frac{3}{2}\frac{1}{t^2} - \frac{1}{12}\frac{1}{t^4} + \cdots \qquad (166.a)$$

$$\chi\ (L;1) = \ln 1 - \frac{3}{2} + \frac{\pi}{2}\frac{1}{t} + \frac{1}{3t^2} \ln \frac{1}{t} - \frac{11}{18}\frac{1}{t^2} - \frac{1}{60}\ \frac{1}{t^4} + \cdots \qquad (166.b)$$

$$F\ (L;1) = \ln 1 - \frac{3}{2} + \frac{\pi}{3}\frac{1}{t} + \frac{1}{6t^2} \ln \frac{1}{t} - \frac{25}{72}\frac{1}{t^2} - \frac{1}{180}\frac{1}{t^4} + \cdots \qquad (166.c)$$

And for $t < 1$:

$$\alpha\ (L;1) = \ln L + \frac{t^2}{12} - \frac{t^4}{60} + \cdots \qquad (167.a)$$

$$\chi\ (L;1) = \ln L - 1 + \frac{\pi}{6} \cdot t - \frac{t^2}{12} + \frac{t^4}{180} + \cdots \qquad (167.b)$$

$$F\ (L;1) = \ln L - \frac{3}{2} + \frac{\pi}{3} \cdot t + \frac{t^2}{6} \ln t - \frac{35}{72} t^2 - \frac{t^4}{180} + \cdot\cdot \qquad (167.c)$$

Now we can try to look on the difference between the linear and de Wijsian models in two dimensions. According to Eqs.(84) and (85) and the definitions of the $F(L)$ and $F(L;1)$ functions (Eqs. 147 and 153) we can write the difference between the variance $\sigma^2(L)$ of the line sample and the variance $\sigma^2(L;1)$ of the rectangular sample in the form:

$$\sigma^2(L) - \sigma^2(L;1) = F(L;1) - F(L) \quad \text{for} \quad t < 1 \qquad (168.a)$$

or

$$\sigma^2(1) - \sigma^2(L;1) = F(L;1) - F(1) \quad \text{for} \quad t > 1 \qquad (168.b)$$

Knowing from equations (161) and (164) the functions $F(L)$ and $F(1)$ we have, for some more general forms of the point-like variograms $\gamma(h)$:

$$\gamma\ (h) = A \cdot |h| \quad \text{for the linear model} \qquad (169.a)$$

$$\gamma\ (h) = 3\alpha \cdot \ln|h| \quad \text{for the logarithmic model} \qquad (169.b)$$

for $t < 1$

$$\frac{\sigma^2(L) - \sigma^2(L;l)}{A \cdot L} = \frac{F(L;l)}{L} - \frac{1}{3} \quad \text{for the linear model, and} \qquad (170.a)$$

$$\frac{\sigma^2(L) - \sigma^2(L;l)}{3\alpha} = F(L;l) - \ln l + \ln t + 1.5 \quad \text{for the de Wijs scheme} \qquad (170.b)$$

and for $t > 1$

$$\frac{\alpha^2(l) - \alpha^2(L;l)}{A \cdot l} = \frac{F(L;l)}{L} \cdot \frac{1}{t} - \frac{1}{3} \quad \text{for the linear model} \qquad (171.a)$$

$$\frac{\alpha^2(l) - \alpha^2(L;l)}{3\alpha} = F(L;l) - \ln l + 1.5 \quad \text{for the de Wijs scheme} \qquad (171.b)$$

The functions given by formulae (170) and (171) are plotted in Figs. 55 and 56. In the following example we are going to explain the practical application of these graphs.

Example

Let us refer to Exercise No 1 (Lecture No 3) where for the samples having the length

$$L = 100 \text{ cm} = 1 \text{ m}$$

taken from the borehole section of the length

$$L^1 = 140 \text{ m}$$

the variance of porosity obtained was

$$\sigma^2(L/L^1) = 12.5\ \%^2$$

We shall calculate the variance of the rectangular samples with the sides:

$$L = 100 \text{ cm}$$

$$l = 50 \text{ cm}$$

assuming the linear or the logarithmic behaviour of the underlying point-like variogram $\gamma(h)$ in the form given by Eqs. (169.a)

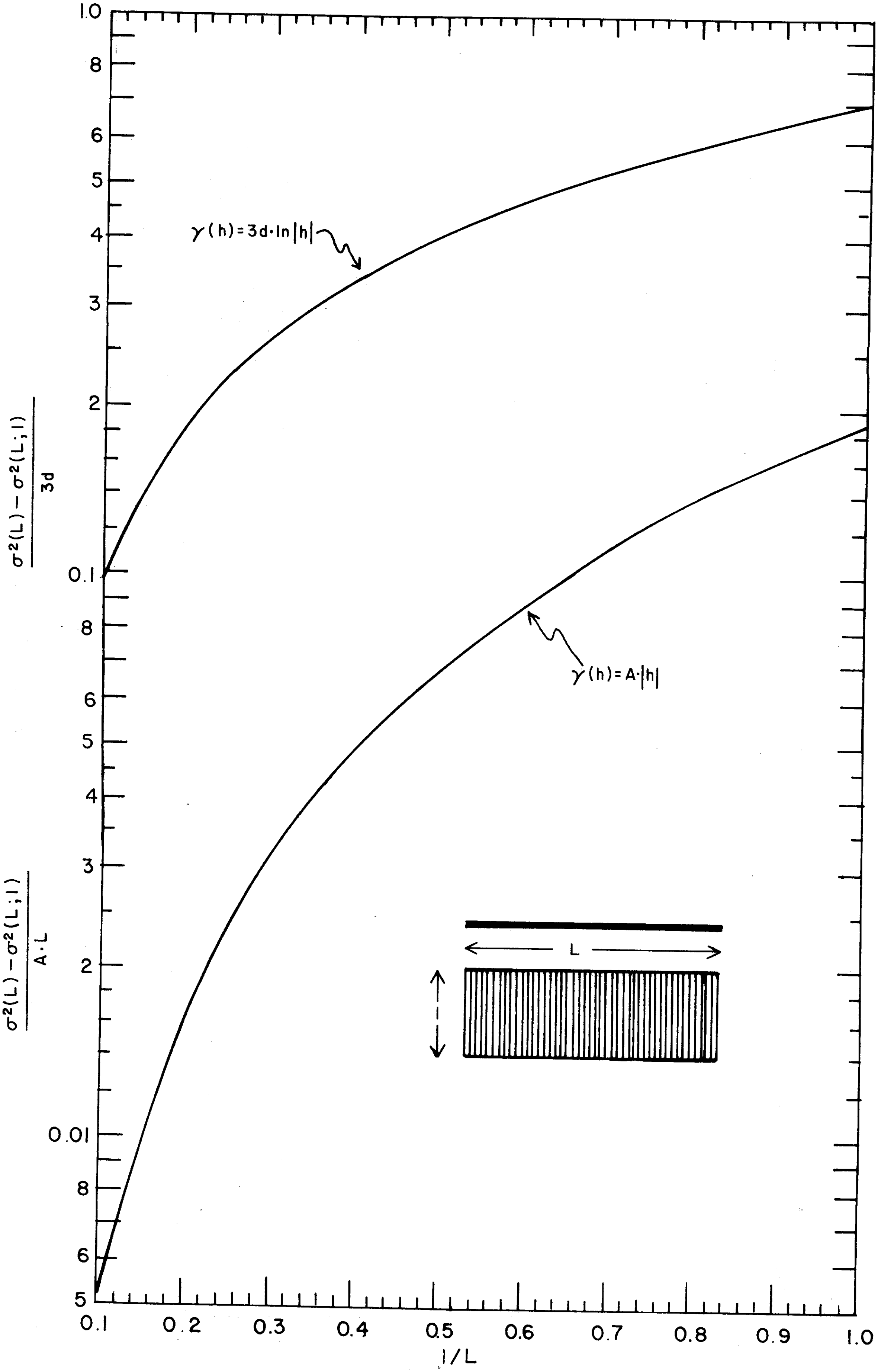

1.0
8
6
5
4
3
2
0.1
8
6
5
4
3
2
0.01
8
6
5
$\frac{\sigma^2(L)-\sigma^2(L;l)}{3d}$
$\frac{\sigma^2(L)-\sigma^2(L;l)}{A\cdot L}$
$\gamma(h)=3d\cdot\ln|h|$
$\gamma(h)=A\cdot|h|$
L
l
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
l/L

Fig. 55

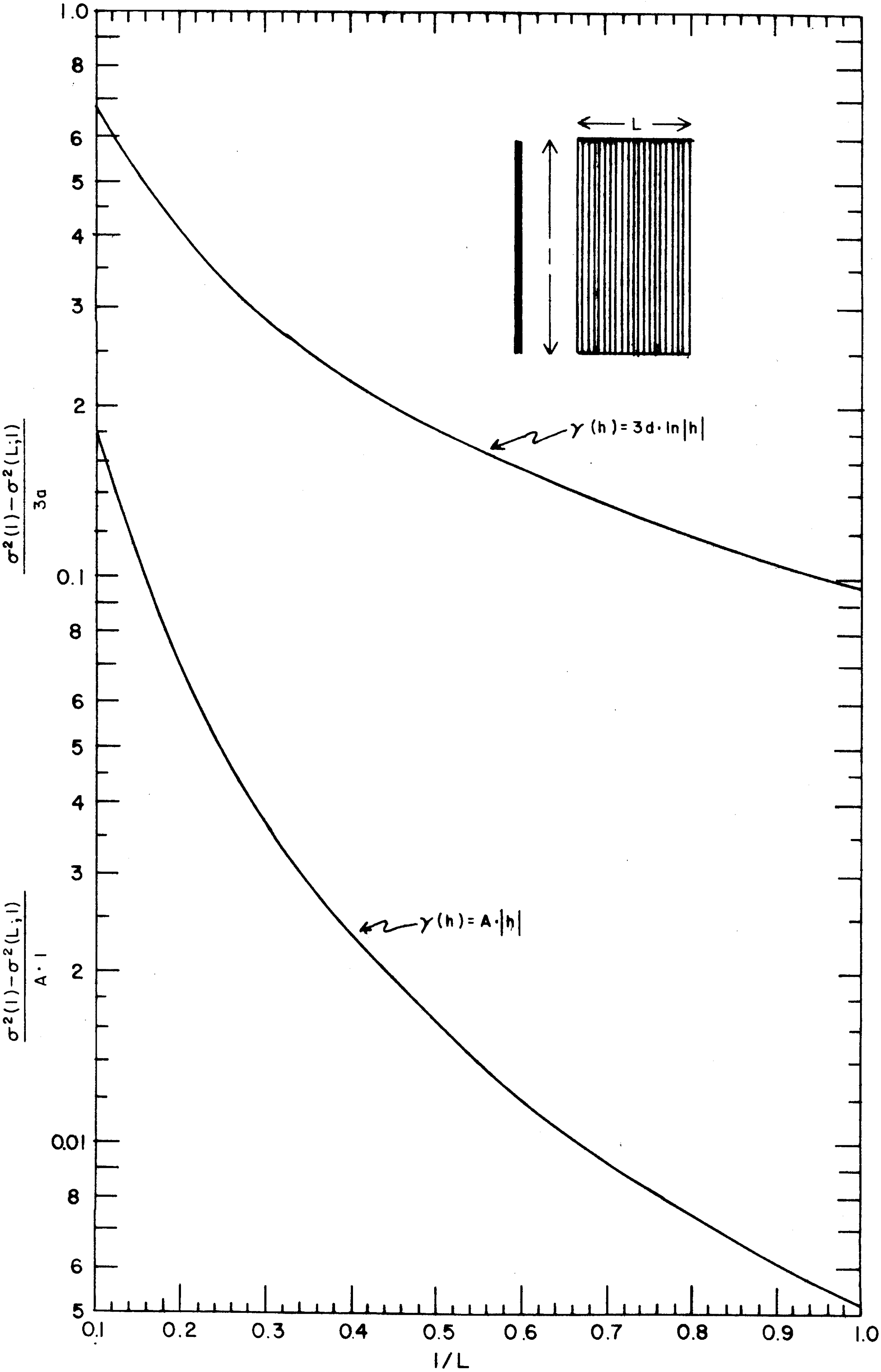

1.0
8
6
5
4
3
2
0.1
8
6
5
4
3
2
0.01
8
6
5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
l/L
$\frac{\sigma^2(l)-\sigma^2(L;l)}{3a}$
$\frac{\sigma^2(l)-\sigma^2(L;l)}{A\cdot l}$
$\gamma(h)=3d\cdot\ln|h|$
$\gamma(h)=A\cdot|h|$
L
l

Fig. 56

and (169.b).

First, we have to calculate the constants A and 3α for the variograms.

Because the initial samples are linear, their variance is (according to Eq. 168.a)

$$\sigma^2(L/L^1) = F(L^1) - F(L) = 12.5\%^2$$

Thus, for both models we have

Linear model:

$$\sigma^2(L/L^1) = A\left[\frac{L^1}{3} - \frac{L}{3}\right]$$

$$A = \frac{3 \cdot \sigma^2(L/L^1)}{L^1 - L} =$$

$$= \frac{3 \cdot 12.5\%^2}{(140 - 1)\,m} = 0.26978\ \%^2\ m^{-1}$$

$$\boxed{A = 0.26978\%^2\ m^{-1}}$$

Logarithmic model:

$$\sigma^2(L/L^1) = 3\alpha\ \ln \frac{L^1}{L} =$$

$$3\alpha = \frac{\sigma^2(L/L^1)}{\ln \frac{L^1}{L}} =$$

$$= \frac{12.5\%^2}{\ln\ 140} = 2.5295\%^2$$

$$\boxed{3\alpha = 2.5295\%^2}$$

Now, for the sample configurations depicted in Fig. 57 for $l = 50\ cm = 0.5\ m$ one has:

$$\text{for}\quad t = \frac{l}{L} = \frac{0.5\ m}{1\ m} = 0.5$$

$$\sigma^2(L;l) = \sigma^2(L) - \frac{\sigma^2(L) - \sigma^2(L;l)}{A \cdot L} \cdot A \cdot L$$

$$\sigma^2(L;l) = \sigma^2(L) - \frac{\sigma^2(L) - \sigma^2(L;l)}{3\alpha} \cdot 3\alpha$$

here one has $\sigma^2(L) = \sigma^2(L/L^1) = 12.5\%^2$, and from Fig. 55 for $t = 0.5$ one has :

$$\frac{\sigma^2(L) - \sigma^2(L;l)}{A \cdot L} = 0.069$$

thus $A \cdot L = 0.26978\%^2$ and

$$\sigma^2(L;l) = 12.5 - 0.069 \cdot 0.26978 =$$

$$= 12.481\ \%^2$$

$$\frac{\sigma^2(L) - \sigma^2(L;l)}{3\alpha} = 0.405$$

thus

$$\sigma^2(L;l) = 12.5 - 0.405 \cdot 2.5295 =$$

$$= 11.475\ \%^2$$

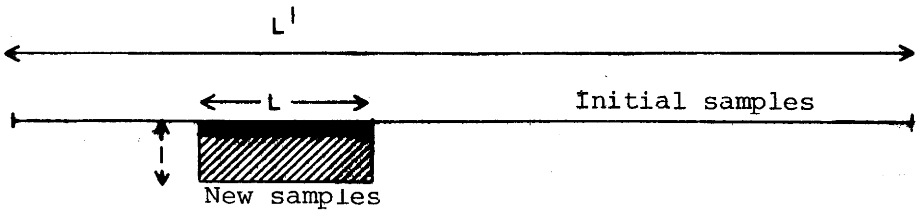


Fig. 57

Finally, for the initial model (linear samples) was:

$$\sigma(L)=\sigma(L/L^1) = \sqrt{12.5}\ \% = 3.535\ \%$$

whereas now, with the new samples 50x100 cm there is

$\sigma(L;1) = \sqrt{12.481} = 3.533\ \%$	$\sigma(L;1) = \sqrt{11.475} = 3.387\ \%$

One can repeat now the calculations for two other values of 1 , for example:

$1 = 100$ cm $\Rightarrow t = 1$ and $1 = 150$ cm $\Rightarrow t = 1.5$

which gives:

for 1 = 100 cm	for 1 = 100 cm
$\frac{\sigma^2(L)-\sigma^2(L;1)}{A \cdot L} = 0.188$	$\frac{\sigma^2(L)-\sigma^2(L;1)}{3\alpha} = 0.695$
$\sigma^2(100;100) = 12.449\ \%^2$	$\sigma^2(100;100) = 10.742\ \%^2$
$\sigma(100;100) = 3.528\ \%$	$\sigma(100;100) = 3.277\ \%$
for 1 = 150 cm	for 1 = 150 cm
$\frac{\sigma^2(1) - \sigma^2(L;1)}{A \cdot L} = 0.1075$	$\frac{\sigma^2(1) - \sigma^2(L;1)}{3\alpha} = 0.515$
thus	thus
$\frac{\sigma^2(L) - \sigma^2(L;1)}{A \cdot L} = 0.3279$	$\frac{\sigma^2(L) - \sigma^2(L;1)}{3\alpha} = 0.920$
$\sigma^2(100;150) = 12.5-0.3279 \times$	$\sigma^2(100;150) = 12.5-0.920\times2.5295 =$
$\times\ 0.26978 = 12.411\ \%^2$	$= 10.173\ \%^2$

and

$$\sigma(100;150) = 3.523\ \%$$

and

$$\sigma(100;150) = 3.189\ \%$$

Here, for the last example (t=1.5) because we know the $\sigma^2(L)$ value and we do not know the $\sigma^2(1)$ we had to transform the

$$\frac{\sigma^2(1)-\sigma^2(L;1)}{A\ .\ 1}$$

value found in Fig. 50 onto the $\frac{\sigma^2(L)-\sigma^2(L\ 1)}{A\ .\ L}$ value according to the obvious formula resulting from the symmetry condition of the function F: $F(L;1)= F(1;L)$:

$$\frac{\sigma^2(L)-\sigma^2(L;1)}{A\cdot L} = \frac{\sigma^2(1)-\sigma^2(L;1)}{A\cdot 1}\cdot t+\frac{t}{3}-\frac{1}{3} \qquad (172.a)$$

for the linear variogram, and

$$\frac{\sigma^2(L)-\sigma^2(L;1)}{3\alpha} = \frac{\sigma^2(1)-\sigma^2(L;1)}{3\alpha} + \ln t \qquad (172.b)$$

for the logarithmic variogram.

From this example we can see that in the case of the linear variogram the gain in the precision of estimation when increasing the sample from linear to the rectangular, is negligible. This gain for the logarithmic variogram is visible in this particular case, though, also not very exciting if one takes into account how much more rock material has been taken to the analysis! Thus, in any practical case, before entering into the detailed sampling operations one has to find out, using a very rough and low cost sampling procedure, the variogram for a given formation, to be next able, even approximately, to calculate the total costs of the sampling operations in function of the available accuracy of the sampling. This problem has to be treated very carefully in each particular case.

3. THE SPHERICAL MODEL

$$\gamma(h) = \begin{cases} \frac{3}{2}\left|\frac{h}{a}\right| - \frac{1}{2}\left|\frac{h}{a}\right|^3 & \text{for } |h| \leq a \\ 1 = \text{sill} & \text{for } |h| \geq a \end{cases} \tag{173}$$

In one dimension the auxiliary functions become

$$\chi(L) = \begin{cases} \frac{3}{4}\frac{L}{a} - \frac{1}{8}\left(\frac{L}{a}\right)^3 & \text{for } |L| \leq a \\ 1 - \frac{3}{8}\frac{a}{L} & \text{for } L \geq a \end{cases} \tag{174}$$

$$F(L) = \begin{cases} \frac{1}{2}\frac{L}{a} - \frac{1}{20}\left(\frac{L}{a}\right)^3 & \text{for } L \leq a \\ 1 - \frac{3}{4}\frac{a}{L} + \frac{1}{5}\left(\frac{a}{L}\right)^2 & \text{for } L \geq a \end{cases} \tag{175}$$

For two and three dimensions the formulae for the auxiliary functions become very complicated. The graph of the function:

$$\gamma_G(r) = \gamma_G(r=L) = \alpha(L;l)-\alpha(O;l) = \alpha(L;l)-F(l) \tag{176}$$

was given in Fig. 44.

The function $\chi(L;l)$ is given in Fig. 58, the function $H(L;l)$ in Fig. 59, and the function $F(L;l)$ is given in Fig. 60. For the three dimensional case in Fig. 61 the function $F(L;l^2)$ is presented and in Fig. 62 the differences

$$\alpha(L;l^2) - \alpha(O;l^2) = \alpha(L;l^2) - F(l;l)$$

are given. All these graphs are taken from the monograph of Journel and Huijbregts [1]

For the parallelepiped of the sizes $a>b>c$ the function $F(a;b;c)$ for the spherical variogram can also be obtained starting from the model variogram $\gamma(h)$ of the type

$$\gamma(h) = |h|^\lambda \tag{177}$$

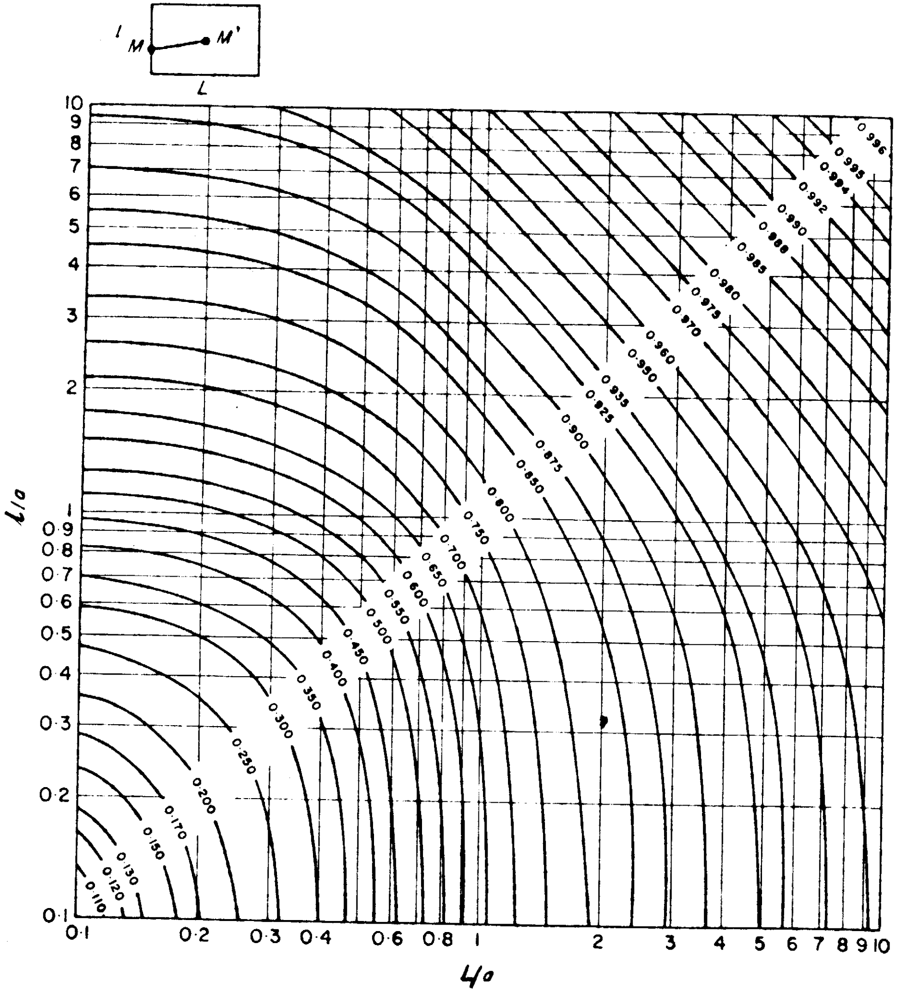


CHART NO. 2. Spherical model. Function $\chi(L; l)$.

Fig. 58

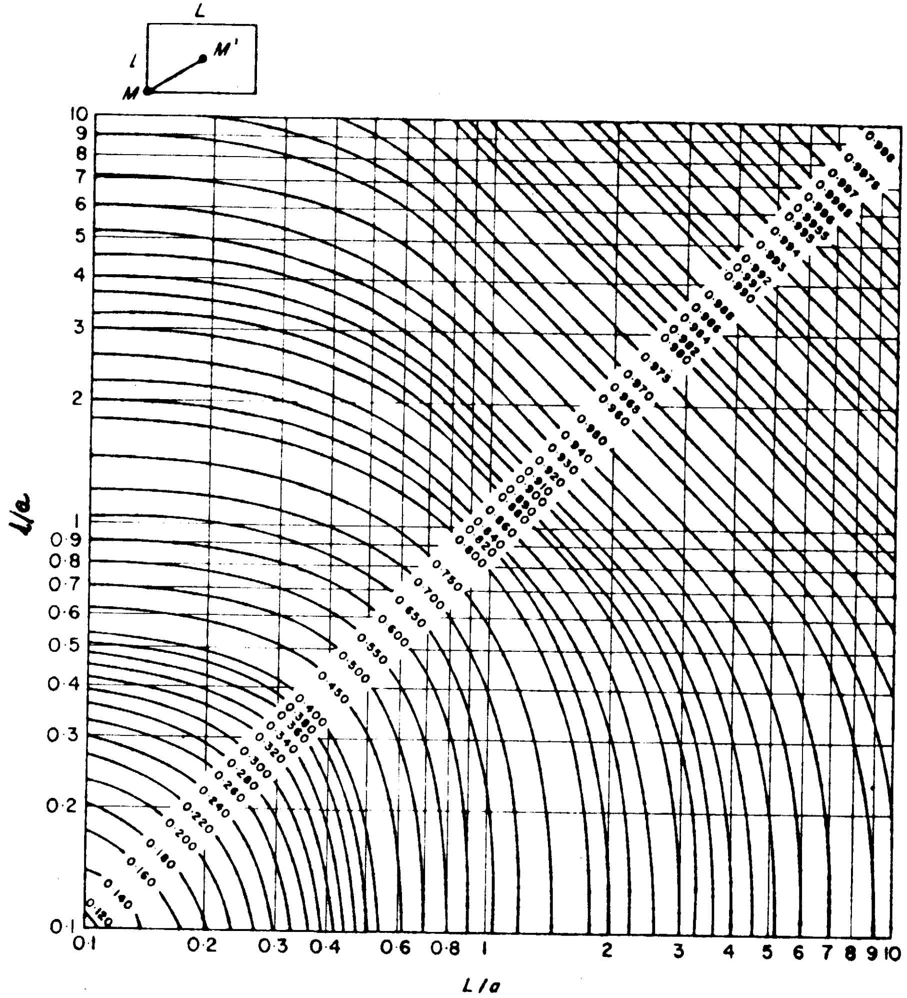


CHART NO. 3. Spherical model. Function $H(L; l)$.

Fig. 59

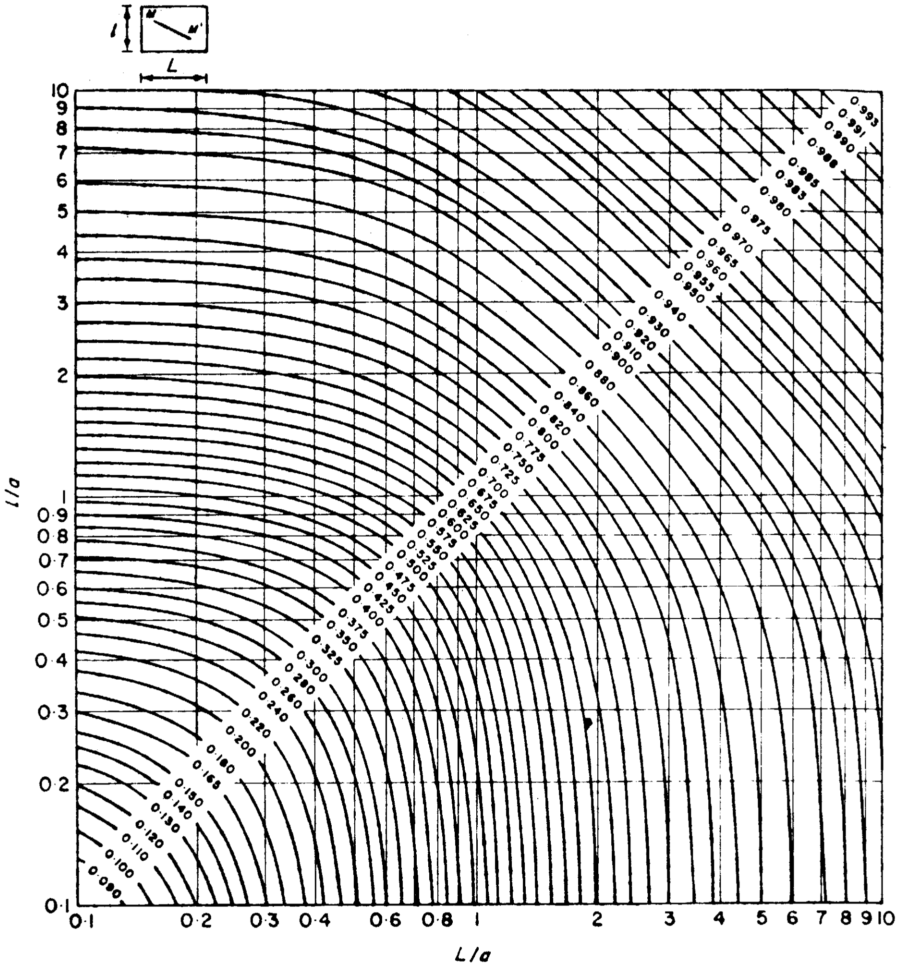


CHART NO. 4. Spherical model. Function $F(L; l)$.

Fig. 60

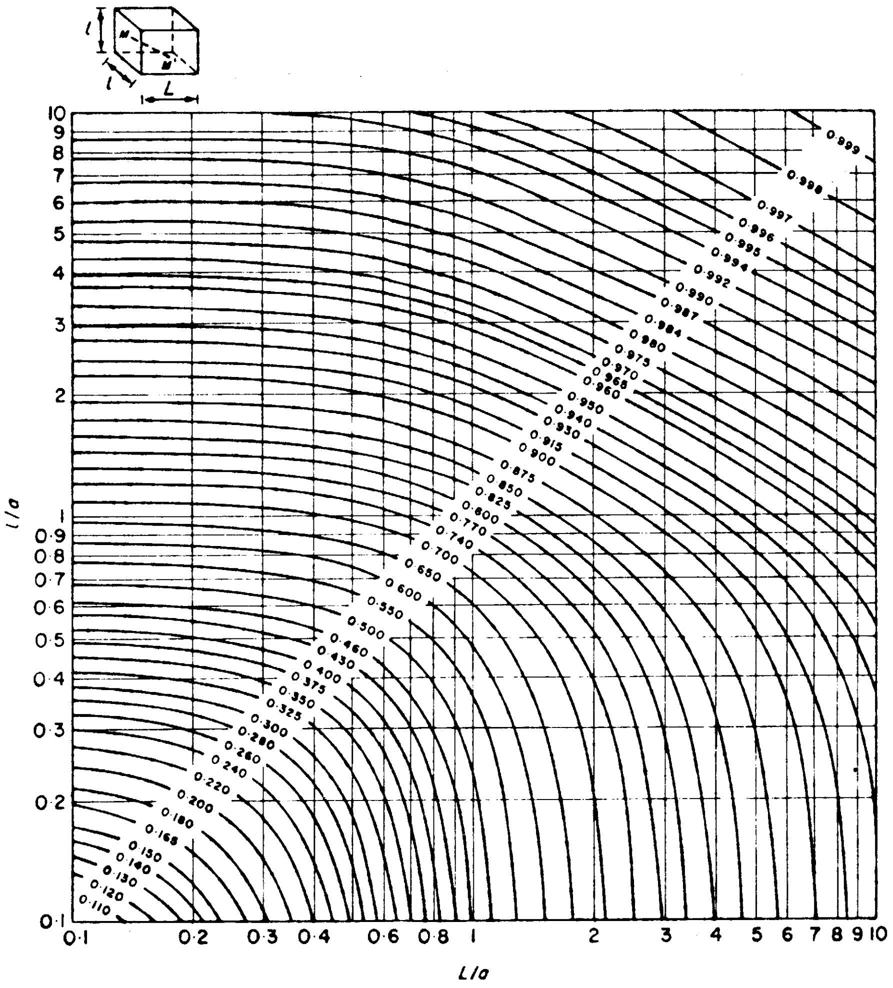


CHART NO. 5. Spherical model. Function $F(L; l^2)$.

Fig. 61

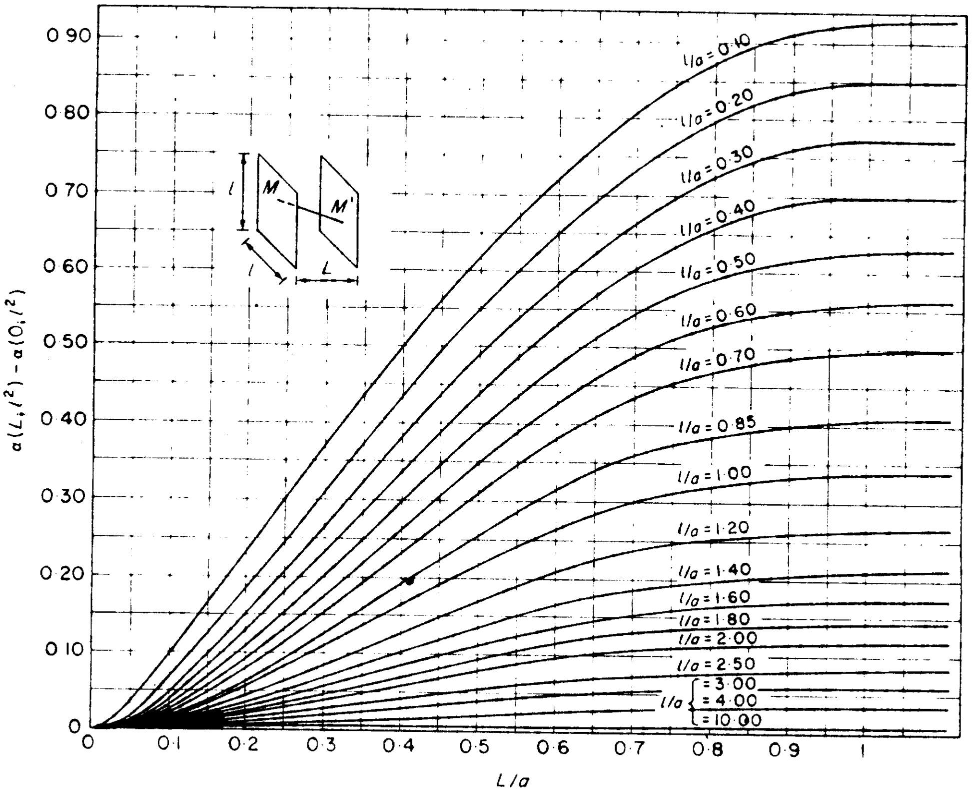


CHART NO. 6. Spherical model. Grading over l^2: $\alpha(L; l^2)-\alpha(0; l^2)$.

Fig. 62

The spherical variogram given by Eq. (173) can be taken as some weighted sum of the two variograms of the type given by Eq.(177) for $\lambda = 1$ and 3. In that case one has to calculate the functions:

$$F_\lambda (a_1,a_2,a_3) = \frac{8}{\tilde{a}_1^2 \tilde{a}_2^2 \tilde{a}_3^2} \int_0^{\tilde{a}_1} \int_0^{\tilde{a}_2} \int_0^{\tilde{a}_3} (\tilde{a}_1-x)(\tilde{a}_2-y)(\tilde{a}_3-z)(x^2+y^2+z^2)^{\lambda/2} dxdydz \qquad (178)$$

which is simply an equation (147) taken for the three dimensional case. Here $\tilde{a}_1$, $\tilde{a}_2$, $\tilde{a}_3$ and the sizes of the parallelepiped given in the units of the range a of the spherical variogram, i.e.:

$$\tilde{a}_1 = a_{1/a} \qquad \tilde{a}_2 = a_{2/a} \quad ; \quad \tilde{a}_3 = a_{3/a} \qquad (179.a)$$

The integral of the form given by equation (178) cannot be expressed by elementary functions, and has to be developed into convergent series. The result, for the spherical variogram and for the special case:

$$a_2 = a_3 \quad \Rightarrow \quad \tilde{a}_2 = \tilde{a}_3 \qquad t = \frac{a_2}{a_1}$$

and for (179.b)

$$\tilde{a}_1 < \sqrt{\frac{1}{1+2t^2}}$$

(which means that the diagonal of the parallelepiped does not exceed the range a of the spherical variogram) is:

$$\frac{F_{sph}(a_1, t)}{C} = \tilde{a}_1[0.5 - t^2(0.5 \ln t - 2.8556905 \cdot 10^{-1}) + + t^3 \cdot 2.4072624 \cdot 10^{-1} - t^4 \cdot 3.54166 \cdot 10^{-2} + t^6 \cdot \cdot 2.1577381 \cdot 10^{-3} - t^8 \cdot 4.6378412 \cdot 10^{-4} \ldots] -$$

$$- \tilde{a}_1^3[0.05 + 8.333317 \cdot 10^{-2} \cdot t^2 + t^4(4.6203835 \cdot 10^{-2} - - 7.083333 \cdot 10^{-2} \ln t) + t^5 \cdot 3.1561269 \cdot 10^{-2} - - t^6 \cdot 4.3154739 \cdot 10^{-3} + t^8 \cdot 2.3189507 \cdot 10^{-4} \ldots] \qquad (180)$$

Here the multiplicative factor C appearing in the definition of the spherical variogram given in Eq. (93) and not depicted in the formula (173) appears as a dividing factor for the function $F\ (\tilde{a}_1, t)$.

The plot of the function given by equation (180) is presented in Fig. 63 . Eq. 180 could very easily be programmed onto the pocket programmable calculator even of the type TI-56 or HP-25.

EXAMPLE

For the example discussed above we assume now the existance of the spherical variogram with the parameters reported in Fig. 37 , i.e.

$C_o = 0 \qquad C = 14.16\ \%^2 \qquad a = 5$ m $\qquad \sigma^2(L/L^1) = 12.5\ \%^2$

L = 100 cm.

The general formula to calculate the variance $\sigma^2(L;1)$ or $\sigma^2(L;1^2)$ of the two or three dimensional samples will be, according to the general Eq. (31):

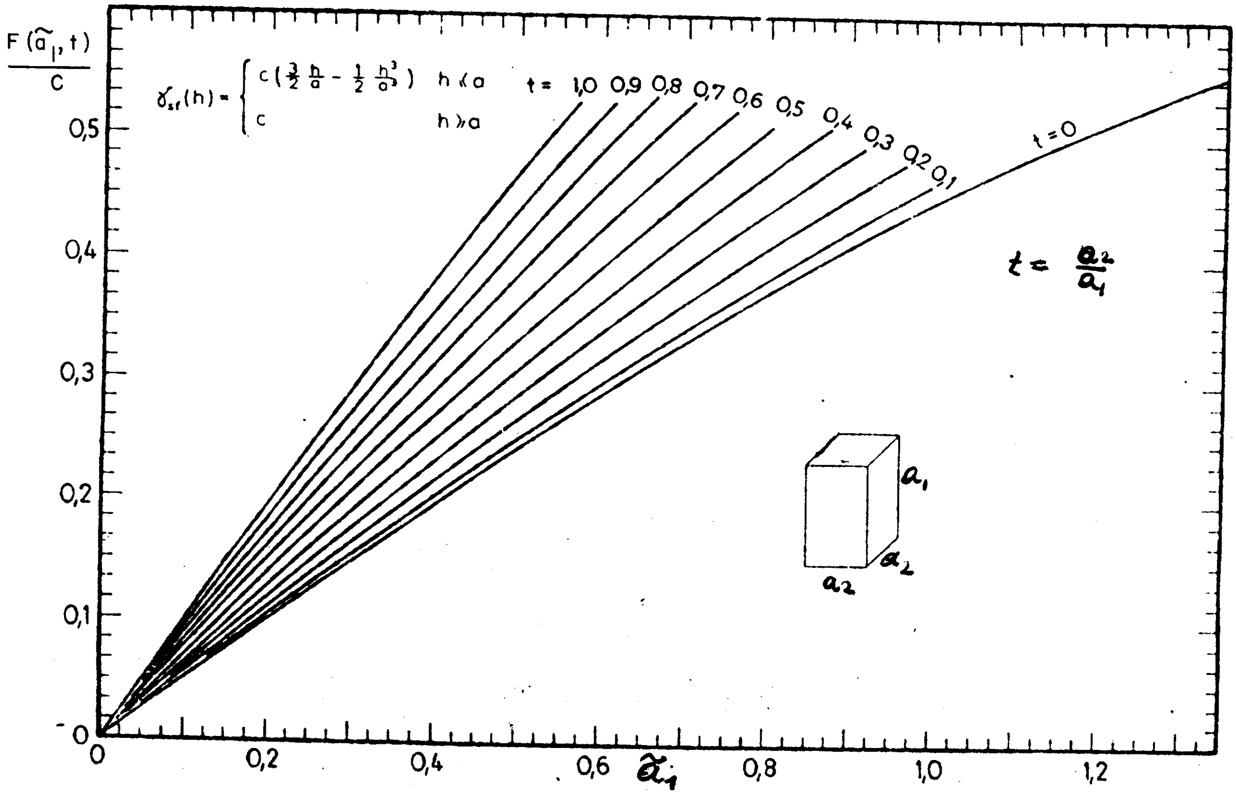


Fig. 63

$$\sigma^2(L;l) - \sigma^2(L;l) = F(L) - F(L;l)$$

or (181)

$$\sigma^2(L;l^2) - \sigma^2(L/L^1) = F(L) - F(L;l^2)$$

in our case, for L=100 cm and a=5 m , the value of F(L) according to Eq. 175 is:

$$F(L) = 0.0996$$

For the two dimensional samples Lxl one finds in Fig. 60 the values of the F(L;l) functions:

l [cm]	$\frac{l}{a}$	$\frac{L}{a}$	F(L;l)
50	0.1	0.2	0.135
100	0.2	0.2	0.200
150	0.3	0.2	0.265

and according to the first formula in Eq. (181) one has:

$$\sigma^2(L;l) = 12.5 - 14.16 \times [F(L;l) - 0.0996] \qquad (181.a)$$

and one has the results:

l (cm)	$\sigma^2(L;l)$ %^{2}	$\sigma(L;l)$ %
50	11.998	3.464
100	11.078	3.328
150	10.158	3.187

Now let us try the three dimensional case. As a matter of fact the porosity data in Table 1 used in Exercise 1 has been obtained using the neutron probe. Assuming for the average porosity there $\phi = 5.6\%$ that the migration length M of thermal neutron is

$$M = 19 \text{ cm}$$

one can assume, that 95% of the neutron signal is from the rock layer distant not more than two migration lengths from the probe axis, thus one has a right cylinder of the hight L=100 cm and the radius equal to 2M = 38 cm , as the rock sample. Taking the same area of the square as the base of this cylinder, one has the size of this square:

$$a_2 = \sqrt{s} = \sqrt{\pi (2M)^2} \cong 67 \text{ cm}$$

Now the sizes of parallelepiped are (with a=5 m):

$$\tilde{a}_1 = \frac{100}{500} = 0.2 \qquad \tilde{a}_2 = \frac{0.67}{5} = 0.134 \text{ , and the } t = 0.67$$

and from Fig. 63 or Eq. 180 are now:

$\frac{1}{\tau} F_{sph} (L,l^2) = 0.158$. Introducing this value to the formula (181.a) instead of the $F(L;l)$ value one has: $\sigma^2(L;l^2) = 11.673$ % and $\sigma(L;l^2) = 3.416\%$, which again, in comparison with the value $\sigma(L/L^1) = \sqrt{12.5} = 3.535\%$ does not give any real improvement in the precision of the estimation.

4. THE EXPONENTIAL MODEL

$$\gamma(h) = 1- \exp\ (-|h|/a_1) \qquad (182)$$

In one dimension the auxiliary function becomes:

$$\chi(L) = 1- \frac{a_1}{L}\ (1 - e^{-L/a_1}) \qquad (183)$$

and

$$F(L) = 1- \frac{2a_1}{L}\ [1 - \frac{a_1}{L}(1 - e^{-L/a_1})] \qquad (184)$$

For the two and three dimensions, similarity to the case of the spherical model, one uses the graphs of the auxiliary functions calculated at the Fontainebleau Center of Geostatistics. The graph of the function

$$\gamma_G(r) \equiv \gamma_G\ (r=L) = \alpha(L;l) - \alpha(O;l) = \alpha\ (L;l) - F(l) \qquad (176)$$

was presented in Fig. 45.

The functions $\chi(L;l)$, $H(L;l)$, $F(L;l)$, $F(L;l^2)$
and

$$\alpha(L;l^2) - \alpha(O;l^2) = \alpha(L;l^2) - F(l;l)$$

are given in figures 64, 65, 66, 67 and 68 , respectively.

Example (continued)

Now, following the example discussed above, we assume that the underlying variogram is exponential with the constants:

$$C = 14.16\% \quad \text{and} \quad a_1 = \frac{a}{3} = \frac{5\ m}{3} = 1.6667\ m.$$

For $L = 1$ m the function $F(l)$ given by Eq. (181) is:
$L/a_1 = 0.6$, $F(L) = 0.1733$ and according to Eq. 181 one has:
(always for $\sigma^2(L/L^1) = 12.5\%^2$)

$$\alpha^2(L;l) = 12.5 - 14.16 \times [F(L;l) - 0.1733] \qquad (177)$$

which for the rectangles with $l = 50;100;$ and 150 cm and for the right cylinder, or rather parallelepiped discussed previously gives (using the charts in Fig. 66 and 67):

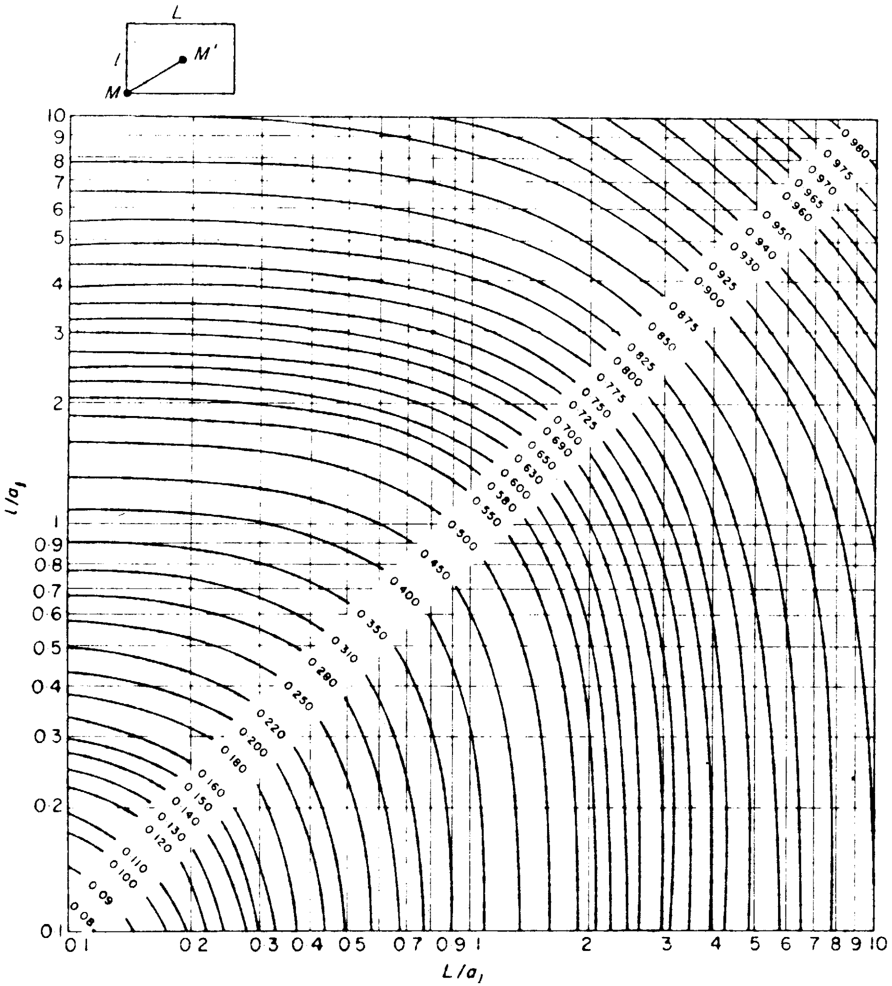


CHART NO. 13. Exponential model. Function $H(L; l)$.

Fig. 65

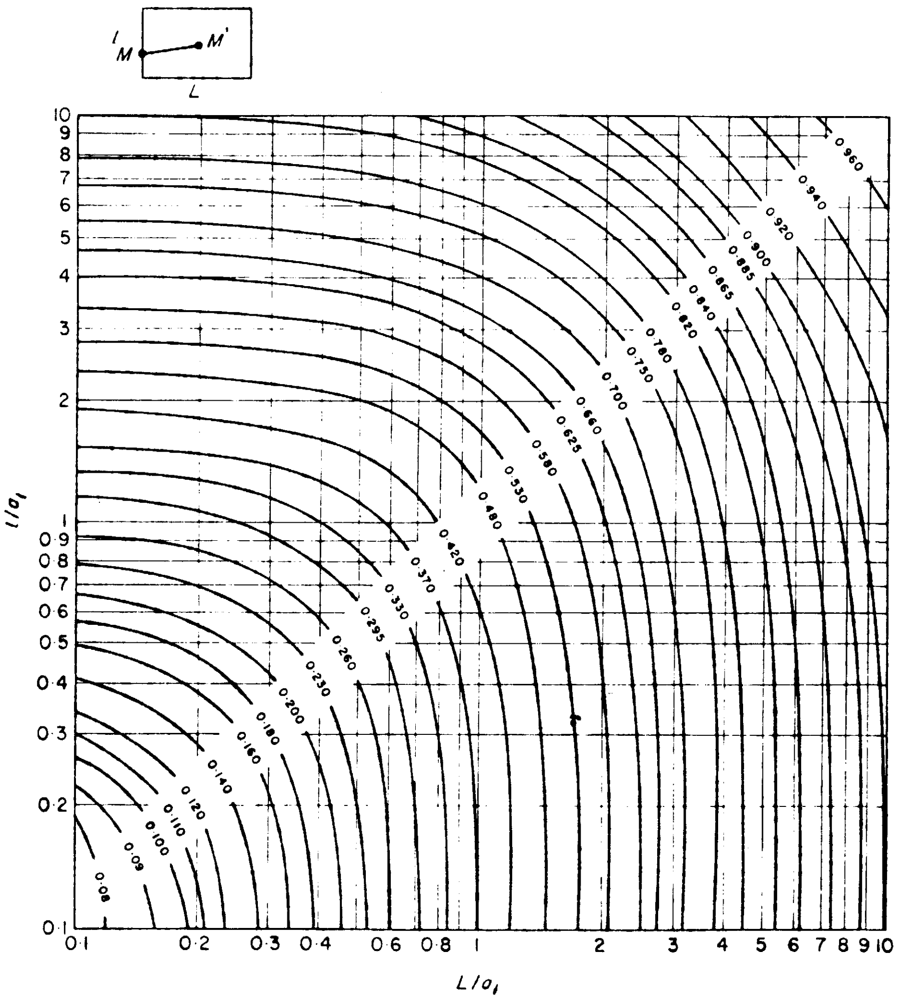


CHART NO. 12. Exponential model. Function $\chi(L; l)$.

Fig. 64

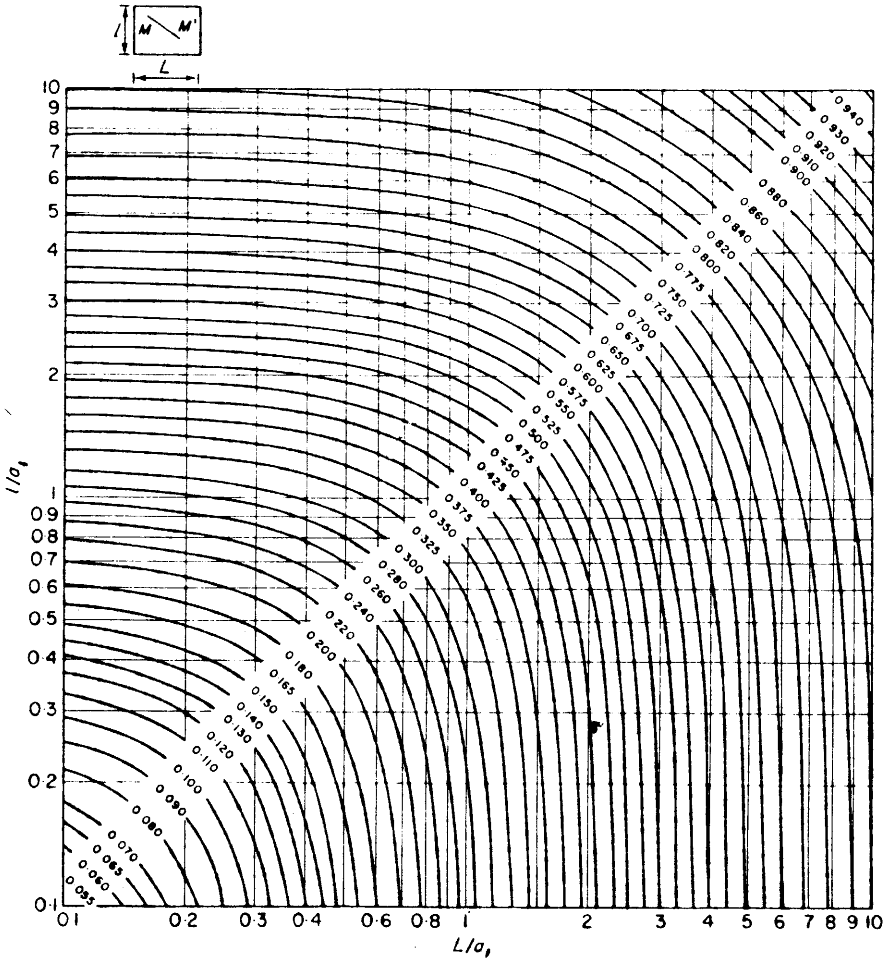


CHART NO. 14. Exponential model. Function $F(L; l)$.

Fig. 66

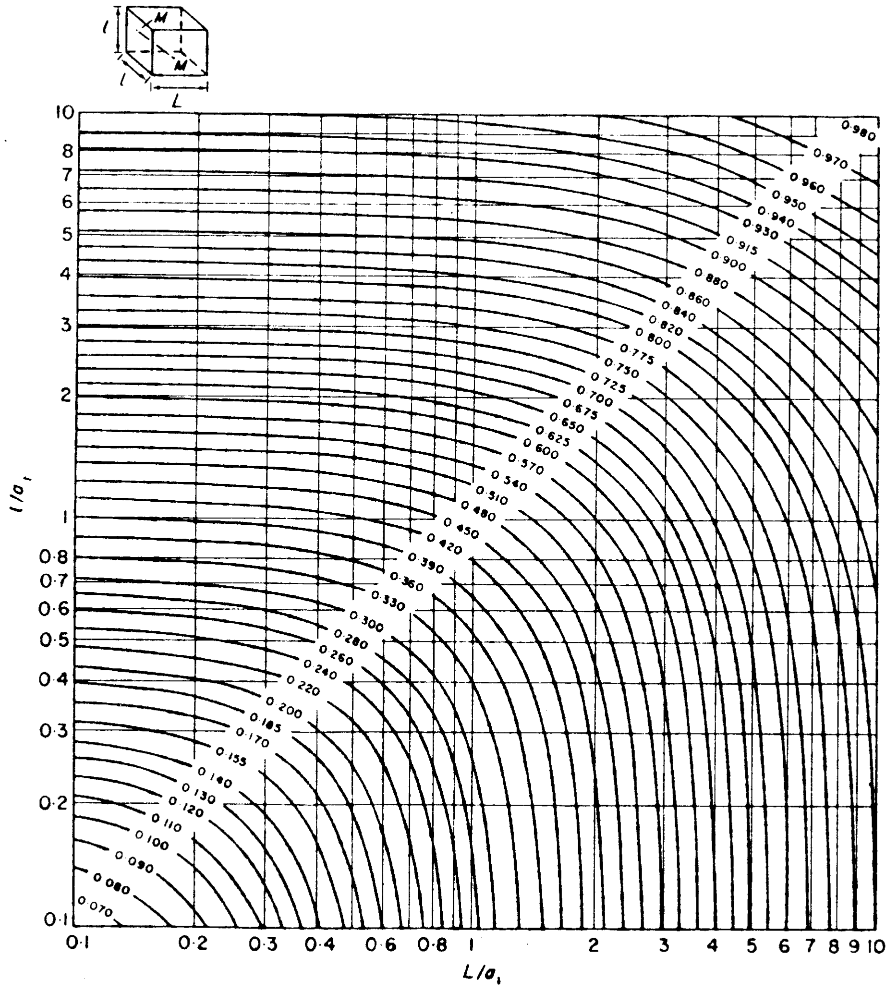


CHART NO. 15. Exponential model. Function $F(L; l^2)$.

Fig. 67

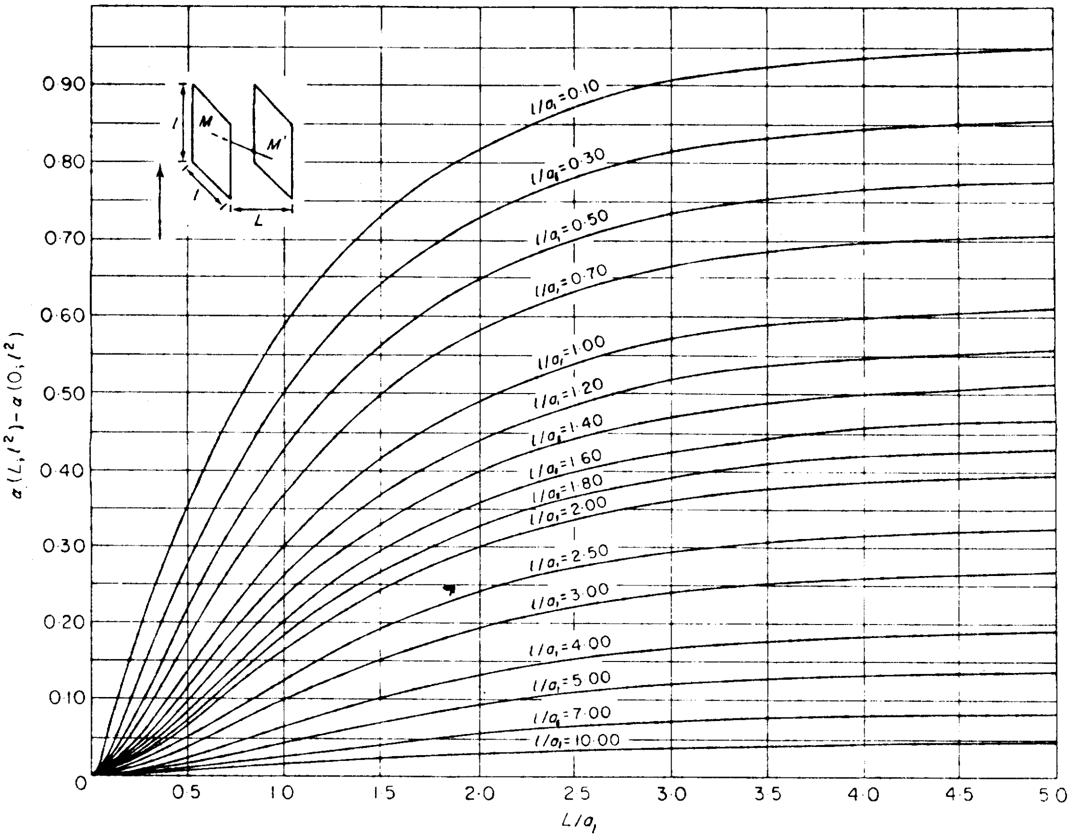


CHART NO. 16. Exponential model. Grading over l^2: $\alpha(L; l^2)-\alpha(0; l^2)$.

Fig. 68

l [cm]	$\frac{l}{a_1}$	$\frac{L}{a_1}$	$F(L;l)$	$\sigma^2(L,l)$	$\sigma(L,l)$
50	0.3	0.6	0.210	11.980	3.461
100	0.6	0.6	0.260	11.272	3.357
150	0.9	0.6	0.312	10.536	3.246
$l \times l=$			$F(L;l^2)$	$\sigma^2(L,l^2)$	$\sigma(L,l^2)$
67^2	0.402	0.6	0.260	11.272	3.357

Now we can present some resumée of the $\sigma(L,l)$ calculations for the some shapes of the samples, but assuming different variograms (which are, however, compatible more or less with the real situation, as far as concerns the variance of the linear one meter samples, which is assumed to be constant for all cases).

Standard deviations of porosity $\alpha_{\phi}(L;l)$:%

Sample:	L	L x l	L x l	L x l	L x l^2
	100 cm	100x50	100x100	100x150	100x67^2
Observed	3.535				
Linear model		3.533	3.528	3.523	
de Wijs model		3.387	3.277	3.189	
Spherical model		3.464	3.328	3.187	3.416
Exponential model		3.461	3.357	3.246	3.357

One can see from the results presented above, that there is not a big difference in them when different models for the theoretical variogram are used (when they are well normalized to each other, of course!) and that the gain in the precision when increasing the volume of the samples is almost not observable. This result is valid just for the example being discussed here. A similar discussion one has to perform in each individual case, when one has to decide upon the sampling system for a given parameter (porosity, thickness, etc.) in a given geological formation.

LECTURE 8

MORE ABOUT THE ESTIMATION VARIANCE

In LECTURE 4 we have introduced the concept of the estimation variance

$$\sigma_E^2[V(x),\ v(x^1)]$$

of the volume V situated at x by the volume v centered at x^1 (cf. Eq. 78). Equation 78 obtained for this estimation variance demands an integration over the volumes V and v. Let us assume now, that the volume $v(x^1)$ - there is a set of point samples situated at $x_1, x_2 \ldots x_k$. How to transform Eq. 78 for the calculations valid for this new situation? Let us repeat first Eq. 78 here:

$$\sigma_E^2[V(x),\ v(x^1)]= \frac{2}{V\cdot v} \int_{V(x)} dx_1 \int_{v(x^1)} dx_2^1\ \gamma\ (x_1 - x_2^1) -$$

$$- \frac{1}{V^2} \int_{V(x)} dx_1 \int_{V(x)} dx_2\ \gamma\ (x_1 - x_2) - \frac{1}{v^2} \int_{v(x^1)} dx_1^1 \int_{v(x^1)} dx_2^1\ \gamma\ (x_1^1 - x_2^1) \qquad (78)$$

Now the set of k point-like samples situated at points x_1^1 , x_2^1 , x_3^1 , ... x_k^1 can be given as a distribution of the samples density:

$$\frac{dv(x^1)}{d x^1} = \sum_{i=1}^{k} \delta(x^1 - x_i^1) \qquad (185)$$

and the volume of the sample $v(x^1)$ is now simply:

$$v = v(x^1) = \int_{x^1} \frac{dv(x^1)}{dx^1} \cdot dx^1 = \int_{x^1} \sum_{i=1}^{k} \delta(x^1 - x_i^1)\, dx^1 = k \qquad (186)$$

Thus, it is simply a total number k of the point samples, and

$$\int_{V(x)} dx_1 \int_{x^1} dx_2^1 \sum_{i=1}^{k} \delta(x_2^1 - x_i^1)\ \gamma\ (x_1 - x_2^1) =$$

$$= \sum_{i=1}^{k} \int_{V(x)} dx_1 \cdot \gamma\ (x_1 - x_i^1) \qquad (187)$$

and

$$\int_{v(x^1)} dx_1^1 \int_{v(x^1)} dx_2^1 \; \gamma \; (x_1^1-x_2^1) =$$

$$= \int_{x^1} dx_1^1 \sum_{j=1}^{k} \delta(x_1^1-x_j^1) \int_{x^1} dx_2^1 \sum_{i=1}^{k} \delta \; (x_2^1-x_i^1) \; \gamma \; (x_1^1-x_2^1) =$$

$$= \int_{x^1} dx_1^1 \sum_{j=1}^{k} \delta(x_1^1-x_j^1) \cdot \sum_{i=1}^{k} \gamma(x_1^1-x_i^1) =$$

$$= \sum_{j=1}^{k} \sum_{i=1}^{k} \gamma(x_j^1-x_i^1) \tag{188}$$

Now, setting equations (187) and (188) into Eq. (78) gives:

$$\sigma_E^2 [V(x) ; v(x^1)] = \frac{2}{V \cdot k} \sum_{k=1}^{k} \int_{V(x)} dx_1 \; \gamma(x_1-x_i^1) -$$

$$- \frac{1}{V^2} \int_{V(x)} dx_1 \int_{V(x)} dx_2 \; \gamma \; (x_1-x_2) - \frac{1}{k^2} \sum_{j=1}^{k} \sum_{j=1}^{k} \gamma(x_j^1-x_i^1) \tag{189}$$

Denoting the set of the point-like samples as ε we can rewrite Eq. (189) in the shortened form:

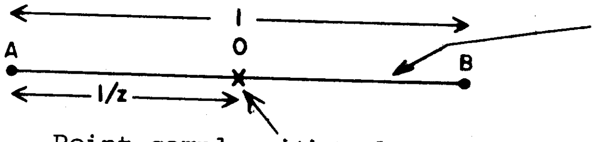
$$\sigma_E^2 \; [V;\varepsilon] = 2\bar{\gamma} \; (V,\varepsilon) - \bar{\gamma} \; (V,V) - \bar{\gamma}(\varepsilon,\varepsilon) \tag{190}$$

where V can again have a meaning of the set of the point samples, line samples, surface or three-dimensional samples.

Now, using Eq. (190) or (189) we shall try to calculate some estimation variances using the auxiliary functions just learned.

ESTIMATION VARIANCE PROBLEMS

1. One has a segment AB of length 1 and one point-like sample has been taken in the middle of the segment



Line segment with value z(1)

Point sample with value z

Fig. 69

We extend the obtained value z on the whole segment AB which really has some value z(1).
What is the extension variance

$$\sigma^2_{E_1} = D\{z-z(1)\} \qquad ?$$

This is just done by equation (190) which in this case is written:

$$\sigma^2_E = 2\ \bar{\gamma}(AB,O) - \bar{\gamma}\ (AB,AB) - \bar{\gamma}(O,O) \tag{191}$$

with

$$\bar{\gamma}(AB,O) = \frac{1}{1} \int_1 dx_1 \cdot \gamma\ (|x_1 - \frac{1}{2}|) = \frac{1}{1/2} \int_{1/2} dx\ \gamma\ (x) =$$

$$= \chi(1/2) = \bar{\gamma}\ (OA,O) = \bar{\gamma}\ (OB,O) \tag{192}$$

and

$$\bar{\gamma}(AB,AB) = F(1) \quad ; \quad \bar{\gamma}(O,O) = O \tag{193}$$

Thus in this case:

$$\sigma^2_{E_1} = 2\chi(1/2) - F(1) \tag{194}$$

2. One has the same segment AB as in 1 (Fig. 69) but now the point-like sample is situated at the point A (i.e. at the extremity of the segment).
Thus now is:

$$\bar{\gamma}(AB,O) = \bar{\gamma}\ (AB,A) = \chi(1)$$

and the estimation variance is:

$$\sigma^2_{E_2} = 2\chi(l) - F(l) \qquad (195)$$

3. We have again the segment AB but now two point-like samples are situated at both extremities, i.e. at the points A and B. As an estimated value of z we take an average of z_A and z_B and the estimation variance, according to Eq. (190), is:

$$\sigma^2_{E_3} = 2\cdot\bar{\gamma}(\varepsilon,AB) - \bar{\gamma}(AB,AB) - \bar{\gamma}(\varepsilon,\varepsilon) \qquad (196)$$

where by ε we have denoted the ensemble of the samples.
Now: $k = 2$ and

$$\bar{\gamma}(\varepsilon,AB) = \frac{1}{l\cdot 2}\{\int_l dx_1\,\gamma(|x_1-0|) + \int_l dx_1\,\gamma(|x_1-l|)\} =$$

$$= \frac{1}{l}\int_l dx\;\gamma(x) = \bar{\gamma}(A,AB) = \bar{\gamma}(B,AB) = \chi(l) \qquad (197)$$

$$\bar{\gamma}(\varepsilon,\varepsilon) = \frac{1}{2^2}\{\gamma(0-0)+\gamma(0-l)+\gamma(l-0)+\gamma(l-l)\} = \frac{1}{2}\gamma(l) \qquad (198)$$

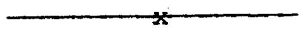
and finally:

$$\sigma^2_{E_3} = 2\chi(l) - F(l) - \frac{1}{2}\gamma(l) \qquad (199)$$

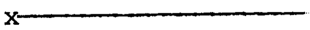
Example

In the case of the spherical variogram the functions $\chi(l)$ and $F(l)$ are given by Eq. (174) and (175) which gives, for $l<a$

$$\sigma^2_{E_1} = \frac{1}{4}\frac{l}{a} + \frac{3}{160}\left(\frac{l}{a}\right)^3 \qquad (194.a)$$



$$\sigma^2_{E_2} = \frac{l}{a} - \frac{1}{5}\left(\frac{l}{a}\right)^3 \qquad (195.a)$$



$$\sigma^2_{E_3} = \frac{1}{4}\frac{l}{a} + \frac{1}{20}\left(\frac{l}{a}\right)^3 \qquad (199.a)$$

and for $l\geq a$ one has:

$$\sigma^2_{E_1} = -1 + \frac{3}{4}\frac{l}{a} - \frac{1}{32}\left(\frac{l}{a}\right)^3 + \frac{3}{4}\frac{a}{l} - \frac{1}{5}\left(\frac{a}{l}\right)^2 \quad \text{for} \quad a\leq l\leq 2a \qquad (194.b)$$

$$\sigma^2_{E_1} = 1 - \frac{3}{4}\frac{a}{l} - \frac{1}{5}\left(\frac{a}{l}\right)^2 \qquad \text{for} \quad l\geq 2a \qquad (194.c)$$

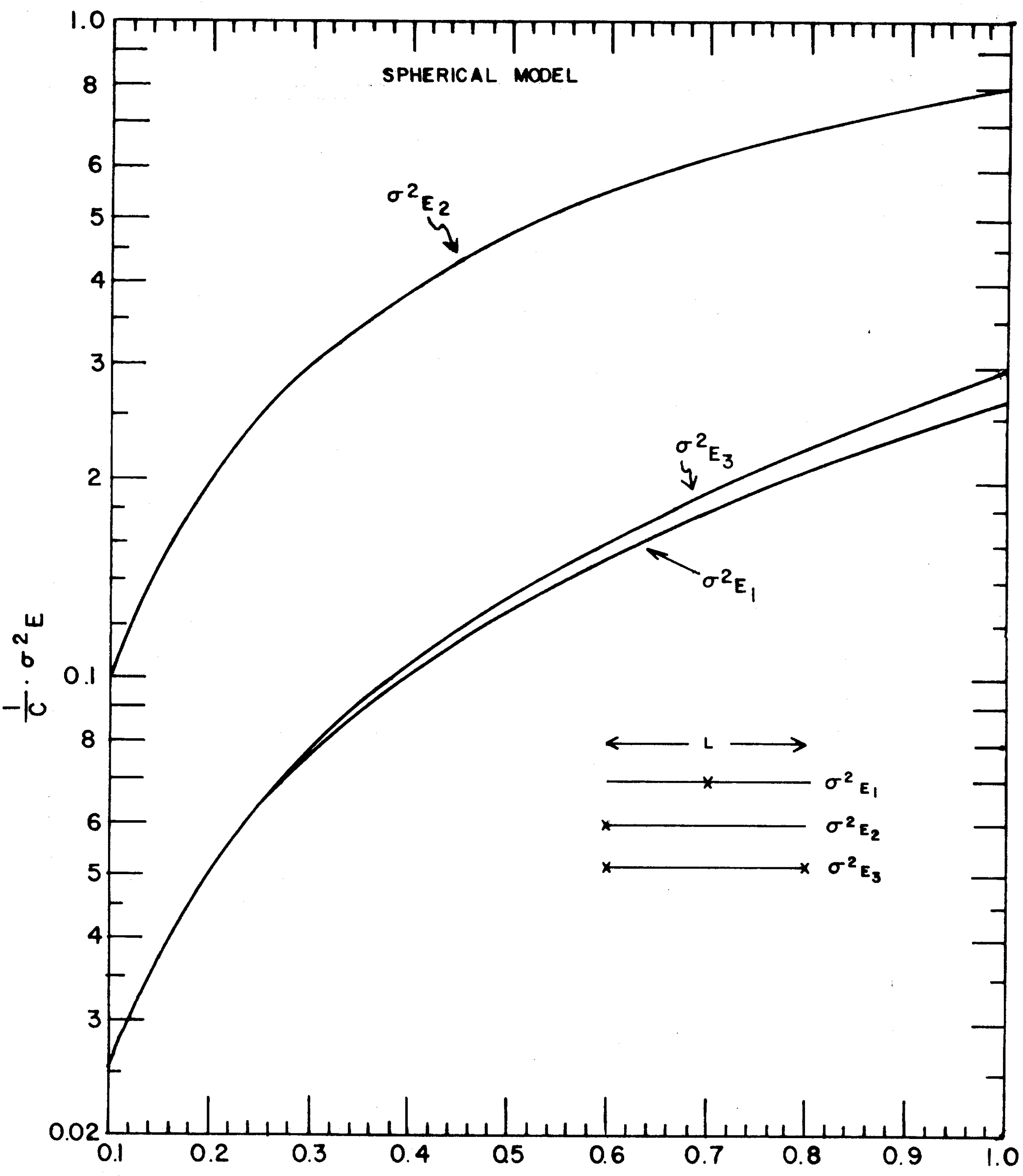


Fig. 70.

$$\sigma^2_{E_2} = 1 - \frac{1}{5}\left(\frac{a}{l}\right)^2 \tag{195.b}$$

$$\sigma^2_{E_3} = \frac{1}{2} - \frac{1}{5}\left(\frac{a}{l}\right)^2 \tag{199.b}$$

The graphs of extension variances $\sigma^2_{E_1}$, $\sigma^2_{E_2}$ and $\sigma^2_{E_3}$ are shown in Figs. 70 and 71. It is surprising at a glance that the variance $\sigma^2_{E_3}$ is in some region bigger than the variance $\sigma^2_{E_1}$ which is obtained with only one sample at the middle of the segment. The

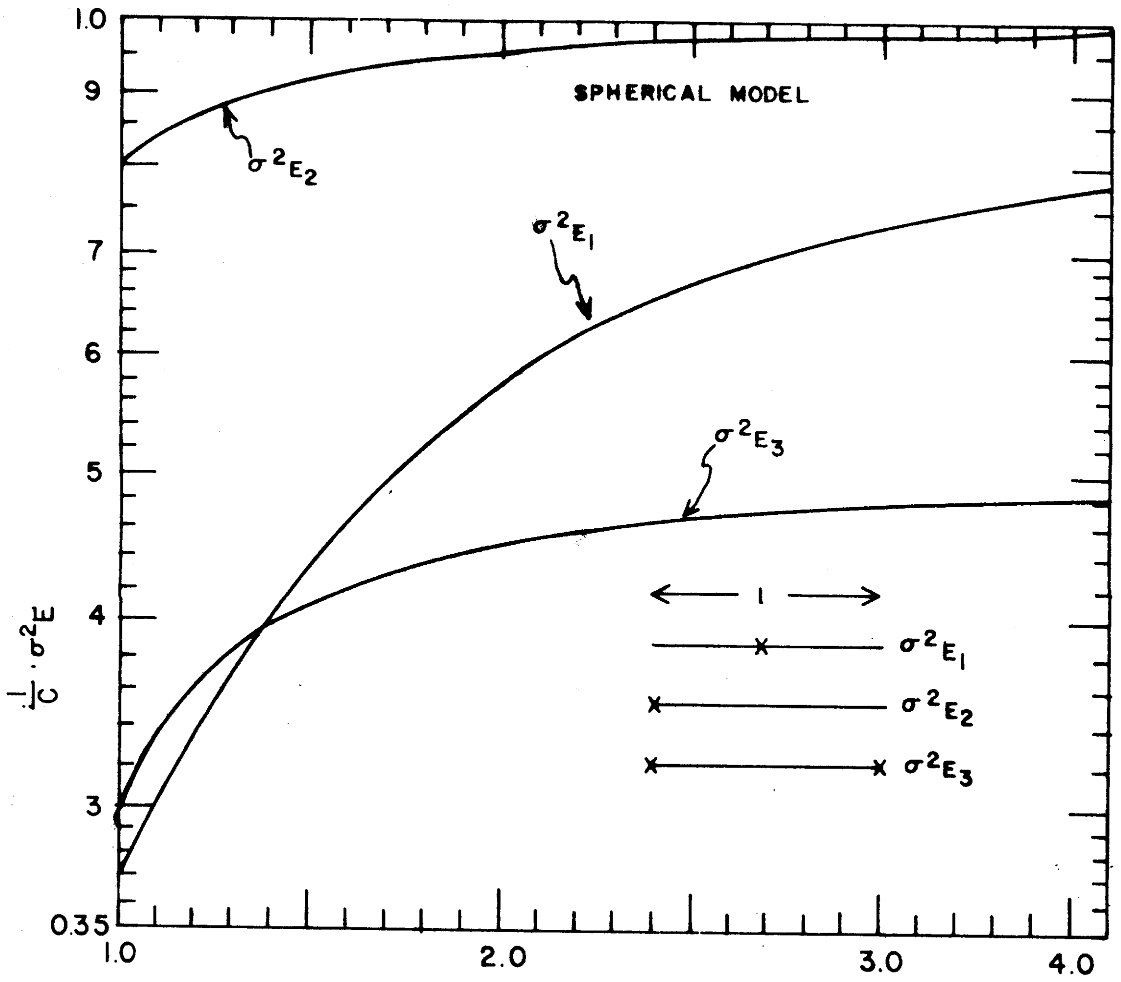


Fig. 71

reason is that in the case of the $\sigma^2_{E_3}$ variance the two samples situated at the two ends of the segment "are seeing" also the variability of the parameter z being measured outside the segment of question. It is only when $l>1.4\,a$ that the variance $\sigma^2_{E_3}$ starts to be lower than the variance $\sigma^2_{E_1}$ and is asymptotically going towards the value of 0.5 which is quite natural, because in this case both samples are measuring the same value z, thus the statistics of measurements is doubled.

The variance $\sigma^2_{E_2}$ is much higher than the variance $\sigma^2_{E_1}$ just because the sample situated at the extremity of the segment l is influenced by the neighbour sample. It is asymptotically reaching (and rather quickly) value 1. The variance for the case when the sample is situated at the middle of the segment l ($\sigma^2_{E_1}$) is also going to this asymptotic value, but not so quick.

When $l >> a$ all samples (belonging to the spherical scheme, however), are independent of each other and they reach simply the value of the sill of the variogram, i.e. the value C.

Formulae (194),(195) and (199) do not contain the sill value C. Thus, to get the real variance one has to multiply the values obtained by the sill value C. For example if one treats the porosity data used in EXERCISE No 1 as the point like data with the variogram (spherical one) described by

$$a = 5 \text{ m}$$

and

$$C = 14.16\%^2$$

one has (for $l = 1$ m in this case) the values

$$\sigma^2_{E_1} = 0.711\%^2 \qquad \sigma_{E_1} = 0.84\%$$

$$\sigma^2_{E_2} = 2.832\%^2 \qquad \sigma_{E_2} = 1.68\%$$

$$\sigma^2_{E_3} = 0.711\%^2 \qquad \sigma_{E_3} = 0.84\%$$

Thus, just in this case, the fact of taking as an estimator for the segment of the length $l = 1$ m the average value of data obtained at its two extremities do not improve the result in comparison with the situation, when only one sample is situated in the middle of the segment and we take its value as a representative for the whole segment.

4. We have now a square ABCD as in Fig. 72

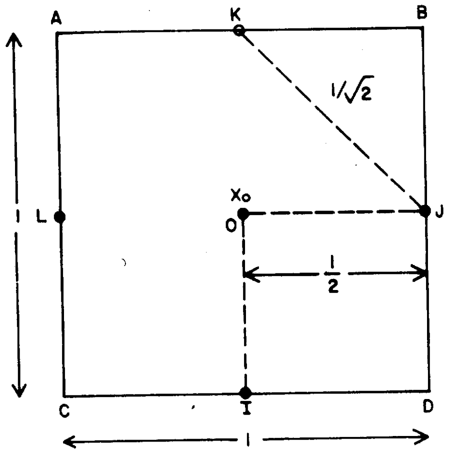


Fig. 72

of side l and one sample at its center O. The z value measured at O we give as an estimate of z for the whole square ABCD. The estimation variance in this case, according to Eq. (189) or (190) is:

$$\sigma^2_{E_4} = 2\bar{\gamma}\,(O,ABCD(- \bar{\gamma}(ABCD,ABCD) - \bar{\gamma}\,(O,O) \qquad (200)$$

with:

$$\bar{\gamma}\,(O,ABCD) = \frac{1}{l^2} \int_{l^2} \gamma\,(|x-x_o|)\,d^2x$$

$$= \frac{4}{l^2} \int_{OJDI} \gamma(\ x-x_o\ )\,d^2x = \cdot\ \bar{\gamma}\,(O;OJDI) =$$

$$= H\,(l/2\,;\,l/2) \qquad (201)$$

$$\bar{\gamma}\,(ABCD\,;\,ABCD) = F\,(l;l) \qquad (202)$$

and

$$\bar{\gamma}(O.O) = \gamma\,(O) = 0 \qquad (203)$$

(nugget effect does not exist)

which finally gives

$$\sigma^2_{E_4} = 2\cdot H(l/2\,;\,l/2) - F(l\,;\,l) \qquad (204)$$

5. Taking the same square ABCD as in Fig. 72 we localize now four point samples at its corners A, B, C and D , and as an estimation for the square we take the arithmetic mean of these four samples. The ensemble of samples we call ε. The estimation (extension) variance will be now (always according to Eq. 189):

$$\sigma^2_{E_5} = 2\,\bar{\gamma}(\varepsilon;ABCD) - \bar{\gamma}\,(ABCD,ABCD) - \bar{\gamma}\,(\varepsilon,\varepsilon) \qquad (205)$$

where

$$\bar{\gamma}\,(\varepsilon;ABCD) = \frac{1}{4}\,[\,\bar{\gamma}\,(A;ABCD) + \bar{\gamma}\,(B;\,ABCD) +$$

$$+\,\bar{\gamma}\,(C\,;ABCD) + \bar{\gamma}\,(D\,;ABCD)\,] =$$

$$= \bar{\gamma}\,(A\,;ABCD) = H(1;1) \qquad (206)$$

here we have used a property of symmetry in the values of $\bar{\gamma}$ (A, ABCD) , $\bar{\gamma}$ (B , ABCD) etc.

For the correlation between the samples we have:

$$\bar{\gamma}\,(\varepsilon\,;\varepsilon) = \frac{1}{4^2}\,\{\gamma\,(x_A - x_A) + \gamma\,(x_A - x_B\,) + \gamma\,(x_A - x_C\,) + \gamma(x_A - x_D) +$$

$$+\gamma\,(x_B - x_B) + \gamma\,(x_B - x_C\,) + \gamma(x_B - x_D) + \gamma\,(x_B - x_A\,) +$$

$$+\gamma\, x_C - x_C\,) + \gamma\,(x_C - x_D\,) + \gamma\,(x_C - x_A\,) + \gamma\,(x_C - x_C\,) +$$

$$+\gamma\,(x_D - x_D\,) + \gamma(x_D - x_A\,) + \gamma\,(x_D - x_B) + \gamma\,(x_D - x_C\,)\,\} =$$

$$= \frac{1}{4}\,\{\gamma(x_A - x_A) + \gamma(x_A - x_B) + \gamma\,(x_A - x_C\,) + \gamma\,(x_A - x_D)\} =$$

$$= \frac{1}{4}\,\{\,\gamma\,(0) + \gamma\,(1) + \gamma(1) + \gamma\,(1\cdot\sqrt{2})\,\} = \frac{1}{2}(1) + \frac{1}{4}\,\gamma\,(1\sqrt{2}) \qquad (207)$$

and finally one has:

$$\sigma^2_{E_5} = 2\,H(1;1) - F(1;1) - \frac{1}{2}\,\gamma\,(1) - \frac{1}{4}\,\gamma\,(1\cdot\sqrt{2}) \qquad (208)$$

6. We take again the square of Fig. 72 but now the four samples are situated at the half length of the sides 1 at the points I, J, K, L.

   For the estimation variance one has:

$$\sigma^2_{E_6} = 2\,\bar{\gamma}\,(\varepsilon;ABCD) - \bar{\gamma}\,(ABCD\,;\,ABCD) - \bar{\gamma}\,(\varepsilon,\varepsilon) \qquad (209)$$

where

$$\bar{\gamma}\,(\varepsilon\,;ABCD) = \bar{\gamma}\,(K;ABCD) = \frac{1}{2}\,\bar{\gamma}\,(K\,;KBDI) +$$

$$+\frac{1}{2}\,\bar{\gamma}\,(K\,;KICA) = \bar{\gamma}\,(K;KBDI) = H(\frac{1}{2};\,1) \qquad (210)$$

and

$$\bar{\gamma}\,(\varepsilon;\varepsilon) = \frac{1}{4}\,\{2\,\gamma\,(1/\sqrt{2}\,)+\gamma\,(1)\} = \frac{1}{2}\,\gamma\,(1/\sqrt{2}\,)+\frac{1}{4}\,\gamma(1) \qquad (211)$$

and finally

$$\sigma^2_{E_6} = 2\,H(1/2\,;\,1) - F(1,1) - \frac{1}{2}\,\gamma\,(1/\sqrt{2}\,) - \frac{1}{4}\,\gamma\,(1) \qquad (212)$$

Here we have again made profit of the symmetry to calculate the $\bar{\gamma}$ $(\varepsilon;ABCD)$ and $\bar{\gamma}\,(\varepsilon;\varepsilon)$ values.

Example

Let us calculate the $\sigma^2_{E_4}$, $\sigma^2_{E_5}$ and $\sigma^2_{E_6}$ values for the spherical variogram with

$a = 5$ m and $C = 14.16\%^2$

for $1 = 2$ m , thus $1/a = 0.4$ $\quad \frac{1}{a}\cdot\sqrt{2} = 0.5657$

From Figs. 59 and 60 one has: $\quad \frac{1}{a}\cdot\frac{1}{\sqrt{2}} = 0.2828$

H (0.2; 0.2) ≈ 0.230

F (0.4; 0.4) = 0.310

H (0.4; 0.4) = 0.450

H (0.2; 0.4) = 0.345

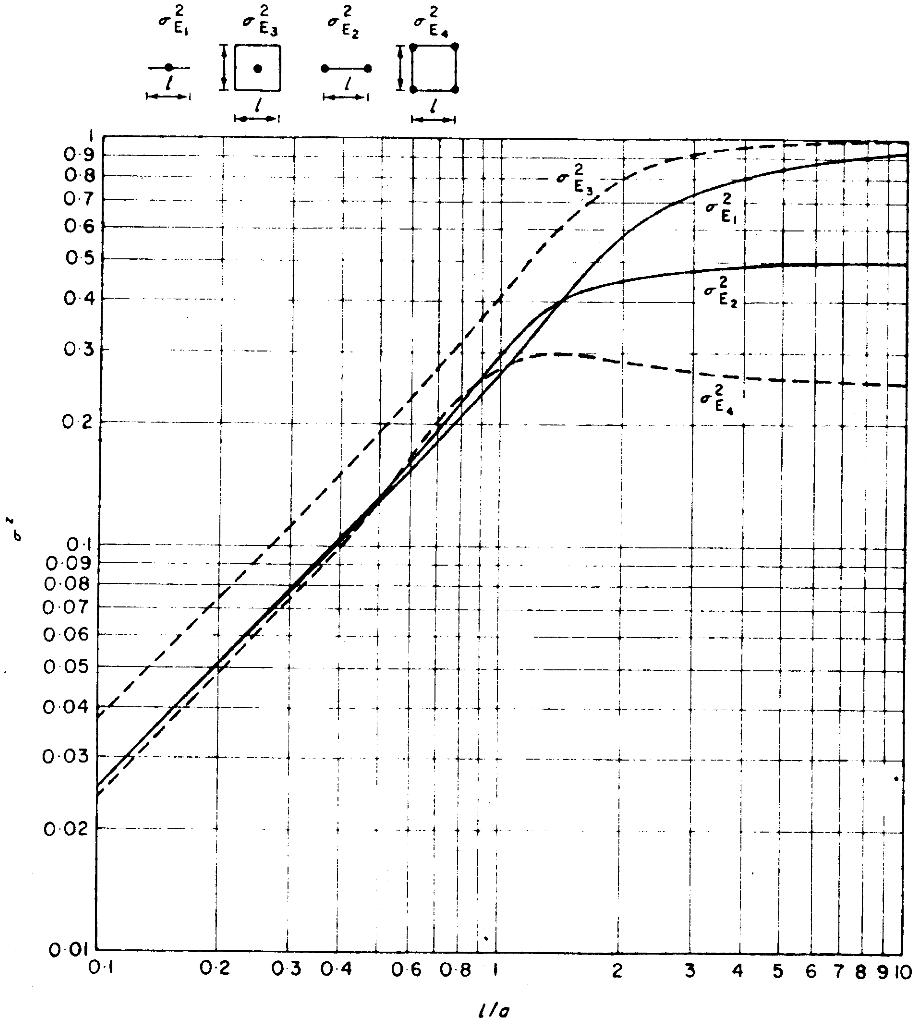


CHART NO. 7. Spherical model. Various extension variances.

Fig. 73

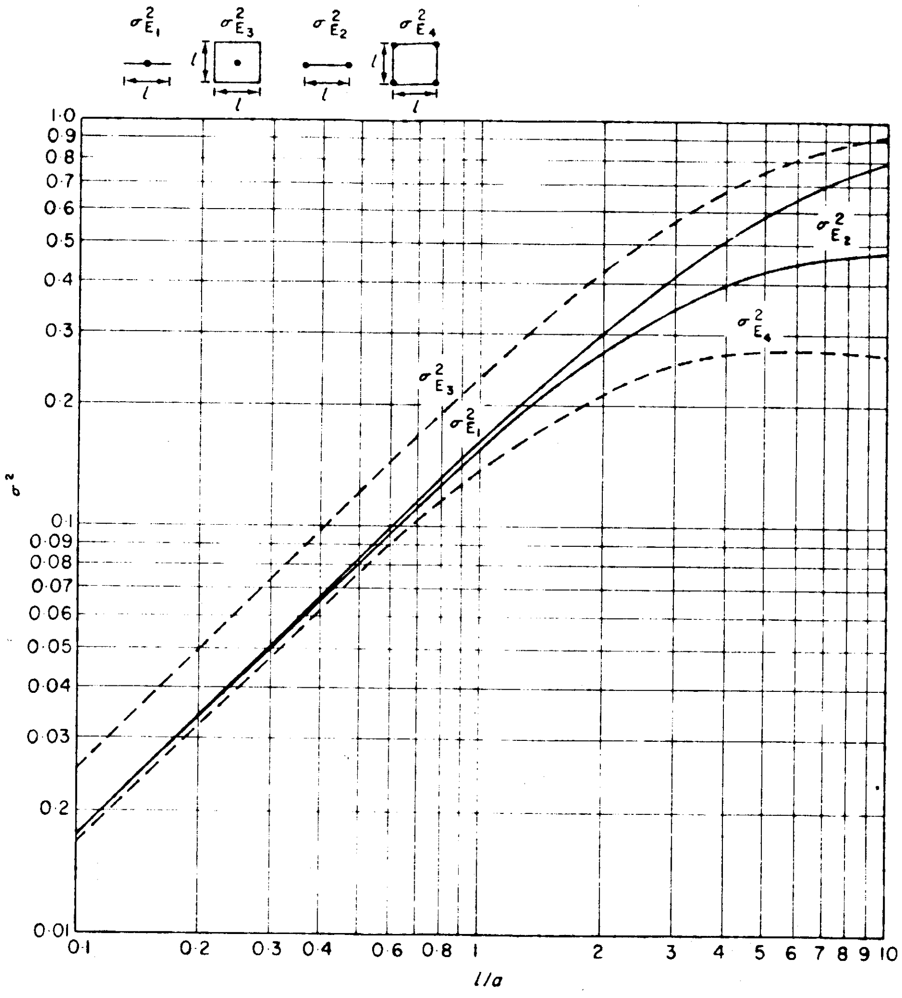


CHART NO. 17. Exponential model. Various extension variances.

Fig. 74

$\gamma(0.4) = 0.568$

$\gamma(0.5657) = 0.7580$

$\gamma(0.2828) = 0.41295$

which gives:

$\sigma^2_{E_4} = (2 \cdot 0.230 - 0.310) \cdot C = 0.15 x 14.16 = 2.124\%^2$

$\sigma_{E_4} = 1.457\%$

$\sigma^2_{E_5} = [2x0.45 - 0.31 - 0.5x0.568 - 0.25x0.758] x C =$

$= 0.1165 x C = 1.6196\%^2$

$\sigma_{E_5} = 1.284\%$

$\sigma^2_{E_6} = [2x0.345 - 0.31 - 0.5x0.41295 - 0.25x0.568] x C =$

$= 0.031525 x C = 0.44639\%^2$

$\sigma_{E_6} = 0.6681\%$

Thus, in this case the scheme No 6 gives the best estimate among all discussed cases.

For some one- and two- dimensional cases the extension variances σ^2_E for the spherical and for the exponential models are also given in Figs 73 and 74 after Journel and Huijbregts [1]. The corresponding schemes for the sample configurations are given above each figure. Note please, that the numbers assigned there for different extensions variances do not always correspond to the numbers in the lecture notes.

LECTURE 9

J.A. Czubek

LECTURE 9 Reykjavík, March 17, 1981

ESTIMATION VARIANCE IN TWO AND THREE DIMENSIONS

All precedent examples dealt with the case of the point-like samples. One can have, however, some more complicated and at the same time more realistic conditions. We are going to discuss them now.

7. We want to estimate the rectangle ABCD of the sides l and L by the value of the z variable known for the median IJ of the rectangle, as it is shown in Fig. 75.

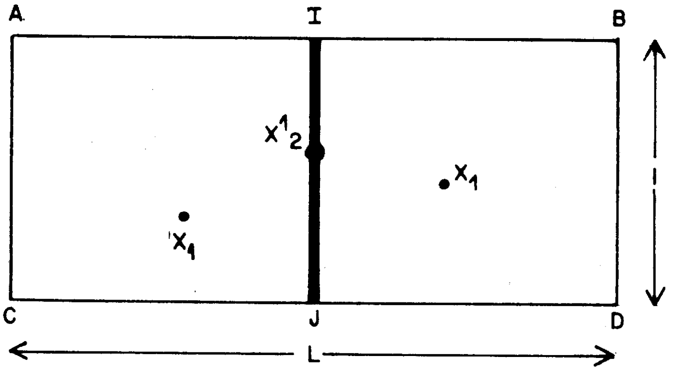


Fig. 75

Now, according to Eq. (78) the estimation variance $\sigma^2_{E_7}$ will be given as:

$$\sigma^2_{E_7} = 2\,\bar{\gamma}\,(IJ;ABCD) - \bar{\gamma}(ABCD;ABCD) - \bar{\gamma}(IJ;IJ) \qquad (213)$$

We have:

$$\bar{\gamma}(IJ;ABCD) = \frac{1}{Lxlxl} \int\limits_{ABCD} dx_1 \int\limits_{1} dx^1_2 \; \gamma(x_1 - x^1_2) =$$

$$= \frac{1}{Lxlxl} \left\{ \int\limits_{1} dx^1_2 \int\limits_{IBDJ} dx_1 \; \gamma\,(x_1 - x^1_2) + \int\limits_{1} dx^1_2 \int\limits_{IJCA} dx_1 \; \gamma(x_1 - x^1_2) \right\} =$$

$$= \frac{1}{2\cdot 1} \left\{ \int\limits_{1} dx^1_2 \cdot \frac{1}{\frac{L}{2}x1} \int\limits_{IBDJ} dx_1 \; \gamma(x_1 - x^1_2) + \int\limits_{1} dx^1_2 \frac{1}{\frac{L}{2}x1} \int\limits_{IJCA} dx_1 \; \gamma(x_1 - x^1_2) \right\} \qquad (214)$$

Both integrals taken over the rectangles IBDJ and IJCA are equal to each other, and the double integral which is left is equal to the average value:

$\bar{\gamma}$ (IJ ; IBDJ) (cf. Fig. 19 and Eq. 151), thus one finally has:

$$\bar{\gamma}\,(IJ\,;\,ABCD) = \bar{\gamma}\,(IJ;IBDJ) = \chi\,(\tfrac{L}{2}\,;\,1) \qquad (215)$$

For $\bar{\gamma}(IJ;IJ)$ we have

$$\bar{\gamma}(IJ;IJ) = F(1) \qquad (216)$$

and the result is:

$$\sigma^2_{E_7} = 2\chi(\tfrac{L}{2}\,;1) - F(L;1) - F(1) \qquad (217)$$

Just to give an idea, how the formulae for $\sigma^2_{E_7}$ look like, using Eqs. (161), (162.b) and (162.c) the variance has been calculated for the linear variogram:

$$\frac{1}{L}\,\sigma^2_{E_7} = \frac{1}{24t^2} + \frac{1}{4}\sqrt{1+4t^2}\,(1-\frac{1}{6t^2}) - \frac{1}{5}\sqrt{1+t^2}\,(1-\frac{1}{3t^2}-\frac{t^2}{3}) -$$

$$- \frac{1}{15}\,(\frac{1}{t^2}+t^3) - \frac{t}{3} + \frac{1}{6t}\,\ln\,\frac{2t+\sqrt{1+4t^2}}{t+\sqrt{1+t^2}} +$$

$$+ \frac{1}{6}\,t^2\,\ln\,\frac{(1+\sqrt{1+4t^2}\,)^2}{4t(1+\sqrt{1+t^2}\,)} \qquad (218)$$

with $t=\frac{1}{L}$, as usually. A similar formula can be found for the logarithmic variogram setting into Eq. 217 the formulae (164), (165.b) and (165.c).

The variance $\sigma^2_{E_7}$ for the spherical and exponential variograms is given in the form of graphs in Figs. 76 and 77 according to Journel and Huijbregts [1]

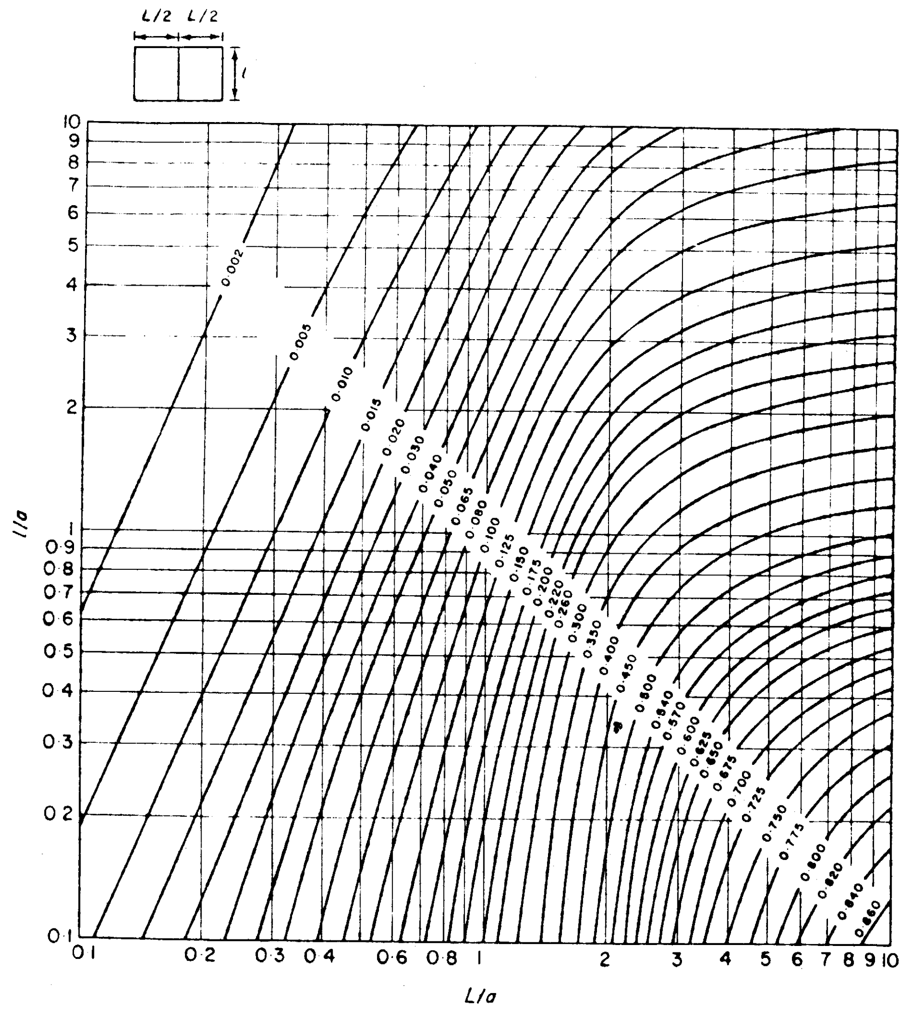


CHART NO. 8. Spherical model. Extension variance $\sigma^2_{E_5}$.

Fig. 76

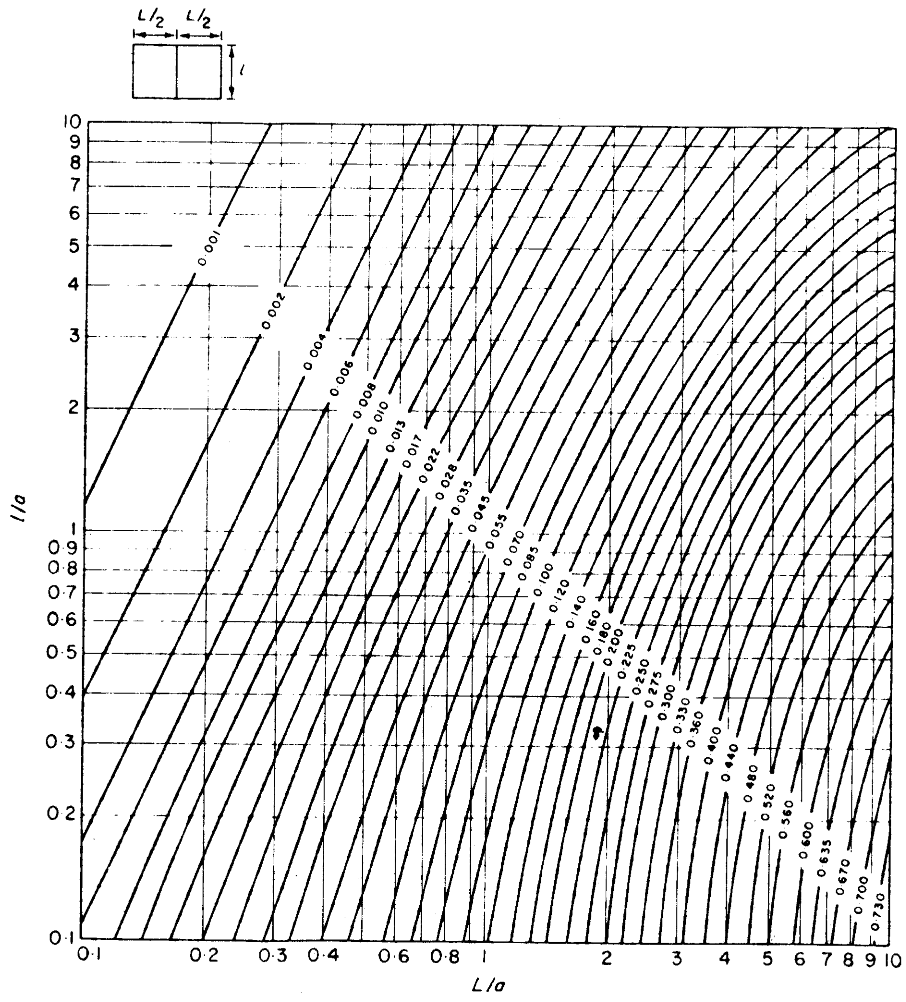


CHART NO. 18. Exponential model. Extension variance $\sigma^2_{E_5}$.

Fig. 77

8. Let us take now a three-dimensional case. The rectangular parallelepiped P with a square base (Lxl^2) is recognized (sampled) by its median square as it is shown in Fig. 78.

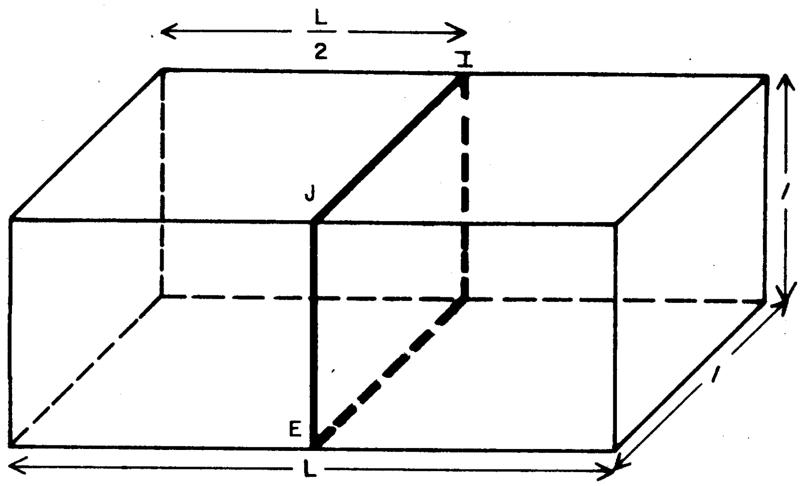


Fig. 78

The estimation variance $\sigma^2_{E_8}$ is:

$$\sigma^2_{E_8} = 2\,\bar{\gamma}(IJFE,\ P) - \bar{\gamma}(IJFE;IJFE) =$$

$$= 2\chi(L/2\ ;\ l^2) - F(L;l^2) - F(l;l) \qquad (219)$$

The graphs of this estimation variance for the spherical and exponential variograms are plotted in Figs 79 and 80 [1]. These graphs can also be used for the computation of the $\chi(L,l^2)$ values using the graphs of the $F(L;l^2)$ and $F(L;l)$ functions given in Figs. 60, 61, 66 and 67 together with Eq. 219.

9. The same rectangular parallelepiped P can be recognized by a single line sample (borhole) situated at its long axis L as it is shown in Fig. 81 . Now the estimation variance $\sigma^2_{E_9}$ becomes:

$$\sigma^2_{E_9} = 2\,\bar{\gamma}(IJ,P) - \bar{\gamma}(P,P) - \bar{\gamma}\ (IJ,\ IJ) =$$

$$= 2\ H[L;(l/2)^2] - F(L;l^2) - F(L) \qquad (220)$$

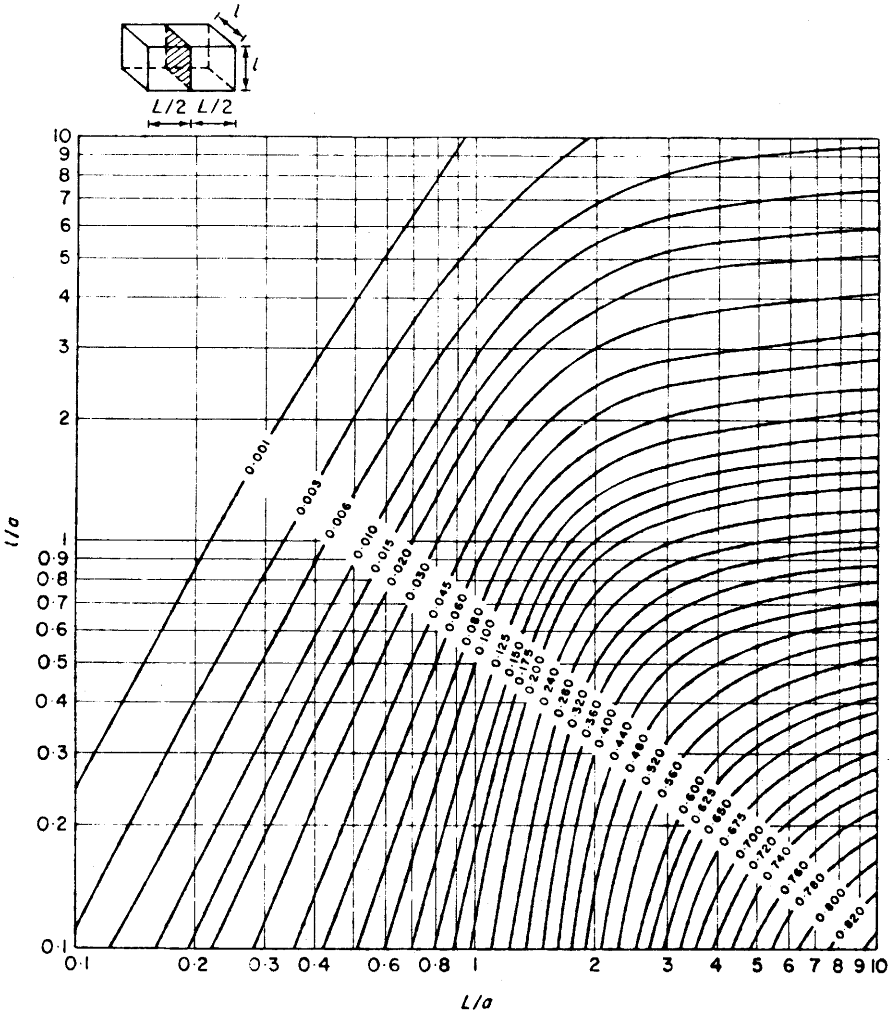


CHART NO. 9. Spherical model. Extension variance $\sigma^2_{E_6}$.

Fig. 79

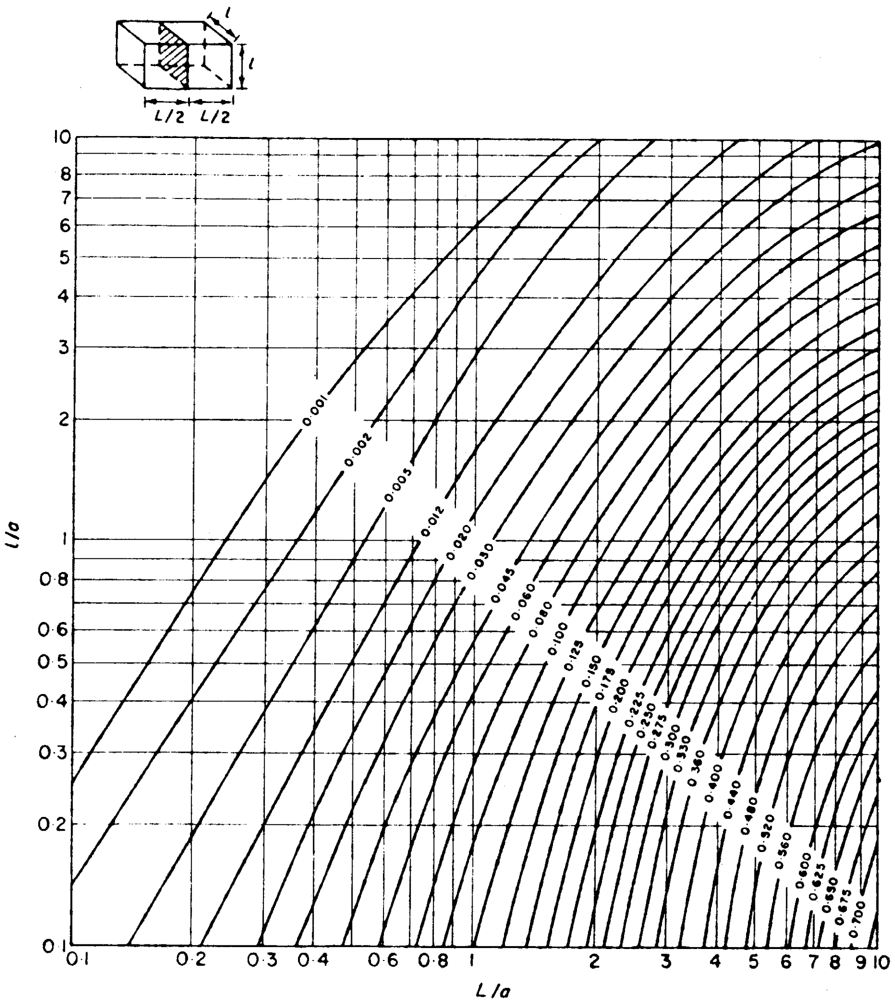


CHART NO. 19. Exponential model. Extension variance $\sigma^2_{E_6}$.

Fig. 80

The graphs of this estimation variance $\sigma^2_{E_9}$ are given by Journel and Huijbregts [1] in Figs. 82 and 83 for the spherical and exponential variograms.

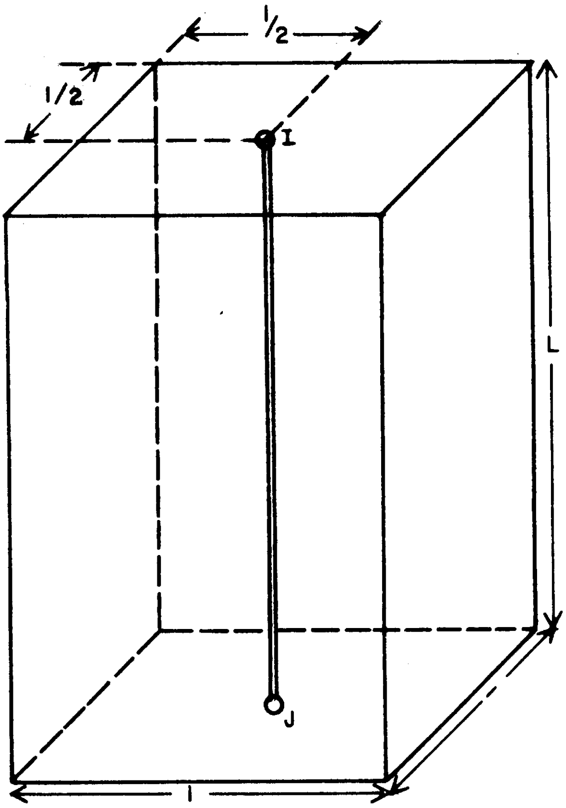


Fig. 81

And again, like for the graphs of the $\sigma^2_{E_8}$ variance, these graphs in Figs 82 and 83 can be used to calculate the values of the three dimensional auxiliary function $H(L;l^2)$.

Example

Let us take again a spherical variogram with $a = 5$ m and $C = 14.16\%^2$ for porosity. Now we want to know what is the estimation (or extension) variance $\sigma^2_{E_9}$ when the rectangular parallelepiped having dimensions 100x67x67 cm^3 we want to estimate using the porosity data obtained from the core taken at the parallelepiped axis L.

We have thus

$$\frac{L}{a} = \frac{1}{5} = 0.2 \qquad \frac{l}{a} = \frac{0.67}{5} = 0.134$$

and from Fig. 82

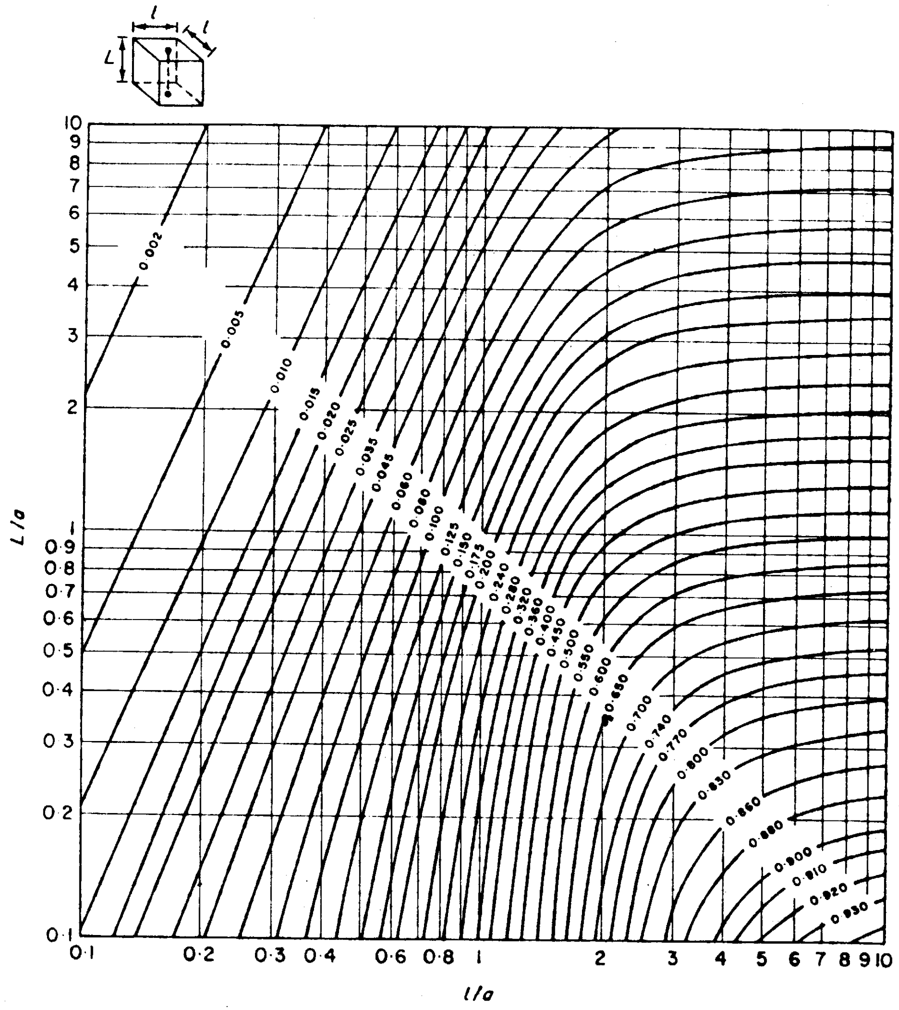


CHART NO. 10. Spherical model. Extension variance $\sigma^2_{E_7}$.

Fig. 82

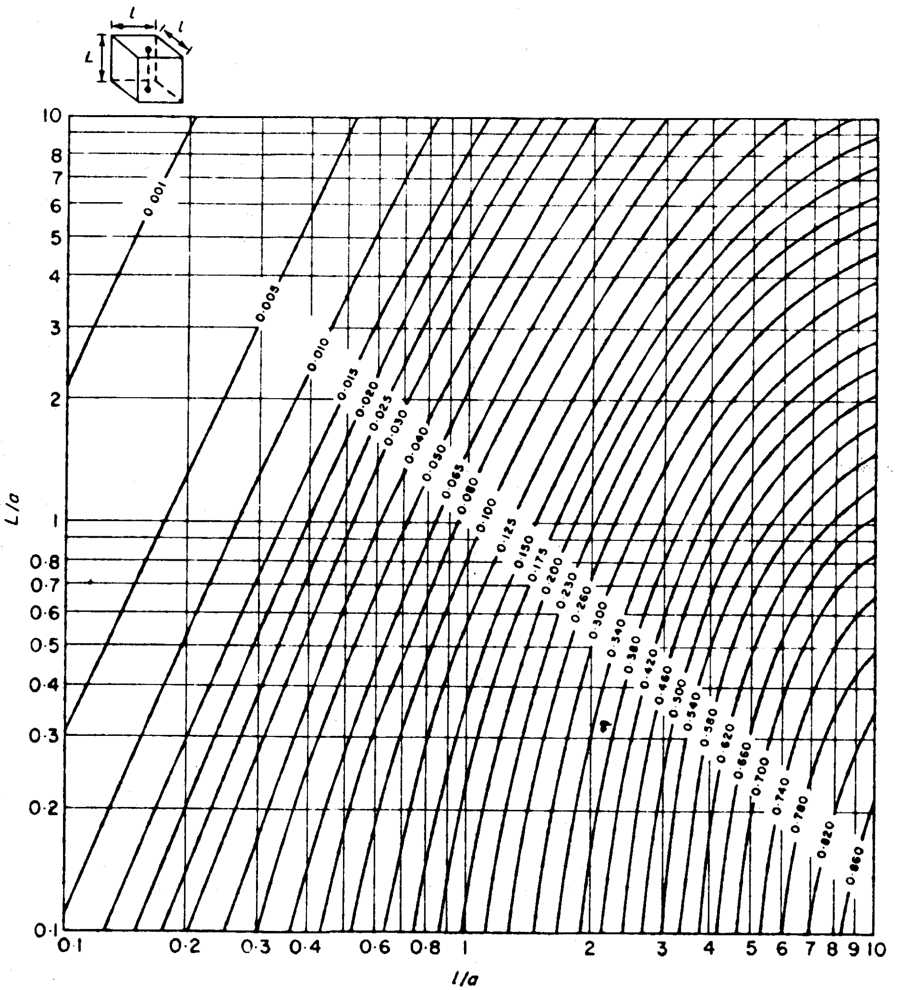


CHART NO. 20. Exponential model. Extension variance $\sigma^2_{E_7}$.

Fig. 83

we have

$$\sigma^2_{E_9} = 0.015 \times C = 0.2124\%^2$$

$$\sigma_{E_9} = 0.461\% \ .$$

Let us understand better what this extension variance really means. Remember that in the example on the dispersion variogram (p. 106) we have calculated the dispersion variance for the same block and for the same spherical variogram and obtaining there

$$\sigma \ (L,l^2) = 3.416\%$$

What is the physical meaning of these two standard deviations ?

$\sigma(L,l^2) = 3.416\%$ it is a standard deviation, that is a mean deviation from the expected (average) value of the random variable Z when the block Lxl^2 is "walking" throughout the whole space occupied by a given geological formation, when the average value of z measured on the whole block Lxl^2 is taken into account.

The σ_{E_9} standard deviation is an average deviation between the true value of Z regularized over the whole block Lxl^2 (i.e. $Z_{Lxl}{}^2$) when instead of the measured value $z_{LXl}2$ over this block we assume that the $z_{Lxl}2$ value is equal to the average z value measured along the central hole only. Thus, it is a standard deviation of the extension procedure, when the knowledge of the axis we extend on the whole block. Let us also remark that this procedure is reciprocal - when one wants to estimate the value (average) along the axial borehole Z_L knowing only the average value of z over the whole block, i.e. $z_{Lxl}2$, the standard deviation of this contraction in this case, will also be given as $\sigma^2{}_{E_9}$. This mutual reciprocity is well visible if one looks at Eq. 78 - it is absolutely symmetrical in V and v values.

ESTIMATION OF RECTANGLE

To get a better idea of how the auxiliary functions can be used in the estimation problems let us work a little on the estimation of rectangle ABCD as it is shown in Fig. 84.

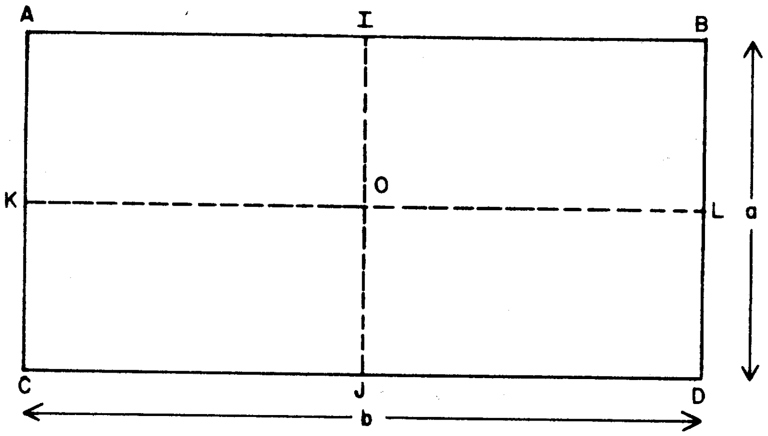


Fig. 84

We have a rectangular panel P with sides a and b which is located in an isotropic formation represented by the point model γ(r), let its thickness be c (not visible in Fig. 84). The panel P = ABCD can be considered as an analogy to the parallelepiped axbxc but graded over the thickness c.

Now we are going to evaluate the unknown value Z of the panel P by each of the following estimators:

1. mean grade z_1 of the perimeter ABCD of panel P (z_1 can be obtained by a continuous horizontal channel sampling along the four sides of the panel P).
2. mean grade of the two parallel sides AB and CD.
3. mean grade of the two medians IJ and KL
4. mean grade of the single median IJ.

For these four different estimators we can write:

1) $$z_1 = \frac{b.z_{AB} + a \cdot z_{BD} + b \cdot z_{CD} + a \cdot z_{AC}}{2a + 2b} \approx z \qquad (221.a)$$

2)

$$z_2 = \frac{1}{2}\,(z_{AB} + z_{CD}) \approx z \qquad (221.b)$$

3)

$$z_3 = \frac{a\cdot z_{IJ} + bz_{KL}}{a+b} \approx z \qquad (221.c)$$

4)

$$z_4 = z_{IJ} \approx z \qquad (221.d)$$

The corresponding estimation variances will be:

1. Rectangle evaluated by its perimeter (Eq. 221.a)

$$\sigma^2_{E_1} = 2\bar{\gamma}(ABCD,\ P) - \bar{\gamma}(P, P) - \bar{\gamma}(ABCD,\ ABCD) \qquad (222.a)$$

here in this formula ABCD means just perimeter i.e. 2a+2b.

According to the experience obtained for example in derivation of Eq. 215 one can write:

$$\bar{\gamma}(ABCD, P) = \bar{\gamma}(AC+AB, P) =$$

$$= \frac{a\cdot\bar{\gamma}(AC,P)+b\cdot\bar{\gamma}(AB,P)}{a+b} = \frac{a\cdot\chi(b;a)+b\chi(a;b)}{a+b} \qquad (223.a)$$

$$\bar{\gamma}(ABCD,ABCD) = \bar{\gamma}\ (AC+AB;ABCD) =$$

$$= \frac{a\cdot\bar{\gamma}(AC;ABCD) + b\,\bar{\gamma}\,(AB;ABCD)}{a+b} \qquad (223.b)$$

but

$$\bar{\gamma}(AC;ABCD) = \frac{a\cdot\bar{\gamma}(AC,AC) + 2b\bar{\gamma}(AC,AB) + a\bar{\gamma}\,(AC,AB)}{2(a+b)}$$

$$= \frac{a\cdot F(a) + 2b\ H(a;b)+a\ \alpha\ (b;a)}{2(a+b)} \qquad (223.c)$$

and

$$\bar{\gamma}(AB;ABCD) = \frac{bF(b)+2a\ H(b;a) + b\cdot\alpha(a;b)}{2(a+b)} \qquad (223.d)$$

Inserting now (223) c and d into (223.b) and (223.b) and (223.a) into (222a) one gives:

$$\sigma^2_{E_1} = 2\ \frac{a\cdot\chi(b;a)+b\chi(a;b)}{a+b} - F(a;b) -$$

$$- \frac{a^2F(a)+4ab\ H(a;b)+b^2F(b)+a^2\alpha(b;a)+b^2\alpha(a;b)}{2(a+b)^2} \qquad (222.b)$$

This formula for a square (a=b) becomes:

$$\sigma^2_{E_1} = 2\chi(a,a) - F(a,a) - \frac{1}{4}\ F(a) - \frac{1}{2}\ H(a,a) - \frac{1}{4}\ \alpha(a,a) \qquad (222.c)$$

If one takes into account the linear model of the isotropic variogram, one has (cf. Eqs. 161 and 163):

$$\left.\begin{aligned} F(a) &= \frac{a}{3}\cdot A \\ \alpha(a,a) &= 1.0766\cdot a\cdot A \\ \chi(a,a) &= 0.65176\ a\cdot A \\ F(a,a) &= 0.5214\ a\cdot A \\ H(a,a) &= 0.7652\ a\cdot A \end{aligned}\right\} (224)$$

When one inserts the above values into Eq. (222.c) one obtains:

$$\sigma^2_{E_1} = 0.0479\cdot a\cdot A \qquad (225.a)$$

for the total sampled length 4a.

For the logarithmic model $\gamma(h) = 3\alpha\ \ln|h|$ one has from Eqs. 164 and 165

$$\left.\begin{aligned} \frac{1}{3\alpha}\ \alpha(a,a) &= \ln a - 0.070796 \\ \frac{1}{3\alpha}\ \chi(a,a) &= \ln a - 0.555087 \\ \frac{1}{3\alpha}\ F(a,a) &= \ln a - 0.805087 \\ \frac{1}{3\alpha}\ H(a,a) &= \ln a - 0.073504 \\ \frac{1}{3\alpha}\ F(a) &= \ln a - 1.5 \end{aligned}\right\} (226)$$

Here the $H(a;a)$ value was obtained from the formula for $H(L,l)$ derived from Eqs.(157) and (165.b) which give:

$$H(L;l)=\chi(L;l)+\frac{1}{2}\left[\frac{2}{3}-\frac{1}{t}\text{ arc tgt}+\frac{1}{3t^2}\ln(1+t^2)+\frac{t}{3}(\frac{\pi}{2}-\text{arc tgt})\right] \quad (227)$$

Inserting now Eq. (226) into Eq. (222.c) yields

$$\sigma^2_{E_1} = 3\alpha \times 0.1244 \quad (225.b)$$

2. <u>Rectangle evaluated by its two parallel sides AB and CD</u>

$$\sigma^2_{E_2} = 2\bar{\gamma}\ (AB+CD,P) - \bar{\gamma}\ (P,P) - \bar{\gamma}\ (AB+CD\,;\,AB+CD) \quad (228)$$

and because:

$$\bar{\gamma}\ (AB+CD,P) = \bar{\gamma}\,(AB,P) = \chi\,(a;b) \quad (229.a)$$

$$\bar{\gamma}\ (AB+CD,AB+CD) = \bar{\gamma}\ (AB,\,AB+CD) =$$

$$= \frac{1}{2}\,[\bar{\gamma}(AB,AB) + \bar{\gamma}\ (AB,CD)] =$$

$$= \frac{1}{2}\,[F(b) + \alpha\,(a;b)] \quad (229.b)$$

Thus we have

$$\sigma^2_{E_2} = 2\chi(a;b) - F(a,b) - \frac{1}{2}\,[F(b) + \alpha\,(a;b)] \quad (230.a)$$

which for $a=b$ gives

$$\sigma^2_{E_2} = 2\chi(a;a) - F(a,a) - \frac{1}{2}\,[F(a) + \alpha\,(a;a)] \quad *230.b($$

For the linear variogram this formula gives:

$$\sigma^2_{E_2} = 0.07715\ a\cdot A \quad (231.a)$$

for a total sampled length $2a$, and for the logarithmic variogram one obtains:

$$\sigma^2_{E_2} = 3\alpha \times 0.4803 \quad (231.b)$$

3. Rectangle evaluated by its two medians IJ and KL

$$\sigma^2_{E_3} = 2\bar{\gamma}\ (IJ+KL;P) - \bar{\gamma}\ (P;P) - \bar{\gamma}(IJ+KL;IJ+KL) \qquad (232)$$

but

$$\bar{\gamma}\ (IJ+KL;P) = \frac{a\cdot\bar{\gamma}(IJ;IBDJ) + b\bar{\gamma}(KL;KABL)}{a+b} \qquad (233)$$

and

$$\left.\begin{aligned} &\bar{\gamma}\ (IJ;IBDJ) = \chi(b/2;a) \\ &\bar{\gamma}\ (KL;KABL) = \chi(a/2;b) \end{aligned}\right\} \qquad (234)$$

Next

$$\bar{\gamma}\ (IJ+KL;IJ+KL) =$$

$$= \frac{a\cdot\bar{\gamma}(IJ;IJ+KL) + b\cdot\bar{\gamma}(KL;IJ+KL)}{a+b} \qquad (235)$$

but

$$\bar{\gamma}\ (IJ;IJ+KL) = \frac{a\cdot\bar{\gamma}(IJ;IJ) + b\bar{\gamma}(IJ;KL)}{a+b} \qquad (236.a)$$

and

$$\bar{\gamma}\ (KL;IJ+KL) = \frac{a\cdot\bar{\gamma}(KL;IJ) + b\cdot\bar{\gamma}(KL;KL)}{a+b} \qquad (236.b)$$

and again:

$$\left.\begin{aligned} &\bar{\gamma}\ (IJ;IJ) = F(a) \\ &\bar{\gamma}\ (KL;KL) = F(b) \\ &\bar{\gamma}\ (IJ;KL) = \bar{\gamma}\ (IO;OL) = H\ (\tfrac{b}{2};\tfrac{a}{2}) = H\ (\tfrac{a}{2};\tfrac{b}{2}) \\ &\bar{\gamma}\ (KL;IJ) = H\ (\tfrac{a}{2};\tfrac{b}{2}) \end{aligned}\right\} \qquad (237)$$

which finally gives, with $\gamma(P,P) = F(a;b)$

the value for $\sigma^2_{E_3}$:

$$\sigma^2_{E_3} = \frac{2a\chi(\frac{b}{2}\,;a) + 2b\cdot\chi(\frac{a}{2}\,;b)}{a+b} - F(a;b) -$$

$$- \frac{a^2\,F(a) + 2\cdot a\cdot b\cdot H(a/2\,;\,b/2) + b^2\,F(b)}{(a+b)^2} \qquad (238.a)$$

This equation for $a=b$ becomes:

$$\sigma^2_{E_3} = 2\chi(a/2;a) - F(a;a) - \frac{1}{2}\,[F(a) + H(a/2\,;\,a/2)] \qquad (238.6)$$

Now, for the linear $\gamma(h)$ model we have from Eq. (162.b)

$$\chi\,(\frac{a}{2};\,a) = 0.45426\cdot a \qquad (239.a)$$

which together with the formulae (224) gives for the estimation variance

$$\sigma^2_{E_3} = 0.02915\ a\cdot A \qquad (238.c)$$

whereas for the logarithmic model of variogram one has from Eq. (165.b)

$$\chi\,(\frac{a}{2}\,;\,a) = \ln a - 0.926149 \qquad (239.b)$$

and together with the formulae (226) it gives for the estimation variance

$$\sigma^2_{E_3} = 0.086115 \times 3\alpha \qquad (238.d)$$

Note please, that here the total sampled length is 2a

4. <u>Rectangle evaluated by a single median IJ</u>

Here is:

$$\sigma^2_{E_4} = 2\bar\gamma\,(IJ;P) - \bar\gamma\,(IJ;IJ) - \bar\gamma(P;P) \qquad (240.a)$$

$$\left.\begin{array}{l}\text{with}\\ \bar\gamma\,(IJ;P) = \bar\gamma\,(IJ;IBDJ) = \chi(\frac{b}{2}\,;\,a)\\ \text{and}\\ \bar\gamma\,(IJ;IJ) = F(a)\\ \text{which yields:}\end{array}\right\} \quad (241)$$

$$\sigma^2_{E_4} = 2\cdot\chi\,(b/2\,;\,a) - F(a;b) - F(a) \qquad (240.b)$$

and for a=b

$$\sigma^2_{E_4} = 2\chi(a/2\,;\,a) - F(a;a) - F(a) \qquad (240.c)$$

for the total sampled length equal to 1·a .

Knowing the values $\chi\ (a/2\,;\,a)$, $F(a;a)$ and $F(a)$

for the linear and logarithmic variograms
(Eqs. 239.a and b , 224, 226) one has finally:

for the linear model:

$$\sigma^2_{E_4} = 0.053787 \cdot a \cdot A \qquad (240.d)$$

and for the logarithmic model:

$$\sigma^2_{E_4} = 0.45279 \cdot 3\alpha \qquad (240.e)$$

These four examples discussed above can be presented in the following table:

	$\sigma^2_{E_3}$	$< \sigma^2_{E_3}$	$< \sigma^2_{E_4}$	$< \sigma^2_{E_2}$
linear model: a·Ax	0.02916	< 0.0479	< 0.05379	< 0.07715
logarithmic model: 3αx	0.08611	< 0.1244	< 0.4528	< 0.4803
total sampled length	2a	4a	a	2a

As we can see from this table the best way to evaluate the rectangle (or the square, exactly in this particular case) is to use the two medians. The sampling along the perimeter does not give any better estimate while the total sampled length is the longest one. Geostatistics gives us a convenient tool to balance the quality (the median information is better located) and the quantity of information (the perimeter has the largest total sampled length but does not correspond to the smallest estimation variance).

LECTURE 10

PRINCIPLES OF KRIGING

We have seen in the last lecture that to evaluate of a rectangle different estimators can be used, each of them characterized by different estimation variance. We have taken into account only a very limited amount of estimators among all possible ones. Maybe there exists another estimator giving better results? Or maybe, there is another way of treating the data (i.e. sampling results) to get better results, to be closer to the true (an unfortunately unknown) value for the estimated body?

The idea of presenting the problem in that way and the first experimental attempts to solve it is due to Danie G. Krige from the Anglo-Transvaal Consolidated Investment Co. Ltd. in South Africa, but the mathematical solution and especially justification of the problem belong to George Matheron from the Center of Geostatistics in Fontainebleau, France, who gave the name Kriging to this procedure, just to honour the first inventor D.G. Krige.

The problem

Inside some geological formation at the points x_i $(i = 1,2, \ldots .n)$ the values z_i of the random variable $Z(x)$ are known. We can consider also that the z_i values are not necessarily taken on the point-like samples, but each sample has some volume $v(x_i)$ centered over the point x_i . Thus the z_i data can be considered as a regularized over the volumes $v_i \equiv v(x_i)$ values of a given realization $z(x)$ of the random variable $Z(x)$. At the point x_o is centered another volume $V(x_o)$ and the problem is to find the regularized value $Z_V(x_o)$ of the random variable $Z(x)$, thus to find

$$Z_V(x_o) = \frac{1}{V} \int_{V(x_o)} Z(x)\,dx \tag{242}$$

The random variable $Z(x)$ is defined on a point support and is weakly (second-order) stationary with the expectation

$$E\{Z(x)\} = m \tag{243}$$

where the constant m is generally unknown and has a covariance C(h):

$$C(h) = \{Z(x+h) \cdot Z(x)\} - m^2 \tag{244}$$

or variogram

$$2\gamma(h) = E\{[Z(x+h) - Z(x)]^2\} \tag{245}$$

Because of the second order stationarity we are limited in the classes of estimators to the linear estimators only. Thus our task is to find a linear estimator Z^* of the unknown value Z_V using a set of the known data $\{z_i, i = 1 \text{ to } n\}$.

The conditions for such estimator are: the non-bias and the minimum estimation variance, which we consider as the optimum estimation conditions. An estimator fulfilling these optimum conditions will be denoted as Z_K^* (for Kriging). Thus, we have the following relations:

Linearity conditions gives the general form of the Z_K^* :

$$Z_K^*[V(x_o)] \equiv Z_K^* = \sum_{i=1}^{n} \lambda_i \cdot z_i = \sum_{i=1}^{n} \lambda_i \cdot z(x_i) \tag{246}$$

The non-bias conditions gives:

$$E\{Z_K^*\} = E\{\sum_{i=1}^{n} \lambda_i z_i\} = \sum_{i=1}^{n} \lambda_i E(z_i) =$$

$$= m \cdot \sum_{i=1}^{n} \lambda_i = m = E\{Z_V\} \tag{247}$$

which entails that

$$E\{Z_V - Z_K^*\} = 0 \tag{248.a}$$

and

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{248.b}$$

Thus, the sum of the weighting coefficients should be normalized to the value 1.0.

The minimum estimation variance requires the value

$$E\{[Z_V - Z_K^*]^2\}$$

to reach the minimum. We can develop the above expression:

$$E\{[Z_V - Z_K^*]^2\} = E\{Z_V^2\} - 2\cdot E\{Z_V\cdot Z_K^*\} + E\{Z_K^{*2}\} \qquad (249)$$

with (cf. Eqs. 29 to 31 and 77):

$$E\{Z_V^2\} = \frac{1}{V^2} \int_V dx \int_V E\{Z(x)\cdot Z(x^1)\}\, dx^1 =$$

$$= \bar{C}(V;V) + m^2 = C(0) + m^2 - \bar{\gamma}(V;V) \qquad (250)$$

$$E\{Z_V\cdot Z_K^*\} = \sum_{i=1}^{n} \lambda_i \frac{1}{V\cdot v_i} \int_V dx \int_{v_i} E\{Z(x)\cdot Z(x^1)\, dx^1 =$$

$$= \sum_{i=1}^{n} \lambda_i \bar{C}(V;v_i) + m^2 = C(0) + m^2 - \cdot \sum_{i=1}^{n} \lambda_i \bar{\gamma}(V;v_i) \qquad (251)$$

$$E\{Z_K^{*2}\} = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \frac{1}{v_i v_j} \int_{v_i} dx \int_{v_j} E\{Z(x)\cdot Z(x^1)\cdot dx^1 =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \bar{C}(v_i;v_j) + m^2 =$$

$$= C(0) + m^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \bar{\gamma}(v_i;v_j) \qquad (252)$$

Inserting Eqs. (250) to (252) into Eq. (249) yields:

$$\sigma_E^2 = E\{[Z_V - Z_K^*]^2\} = \bar{C}(V;V) - 2 \sum_{i=1}^{n} \lambda_i \bar{C}(V;v_i) +$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \bar{C}(v_i;v_j) \qquad (253.a)$$

or

$$E\{[Z_V - Z_K^*]^2\} = 2 \sum_{i=1}^{n} \lambda_i \bar{\gamma}(V;v_i) - \bar{\gamma}(V,V) -$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i\cdot\lambda_j \bar{\gamma}(v_i;v_j) \qquad (253.b)$$

Derivation of Kriging equations

Comparing Eqs. (253) with Eq. (78) we can see that Eqs. (253) are, simply speaking, the estimation variances subject of the constraint C (only one) given by Eq. (248.b). The geometrical situation of the volume V and v_i is exactly the same as in Fig. 25. And the problem of kriging is to minimize the function $E\{[Z_V - Z_K^*]^2\}$ subject of the constraint C. This is a typical problem which is solved using the method of Lagrange multipliers (cf. for example KORN AND KORN 1961[8]). In this method, when the extremum of the multivariable function $F(x_1, x_2 \ldots x_n)$ whould be found, subject to the k constraints C_i (i.e. 1, 2 ... k) with $k < n$, where $C_i(x_1, x_2 \ldots x_n) = 0$, one takes a new functional

$$\phi\ (x_1, x_2 \ldots x_n\ , \mu_1, \mu_2 \ldots \mu_k) =$$

$$= F(x_1, x_2 \ldots x_n) + \sum_{i=1}^{k} \mu_i \cdot C_i\ (x_1, x_2 \ldots x_n) \qquad (254)$$

and solves simultaneously (n+k) equations:

$$\frac{\partial \phi}{\partial x_j} = 0 \qquad j = 1,\ 2 \ldots n \qquad - n \text{ equations}$$

$$C_i = 0 \qquad i = 1,\ 2 \ldots k \qquad - k \text{ equations} \qquad (255)$$

which have the solution:

$$x_1^o\ ,\ x_2^o \ldots x_n^o\ ,\ \mu_1\ ,\ \mu_2 \ldots \mu_k$$

at which the functional (254) reaches its extremum when the conditions $C_i = 0$ are fulfilled.

In our case we have only one constraint

$$\sum_{i=1}^{n} \lambda_i - 1 = 0 \qquad (256)$$

thus $k = 1$, the set of unknowns $\{x_i\}$ is just $\{\lambda_i\}$ - - thus n unknowns and the (n+1)-st unknown μ_1 which is convenient to take

$$\mu_1 \equiv -2\mu \qquad (257)$$

---

[8]. KORN G.A. and KORN T.M. Mathematical Handbook for Scientists and Engineers. Mc-GRAW-HILL Book Co. INC., New York, 1961.

Now, to calculate the $\partial\phi/\partial x_j$ we have

$$\frac{\partial}{\partial\lambda_i}\{E[Z_V - Z_K^*]^2] - 2\mu\ (\sum_{j=1}^{n}\lambda_j - 1)\} = 0 \qquad i=1,\ 2\ \ldots n \tag{258}$$

which, after insertion Eq. (253.a) or Eq. (253.b) yields:

$$-2\bar{C}\ (V;v_i) + 2\sum_{j=1}^{n}\lambda_i\cdot\bar{C}(v_i;v_j) - 2\mu = 0 \qquad i=1,\ 2\ \ldots\ n \tag{259.a}$$

or

$$2\bar{\gamma}\ (V;v_i) - 2\sum_{j=1}^{n}\lambda_i\cdot\bar{\gamma}\ (v_i;\ v_j) - 2\mu = 0 \qquad i=1,\ 2\ \ldots\ n \tag{259.b}$$

with the (n+1) of equation

$$\sum_{j=1}^{n}\lambda_j - 1 = 0 \tag{259.c}$$

Thus we have to solve the linear system:

$$\sum_{j=1}^{n}\lambda_i\cdot\bar{C}(v_i;v_j) - \mu = \bar{C}(v_i;V) \qquad i=1,\ 2\ \ldots\ n$$

$$\sum_{j=1}^{n}\lambda_j = 1 \tag{260.a}$$

or

$$\sum_{i=1}^{n}\lambda_j\ \bar{\gamma}(v_i;v_j) + \mu = \bar{\gamma}\ (V;v_i) \qquad i=1,\ 2\ \ldots\ n \tag{260.b}$$

$$\sum_{j=1}^{n}\lambda_j = 1$$

to find λ_1 , λ_2 ... λ_n , μ unknowns (n+1 unknowns). This solution, when inserted into Eqs. (253.a) or (253.b) yields the value of the minimum estimation variance, which in this case is called the "Kriging variance" σ_K^2 . To calculate the Kriging variance σ_K^2 , let us take the first of Eq. (260.b) in the form

$$\sum_{j=1}^{n}\lambda_j - \bar{\gamma}\ (v_i;v_j) = \bar{\gamma}\ (V;v_i) - \mu \tag{260.c}$$

We multiply both sides of Eq. (260.c) by λ_i :

$$\sum_{j=1}^{n} \lambda_i\lambda_j\bar{\gamma}(v_i;v_j) = \lambda_i\ \bar{\gamma}(V;v_i) - \mu\cdot\lambda_i \tag{261}$$

and having n such equations (for i=1, 2 ... n) we are doing the summation of the left and right sides, which gives:

$$\sum_{j=1}^{n}\ \sum_{j=1}^{n} \lambda_i\lambda_j\bar{\gamma}(v_i;v_j) = \sum_{i=1}^{n} \lambda_i\bar{\gamma}(V;v_i) - \mu \sum_{i=1}^{n} \lambda_i \tag{262.a}$$

but because of the constrain given by Eq. (259.c) we have

$$\sum_{i=1}^{n}\ \sum_{i=1}^{n} \lambda_i\lambda_j\bar{\gamma}(v_i;v_j) = \sum_{i=1}^{n} \lambda_i\bar{\gamma}(V;v_i) - \mu \tag{262.b}$$

which setted into Eq. (253.b) gives the Kriging variance

$$\sigma_K^2 = \sum_{i=1}^{n} \lambda_i\bar{\gamma}(V;v_i) + \mu - \bar{\gamma}(V;V) \tag{263.a}$$

or, when using the covariance notation:

$$\sigma_K^2 = \bar{C}\,(V;V) + \mu - \sum_{i=1}^{n} \lambda_i\ \bar{C}\,(V;v_i) \tag{263.b}$$

The system of the Kriging equations (260.a) or (260.b) is usually used in a matrix form

$$[K]\cdot[\lambda] = [M2] \tag{264}$$

to solve them easier by using a computer. Here the Kriging matrix [K] contains all $\bar{C}\ (v_i;v_j)$ or $\bar{\gamma}(v_i;v_j)$ terms with the terms $\bar{C}(v_i;v_i)$ or $\bar{\gamma}(v_i;v_i)$ on its main diagonal. The last column and the last row of this matrix contain 1 except on the main diagonal where 0 is putted. The column vector [λ] contains the λ_i values (unknowns) and at the last place the Lagrange multiplier μ (unknown). [M2] is the column matrix containing the second member of the Kriging equations, i.e. the $\bar{C}(v_i;V)$ or $\bar{\gamma}(V;v_i)$ values with the value 1 at the bottom. Note that in the Kriging matrix the off main diagonal terms are symmetrical, i.e.

$$\bar{C}(v_i;v_j) = \bar{C}(v_j;v_i) \quad \text{or} \quad \bar{\gamma}(v_i;v_j) = \bar{\gamma}(v_j;v_i)$$

and in the $[\lambda]$ one column matrix the last term is $-\mu$ when the covariances $\bar{C}(v_i;v_j)$ are used and $+\mu$ when the variograms $\bar{\gamma}(v_i;v_j)$ are used.

Thus, in terms of variogram one has:

$$[K] = \begin{bmatrix} \bar{\gamma}(v_1;v_1) & \cdots\cdots & \bar{\gamma}(v_1;v_j) & \cdots\cdots & \bar{\gamma}(v_1;v_n) & 1 \\ \vdots & \ddots & \vdots & & \vdots & \\ \bar{\gamma}(v_j;v_1) & \cdots\cdots & \bar{\gamma}(v_j;v_j) & \cdots\cdots & \bar{\gamma}(v_j;v_n) & 1 \\ \vdots & & \vdots & \ddots & \vdots & \\ \bar{\gamma}(v_n;v_1 & & \bar{\gamma}(v_n;v_j & & \bar{\gamma}(v_n;v_n) & 1 \\ & & & & & \searrow \\ 1 & & 1 & & 1 & 0 \end{bmatrix} \tag{264.a}$$

Main diagonal

$$[\lambda] = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \lambda_j \\ \cdot \\ \cdot \\ \lambda_n \\ \mu \end{bmatrix} \quad [M2] = \begin{bmatrix} \bar{\gamma}(V;v_1) \\ \bar{\gamma}(V;v_2) \\ \cdot \\ \cdot \\ \bar{\gamma}(V;v_j) \\ \cdot \\ \cdot \\ \bar{\gamma}(V;v_n) \\ 1 \end{bmatrix} \tag{264.b}$$

When the matrix $[K]$ is positive defined the solution to Eq. (266) is in the standard form:

$$[\lambda] = [K]^{-1} \cdot [M2] \tag{265}$$

and the Kriging variance is given as:

$$\sigma_K^2 = [\lambda]^t \cdot [M2] - \bar{\gamma}\ (V;V) \tag{266}$$

where $[\lambda]^t$ is a one raw matrix transposed to $[\lambda]$.

For the reason of technical problems with the numerical calculations the set of Kriging equations in the form of Eqs. (260.a) is usually preferred, even when the covariance is not defined. For this purpose it is enough to define the "pseudo-covariance" $C(h)$ such that

$$\gamma(h) = A - C(h) ,$$

the constant A being any positive value greater than the greatest mean value $\bar{\gamma}$ used in the Kriging system (260.b). The non-bias condition

$$\sum_i \lambda_i = 1$$

allows to eliminate this constant from the system (260.b) which in that case becomes of the type (260.a) written in the "pseudo-covariance" $C(h)$.

The co-variance system in the matrix form is:

$$[\lambda] = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_j \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ -\mu \end{bmatrix} \qquad [M2] = \begin{bmatrix} \bar{C}(V;v_1) \\ \bar{C}(V;v_2) \\ \cdot \\ \cdot \\ \cdot \\ \bar{C}(V;v_j) \\ \cdot \\ \cdot \\ \cdot \\ \bar{C}(V;v_n) \\ 1 \end{bmatrix} \qquad (264.c)$$

and the Kriging matrix $[K]$ (which is simply a co-variance matrix) is done as:

$$[K]=\begin{bmatrix} \bar{C}(v_1;v_1) & \ldots\ldots & \bar{C}(v_1;v_j) & \ldots\ldots & \bar{C}(v_1;v_n) & 1 \\ \vdots & \ddots & \vdots & & \vdots & \\ \bar{C}(v_j;v_1) & \ldots\ldots & \bar{C}(v_j;v_j) & \ldots\ldots & \bar{C}(v_j;v_n) & 1 \\ \vdots & & \vdots & \ddots & \vdots & \\ \bar{C}(v_n;v_1) & \ldots\ldots & \bar{C}(v_n;v_j) & \ldots\ldots & \bar{C}(v_j;v_j) & 1 \\ 1 & & 1 & & 1 & 0 \end{bmatrix} \searrow \qquad (264.d)$$

main diagonal

where on the main diagonal are, in this case, the dispersion variances of samples in the field.

In the case when some of the volumes v_i of the samples are zero (and in the limit case all of them can be zero i.e. - point-like samples) and when the volume V is also reduced to zero we are talking about punctual Kriging and in that case the averaged values of the variogram $\bar{\gamma}(v_i;v_j)$ or of the covariance $\bar{C}(v_i;v_j)$ become simply the values of the point variogram $\gamma(x_i-x_i)$ or covariance $C(x_i-x_j)$. This is just the case of the contour map design using the Kriging method, when for a given implantation of the point-like samples $\{v_i\}$ the contour of a constant value of $Z_K^*(x)$ is sought.

LECTURE 11

J.A. Czubek

LECTURE 11 Reykjavík, March 22, 1981

SOME REMARKS ABOUT KRIGING

1. The system of the Kriging equations (260.a) has a unique solution when the matrix [K] given by Eq. (264.d) has a positive determinant. It also means that no data support v_i coincides completely with another one, because when for $i \neq k$ one has $\bar{C}(v_i;v_j) = \bar{C}(v_k;v_j)$ for all j, then the determinant $|\bar{C}(v_i;v_j|$ becomes zero.
This condition of the uniqness of the solution entails that the Kriging variance is non-negative.

2. Kriging is an un-biased estimator and it is also an exact linear interpolator. When the volume V coincides with one of the volumes v_i (samples) the Kriging estimator Z_K^* becomes identical with the value z_i measured at this sample with the zero Kriging variance $\sigma_K^2 = 0$.

3. The Kriging system and the Kriging variance depend upon the structural model C(h) or intrinsic model $\gamma(h)$ and on the relative positions at the samples v_i and the volume V being Kriged.
The solution of this system (i.e. the values λ_i and σ_K^2) do not depend upon the values z_i being measured on the samples. This permits, for a given configuration of samples and that of the volume V to solve the Kriging system in advance, before the field work is done (knowing, however, the intrinsic parameters $\gamma(h)$).

4. The Kriging matrix [K] depends on the sample configurations and not on the volume V. Thus, for different positions of the volume V, having once the matrix [K] inversed, the solution $[\lambda]$ is found by changing the matrix [M2] only, i.e.

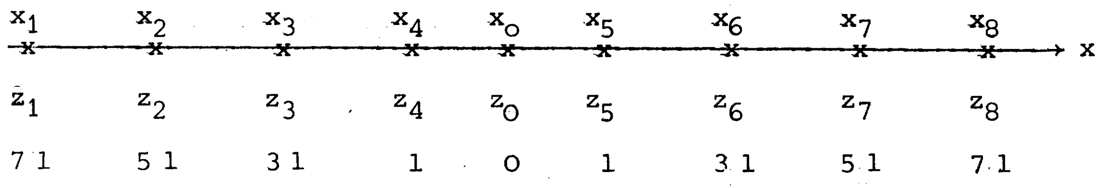
$$[\lambda] = [K]^{-1}\cdot[M2] \text{ and } [\lambda^1] = [K]^{-1}\cdot[M2^1]$$

where $[\lambda^1]$ and $[M2^1]$ are the solution and the column matrix M2 obtained for the new position of the volume V towards the sample system.

EXAMPLE No 1

Let us apply the Kriging system given by Eqs. 264 to the following problem:

One has a set of the point-like data situated along the straight line at the regular distances, as it is shown in Fig. 85



$|x_{i+1} - x_i| = 2 \cdot 1$

Fig. 85

We want to calculate, using the Kriging method the value z_o at the point x_o when the values $z_1, z_2 \ldots z_8$ at the points $x_1, x_6 \ldots x_8$ are known and the variogram is $\gamma(h)$.

The Kriging estimator will obviously be:

$$z_o^* = \lambda_1 \cdot \frac{1}{2} (z_1 + z_8) + \gamma_2 \cdot \frac{1}{2} (z_2 + z_7) + \lambda_3 \cdot \frac{1}{2} (z_3 + z_6) + \lambda_4 \frac{1}{2}(z_4 + z_5) \quad (267)$$

Here, in practice, we have 4 samples

$$\left.\begin{aligned} S_1 &= \frac{1}{2} (z_1+z_8) \\ S_2 &= \frac{1}{2} (z_2+z_7) \\ S_3 &= \frac{1}{2} (z_3+z_6) \\ S_4 &= \frac{1}{2} (z_4+z_5) \end{aligned}\right\} \quad (268)$$

for which the weights λ_1, λ_2, λ_3 and λ_4 have to be found.

Now, to be able to calculate the mean $\bar{\gamma}$ values for the point-like samples $z_1, z_2 \ldots z_8$ we have to use formula 189 (lecture No 8) which explains, how to calculate the mean values of variogram between the point samples.

The system of the Kriging equations will be:

$$\left.\begin{aligned}
&\lambda_1\cdot\bar{\gamma}(S_1;S_1) + \lambda_2\cdot\bar{\gamma}(S_1;S_2) + \lambda_3\cdot\bar{\gamma}(S_1;S_3(+ \lambda_4\cdot\bar{\gamma}(S_1;S_4) + \mu = \bar{\gamma}(z_o;S_1)\\
&\lambda_1\cdot\bar{\gamma}(S_2;S_1) + \lambda_2\bar{\gamma}(S_2;S_2) + \lambda_3\bar{\gamma}(S_2;S_3) + \lambda_4\bar{\gamma}(S_2;S_4) + \mu = \bar{\gamma}(z_o;S_2)\\
&\lambda_1\cdot\bar{\gamma}(S_3;S_1) + \lambda_2\bar{\gamma}(S_3;S_2) + \lambda_3\bar{\gamma}(S_3;S_3) + \lambda_4\bar{\gamma}(S_3;S_4) + \mu = \bar{\gamma}(z_o;S_3)\\
&\lambda_1\cdot\bar{\gamma}(S_4;S_1) + \lambda_2\bar{\gamma}(S_4;S_2) + \lambda_3\bar{\gamma}(S_4;S_3) + \lambda_4\cdot\bar{\gamma}(S_4;S_4) + \mu = \bar{\gamma}(z_o;S_4)\\
&\lambda_1 \qquad + \lambda_2 \qquad + \lambda_3 \qquad + \lambda_4 \qquad + o = 1
\end{aligned}\right\}\quad(269)$$

Let us start with the variance terms:

$\bar{\gamma}\ (S_1;S_1)$

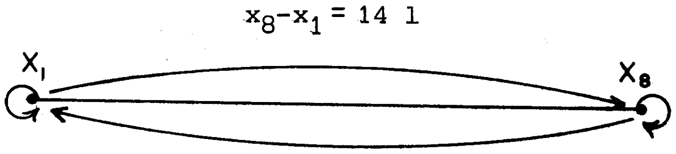


Fig. 86

According to the last term in Eq. (189) one has k=2 and

$$\bar{\gamma}(S_1;S_1) = \frac{1}{2^2}\left[\ \gamma(x_1-x_1) + \gamma\ (x_1-x_8) + \gamma(x_8-x_8) + \gamma\ (x_8-x_1)\ \right]$$

$$= \frac{1}{2}\ [\gamma\ (o) + \gamma(x_8-x_1)] = \qquad (270.a)$$

$$= \frac{1}{2}\ [\gamma(o) + \gamma(14\cdot 1)]$$

Similarily:

$$\bar{\gamma}(S_2;S_2) = \frac{1}{2}\ [\gamma(o) + \gamma(x_7-x_2)] =$$

$$= \frac{1}{2}\ [\gamma(o) + \gamma(10\cdot 1)] \qquad (270.b)$$

$$\bar{\gamma}(S_3;S_3) = \frac{1}{2}\ [\gamma(o) + \gamma(x_6-x_3)] =$$

$$= \frac{1}{2}\ [\gamma(o) + \gamma(6\cdot 1)] \qquad (270.c)$$

$$\bar{\gamma}(S_4;S_4) = \frac{1}{2}\,[\gamma(o) + \gamma(x_5-x_4)] =$$

$$= \frac{1}{2}\,[\gamma(o) + \gamma(2\cdot 1)] \qquad 270.d)$$

For the mixed terms in Eq. (269) the situation is a little more complicated (cf. Fig. 87)

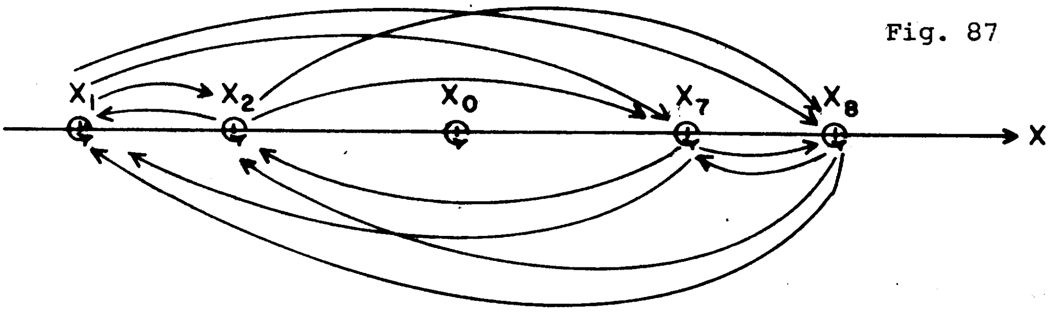


Fig. 87

Always according to the last term in Eq. 189 (we have now k=4):

$$\bar{\gamma}(S_1;S_2) = \frac{1}{4^2}\,[\gamma(x_1-x_1) + \gamma(x_1-x_2) + \gamma(x_1-x_7) + \gamma(x_1-x_8) +$$

$$+ \gamma(x_2-x_1) + \gamma(x_2-x_2) + \gamma(x_2-x_7) + \gamma(x_2-x_8) +$$

$$+ \gamma(x_7-x_1) + \gamma(x_7-x_2) + \gamma(x_7-x_7) + \gamma(x_7-x_8) +$$

$$+ \gamma(x_8-x_1) + \gamma(x_8-x_2) + \gamma(x_8-x_7) + \gamma(x_8-x_8)] =$$

$$= \frac{1}{4^2}\,[\,4\,\gamma(o) + 4\gamma\,(2\cdot 1) + 4\gamma\,(12\cdot 1) + 2\gamma(10\cdot 1) + 2\gamma(14\cdot 1)] =$$

$$= \frac{1}{8}\,[2\gamma(o) + 2\gamma\,(2\cdot 1) + 2\gamma(12\cdot 1) + \gamma(10\cdot 1) + \gamma(14\cdot 1)] \qquad (271.a)$$

Similarily:

$$\bar{\gamma}(S_1;S_3) = \frac{1}{8}\,[2\gamma(o) + 2\gamma(4\cdot 1) + 2\gamma(10\cdot 1) + \gamma(6\cdot 1) + \gamma(14\cdot 1)] \qquad (271.b)$$

$$\bar{\gamma}(S_1;S_4) = \frac{1}{8}\,[2\gamma(o) + 2\gamma(6\cdot 1) + 2\gamma(8\cdot 1) + \gamma(2\cdot 1) + \gamma(14\cdot 1)] \qquad (271.c)$$

$$\bar{\gamma}(S_2;S_3) = \frac{1}{8}\ [2\gamma(0) + 2\gamma(2\cdot 1) + 2\gamma(8\cdot 1) + \gamma(6\cdot 1) + \gamma(10\cdot 1)] \qquad (271.d)$$

$$\bar{\gamma}(S_2;S_4) = \frac{1}{8}\ [2\gamma(o) + 2\gamma(4\cdot 1) + 2\gamma(6\cdot 1) + \gamma(2\cdot 1) + \gamma\ (10\cdot 1)] \qquad (271.e)$$

$$\bar{\gamma}(S_3;S_4) = \frac{1}{8}\ [2\gamma(o) + 2\gamma(2\cdot 1) + 2\gamma(4\cdot 1) + \gamma(2\cdot 1) + \gamma(6\cdot 1)\ ] \qquad (271.f)$$

and because it is always

$$\bar{\gamma}(S_i;S_j) = \bar{\gamma}(S_j;S_i) \qquad (272)$$

thus all terms in the Kriging matrix [K] are found.
Now, to determine the column matrix [M2], one has to use the first term in formula (189) for the volume V reduced to the point x_o (cf. Fig. 85). Thus, one has:

$$\bar{\gamma}(z_o,S_1) = \frac{1}{2}\ [\gamma(x_o-x_1) + \gamma(x_o-x_8)] = \gamma\ (7\cdot 1) \qquad (273.a)$$

$$\bar{\gamma}(z_o,S_2) = \gamma(x_o-x_2) = \gamma(5\cdot 1) \qquad (273.b)$$

$$\bar{\gamma}(z_o,S_4) = \gamma(x_o-x_4) = \gamma(1) \qquad (273.d)$$

Now, because we are treating the point-like samples, let us take the spherical variogram with the parameters $a = 6$ m and $C = 12.53\%$ and $l = 1$ m. Thus, the model variogram is of the form:

$$\gamma(h) = 12.53\ [1.5\cdot(\frac{h}{6}) - 0.5(\frac{h^3}{6})] \quad \%^2 \qquad (274)$$

[H in m]

which gives immediately:

$$\left.\begin{array}{llll} \gamma(o) = 0 & & \gamma\ (5) = 12.04 \\ \gamma(1) = 3.10 & & \gamma\ (6) = 12.53 \\ \gamma(2) = 6.032 & & \gamma(h\geq 6) = 12.53 \\ \gamma(3) = 8.61 & & \\ \gamma(4) = 10.67 & & \end{array}\right\} \qquad (275)$$

Now, setting the data from Eq. (275) into Eqs. (270),(271) and (273) one has:

$$\left.\begin{aligned}
\bar{\gamma}(S_1;S_1) &= 6.265\\
\bar{\gamma}(S_2;S_2) &= 6.265\\
\bar{\gamma}(S_3;S_3) &= 6.265\\
\bar{\gamma}(S_4;S_4) &= 3.016\\
\bar{\gamma}(S_1;S_2) &= 7.773\\
\bar{\gamma}(S_1;S_3) &= 8.9325\\
\bar{\gamma}(S_1;S_4) &= 8.585\\
\bar{\gamma}(S_2;S_3) &= 7.773\\
\bar{\gamma}(S_2;S_4) &= 8.120\\
\bar{\gamma}(S_3;S_4) &= 6.496\\
\bar{\gamma}(z_o;S_1) &= 12.53\\
\bar{\gamma}(z_o;S_2) &= 12.04\\
\bar{\gamma}(z_o;S_3) &= 8.61\\
\bar{\gamma}(z_o;S_4) &= 3.10
\end{aligned}\right\} \qquad (276)$$

Which, together with Eq. (272) gives the Kriging matrix:

$$[K] = \begin{bmatrix} 6.265 & 7.773 & 8.9325 & 8.585 & 1\\ 7.773 & 6.265 & 7.773 & 8.120 & 1\\ 8.9325 & 7.773 & 6.265 & 6.496 & 1\\ 8.585 & 8.120 & 6.496 & 3.016 & 1\\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \qquad (277.a)$$

and the column matrix [M2] is:

$$[M2] = \begin{bmatrix} 12.53 \\ 12.04 \\ 8.61 \\ 3.10 \\ 1.0 \end{bmatrix} \qquad (277.6)$$

This matrix equation can be solved, using for example the TI-59 calculator (Master Library program 02), which gives the solution:

$$[\lambda] = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \mu \end{bmatrix} = \begin{bmatrix} -0.208 \\ -0.393 \\ 0.034 \\ 1.567 \\ 3.131 \end{bmatrix} \qquad (277.c)$$

and the Kriging variance (cf. Eq. 263.a or 266) for $\bar{\gamma}(V;V) = \gamma(o) = 0$ is:

$$\sigma_K^2 = [\lambda]^t \cdot [M2] = \sum_{i=1}^{4} \lambda_i \cdot \bar{\gamma}(z_o, s_i) + \mu =$$

$$= 0.94349 \, \%^2 \qquad (277.d)$$

and the Kriging standard deviation is:

$\sigma_K = 0.971$ %

What surprising result! The nearest samples to the point z_o have the biggest contribution to the Z_o^* value, whereas the next one situated at the points x_3 and x_6 have almost zero contribution and further samples have even the negative λ values! This is a well known in geostatistics effect of screen. The further samples are in the "shadow" of the nearest one. They contribute, however, in the Kriging variance. If the samples, for example, z_1, z_2, z_7 and z_8 are absent, the system of equations (269) will be reduced to the three last ones with the values $\gamma_1 = \lambda_2 = 0$ and its solution will be

$\lambda_3 = -0.547$
$\lambda_4 = 1.547$
$\mu = 1.988$

which will give the Kriging variance (cf. Eq. 277.d)

$\sigma_K^2 = 2.07366$ and $\sigma_K = 1.44$, thus the estimation error in this case will be greater. Here the screening effect is even more emphasized - the second sample has still a negative λ value.

LECTURE 12

J.A. Czubek

LECTURE 12 Reykjavík, March 23, 1981

EXAMPLE No 2

Let us now take again 8 samples z_1, z_2 ...z_8 having the same distances from the point x_o but now they are distributed in the plane x,y as in Fig. 88,

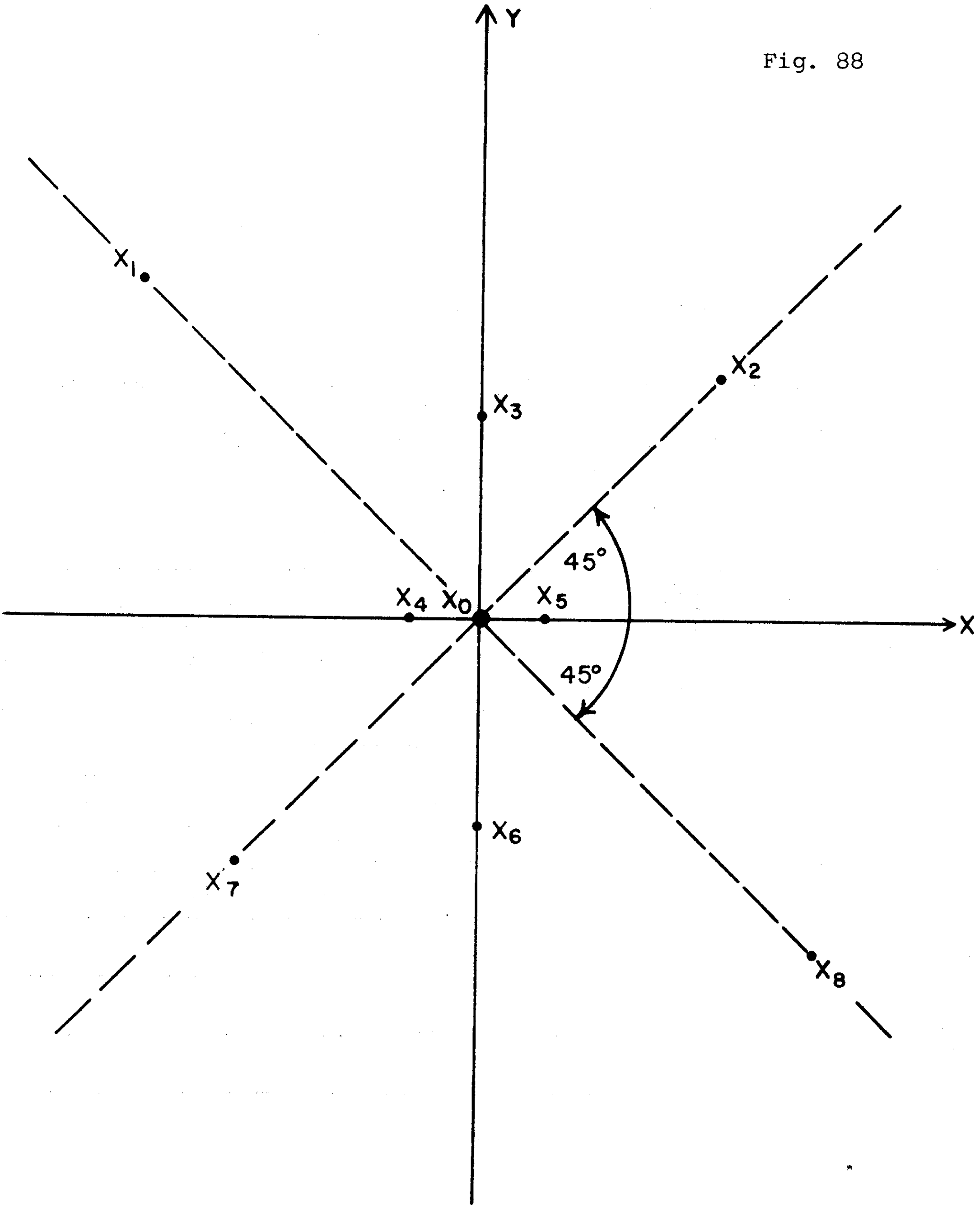


Fig. 88

where the x,y coordinates of the points are as below:

(in the l units):

Point	x	y
x_o	0	0
x_1	- 4.9497	+ 4.9497
x_2	+ 3.5355	+ 3.5355
x_3	0	+3
x_4	- 1	0
x_5	+ 1	0
x_6	0	- 3
x_7	- 3.5355	- 3.5355
x_8	+ 4.9497	- 4.9497

and the relative distances between the points (always in the units of l) are, according to their coordinates, given below

	x_o	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_o	0	7.0	5.0	3.0	1.0	1.0	3.0	5.0	7.0
x_1		0	8.6022	5.3198	6.3324	7.7394	9.3674	8.6022	14.00
x_2			0	3.5758	5.7507	4.3507	7.4305	10.00	8.6022
x_3				0	3.1623	3.1623	6.00	7.4305	9.3647
x_4					0	2.00	3.1623	4.3507	7.7394
x_5						0	3.1623	5.7507	6.3324
x_6							0	3.5758	5.3198
x_7								0	8.6022
x_8									0

The general form of the Kriging equations for this example is the same as for the previous one (i.e. Eq. 269), but the particular $\bar{\gamma}(S_i\,;\,S_j)$ values are different now. The $\bar{\gamma}(S_i;S_i)$ values are the same as previously, as well as the whole [M2] matrix and, referring to Fig. 88, one has:

$$\bar{\gamma}(S_1;S_2) = \frac{1}{4^2}\,[4\gamma(0) + 2\gamma(\bar{x}_1-\bar{x}_2) + 2\gamma(\bar{x}_1-\bar{x}_7) + 2\gamma(\bar{x}_1-\bar{x}_8) +$$

$$+\ 2\gamma(\bar{x}_2-\bar{x}_7) + 2\gamma(\bar{x}_2-\bar{x}_8) + 2\bar{\gamma}(\bar{x}_7-\bar{x}_8)\,] =$$

$$= \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_1-\bar{x}_2) + \gamma(\bar{x}_1-\bar{x}_7) + \gamma(\bar{x}_1-\bar{x}_8)\ +$$

$$+\gamma(\bar{x}_2-\bar{x}_7) + \gamma(\bar{x}_2-\bar{x}_8) + \gamma(\bar{x}_7-\bar{x}_8)\,] \qquad (278.a)$$

which referred to the table with the mutual distances between the points becomes:

$$\bar{\gamma}(S_1;S_2) = \frac{1}{8}\,[2\gamma(0) + \gamma(8.6022) + \gamma(8.6022) + \gamma(14) + \gamma(10) +$$

$$+\ \gamma(8.6022) + \gamma(8.6022)\,] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 4\gamma(8.6022) + \gamma(10) + \gamma(14)\,] \qquad (278.b)$$

$$\bar{\gamma}(S_1;S_3) = \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_1-\bar{x}_3) + \gamma(\bar{x}_1-\bar{x}_6) + \gamma(\bar{x}_1-\bar{x}_8) +$$

$$+\ \gamma(\bar{x}_3-\bar{x}_6) + \gamma(\bar{x}_3-\bar{x}_8) + \bar{\gamma}(\bar{x}_6-\bar{x}_8)\,] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 2\gamma(5.3198) + 2\gamma(9.3647) + \gamma(6) + \gamma(14)\,] \qquad (278.c)$$

$$\bar{\gamma}(S_1;S_4) = \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_1-\bar{x}_4) + \gamma(\bar{x}_1-\bar{x}_5) + \gamma(\bar{x}_1-\bar{x}_8)\ +$$

$$+\ \gamma(\bar{x}_4-\bar{x}_5) + \gamma(\bar{x}_4-\bar{x}_8) + \gamma(\bar{x}_5-\bar{x}_8)\,] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 2\gamma(6.3324) + 2\gamma(7.7394) + \gamma(2) + \gamma(14)\,] \qquad (278.d)$$

$$\bar{\gamma}(S_2;S_3) = \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_2-\bar{x}_3) + \gamma(\bar{x}_2-\bar{x}_6) + \gamma(\bar{x}_2-\bar{x}_7)\ +$$

$$+\ \gamma(\bar{x}_3-\bar{x}_6) + \gamma(\bar{x}_3-\bar{x}_7) + \gamma(\bar{x}_6-\bar{x}_7)\,] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 2\gamma(3.5758) + 2\gamma(7.4305) + \gamma(6) + \gamma(10)\,] \qquad (278.e)$$

$$\bar{\gamma}(S_2;S_4) = \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_2-\bar{x}_4) + \gamma(\bar{x}_2-\bar{x}_5) + \gamma(\bar{x}_2-\bar{x}_7) +$$

$$+ \gamma(\bar{x}_4-\bar{x}_5) + \gamma(\bar{x}_4-\bar{x}_7) + \gamma(\bar{x}_5-\bar{x}_7)] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 2\gamma(4.3507) + 2\gamma(5.7507) + \gamma(2) + \gamma(10)] \qquad (278.f)$$

$$\bar{\gamma}(S_3;S_4) = \frac{1}{8}\,[2\gamma(0) + \gamma(\bar{x}_3-\bar{x}_4) + \gamma(\bar{x}_3-\bar{x}_5) + \gamma(\bar{x}_3-\bar{x}_6) +$$

$$+ \gamma(\bar{x}_4-\bar{x}_5) + \gamma(\bar{x}_4-\bar{x}_6) + \gamma(\bar{x}_5-\bar{x}_6)] =$$

$$= \frac{1}{8}\,[2\gamma(0) + 4\gamma(3.1625) + \gamma(2) + \gamma(6)] \qquad (278.g)$$

Taking again the spherical variogram given by Eq. (274) one has additionally to the values given in Eq. 275):

$$\left.\begin{array}{l} \gamma\ (3.1623) = 8.9887 \\ \gamma\ (3.5758) = 9.8750 \\ \gamma\ (4.3507) = 11.2400 \\ \gamma\ (5.3198) = 12.2976 \\ \gamma\ (5.7507) = 12.4980 \end{array}\right\} \qquad (279)$$

which setted together with the formulae (275) into Eqs. 278 gives:

$$\left.\begin{array}{l} \bar{\gamma}\ (S_1;S_2) = 9.3975 \\ \bar{\gamma}\ (S_1;S_3) = 9.3394 \\ \bar{\gamma}\ (S_1;S_4) = 8.5854 \\ \bar{\gamma}\ (S_2;S_3) = 8.7337 \\ \bar{\gamma}\ (S_2;S_4) = 8.2549 \\ \bar{\gamma}\ (S_3;S_4) = 6.8147 \end{array}\right\} \qquad (280)$$

Now the Kriging matrix [K] is:

$$[K] = \begin{vmatrix} 6.265 & 9.3975 & 9.3394 & 8.5854 & 1 \\ 9.3975 & 6.265 & 8.7337 & 8.2549 & 1 \\ 9.3394 & 8.7337 & 6.265 & 6.8147 & 1 \\ 8.5854 & 8.2549 & 6.8147 & 3.016 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix} \qquad (281)$$

the determinant of which is:

$$\| K \| = 98.31341151$$

The Kriging matrix (281) together with the one column matrix [M2] given by (277.b) give the solution for the [λ] matrix in the form:

$$[\lambda] = \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \mu \end{vmatrix} = \begin{vmatrix} -0.330 \\ -0.362 \\ 0.114 \\ 1.578 \\ 3.3867 \end{vmatrix} \qquad (282)$$

which gives the Kriging variance σ_k^2 according to Eq. (277.d):

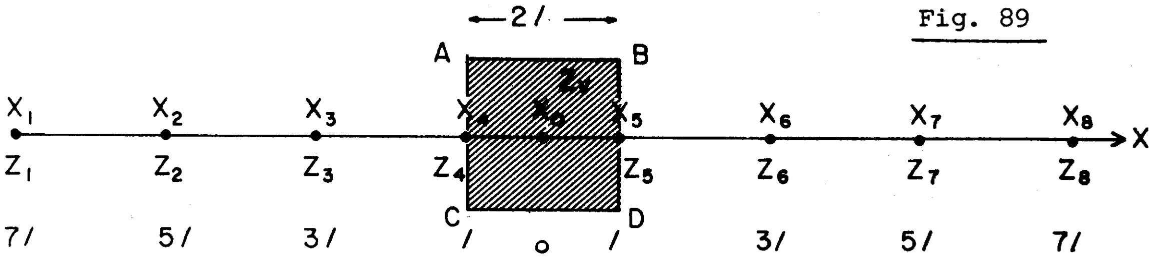
$$\sigma_k^2 = 0.766677 \,\%^2$$
$$\sigma_k = 0.876 \,\% \qquad (283)$$

Here again we have the screen effect visible for the samples No 1 and 2 , but it is not so strong (in per cent) as in Example No 1. Moreover the better spatial distribution of samples gives here a lower Kriging variance σ_k^2 than in the case of the aligned data.

Example No 3

We go back to Example No 1 and for the sample configuration presented in Fig. 85 we shall try to define the value Z_v of the square ABCD located around the point z_o with the side 2·1 , cf. Fig. 89

Fig. 89



The Kriging estimator z_v^* is essentially the same as in Eq. (267), but with different λ_i values. The Kriging system given by Eq. 269 is still valid, with the same values of the $\bar{\gamma}(S_i;S_j)$ coefficients (the sample configuration has not been changed). The only difference is in the [M2] matrix. Here the $\bar{\gamma}(Z_o,S_i)$ means now $\bar{\gamma}(Z_v,S_i)$ and it is an averaged value of the variogram, when one extremity of the segment h in the variogram $\gamma(h)$ is "walking" around the square ABCD, and another extremity is fixed at the sample S_i this is just the first term of the right hand side of Eq. 189 (cf. Lecture 8) divided by 2.

For the S_1 sample one has to calculate:

$$\bar{\gamma}(Z_v;S_1) = \frac{1}{S\cdot 2}\,[\int\limits_{ABCD} ds\,\gamma(x-x_1) + \int\limits_{ABCD} ds\,\gamma(x-x_8)] =$$

$$= \frac{1}{S}\int\limits_{ABCD}\gamma(x-x_1)\,ds \qquad (284)$$

where the point x is moving inside the square ABCD and the point x_1 is fixed, and $x-x_1$ means the distance between the two points. The situation given by Eq. (284) is presented with more details in Fig. 90 for the arbitrary rectangle ABCD with sides L and 2·1 .

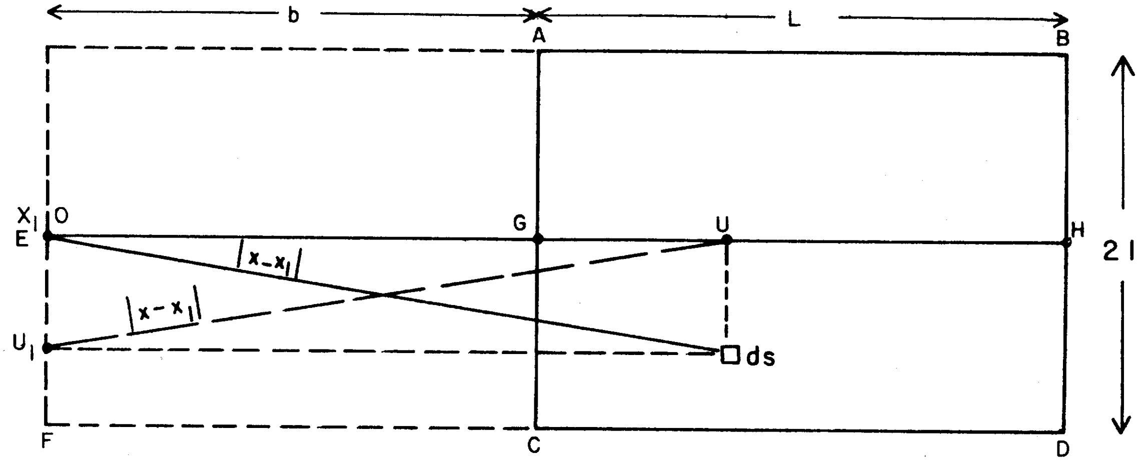


Fig. 90

We have:

$$\frac{1}{S}\int_{ABCD}\gamma\,(x-x_1)\,ds = \frac{1}{L\cdot 2l}\left[\int_{ABGH}\gamma(x-x_1)\,ds + \int_{GHCD}\gamma\,(x-x_1)\,ds\right] =$$

$$= \frac{1}{L\cdot l}\int_{GHCD}\gamma(x-x_1)\,ds = \frac{1}{L\cdot l}\int_E^F du_1\int_G^H du\;\gamma(\sqrt{u^2+u_1^2}) =$$

$$= \frac{1}{L\cdot l}\int_0^l du_1\int_b^{b+L} du\;\gamma(\sqrt{u^2+u_1^2}) =$$

$$= \frac{1}{L\cdot l}\int_0^l du_1\left[\int_0^{b+L} du\;\gamma(\sqrt{u^2+u_1^2}) - \int_0^b du\;\gamma(\sqrt{u^2+u_1^2})\right]$$

$$= \frac{b+L}{L}\;\frac{1}{(b+L)\,l}\int_0^l du_1\int_0^{b+L} du\;\gamma(\sqrt{u^2+u_1^2}) -$$

$$- \frac{b}{L}\cdot\frac{1}{b\cdot l}\int_0^l du_1\int_0^b du\;\gamma(\sqrt{u^2+u_1^2}) =$$

$$= \frac{b+L}{L}\cdot H[(b+L);l] - \frac{b}{L}\cdot H(b;l) \qquad (285)$$

Here we have used the definition of the auxiliary function H(L;l) given in Eq. 154.

Adjusting now Eq. 285 to the conditions of Eq. 284 we have: ($L=2\,l$, $b=6\,l$)

$$\bar{\gamma}(Z_V;S_1) = 4H[8\,l;l] - 3H[6\,l;\,l] \qquad (286.a)$$

and similarily:

$$\bar{\gamma}(Z_V;S_2) = 3H[6\,l;l] - 2\,H[4\,l;\,l] \qquad (286.b)$$

$$\bar{\gamma}(Z_V;S_3) = 2H[4\,l;l] - H[2\,l;l] \qquad (286.c)$$

$$\bar{\gamma}(Z_V;S_4) = H[2\,l;l] \qquad (286.d)$$

Now we can solve the Kriging system given by Eq. (269) and for the Kriging variance (Eq. 273.a) one needs more the value

$$\bar{\gamma}(V,V) = \bar{\gamma}(ABCD;ABCD) = F(2\,l;2\,l) \qquad (287)$$

Taking again $l = 1$ m and the spherical variogram given by Eq. (274) we have:

$$\left.\begin{aligned} &\frac{l}{a} = \frac{1}{6} = 0.1667 \\ &\frac{2\,l}{a} = \frac{1}{3} = 0.3333 \\ &\frac{4\,l}{a} = \frac{2}{3} = 0.6667 \\ &\frac{6\,l}{a} = 1.000 \\ &\frac{8\,l}{a} = \frac{4}{3} = 1.3333 \end{aligned}\right\} \qquad (288)$$

The values $\frac{1}{C} \cdot H(L;l)$ for the spherical variogram can be found on the graph in Fig. 59. One has for the values in (288):

$$\left.\begin{aligned} &\frac{1}{C} \cdot H(2\,l;l) = 0.292 \qquad && \frac{1}{C}\,H(8\,l;\,l) = 0.730 \\ &\frac{1}{C}\,H(4\,l;l) = 0.445 && \\ &\frac{1}{C}\,H(6\,l;l) = 0.637 && \end{aligned}\right\} \qquad (289)$$

The value F(21;21) is taken from Fig. 60:

$$\frac{1}{C} \cdot F(21;21) = 0.255 \qquad (289.a)$$

Now, for $C = 12.53\%^2$ one has from Eqs. (286) and (289):

$$\begin{aligned} \bar{\gamma}(Z_V;S_1) &= 12.53 \\ \bar{\gamma}(Z_V;S_2) &= 12.53 \\ \bar{\gamma}(Z_V;S_3) &= 7.49 \\ \bar{\gamma}(Z_V;S_4) &= 3.66 \end{aligned} \qquad (290)$$

Now one solves the system

$$[K] \cdot [\lambda] = [M2] \qquad (291)$$

With the matrix [K] given in Eq. (277.a) and the matrix [M2] as below:

$$[M2] = \begin{vmatrix} 12.53 \\ 12.53 \\ 7.49 \\ 3.66 \\ 1.0 \end{vmatrix} \qquad (292)$$

which gives the vector $[\lambda]$:

$$[\lambda] = \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \mu \end{vmatrix} = \begin{vmatrix} 0.0106 \\ -1.1398 \\ 1.0722 \\ 1.0570 \\ 2.6714 \end{vmatrix} \qquad (293)$$

The determinant of the matrix [K] is:

$$\|K\| = 31.8663571 \qquad (277.e)$$

and the Kriging variance (cf. Eq. 263.a):

$$\sigma_K^2 = \sum_{i=1}^{4} \lambda_i \cdot \bar{\gamma}(Z_V;S_i) + \mu - \bar{\gamma}(Z_V;Z_V) = 0.4219 - 0.255 \times 12.53 =$$

$$= -2.773 \qquad (294)$$

Thus, the value of the Kriging variance is negative, which is, of course, not possible. This is a sign that something is wrong in the calculation. The most probable error is in the matrix [M2] the elements of which has been obtained as a combination of the H(L;l) functions. The values of the H(L;l) functions have been found on graphs which does not give enough accuracy. The matrix [K] has been obtained using the exact formulae and the numerical calculations enough accurate. In the solution, given by Eq. (293) the value λ_2 seems to be too low whereas the values λ_1 too high. Next, the set of samples S_2 have a higher weight than the sample S_1 which is situated just at the perimeter of the rectangle ABCD.

To solve this problem we have to calculate the exact values of the H(L,l) function or rather, the exact values of the $\bar{\gamma}(Z_V;S_i)$ averaged variograms. This is very often the case in any Kriging matrix, and computer programs for Kriging contain the procedure to calculate all auxiliary functions needed for this kind of calculation.

In our case the $\bar{\gamma}(Z_V;S_i)$ functions can be written as (cf. Eq. 285):

$$\bar{\gamma}(Z_V;S) = \frac{1}{2\,l \cdot l} \cdot \int_o^l du_1 \int_{b_k}^{b_k+2\,l} du \cdot \gamma(\sqrt{u_1{}^2+u^2}) \approx$$

$$\approx \frac{1}{n_1 \cdot n_2} \cdot \sum_{i=o}^{n_1-1} \sum_{j=o}^{n_2-1} \gamma(\sqrt{[b_k+\frac{2\,l}{n_1}(i+0.5)]^2+[\frac{l}{n_2}(j+0.5]^2}) =$$

$$= \mathrm{HFUNC2} \cdot C$$

The program for the calculation of the function HFUNC2 written for the TI-59 calculator is given below. It is written for the spherical variogram.

TITOLO
TITEL
TITEL
TITEL HFUNC 2

PROGRAMMATORE
PROGRAMMERARE
PROGRAMMER
PROGRAMMEUR J.A. CZUBEK

PAG. SIDA SIDE PAG. 1 DI AV AF VAN 3

DATA DATUM DATO DATUM MARCH 24, 1981

TI PROGRAMMABLE
SCHEDA DI PROGRAMMA
PROGRAMDOCUMENTATION
DOKUMENTATION
DOCUMENTATIE

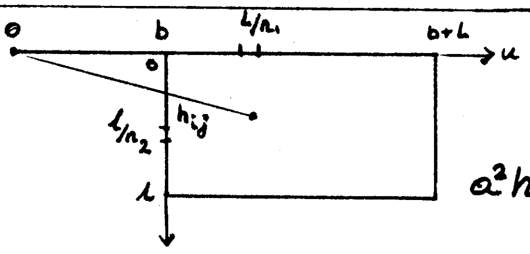
P izione (Op 17)
eining (Op 17)
S ning (Op 17)
Onderverdeling (Op 17)

MODULO BIBLIOTECA
BIBLIOTEKS-MODUL
BIBLIOTEKSMODUL
BIBLIOTHEEKMODUUL

Stampante
Skrivenhet
Printer
Afdrukeenheid

Schede
Kort
Kort
Kaarten

DESCRIZIONE DEL PROGRAMMA ● PROGRAMBESKRIVNING ● PROGRAMBESKRIVELSE ● PROGRAMMABESCHRIJVING



$$HFUNC2=\frac{1}{n_1\cdot n_2}\cdot\sum_{i=0}^{n_1-1}\sum_{j=0}^{n_2-1}\gamma(h_{ij})$$

$$a^2h_{ij}^2=\left[b+\frac{L}{n_1}(i+0.5\right]^2+\left[\frac{\ell}{n_2}(j+0.5)\right]^2$$

$$\gamma(h_{ij})=0.5\left[3\left(\frac{h_{ij}}{a}\right)-0.5\left(\frac{h_{ij}}{a}\right)^3\right]$$

ISTRUZIONI PER L'OPERATORE ● BRUKSANVISNING ● BRUGERVEJLEDNING ● INSTRUKTIES VOOR DE GEBRUIKER

PASSO STEG TRIN STAP	PROCEDURA PROCEDUG PROCEDURE PROCEDURE	IMPOSTARE MATA IN INDTAST VOER IN	PREMERE TRYCK NED TRYK TOETS IN	VISUALIZZATORE SIFFERINDIKATOR LYSPANEL DISPLAY
1	INSERT PROGRAM			
2			CMS	
3	PARAMETER OF VARIOGRAM-RANGE	a	R/S	a
4		b	R/S	b
5		L	R/S	L
6		ℓ	R/S	ℓ
7		n_1	R/S	n_1
8		n_2	R/S	$H(b,L,\ell)$

TASTI DEF. DALL OPERATORE
ANVANDBAR DEF. TANGENTER
BRUGERDEFINERBARE TASTER
DOOR DE GEBRUIKER TE DEF. TOETSEN

Ⓐ $j,\ i(0,n_1-1)=h_{ij}$
Ⓑ $i=0,\ j=j+1$
Ⓒ HFUNC2
Ⓓ $\gamma(h_{ij})\ h_{ij}\leq 1$
Ⓔ $\gamma(h_{ij})\ h_{ij}\geq 1$
A'
B'
C'
D'
E'

REGISTRI DATI
DATA REGISTER
DATAREGISTRE
DATAREGISTERS
(INV List)

0		10	$1/n_1\times L/a$
1	a	11	$\frac{1}{n_2}\cdot\ell/a$
2	b	12	i
3	L	13	j
4	ℓ	14	$n_1\cdot n_2$
5	n_1	15	h_{ij}/a
6	n_2	16	$\sum\gamma(h_{ij})$
7	b/a	17	HFUNC 2
8	L/a	18	$\sum_i\sum_j=n_1\cdot n_2$
9	ℓ/a	9	

ETICHETTE (Op 08)
LABELS (Op 08)
ETIXETTER (Op 08)
LABELS (Op 08)

INV __	lnx __	CE __	CLR __	x⇌t __	x² __
√x __	1/x __	STO __	RCL __	SUM __	yˣ __
EE __	(__	) __	÷ __	GTO __	× __
SBR __	− __	RST __	+ __	R/S __	. __
+/− __	= __	CLR __	INV __	log __	CP __
tan __	Pgm __	P→R __	sin __	cos __	CMs __
Exc __	Prd __	|x| __	Eng __	Fix __	Int __
Deg __	Pause __	x=t __	Nop __	Op __	Rad __
Lbl __	x≥t __	Σ+ __	x̄ __	Grad __	St flg __
If flg __	D.MS __	π __	List __	Write __	Dsz __
Adv __	Prt __				

0	1	2	3	4	5	6	7	8	9

TITOLO TITEL TITEL TITEL: HFUNC 2
PAG. SIDA SIDE PAG.: 2 DI AV AF VAN 3
PROGRAMMATORE PROGRAMMERARE PROGRAMMER PROGRAMMEUR: J.A. CZUBEK
DATA DATUM DATO DATUM: 24 MARCH 1981

TI PROGRAMMABLE
MODULO DI PROGRAMMAZIONE
KOD FORMULAR
PROGRAMBLAD
PROGRAMMABLAD

POS. ADR TRIN LOC.	COD. KOD KODE CODE	TASTO TANGENT TAST TOETS	COMMENTO KOMMENTAR KOMMENTAR TOELICHTING
0	42	STO	
1	01	01	a
2	91	R/S	
3	42	STO	
4	02	02	b
5	91	R/S	
6	42	STO	
7	03	03	L
8	91	R/S	
9	42	STO	
10	04	04	ℓ
1	91	R/S	
2	42	STO	
3	05	05	n_1
4	91	R/S	
5	42	STO	
6	06	06	n_2
7	43	RCL	
8	01	01	
9	35	1/x	
20	65	x	
1	32	x><t	
2	43	RCL	
3	02	02	
4	95	=	
5	42	STO	
6	07	07	b/a
7	32	x><t	
8	65	x	
9	32	x><t	
30	43	RCL	
1	03	03	
2	95	=	
3	42	STO	
4	08	08	L/a
5	32	x><t	
6	65	x	
7	43	RCL	
8	04	04	
9	95	=	
40	42	STO	
1	09	09	ℓ/a
2	43	RCL	
3	08	08	
4	55	÷	
5	43	RCL	
6	05	05	
7	95	=	
8	42	STO	
9	10	10	$1/n_1 \cdot L/a$
50	43	RCL	
1	09	09	
2	55	÷	
3	43	RCL	
4	06	06	

POS. ADR TRIN LOC.	COD. KOD KODE CODE	TASTO TANGENT TAST TOETS	COMMENTO KOMMENTAR KOMMENTAR TOELICHTING
5	95	=	
6	42	STO	
7	11	11	$1/n_2 \cdot ℓ/a$
8	00	0	
9	42	STO	
60	12	12	i=0
1	42	STO	
2	13	13	j=0
3	43	RCL	
4	05	05	
5	65	x	
6	43	RCL	
7	06	06	
8	95	=	
9	42	STO	
70	14	14	$n_1 \cdot n_2$
1	32	x><b	
2	43	RCL	
3	12	12	
4	85	+	
5	01	1	
6	54	)	
7	65	x	
8	53	(	
9	43	RCL	
80	13	13	
1	85	+	
2	01	1	
3	54	)	
4	85	+	
5	01	1	
6	95	=	
7	77	x≥t	
8	13	C	
9	11	A	
90	76	LBL	
1	11	A	
2	43	RCL	
3	05	05	
4	32	x><b	
5	43	RCL	
6	12	12	
7	77	x≥t	
8	12	B	
9	43	RCL	
100	12	12	
1	85	+	
2	93	.	
3	05	5	
4	95	=	
5	65	x	
6	43	RCL	
7	10	10	
8	85	+	
9	43	RCL	

POS. ADR TRIN LOC.	COD. KOD KODE CODE	TASTO TANGENT TAST TOETS	COMMENTO KOMMENTAR KOMMENTAR TOELICHTING
110	07	07	
1	95	=	
2	33	x^2	
3	85	+	
4	53	(	
5	53	(	
6	43	RCL	
7	13	13	
8	85	+	
9	93	.	
120	05	5	
	54	)	
2	65	x	
3	43	RCL	
4	11	11	
5	54	)	
6	33	x^2	
	95	=	
8	34	√	h_{ij}
	42	STO	
130	15	15	
	01	1	
	44	SUM	
	12	12	
4	44	SUM	
	18	18	
	14	D	
	76	LBL	
8	12	B	
9	00	0	
140	42	STO	
	12	12	
	43	RCL	
	06	06	
4	32	x><b	
	43	RCL	
	13	13	
	85	+	
	01	1	
	95	=	
150	77	x≥t	
	13	C	
	01	1	
3	44	SUM	
4	13	13	
5	11	A	
6	76	LBL	
7	13	C	
8	43	RCL	
9	16	16	

CODICI COMPOSTI
KOPPLADE KODER
FLETTEDE KODER
SAMENGEVOEGDE CODES

62	Pgm	Ind	72	STO	Ind	83	GTO	Ind
63	Exc	Ind	73	RCL	Ind	84	Op	Ind
64	Prd	Ind	74	SUM	Ind	92	INV	SBR

TEXAS INSTRUMENTS

TITOLO
TITEL
TITEL
TITEL HFUNC 2
PROGRAMMATORE
PROGRAMMERARE
PROGRAMMER
PROGRAMMEUR J.A.CZUBEK

PAG. SIDA SIDE PAG. 3 DI AV AF VAN 3
DATA DATUM DATO DATUM 24 MARCH 1981

TI PROGRAMMABLE
MODULO DI PROGRAMMAZIONE
KOD FORMULAR
PROGRAMBLAD
PROGRAMMABLAD

POS. ADR TRIN LOC.	COD. KOD KODE CODE	TASTO TANGENT TAST TOETS	COMMENTO KOMMENTAR KOMMENTAR TOELICHTING	POS. ADR TRIN LOC.	COD KOD KODE CODE	TASTO TANGENT TAST TOETS	COMMENTO KOMMENTAR KOMMENTAR TOELICHTING
160	55	÷		5			
1	43	RCL		6			
2	14	14		7			
3	95	=		8			
4	42	STO		9			
5	17	17		0			
6	91	R/S		1			
7	81	RST		2			
8	76	LBL		3			
9	14	D		4			
170	01	1		5			
1	32	x><t		6			
2	43	RCL		7			
3	15	15		8			
4	77	x≥t		9			
5	15	E		0			
6	43	RCL		1			
7	15	15		2			
8	65	x		3			
9	53	(		4			
180	24	CE		5			
1	33	x^2		6			
2	75	−		7			
3	03	3		8			
4	54	)		9			
5	55	÷		0			
6	02	2		1			
7	95	=		2			
8	94	+/−		3			
9	44	SUM		4			
190	16	16		5			
1	11	A		6			
2	92	INV SBR		7			
3	76	LBL		8			
4	15	E		9			
5	01	1		0			
6	44	SUM		1			
7	16	16		2			
8	11	A		3			
9	92	INV SBR		4			
200				5			
1				6			
2				7			
3				8			
4				9			
5				0			
6				1			
7				2			
8				3			
9				4			
0				5			
1				6			
2				7			
3				8			
4				9			

CODICI COMPOSTI
KOPPLADE KODER

FLETTEDE KODER
SAMENGEVOEGDE CODES

62 Pgm Ind	72 STO Ind	83 GTO Ind
63 Exc Ind	73 RCL Ind	84 Op Ind
64 Prd Ind	74 SUM Ind	92 INV SBR

TEXAS INSTRUMENTS

The values of the HFUNC2 $(\frac{b}{a};\frac{b}{a};\frac{1}{a})$ are now:

$$\text{HFUNC2}\ (\tfrac{4}{6};\tfrac{2}{6};\tfrac{1}{6}) = 0.9517 \qquad (n_1=20;n_2=10)$$

$$\text{HFUNC2}\ (\tfrac{2}{6};\tfrac{2}{6};\tfrac{1}{6} = 0.6912 \qquad (n_1=20;n_2=10)$$

$$\text{HFUNC2}\ (0;\tfrac{2}{6};\tfrac{1}{6}) = 0.2906 \qquad (n_1=50;n_2=40)$$

which gives the new [M2] matrix

$$[M2] = \begin{vmatrix} 12.53 \\ 11.92 \\ 8.66 \\ 3.641 \\ 1.0 \end{vmatrix} \qquad (292.a)$$

and the new set of solutions:

$$[\lambda] = \begin{vmatrix} -0.212 \\ -0.338 \\ 0.113 \\ 1.437 \\ 3.139 \end{vmatrix} = \begin{vmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \mu \end{vmatrix} \qquad (293)$$

which looks more reasonable. The Kriging variance, however is still negative (slightly!). This is probably due to the uncertainity in the knowledge of the $\bar{\gamma}(Z_v;Z_v) = F(2\,1;\,2\,1)$ function, which again has to be computed numerically:

$$\sigma_k^2 = 2.664 - 12.53 \cdot F(2\,1;2\,1) = -\,0.530 \qquad ?$$

LECTURE 13

J.A. Czubek

LECTURE 13 Reykjavík, 24 March, 1981

KRIGING IN THE PRESENCE OF DRIFT

We assume now that the random variable $Z(x)$ has a drift, i.e. the expected value at x

$$E\{Z(x)\} = m(x) \tag{296}$$

is now a function (usually unknown) of x. This situation can be depicted as in Fig. 91:

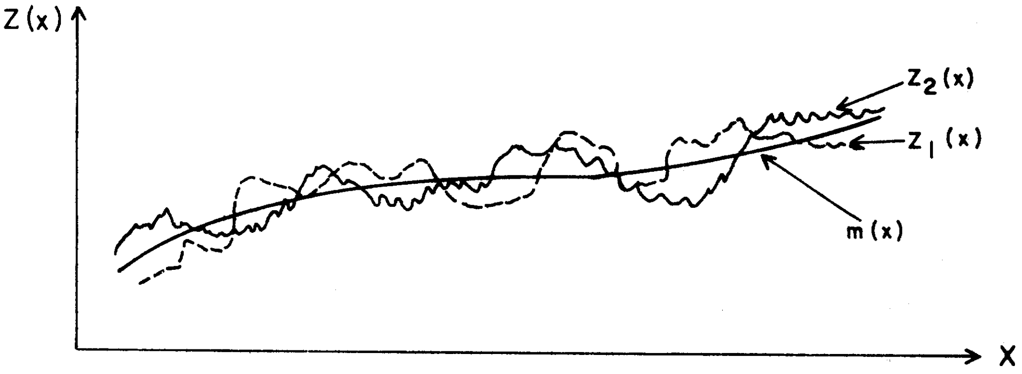


Fig. 91

Around the drift $m(x)$ the particular realizations $z_1(x)$, $z_2(x)$ etc. of the random variable $Z(x)$ are oscillating. We are again limited to the one particular realization, let say $z_1(x) \equiv z(x)$ of the random variable $Z(x)$.

We can consider that the random variable $Z(x)$, according to Fig. 91 is a sum

$$Z(x) = m(x) + Y(x) \tag{297}$$

where $Y(x)$ is a residual term having the zero expectation value

$$E\{Y(x)\} = 0 \tag{298}$$

and we can assume that it has a second order (weak) stationarity or even it has an intrinsic hypothesis i.e. we assume that for the residual $Y(x)$ exists the variogram

$$\gamma(h) = \frac{1}{2} E\{[Y(x+h) - Y(x)]^2\} \tag{298}$$

which together with (297) and (296) gives

$$\gamma(h) = \frac{1}{2} \ E\{[Z(x+h) - Z(x)]^2\} - \frac{1}{2} \ [m(x+h) - m(x)]^2 \tag{299}$$

The value $\frac{1}{2} \ E\{[Z(x+h) - Z(x)]^2\}$ which is only accessible for measurement when the drift is unknown, is simply an experimental variogram which can differ considerably from the underlying variogram $\gamma(h)$. It is not rigourously possible to determine from one single realization $z(x)$ of the random variable $Z(x)$ simultaneously the variogram and the drift. In practical applications, however,one can assume the existance of the intrinsic variogram $\gamma(h)$, which for small values of h (in comparison with the long range variability of the drift) can be presented as a linear type:

$$\left.\begin{aligned} &\gamma(x,y) = \gamma(h) \approx \overline{\omega} \ |h| \\ &\text{where} \quad h = (x-y) \end{aligned}\right\} \tag{300}$$

and the drift of the non-stationary random function $Z(x)$ can be assumed (at least at the vicinity of some point x_o) being of the form:

$$E\{Z(x)\} = m(x) = \sum_{l=1}^{k} a_l \cdot {}^l f(x) \tag{301}$$

where a_l are unknown coefficients and ${}^l f(x)$ are the known functions of x^l (x power l). These functions in the simplest one dimension case are just the monoms x^l , whereas in the two-dimensional case (u,v) there is (for l=2 , for example)

$${}^2 f(x) = {}^2 f(u,v) = {}_2 a_1 \ u^2 + {}_2 a_2 \ u \cdot v + {}_2 a_3 \cdot v^2 \tag{302}$$

etc. Approximation (301) can be regarded as the truncated expansion of the function m(x) into the Taylor's series.

---

[9]. G. GAMBOLATI, G. VOLPI, 1979. Groundwater Contour Mapping in Venice by Stochastic Interpolators. WATER RESOURCES RES. 15, No 2, 281 - part I and II.

[10]. J.P.DELHOMME, 1979, SPATIAL VARIABILITY AND UNCERTAINTY IN GROUNDWATER FLOW PARAMETERS: A GEOSTATISTICAL APPROACH: WATER RESOURCES RES., 15, No 2, 269.

Now, when the n data are available at n points x_i $(i=1,2...n)$, e.g. $z_i = z(x_i)$, and when $D(x_o)$ is neighbourhood centered on x_o, we assume that at this vicinity the co-variance $C(h)$ or the variogram $\gamma(h)$ are known and the Kriging estimator is

$$Z_k^* = \sum_{i=1}^{n} \lambda_i \, z_i \qquad (303)$$

for some volume V. Assuming the non-bias condition:

$$E\{Z_V - Z_K^*\} = E\{Z_V\} - E\{Z_K^*\} = 0 \qquad (304)$$

one has here

$$E\{Z_V\} = \sum_{l=1}^{k} a_l \frac{1}{V} \int_{V(x_o)} {}^{l}f(x)\,dx = \sum_{l=1}^{k} a_l \cdot {}^{l}b_V \qquad (305)$$

and

$$E\{Z_K^*\} = \sum_{i=1}^{n} \lambda_i \; E\{Z_i\} = \sum_{i=1}^{n} \lambda_i \cdot \sum_{l=1}^{k} a_i \cdot \frac{1}{v_i} \int_{v_i} {}^{l}f(x)\,dx =$$

$$= \sum_{i=1}^{n} \lambda_i \cdot \sum_{l=1}^{k} a_l \cdot {}^{l}b_{v_i} = \sum_{i=1}^{k} \lambda_i \; {}^{l}b_{v_i} \qquad (306)$$

Here the ${}^{l}b_v$ coefficients are defined as a mean value of ${}^{l}f(x)$ function inside the volume v :

$$ {}^{l}b_v = \frac{1}{v} \int_{v(x_o)} {}^{l}f(x)\,dx \qquad (307)$$

thus, the coefficients ${}^{l}b_v$ are the functions of the position x_o at which the volume v is centered. Expressions (305) and (306) introduced into the non-bias condition (304) lead to the k constraints:

$$\sum_{i=1}^{n} \lambda_i \; {}^{l}b_{v_i} = {}^{l}b_V \quad \text{for} \quad l = 1,2 \; ... \; k \qquad (308)$$

When the conditions (308) are fulfilled, the estimation variance contains only the covariance terms (without drift):

$$E\{[Z_V - Z_k^*]^2\} = \bar{C}(V;V) - 2 \sum_{i=1}^{n} \lambda_i \; \bar{C}(V_i v_i) + \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \; \bar{C}(v_i;v_j) \qquad (309)$$

For optimizing the Kriging estimator (303) the estimation variance (309) subject to k constraints (308) has to have a minimum value, which leads to the problem of the minimization of this variance in respect to the λ_i values in the presence of μ_l ($l = 1,2 \ldots k$) unknown Lagrange multipliers, i.e.:

$$\left.\begin{array}{l} \sum\limits_{i=1}^{n} \lambda_i \, \bar{C}(v_i;v_j) - \sum\limits_{l=1}^{n} \mu_l \, {}^{l}b_{v_j} = \bar{C}(v_j;V) \\ \qquad\qquad\qquad\qquad j = 1,2 \ldots n \\ \\ \sum\limits_{i=1}^{n} \lambda_i \, {}^{l}b_{v_i} = {}^{l}b_V \, ; l = 1,2 \ldots k \end{array}\right\} \tag{310}$$

is now the Kriging system given by n+k equations with the n+k unknown parameters λ_i and μ_l , with the Kriging variance

$$\sigma_k^2 = \bar{C}(V;V) + \sum_{l=1} \mu_l \cdot {}^{l}b_V - \sum_{i=1}^{n} \lambda_i \, \bar{C}(v_i;V) \tag{311}$$

Note please, that both the Kriging system (310) and the Kriging variance (311) can be presented in the matrix form (see [1] page 319 , ff).

When the measurements z_i are affected by the random errors ε with

$$E(\varepsilon_i) = 0 \quad \text{and} \quad D(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_i^2 \tag{312}$$

and these errors are non spatially correlated, i.e.

$$C(\varepsilon_i;\varepsilon_j) = 0$$

the system (310) becomes [9]:

$$\sum_{i=1}^{n} \lambda_i \, \bar{C}(v_i;v_j) + \sigma_j^2 \, \lambda_j - \sum_{l=1}^{k} \mu_i \cdot {}^{l}b_{v_j} = \bar{C}(v_j;V) \tag{314}$$

$$j = 1,2,\ldots n$$

and for the Kriging variance σ_k^2 one has an additional term

$$\sigma_\varepsilon^2 = \sum_{i=1}^{n} \lambda_i^2 \cdot \sigma_i^2 \, . \tag{315}$$

Kriging in the presence of a drift is not yet a common procedure introduced into the practice and there are even some doubts on its correctness [9]. It seems, however, that especially in small regions (in comparison with the drift variability), it can give very positive practical results [9, 10, 1, 2].

What was not said during the lectures

1. All possibilities given by the idea of the linear equivalents have not been explored entirely.
2. The nugget effect - its behaviour in regularization and Kriging.
3. Coregionalization and co-kriging.
4. Techniques of numerical calculus for auxiliary functions and Kriging.
5. Random kriging.
6. Simulation of spatially distributed data in geology.
7. Non-linear geostatistics, disjunctive kriging. [3, 11].
8. Statistical distributions of the regularized (or graded) data in the scope of non-linear geostatistics [3.11].

[11.] Y.C. KIM, D.E. MYERS, H.P. KNUDSEN: Advanced Geostatistics in Ore Reserve Estimation and Mine Planning. (Practitioner's Guide). GJBX-65(77) Grand Junction Bendix Field Engineering Corp. USA, October 1977.