3. STATISTICAL ANALYSIS

3.1. Statistics of the Conductivity data set.

The results of the measurements, after the mean Conductivity value for each point is calculated, form a data set of 80 values. The statistics of this data set are calculated, by running the computer program VARIO1. The program's modified version VARIO3 is situated in appendix A. The arithmetic mean was found to be 65.79 m/day, and the sample variance $1 \cdot 10^4 (m/day)^2$. The coefficient of skewness was found equal to 1.915, which means that the distribution is skewed to the right. The coefficient of kurtosis is found 6.73 which is much greater than 3, showing a leptokurtic distribution.

The very high value of the variance, shows the big variability of hydraulic conductivity in the studied area. The lowest value observed was $3 \cdot 10^{-4}$ m/day and the highest 488 m/day. This variability should be expected since observation points expand over an area of about 1 square kilometer.

From the above-mentioned moments of the data set, it can be concluded that the conductivity in the studied area is not normally distributed, which is to be further examined, since this information only gives a general estimate of the parameter.

Considering as parameter under study the logarithm of the conductivity (under base 10), we find the same moments for this parameter. The arithmetic mean is found equal to 0.491, with a sample variance of 3.062 (standard deviation 1.75).The skewness coefficient is -0.696 and the coefficient of kurtosis 2.34. Hence, the experimental distribution is found to be sligthly platukurtic and skewed to the left.

A lognormal distribution might fit the data. To get an idea about that, we plot the cummulative relative frequencies, expressed as percentages, against the observations (logK values), in a special probability graph paper (figure 10). The theoretical lognormal distribution, with the same mean and variance as the data, is also shown in the figure. The curve fits the data rather well. The above-mentioned graph, gives only an idea about the distribution, since no statistical test is possible to verify the fit, because the data are spatially correllated. The statistical test of Kolmogorov-Smirnov will be applied in a later chapter.

3.2. Comparison of results with previous studies.

As it was mentioned before, two different studies had taken place in the same region. The second one, timewise (Tan 1986), had a plot of 90x90 meters, situated near the South-East corner of the present study. It included observations of 100 points situated on the nodes of a grid with distances of 10 m.

The first one (Nurul 1984), had a plot of 14x14 meters, which was included in the above-mentioned plot. The samples were

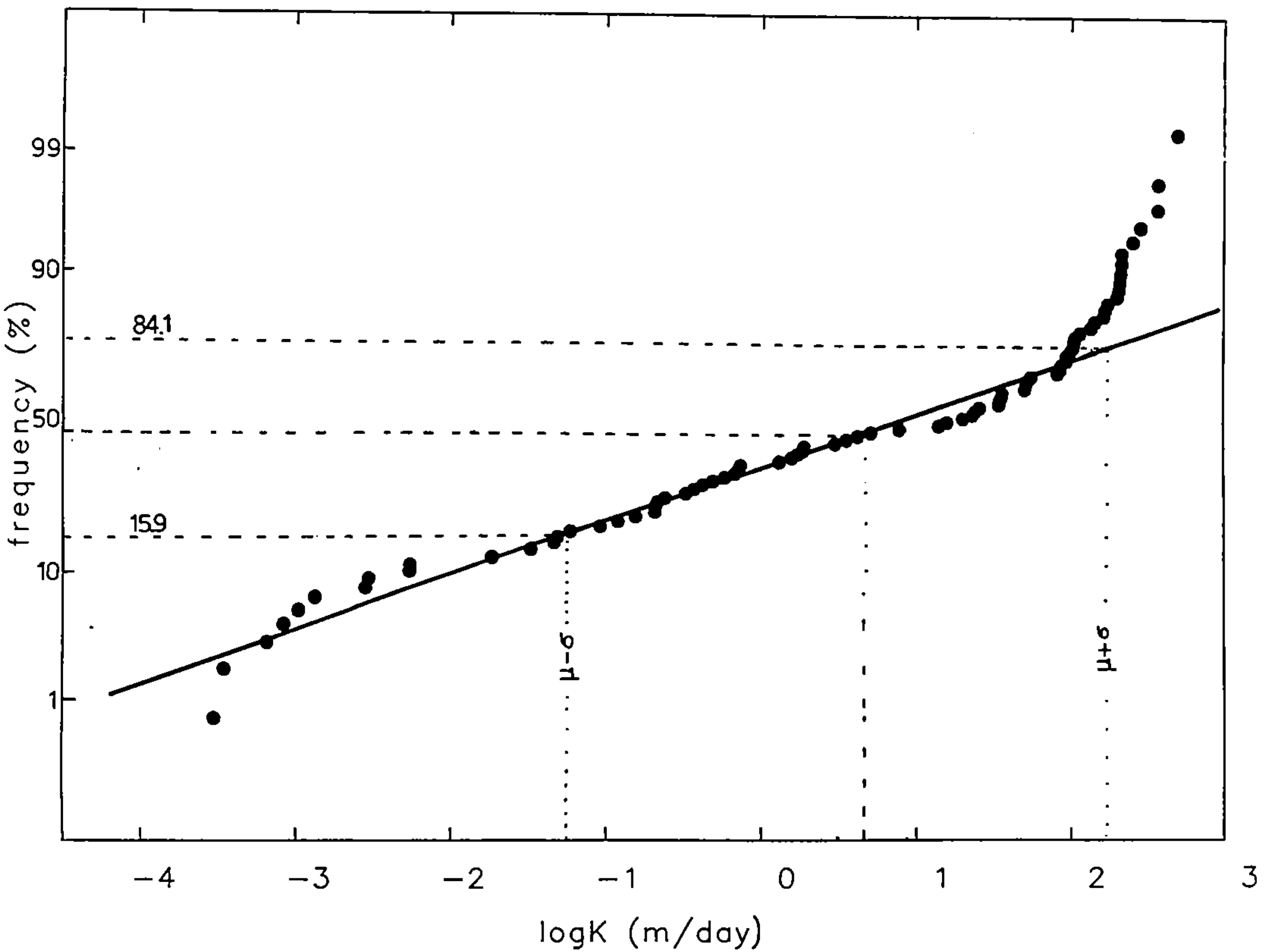


fig. 10. Cumulative probability distribution of logK values

taken at the nodes of a grid, every 2 meters.

In figure 11, the cumulative distributions of the three data sets are shown. One can observe there the gradual increase of the variance of the data set, with the increase of the area. The variance is proportional to the slope of the line of each data set, with respect to the normal distribution axis. The variance increase is better shown in figure 12, where the observations of each data set, have been brought around their mean.

The influence of the study area size on the variance was investigated by plotting the three variances versus the logarithms of the area (figure 13). In the figure it can be seen that a parabolic or exponential relationship appears to exist, unlike the linear relationship expected (A. G. Journel and Ch. J. Huijbregts, 1978).

The statistics of the three studies are compared and shown in table 8. The mean of the new data set, which corresponds to an area much larger than the other ones, is much bigger than the other means. The ratioes of the variances of the new data set with the "older" ones, are also very big.

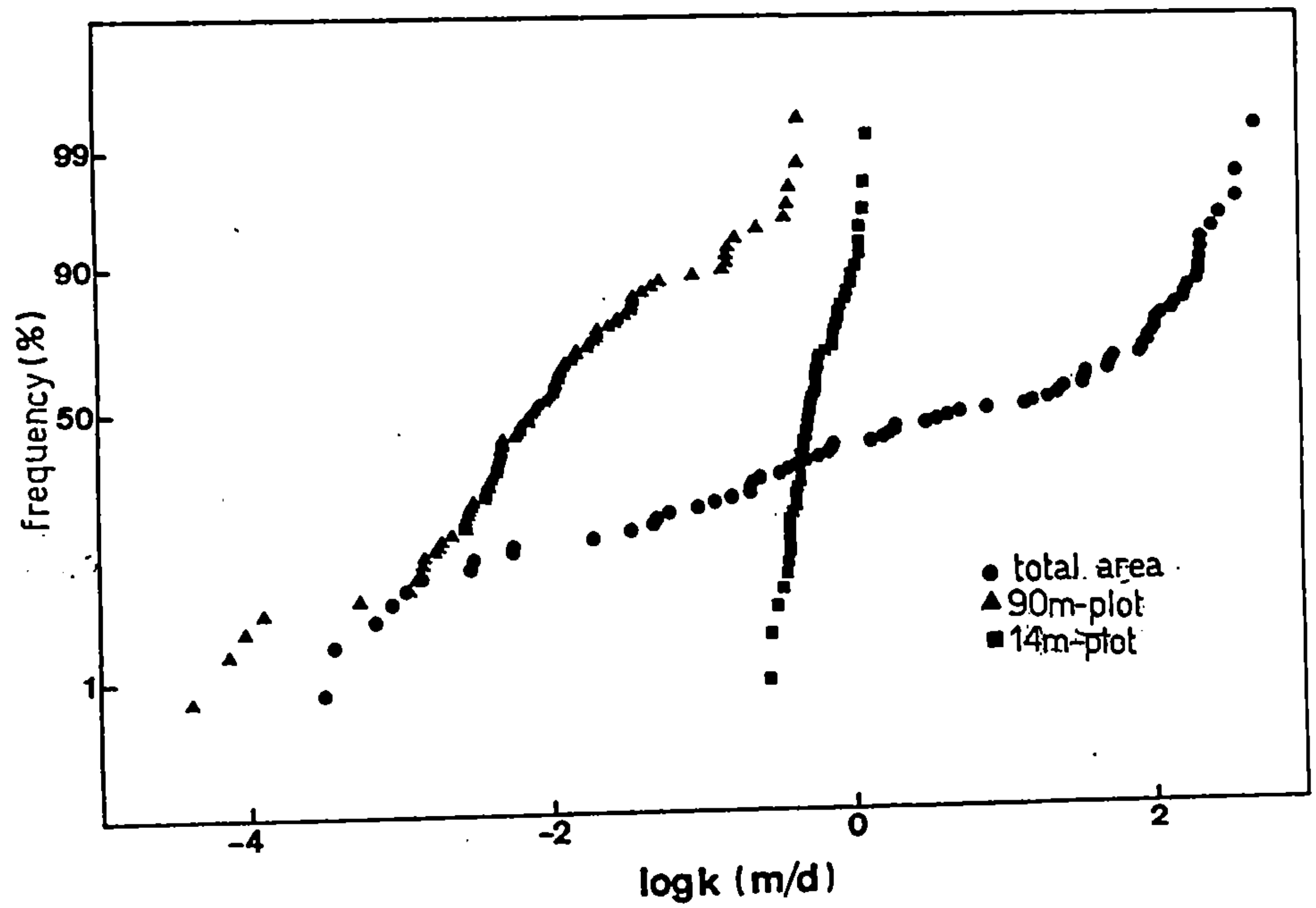


fig. 11. Cumulative distribution of the three studies

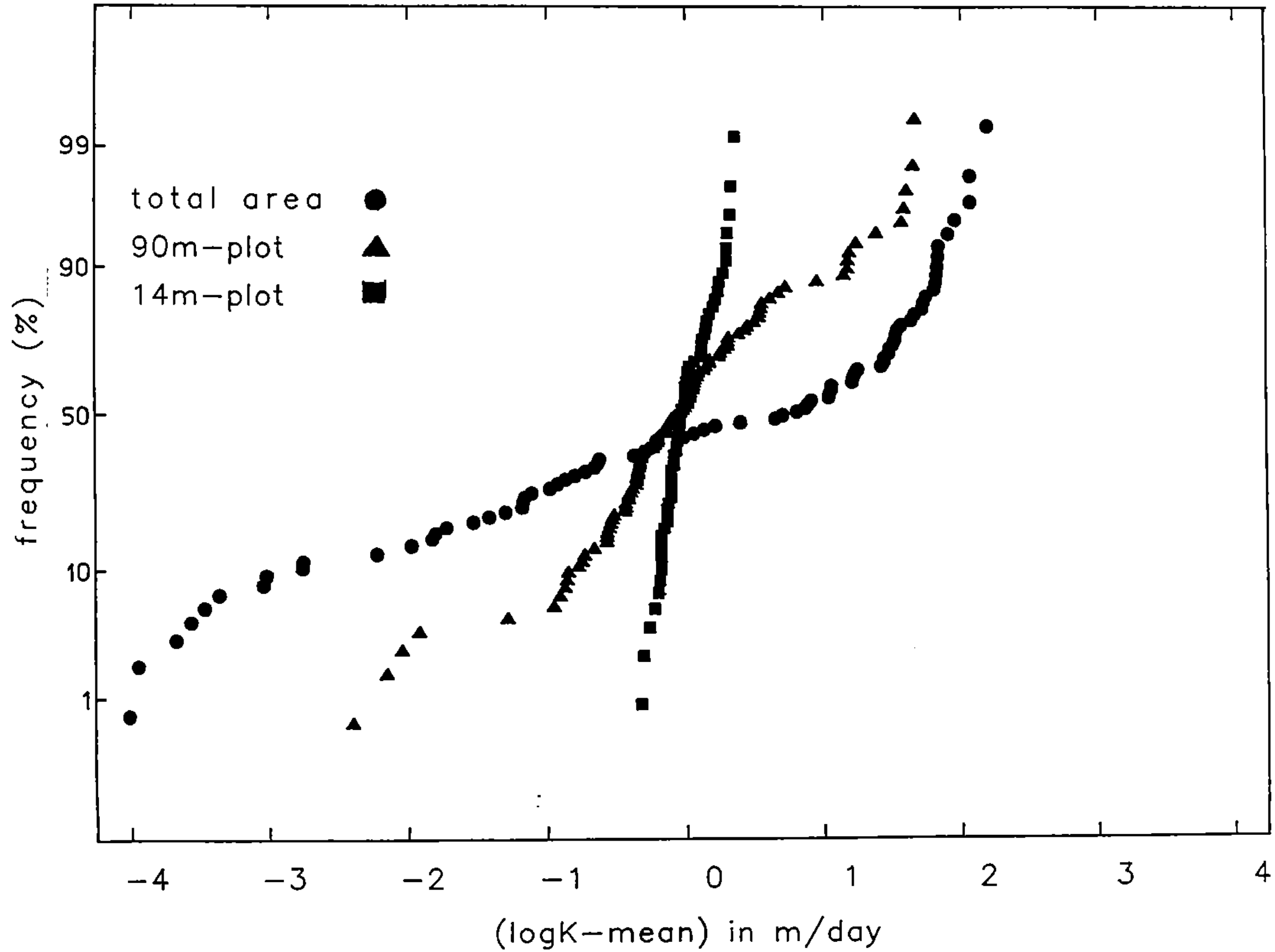
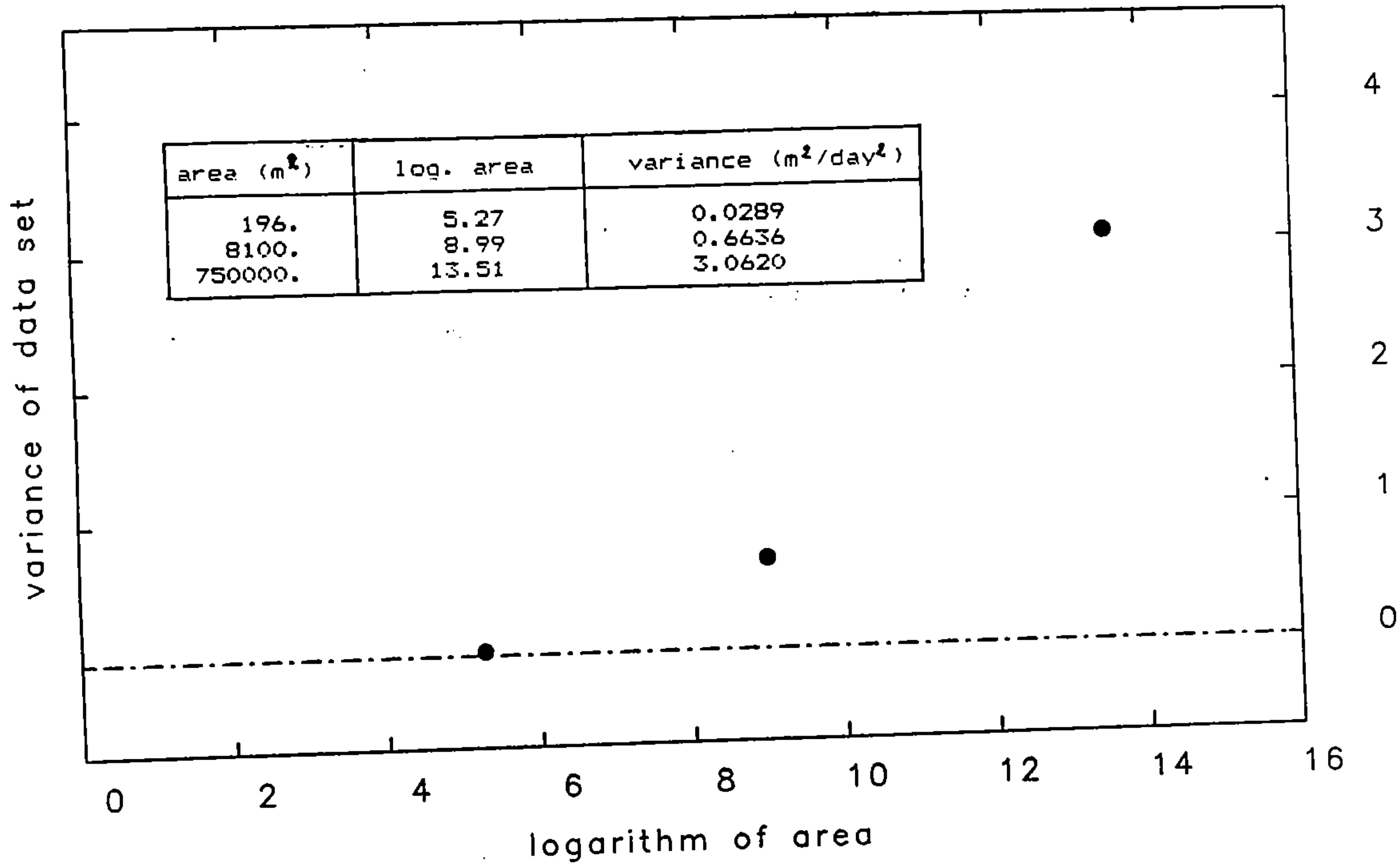


fig. 12. Probability distributions around the means



area (m^2)	log. area	variance (m^2/day^2)
196.	5.27	0.0289
8100.	8.99	0.6636
750000.	13.51	3.0620

fig. 13. Variance versus logarithm of area

Table 6. Comparison of the statistics with previous studies.

Data set :		Nurul	Tan	Present	Pre./Nur.	Pre./Tan
mean	1	0.6155	0.0404	65.788	106.9	1628.4
	2	-0.2451	-2.0070	0.491	-2.0	-0.2
variance	1	0.0688	0.0089	10090.000	146721.0	1133707.9
	2	0.0289	0.6636	3.062	106.1	4.6
st. deviat.	1	0.2622	0.0943	100.455	383.1	1065.3
	2	0.1700	0.8146	1.750	10.3	2.1
skewness	1	1.0480	3.2050	1.915	1.8	0.6
	2	0.4493	0.0589	-0.697	-1.6	-11.8
kurtosis	1	3.0300	12.6000	6.728	2.2	0.5
	2	2.2370	3.9750	2.339	1.0	0.6

Note : 1 stands for K values, 2 for logK values

4. SPATIAL VARIABILITY

4.1. Introduction

Spatial variability analysis is the study of the differences that might exist between the value of a certain variable at a point with another at some distance in the same field area. Geostatistics was introduced in the middle of this century for mining purposes and geological studies. More recently, the technique has been applied in water resourses problems (Delfiner and Delhomme, 1973; Delhomme, 1978). Since then, many papers have been published about the study of spatial variability by the Geostatistical method applied to hydrogeology (Byers and Stephens, 1983; De Marsily, 1984; Gutjahr and Gelhar, 1981; Virdee and Kottegoda, 1984).

4.2. Geostatistics.

Geostatistics can be applied in the study of any phenomenon which can be characterised as a "regionalised phenomenon". It is called as such, a phenomenon that spreads out in space and shows certain structure. A variable which characterises such a phenomenon is termed as a regionalised variable (ReV). In fact all variables that describe properties of the subsurface or the atmosphere, may be considered as regionalised variables.

From both conceptual and practical stand points, it is more convenient to deal with regionalised variables by applying the probabilistic theory of random functions (RF). But it is necessary to reconstitute the distribution law of this random function from the available data. The problem that arises here is that many regionalised variables have a unique existence, that is, one realisation. It is possible to compute the structure based only on this single outcome.

4.3. Hypotheses of stationarity and intrinsic.

Due to the above-mentioned problem, it is necessary to impose further hypotheses about the random function, so that it could be overcome.

The first hypothesis which is usually used in the theory of random functions, is the hypothesis of stationarity. This means that the expectation of the random function Z(x), is constant in space:

$$E(Z(x))=m \qquad (4.1)$$

x being the location vector.

The covariance depends only on the separation vector h:

$$cov(Z(x+h)+Z(x))=E(Z(x+h)-m)(Z(x)-m)=C(h) \qquad (4.2)$$

The concept of ergodicity implies that the unique realisation will behave in space with the same probability density function. In other words, by observing the variation in space of the property, it is possible to determine the probability density function of the random function for all realisations, which is termed as statistical inference.

In many natural phenomena, a finite variance does not exist. This leads to the introduction of the intrinsic hypothesis: Any increment of $(Z(x+h)-Z(x))$, has a finite variance which is independent of x:

$$E(Z(x+h)-Z(x))=0 \qquad (4.3)$$

$$var(Z(x+h)-Z(x))=2\gamma(h) \qquad (4.4)$$

Equation (4.4) defines the variogram, which is a structural function of the parameter under study, and will be examined in the following in a higher extend.

Another hypothesis, the one of quasi-stationarity, is also necessary. It comes out of the fact that the mean, covariance and variogram are, in practise, functions of the size of the investigated area, as well as of the area itself. The structural function, covariance or variogram, is only used for limited distances $|h|\leqslant b$. The limit b can represent for example, the diameter of the neighbourhood of estimation, or, in other cases, the extend of an homogeneous zone. It is allowed for every study area separately to use the locally found mean, covariance and variogram. However, they have no effect outside the area for which they have been calculated.

4.4. Variograms.

The variogram (h) is defined from the equation:

$$2\gamma(h)=E(Z(x+h)-Z(x)) \qquad (4.5)$$

It gives the mean squared difference in value for all pairs of measurements separated by a distance h. The term semivariogram is usually abreviated as "variogram", which may cause confusion to the readers. In the present paper, the term "variogram" will be used, unless otherwise stated.

The variogram considered, refers to point variables. Such a point variogram can be estimated from point measurements Z(x), as follows:

$$\gamma(h)=(1/(2*N(h))\sum_{i}^{N(h)}(Z(x+h)-Z(x))^2 \qquad (4.6)$$

where N(h) is the number of pairs of samples separated by the vector h.

It can be proved that, if the regionalised variable is stationary, the relation between variance and covariance is:

$$\gamma(h)=C(0)-C(h) \qquad (4.7)$$

where C(0) is the covariance at the point itself, and

C(h) is the covariance between points with distance h.

Variograms exhibit certain characteristics with the variation of distances. When the variogram has an increasing function, it starts from the origin, and increases until a certain value C_1, which is the sill and is equal to the sample variance. The distance at which variograms reach the sill, is called the range a. In physical sence, it expresses the fact that beyond this range, the samples are not correlated.

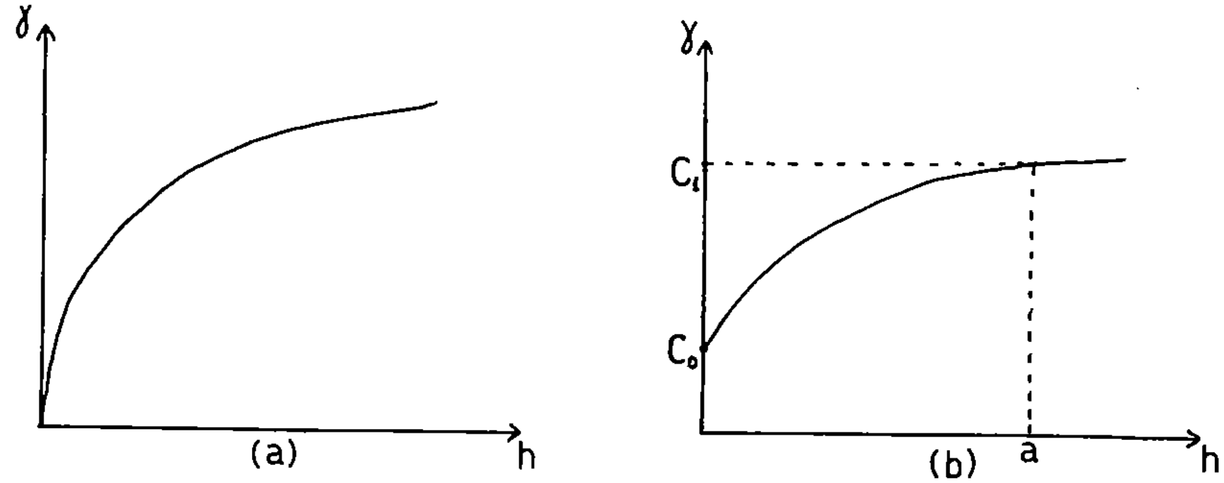


fig. 14. Variogram with sill, range, and nugget

In most cases, the variograms show a discontinuity close to the origin of the axes, which can be mathematically expressed by:

$$\lim_{h \to 0}(\gamma(h)) = C_0 \qquad (4.8)$$

C_0 is called the nugget effect. It represents the non-structural variability which can be spatial, or due to measurement errors.

The computation of a variogram can be carried out in one particular orientation, e.g. in a North-South grid, or in an East-West grid. When the variograms obtained from different orientations are equal, this means that we have isotropic properties. If that is not the case, we have anisotropism, that is, different properties in different directions.

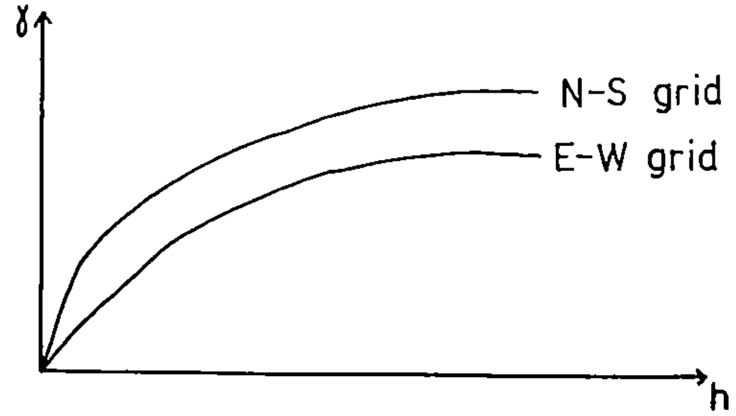


fig. 15. Variograms in different directions showing anisotr.

4.5. Variogram models.

It is usual practise to fix a mathematical description (model) to the variogram. The usual method for obtaining a mathematical description is to choose one of the proposed functional forms, and then calibrate it by using a statistical method, for instance, the least squares method.

The most commonly used models for variograms are:

Linear $\gamma(h)=C_0+ah$ for $h<b$ (4.9a)

$=C_0+ab$ for $h>b$ (4.9b)

Spherical $\gamma(h)=C_0+a(\frac{3h}{2b}-\frac{h^3}{2b^3})$ for $h<b$ (4.10a)

$=C_0+a$ for $h>b$ (4.10b)

Power $\gamma(h)=C_0+ah^b$ (4.11)

Logarithmic $\gamma(h)=C_0+a\log(1+bh)$ (4.12)

Exponential $\gamma(h)=C_0+a(1-\exp(-bh))$ (4.13)

Gaussian $\gamma(h)=C_0+a(1-\exp(-bh^2))$ (4.14)

4.6. Computation of a variogram.

Each sample is considered as a point in the field. Equation (4.6) is used. Hence, the contribution of the pair of points x1 and x2 which are in a distance h apart, is:

$$\gamma(h)=(1/2)(Z(x_1)-Z(x_2))^2 \qquad (4.15)$$

As the computations for a set of data can be long and tedious, it is always easier to employ a computer for it. The program that was used here, is adopted from David, 1978 (refer to the appendices).

Algorithm of the program.

The main step is to sort all pairs of points available to a particular class, with respect to direction and distance. The classification pattern for a given direction, is as shown in figure 16.

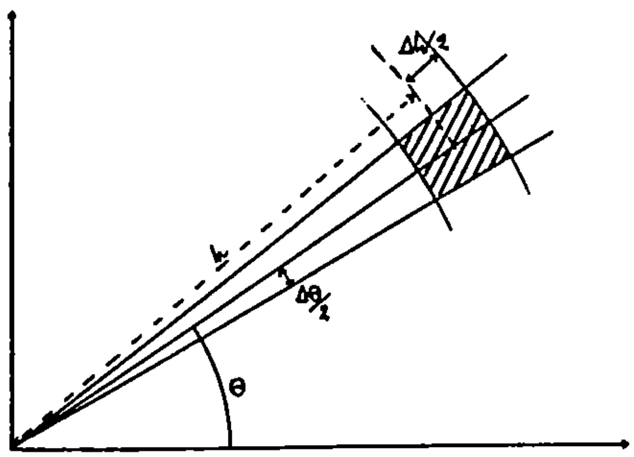


fig. 16. Classification pattern for the variogram computation

In the program, the direction θ is refered to as PHI and $\Delta\theta$ is refered to as PSI (the angular regularisation). Directional classification of the line joining two sample points x_1 and x_2, is done by computing the scalar product

$$s=(\overrightarrow{x1x2})/|\overrightarrow{x1x2}|*\vec{u} \qquad (4.16)$$

where $\vec{u}$ is the unit vector of the direction selected.

The distance $x_1 \rightarrow x_2$ is classified according to Δh and its combination.

$$C_{x_1x_2}=(Z(x_1)-Z(x_2))^2 \qquad (4.17)$$

When all pairs have been tested and grouped against a certain direction class, the smoothed or average variogram along the direction, is expressed by:

$$\gamma(h,\theta)=(1/2n_i)\sum_{n_i}(Cx_1x_2) \qquad (4.18)$$

where $h=\sum_{n}|x_1x_2|/n_i$, and n_i=number of pairs grouped in the same class (i).

What is finally obtained is a set of one-dimensional smoothed variogram values along a selected direction. The calculation of the drift:

$$D(h_i,\theta)=(1/n_i)\,(\Sigma(x_k)-Z(x_l)) \qquad (4.19)$$

enables us to detect the presence of trends.

The procedure of using the program VARIO1 is given in Appendix A.

4.7. Results.

The variograms of the K values data set and of the logK values data set were calculated. Originally, orientation PHI is taken 45 degrees and the angular regularisation PSI equal to 180 degrees, in order to have all possible pairs.

The maximum distance until which the variograms are calculated, is 500 m. Intervals of 100 meters were used. The results of the computations are situated in tables 9 and 10, and the respective plotts in figures 17 and 18.

The computations were done by using the FORTRAN program VARIO1 (M. David, 1977), with some alterations. The output includes some statistical parameters, and, for each class, the number of pairs, the drift and variogram values, the average pair distance, and the "maxvar pair", that is, the pair with the maximum contribution to the variogram value of the specific class. The results will be further discussed and explained, after examining the isotropy of the parameter.

4.7.1. Directional analysis.

In order to check whether the parameter (logK) has the same statistical properties in different directions, we calculate the variograms for the directions of 0, 45 and 90 degrees, with angular regularisation of 5 and 20 degrees. The results can be seen in tables 11 and 12, and their graphs in figures 19 and 20. In the first graph (5°), the variograms are very irregular for table 9. The variogram of the K data set

1

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

K

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .4883E+03
GENERAL MEAN OF Z = .6579E+02
GENERAL VARIANCE OF Z = .1009E+05
GENERAL SKEWNESS OF Z = .1915E+01
GENERAL KURTOSIS OF Z = .6728E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE	MAXVAR PAIR
0 —— 100	120	.505E+01	.9001E+04	76.4	73—— 74
100 —— 200	273	-.929E+01	.7892E+04	154.4	36—— 73
200 —— 300	401	.215E+01	.7848E+04	252.2	4—— 73
300 —— 400	448	.862E+00	.6648E+04	349.9	6—— 73
400 —— 500	461	-.151E+02	.6914E+04	448.4	7—— 73
500 —— 600	426	-.372E+01	.7082E+04	548.2	53—— 73
600 —— 700	385	.681E+01	.8621E+04	650.2	55—— 73
700 —— 800	276	.555E+01	.8870E+04	745.9	26—— 73
800 —— 900	171	-.172E+02	.9827E+04	844.8	46—— 73
900 ——1000	91	-.408E+01	.1082E+05	943.7	65—— 73
1000 ——1100	29	-.625E+02	.1196E+05	1037.1	42—— 47

99.99

14.00.08.UCLP, 5B, DEFTERM, 0.128KLNS.

table 10. The variogram of the logK data set

1

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE	MAXVAR PAIR
0 —— 100	120	-.684E-01	.2575E+01	76.4	69—— 79
100 —— 200	273	-.852E-01	.3570E+01	154.4	6—— 59
200 —— 300	401	-.624E-01	.2967E+01	252.2	6—— 71
300 —— 400	448	-.388E-01	.3360E+01	349.9	6—— 73
400 —— 500	461	-.271E+00	.3429E+01	448.4	55—— 58
500 —— 600	426	-.292E+00	.3087E+01	548.2	55—— 71
600 —— 700	385	-.246E+00	.2860E+01	650.2	55—— 73
700 —— 800	276	-.209E+00	.2776E+01	745.9	26—— 73
800 —— 900	171	-.509E+00	.2406E+01	844.8	26—— 63
900 ——1000	91	-.383E+00	.1865E+01	943.7	42—— 46
1000 ——1100	29	-.652E+00	.1309E+01	1037.1	63—— 65

99.99

14.01.13.UCLP, 5B, OEFTERM, 0.128KLNS.

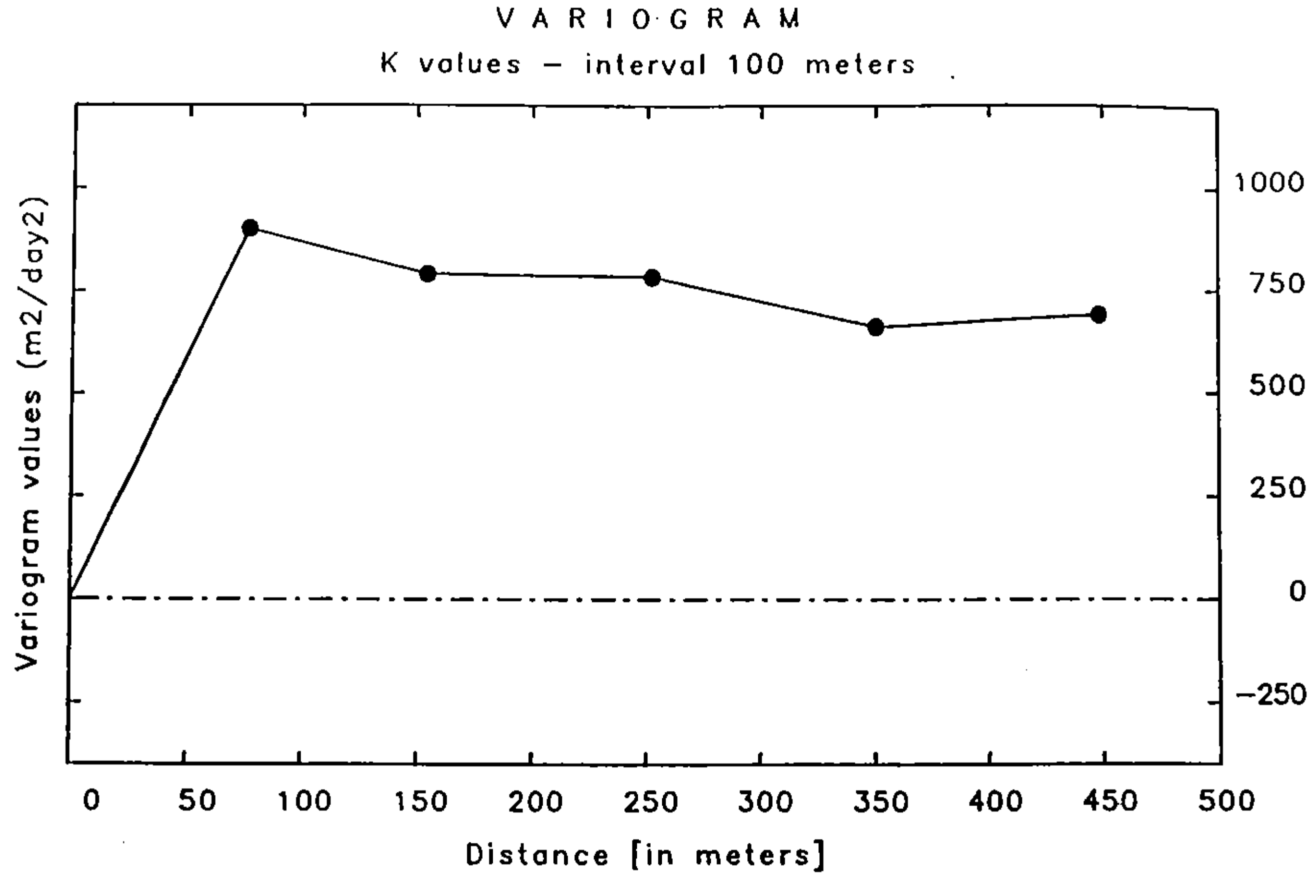


fig. 17. Variogram of the K data set

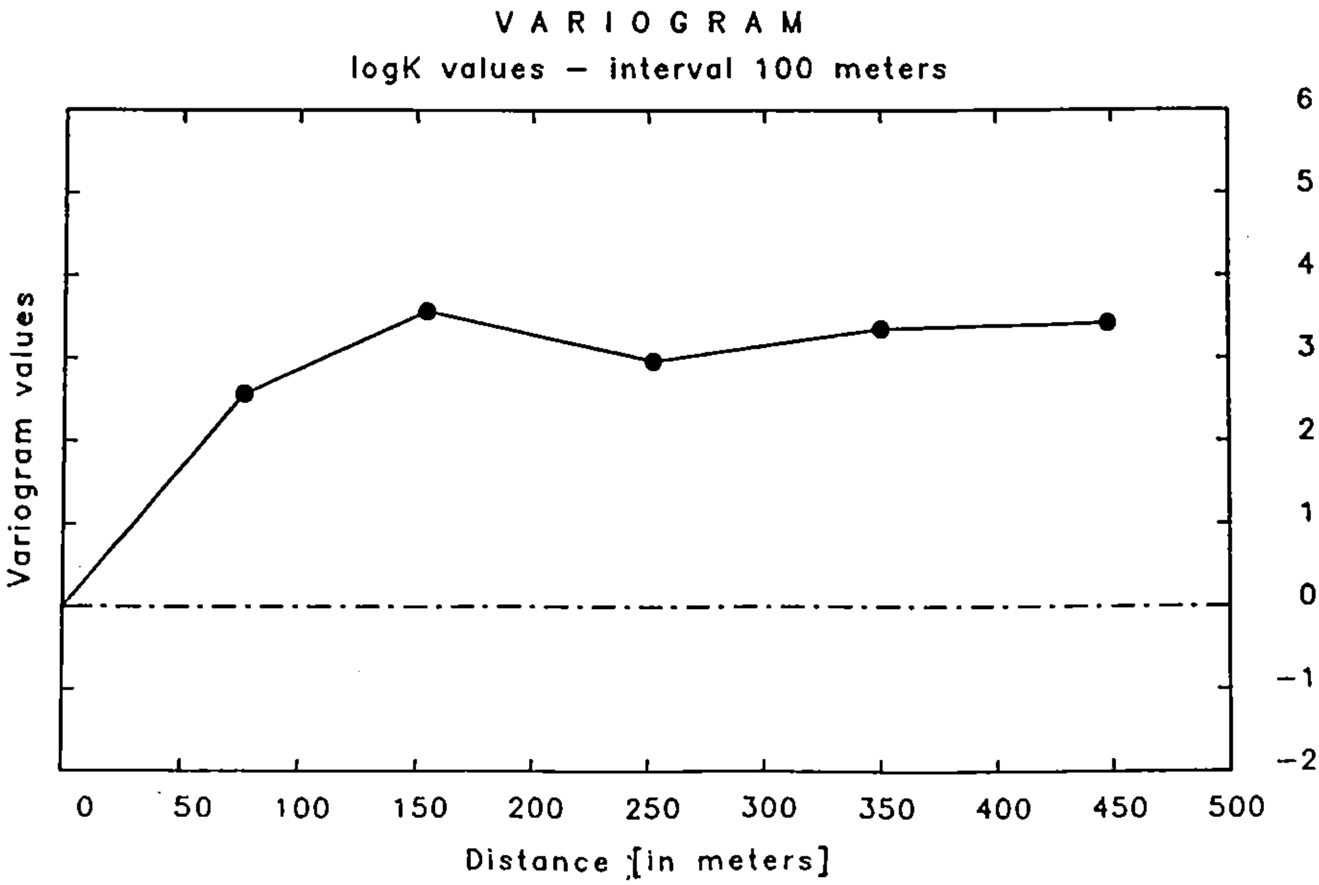


fig. 18. Variogram of the logK data set

table 11. Directional analysis (PSI=5)

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 5. DEGREES IN EACH DIRECTION)

LOOK: 0.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	7	-.108E+01	.4061E+01	67.0
100 ---- 200	9	-.660E+00	.2519E+01	158.7
200 ---- 300	17	.810E+00	.4931E+01	267.1
300 ---- 400	17	-.469E+00	.3667E+01	356.0
400 ---- 500	20	-.234E+00	.2452E+01	440.6
500 ---- 600	16	-.179E+00	.4594E+01	542.7
600 ---- 700	18	.479E+00	.2503E+01	644.2
700 ---- 800	12	-.807E-01	.1928E+01	748.8
800 ---- 900	8	-.134E+00	.1003E+01	857.0
900 ----1000	3	-.565E+00	.1955E+01	933.3

99.99

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 5. DEGREES IN EACH DIRECTION)

LOOK: 45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	1	.130E+01	.8513E+00	61.2
100 ---- 200	10	-.733E+00	.3641E+01	162.1
200 ---- 300	5	.256E+00	.2356E+01	256.2
300 ---- 400	14	.106E+00	.3262E+01	351.2
400 ---- 500	9	.428E+00	.4298E+01	456.8
500 ---- 600	12	.634E+00	.3401E+01	559.1
600 ---- 700	8	-.807E+00	.2573E+01	640.0
700 ---- 800	7	-.883E+00	.5083E+01	735.8
800 ---- 900	9	.691E+00	.2231E+01	838.6
900 ----1000	3	-.207E+00	.9569E+00	931.1
1000 ----1100	1	.307E+00	.4698E-01	1020.8

99.99

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 5. DEGREES IN EACH DIRECTION)

LOOK: 90.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	1	-.319E+01	.5084E+01	94.6
100 ---- 200	9	.577E+00	.9612E+00	175.8
200 ---- 300	15	.879E-01	.2456E+01	255.0
300 ---- 400	8	.330E+00	.6192E+01	362.5
400 ---- 500	10	-.183E+00	.1133E+01	451.3
500 ---- 600	11	.786E+00	.3752E+01	538.5
600 ---- 700	10	-.151E+01	.5834E+01	642.7
700 ---- 800	2	.287E+00	.1518E+01	728.0

99.99

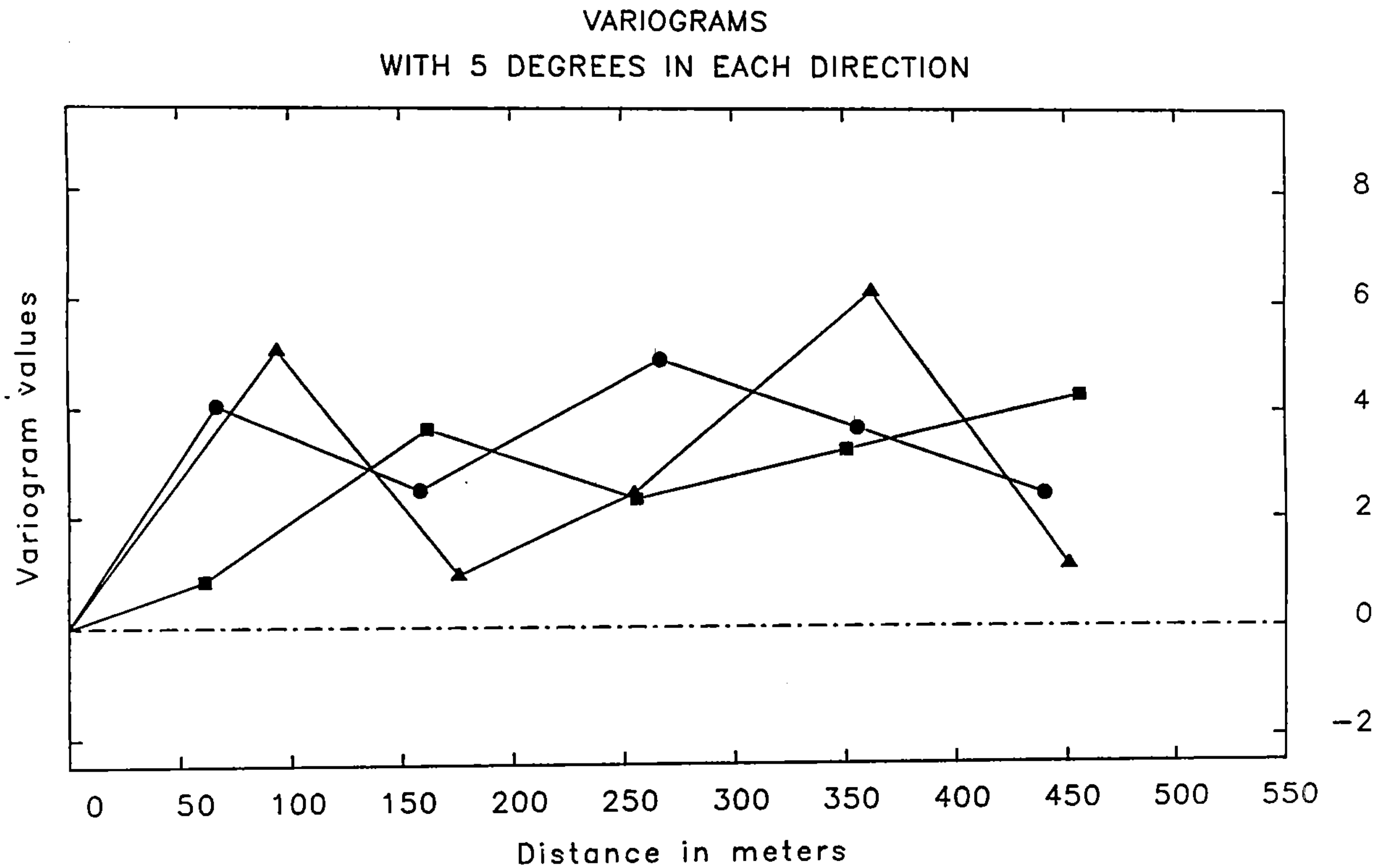


fig. 19. Variograms for different directions (PSI=5)

table 12. Directional analysis (PSI=20)

VARIOGRAM

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 20 DEGREES IN EACH DIRECTION)

LOCK

0.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	20	-.866E-01	.3586E+01	77.1
100 ---- 200	45	.508E+00	.4371E+01	160.6
200 ---- 300	58	.443E+00	.3761E+01	258.6
300 ---- 400	65	-.154E+00	.4080E+01	354.6
400 ---- 500	64	-.268E+00	.3410E+01	451.0
500 ---- 600	59	-.494E+00	.3708E+01	545.8
600 ---- 700	59	.868E-01	.1886E+01	653.2
700 ---- 800	40	.198E+00	.1886E+01	748.9
800 ---- 900	32	-.859E+00	.2501E+01	850.5
900 ----1000	13	.301E+00	.2886E+01	937.0
1000 ----1100	1	.184E+01	.1691E+01	1005.3

99.99

VARIOGRAM

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 20. DEGREES IN EACH DIRECTION)

LOCK

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	8	.924E+00	.1970E+01	65.2
100 ---- 200	31	-.181E+00	.2666E+01	159.9
200 ---- 300	43	-.300E+00	.1992E+01	253.5
300 ---- 400	59	.465E+00	.3213E+01	352.1
400 ---- 500	56	.935E-01	.4164E+01	449.0
500 ---- 600	59	.225E+00	.3064E+01	550.6
600 ---- 700	39	-.178E+00	.3177E+01	652.6
700 ---- 800	35	-.587E+00	.5118E+01	745.2
800 ---- 900	30	.141E+00	.2540E+01	841.4
900 ----1000	18	-.847E+00	.1393E+01	937.0
1000 ----1100	5	-.116E+01	.1194E+01	1040.8

99.99

VARIOGRAM

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 20. DEGREES IN EACH DIRECTION)

LOCK

90.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6965E+00
GENERAL KURTOSIS OF Z = .2339E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
0 ---- 100	11	-.325E+00	.1572E+01	85.2
100 ---- 200	24	-.129E+00	.2970E+01	153.7
200 ---- 300	48	-.118E+00	.2985E+01	253.6
300 ---- 400	51	.161E-01	.3928E+01	354.8
400 ---- 500	41	-.268E+00	.4270E+01	451.9
500 ---- 600	48	-.334E-01	.3299E+01	548.5
600 ---- 700	37	-.834E+00	.4114E+01	648.6
700 ---- 800	10	-.855E+00	.1519E+01	730.3

99.99

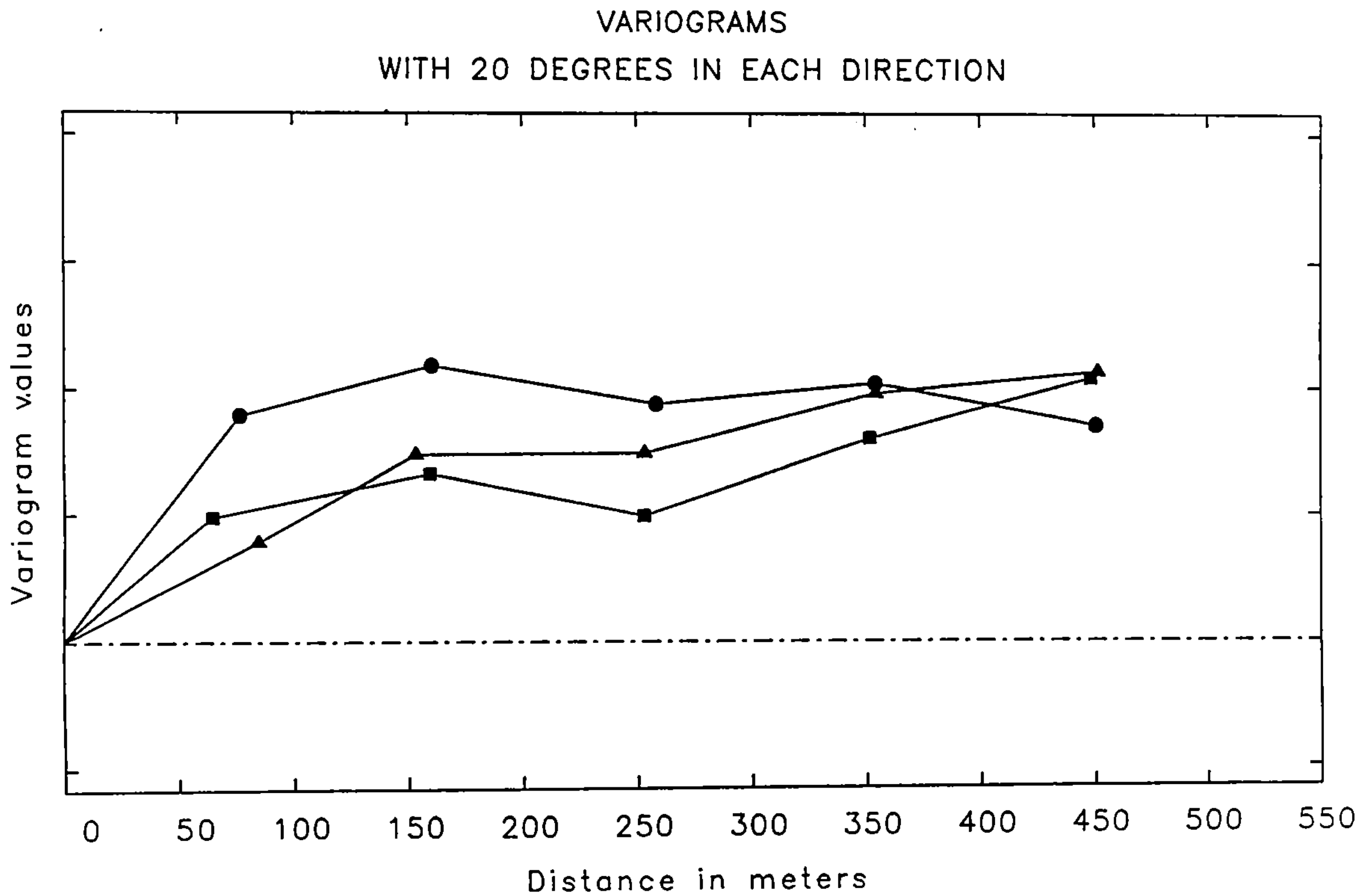


fig. 20. Variograms for different directions (PSI=20)

every direction. This is due to the fact that the number of observations is extremely low. In the second graph, which represents more observations, there are less irregularities, and there does not seem to appear any influence of the direction. Thus we may say that the parameter (logK) shows isotropic properties. This was found to be true in the previous studies too.

4.7.2. Search for observational errors.

Through the "maxvar" option of the program VARIO1, a search for observational errors was tried. The idea was that, if a point appeared at the same time at the "maxvar pair" of many classes, it could carry an observational error in it.

In the variogram of logK values (table 10), we observe that the points 6,55,73 appear three times each in the maxvar pairs. The above-mentioned points, happen to have very low or very high values of the parameter, which could alone explain the fact that they appeared at the maxvar pairs. Nevertheless, we construct the variogram of the remaining observations (table 13). Its plott is reffered to as "second variogram" in figure 21, where it can be seen that the variogram, having not changed in shape, just shifted towards the distance axis, showing lower sill or sample variance.

We then calculate a "third" variogram (table 14), after having excluded the point number 25, which appears 4 times in the table 13. Variogram excluding points 6,55,73

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .5678E+00
GENERAL VARIANCE OF Z = .2703E+01
GENERAL SKEWNESS OF Z = -.6675E+00
GENERAL KURTOSIS OF Z = .2304E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE	MAXVAR PAIR
0 —— 100	111	-.116E+00	.2446E+01	76.6	67—— 76
100 —— 200	249	-.108E+00	.3071E+01	135.4	67—— 73
200 —— 300	371	-.234E+00	.2614E+01	251.9	4—— 73
300 —— 400	403	-.185E+00	.2711E+01	349.8	4—— 40
400 —— 500	417	-.341E+00	.2976E+01	447.9	5—— 41
500 —— 600	396	-.328E+00	.2704E+01	548.8	25—— 41
600 —— 700	362	-.423E+00	.2574E+01	650.9	25—— 69
700 —— 800	262	-.361E+00	.2652E+01	745.6	25—— 73
800 —— 900	166	-.575E+00	.2409E+01	845.0	25—— 61
900 ——1000	84	-.529E+00	.1896E+01	943.8	41—— 45
1000 ——1100	29	-.652E+00	.1309E+01	1037.1	61—— 63

99.99

table 14. Variogram excluding points 6,26,55,73

VARIOGRAM

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .6130E+00
GENERAL VARIANCE OF Z = .2380E+01
GENERAL SKEWNESS OF Z = -.6798E+00
GENERAL KURTOSIS OF Z = .2359E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE	MAXVAR PAIR
0 —— 100	108	-.145E+00	.2351E+01	76.5	66—— 73
100 —— 200	245	-.103E+00	.2985E+01	155.3	66—— 72
200 —— 300	369	-.255E+00	.2586E+01	251.9	4—— 72
300 —— 400	393	-.230E+00	.2532E+01	349.3	4—— 39
400 —— 500	402	-.449E+00	.2804E+01	447.6	5—— 40
500 —— 600	379	-.480E+00	.2470E+01	548.7	53—— 68
600 —— 700	350	-.522E+00	.2514E+01	650.9	40—— 53
700 —— 800	255	-.465E+00	.2527E+01	745.7	53—— 72
800 —— 900	161	-.693E+00	.2311E+01	844.7	53—— 60
900 ——1000	84	-.529E+00	.1896E+01	943.8	40—— 44
1000 ——1100	29	-.652E+00	.1309E+01	1037.1	60—— 62

99.99

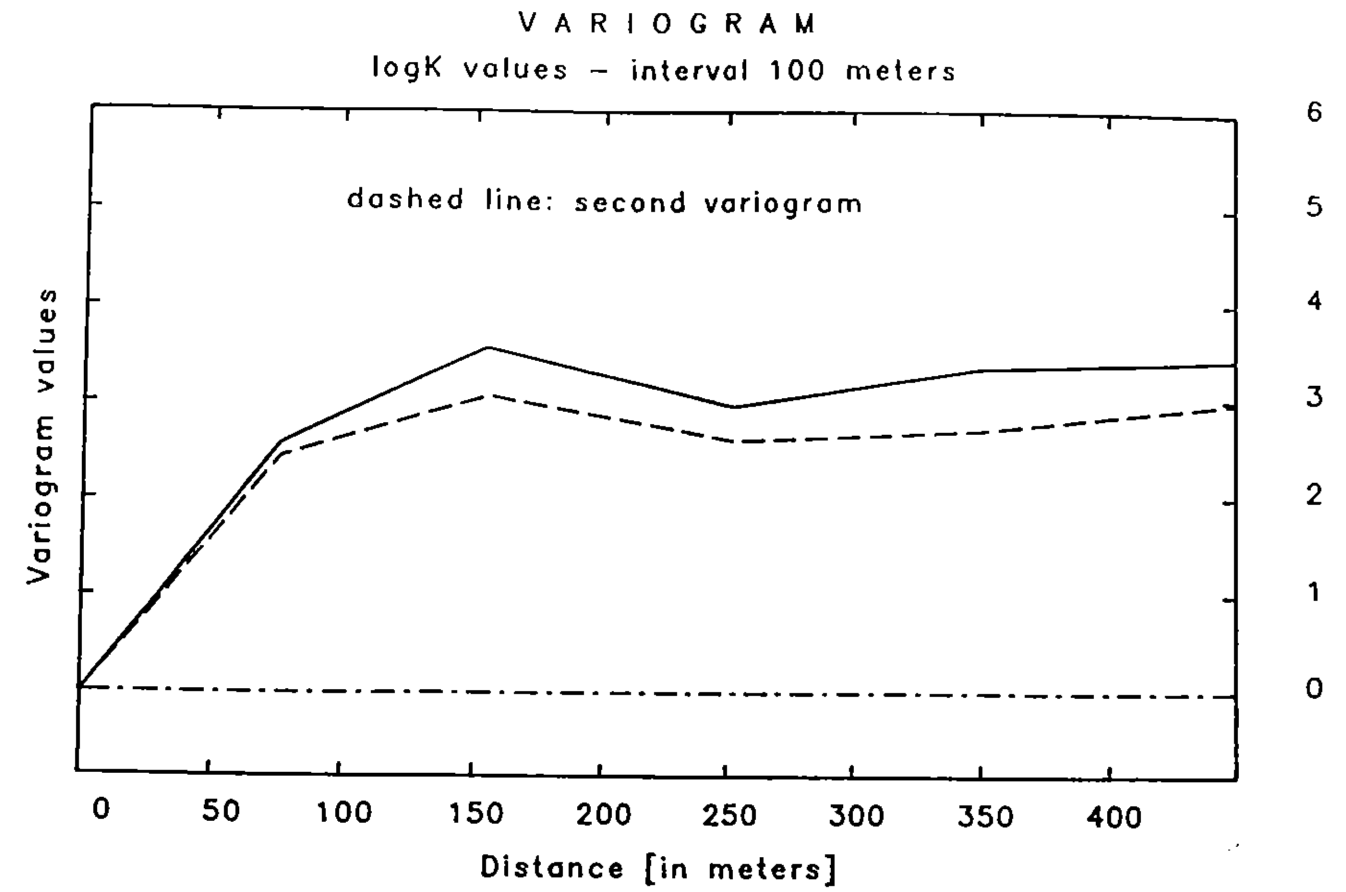


fig. 21. Variogram without observations 6,55,73

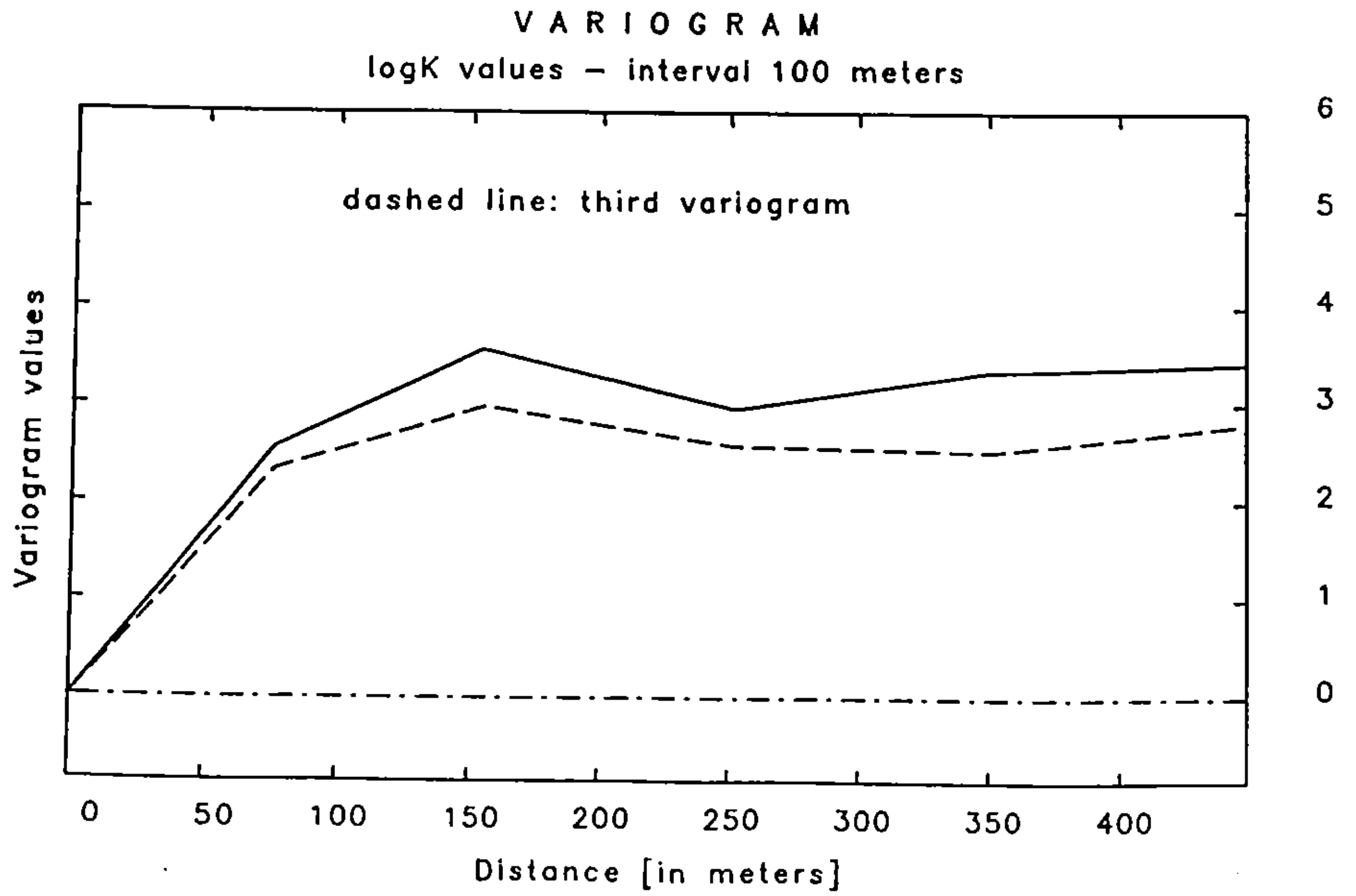


fig. 22. Variogram without observations 6,26,55,73

maxvar pairs of the second one. The resulting plott (figure 22), is very similar to the second one, showing that no observational error was detected.

4.7.3. Conclusion.

Considering the two previous conclusions, that is, isotropism and no observational error detection, we interprete the variograms in figures 17 and 18. Both of them, show a pure nugget effect, which can be explained by the fact that, variations in the data exist at a scale smaller than the sampling distances.

4.7.4. Comparison of the variogram with previous studies.

The variogram of the present study (logK values), was plotted together with the variograms of the two previous studies (figure 23). There appear to be a lot of differences in the structure of the three variograms. The effect of the plot size, is responsible for these differences.

4.7.5. Total variogram.

The coordinates of the points of observation of the two previous studies, were transformed into the system of the present one. Two points of the study of Tan were dropped, since they corresponded to 0 conductivity values. The rest, a total of 242 points, are shown in figure 24.

The total variogram is then calculated, for the 242 points (table 15). The results are plotted in fig. 25. Apart from the points, the line of the proposed model for the variogram is plotted there. It is a model with a sill of C_1 =3.79 and a range of a=142.9m. Since there are no points between, the model is increasing linearly until the point (142.9,3.79). This model will be used in the application of the kriging estimation technique in chapter 6.

4.7.6. One-point variograms of each study.

Considering the points of each study separately, as belonging to the same class, we calculate one variogram value for each study. The results are situated in table 16, and plotted in fig. 26. No model variogram can be fitted to the three points, since the variogram value of the first study seemes relatively very small with respect to the average distance.

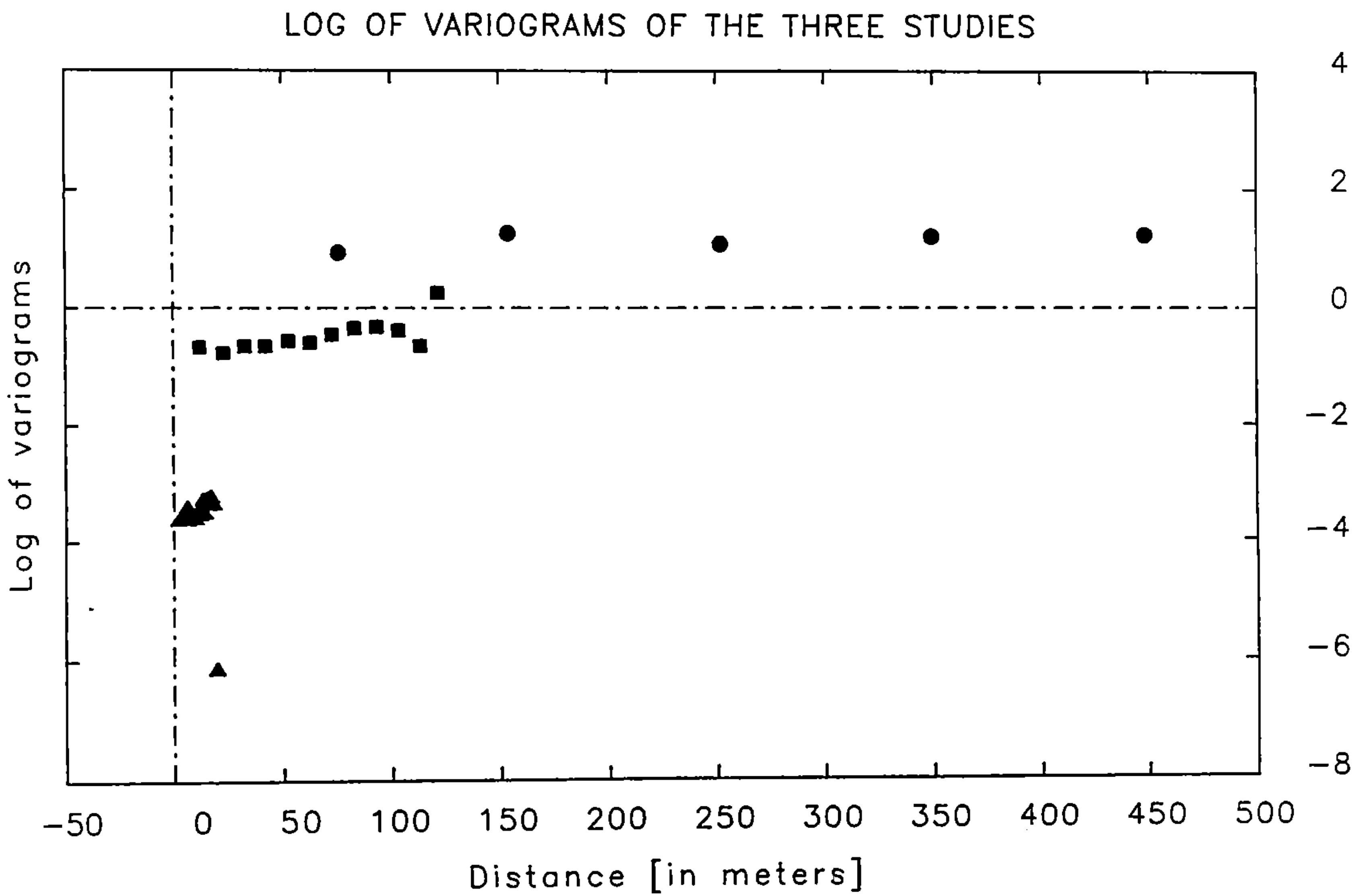


fig. 23. Log of variograms of the three studies

scale 1:5000

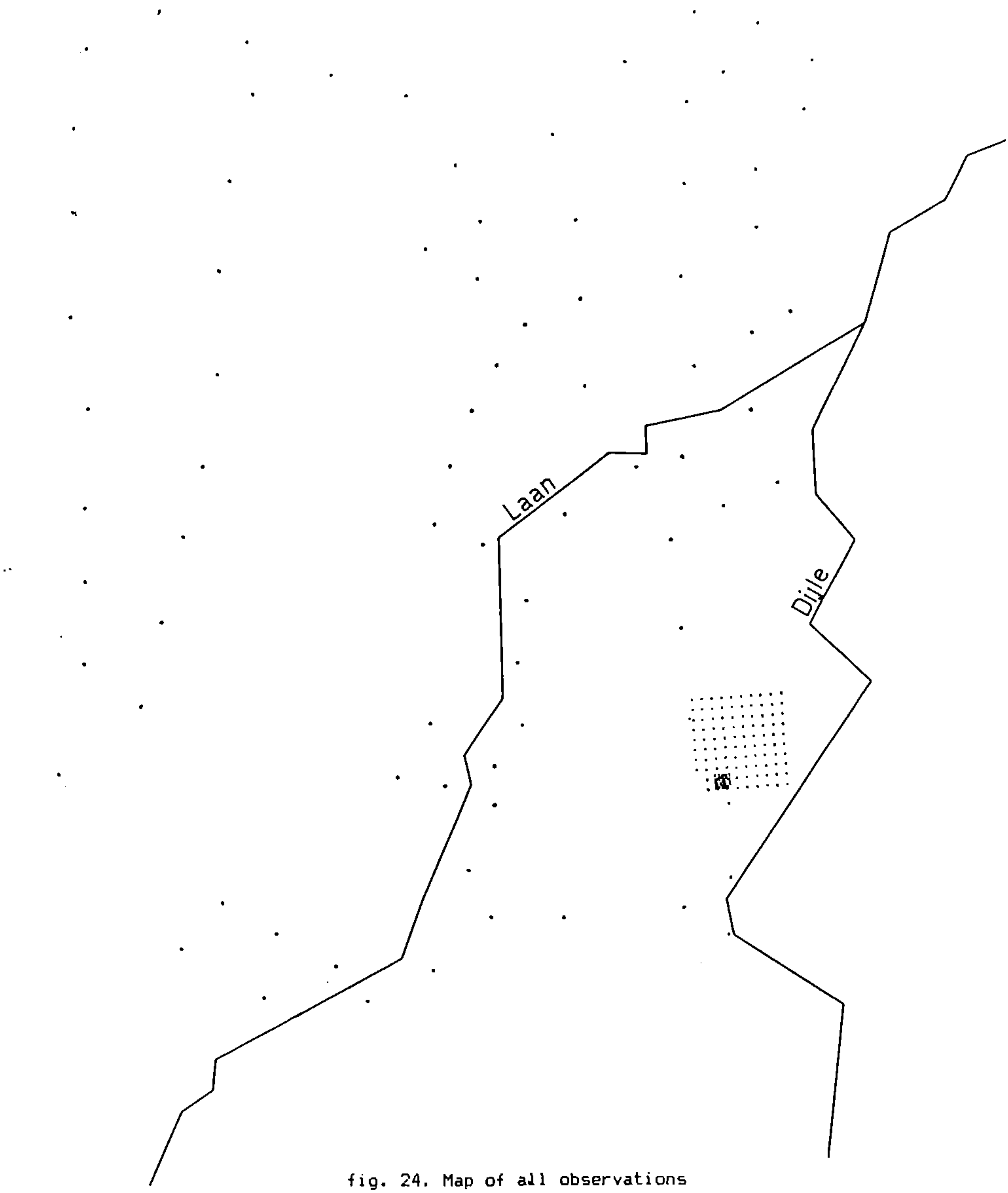


fig. 24. Map of all observations

table 15. The total variogram

1

VARIOGRAM

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK

45.

STEP IN METER = .1000E+03
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = -.7318E+00
GENERAL VARIANCE OF Z = .2514E+01
GENERAL SKEWNESS OF Z = .2998E+00
GENERAL KURTOSIS OF Z = .2429E+01

DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE	MAXVAR PAIR
0 —— 100	13342	-.703E+00	.1173E+01	46.3	69—— 79
100 —— 200	969	-.623E+00	.3792E+01	142.9	32——129
200 —— 300	2339	-.677E+00	.3927E+01	253.6	71——172
300 —— 400	2062	-.440E+00	.3587E+01	344.5	79——129
400 —— 500	2238	-.198E+00	.2875E+01	448.5	79——172
500 —— 600	2295	-.952E-01	.3714E+01	553.7	76——172
600 —— 700	2671	.596E+00	.3366E+01	649.6	73——172
700 —— 800	1762	.468E+00	.4046E+01	740.8	51——172
800 —— 900	688	-.457E+00	.4407E+01	847.1	41——172
900 ——1000	521	.108E+01	.2937E+01	942.6	42——172
1000 ——1100	33	-.518E+00	.1179E+01	1033.5	63—— 65

99.99

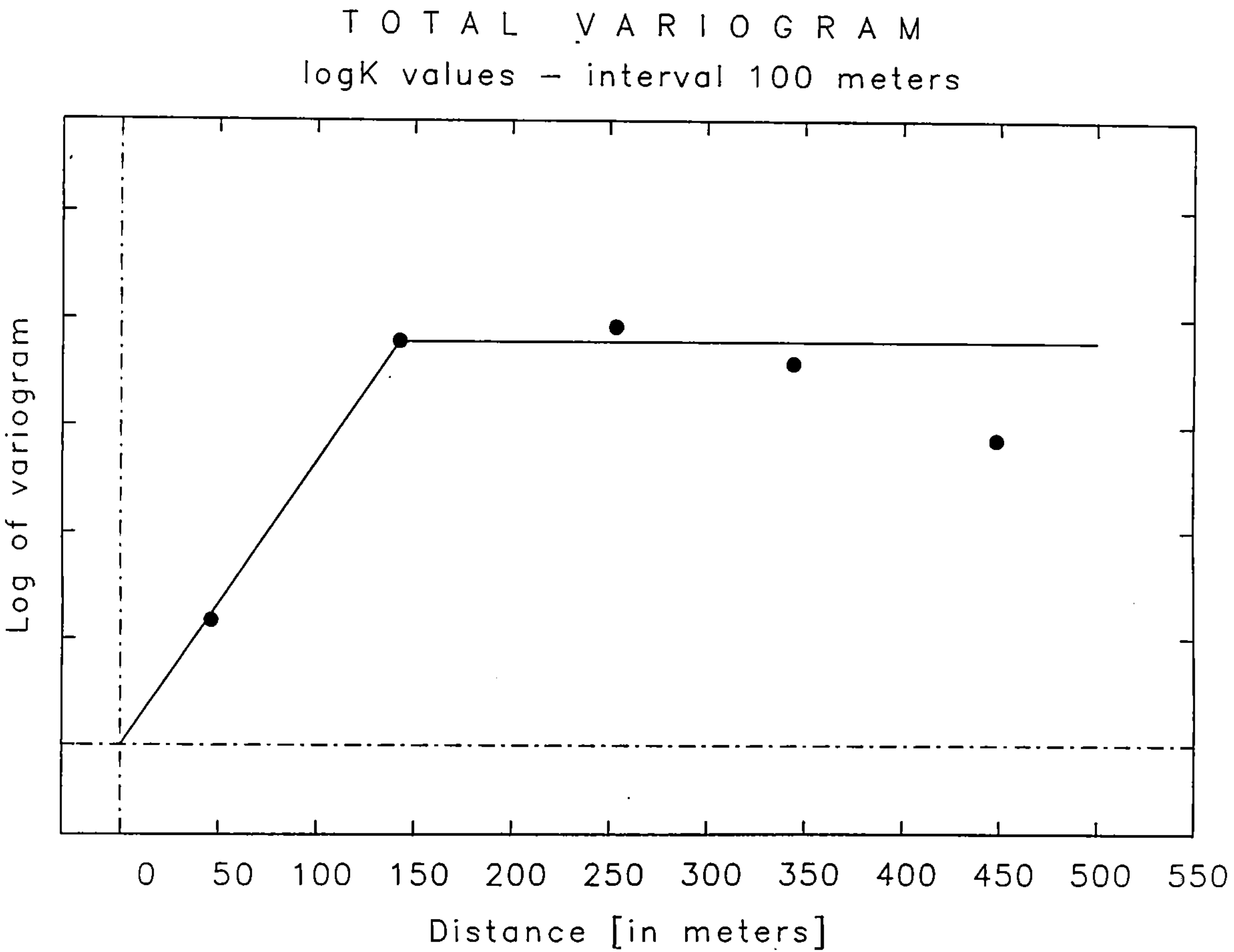


fig. 25. The total variogram

table 16. One point variograms for each study

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER
(HYDRAULIC CONDUCTIVITY K-M/DAY)
(WITH A FIELD OF 180 DEGREES IN EACH DIRECTION)

LOOK 45

STEP IN METER = .2000E+02
UPPER LIMIT FOR Z = .1129E+00
GENERAL MEAN OF Z = -.2451E+00
GENERAL VARIANCE OF Z = .2806E-01
GENERAL SKEWNESS OF Z = .4493E+00
GENERAL KURTOSIS OF Z = 2327E+01

	DISTANCE IN METER	NO OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
99.99	0 ---- 20	1953	.259E-01	.2722E-01	8.4

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER
(HYDRAULIC CONDUCTIVITY K-M/DAY)
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK 45.

STEP IN METER = .2000E+03
UPPER LIMIT FOR Z = -.3413E+00
GENERAL MEAN OF Z = -.2007E+01
GENERAL VARIANCE OF Z = .6636E+00
GENERAL SKEWNESS OF Z = .8893E-01
GENERAL KURTOSIS OF Z = .3975E+01

	DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
99.99	0 —— 200	4753	-.128E+00	.5926E+00	51.7

V A R I O G R A M

STUDY OF SPATIAL VARIABILITY OF FIELD PARAMETER:
HYDRAULIC CONDUCTIVITY IN M/DAY
(WITH A FIELD OF 180. DEGREES IN EACH DIRECTION)

LOOK 45.

STEP IN METER = .2000E+04
UPPER LIMIT FOR Z = .2689E+01
GENERAL MEAN OF Z = .4913E+00
GENERAL VARIANCE OF Z = .3062E+01
GENERAL SKEWNESS OF Z = -.6963E+00
GENERAL KURTOSIS OF Z = .2339E+01

	DISTANCE IN METER	NO. OF PAIRS	DRIFT	VARIOGRAM	AVERAGE DISTANCE
99.99	0 ——2000	3081	-.200E+00	.3038E+01	475.8

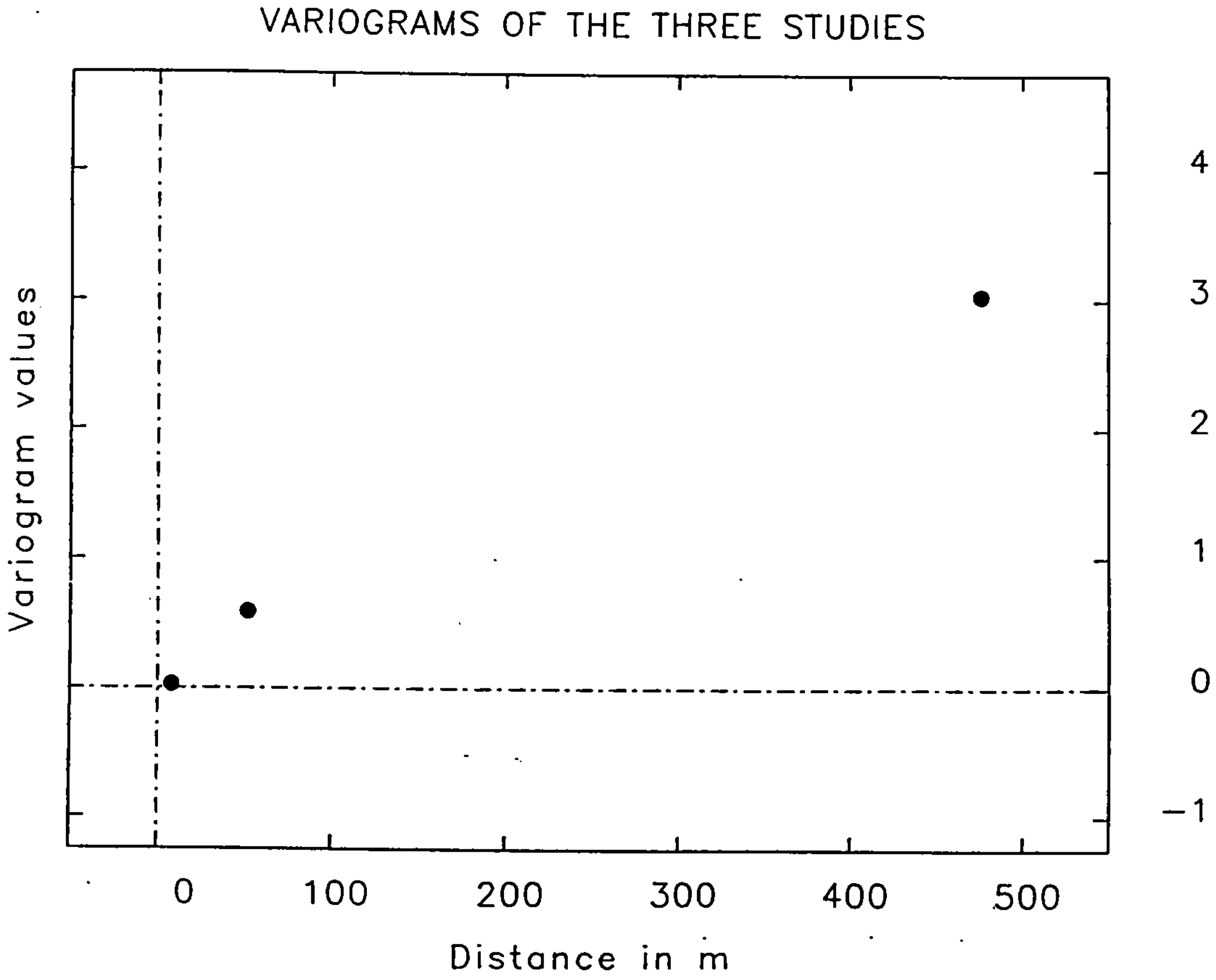


Fig. 26. One point variograms for each study

4.8. Kolmogorov-Smirnov goodness of fit test.

Looking for the distribution of the data sets, we will test the hypothesis that it is normal. In order to test this hypothesis, we will use the Kolmogorov-Smirnov goodness of fit test, which, it should be stressed, cannot be applied in cases where the data is correllated.

The theoretical probabilities of an observation being smaller or equal to each specific one (under the hypothesis of normal distribution), are first calculated. This is done by first normalising the values with the formula:

$$Z=(K-m)/s \qquad (4.20)$$

where m is the mean, and s the standard deviation, and second, reading the probability from the standard normal curve table:

$$F(K)=P(Z<z) \qquad (4.21)$$

The corresponding experimental probabilities are then also calculated. The observed values have to be in ascending order. Then, as the experimental probability of a measurement being smaller or equal to the Ith observation with value K, we consider:

$$Fe(K)=(I-0.5)/n \qquad (4.22)$$

where n is the total number of observations.

For each specific observation K, the difference between the experimental probability is calculated:

$$di=Fe(Ki)-F(Ki) \qquad (4.23)$$

The greatest value of the dis is then compared with standard functions, depending upon the significance level considered:

for significance level 5%, dmax allowed is $(1.36/\sqrt{n})$

for significance level 10%, dmax allowed is $(1.22/\sqrt{n})$

4.9. Application of the Kolmogorov-Smirnov test.

In the present study, the distances between the observation points are of the order of magnitude of 50 to 120 meters. In order to apply the goodness of fit test for the normality of the observations, we have first to ensure the non-correllatedness among them. The model of the total variogram, has a range of 142m. So, to ensure the non-correlatedness we would have to exclude all points with distances smaller than 140m. But in that case, we would have been left with very few observations. For this reason, we select 35 observations between which there is no distance smaller than 90 meters, since hydraulic conductivities of points with smaller distances, can be considered as being correllated.

A map of these points is situated in the following page. The program VARIO3 is then used to make the test, and the results can be seen in table 18 for the logK values, and in table 19 for the K values. Figure 28 shows the theoretical and experimental probabilities versus the logK values, and figure 29 versus the K values. The results show that both the normality of logK and K values are accepted at a significance level of 5%, and rejected at a level of 10%.

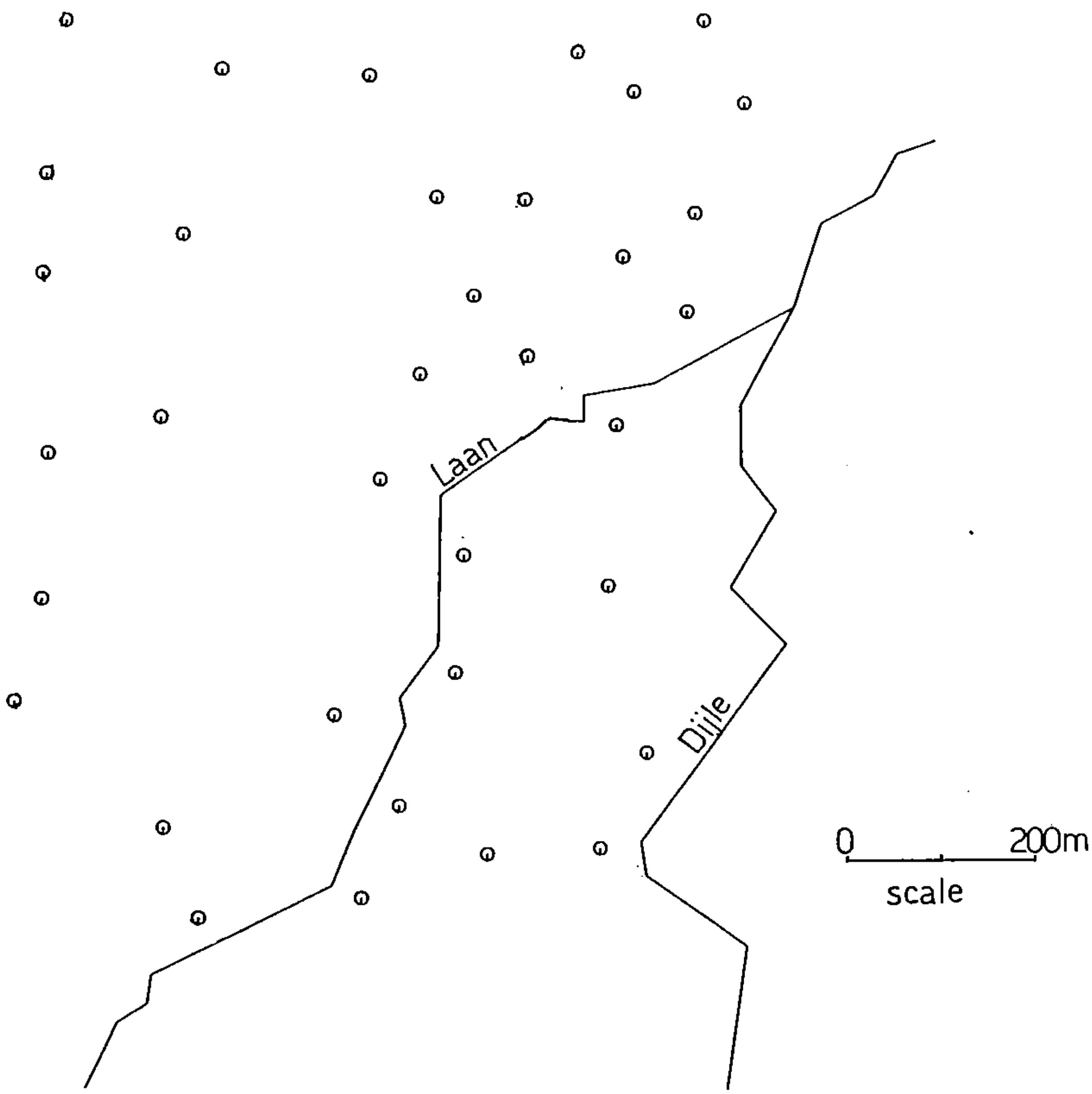


fig. 27. Observations considered for the K-S test

table 17. Kolmogorov-Smirnov test for logK values

NUMBER OF SAMPLES= 35

RANK	Z VAL	NOR. VAL	EXPT. PROB	THEO. PROB	D1	D2	D
1	-3.456	-2.622	.0142857	.0043696	.010	.000	.010
2	-3.071	-2.387	.0428571	.0085033	.034	.006	.034
3	-1.477	-1.412	.0714286	.0789338	.008	.036	.036
4	-1.328	-1.321	.1000000	.0931946	.007	.022	.022
5	-1.304	-1.307	.1285714	.0956069	.033	.004	.033
6	-1.225	-1.258	.1571429	.1041386	.053	.024	.053
7	-.918	-1.071	.1857143	.1420823	.044	.015	.044
8	-.681	-.926	.2142857	.1772225	.037	.008	.037
9	-.428	-.771	.2428571	.2202541	.023	.006	.023
10	-.229	-.650	.2714286	.2578152	.014	.015	.015
11	.121	-.436	.3000000	.3315578	.032	.060	.060
12	.278	-.340	.3285714	.3669936	.038	.067	.067
13	.895	.037	.3571429	.5149216	.158	.186	.186
14	1.150	.193	.3857143	.5765070	.191	.219	.219
15	1.202	.225	.4142857	.5889165	.175	.203	.203
16	1.305	.288	.4428571	.6133477	.170	.199	.199
17	1.385	.336	.4714286	.6317391	.160	.189	.189
18	1.409	.351	.5000000	.6373160	.137	.166	.166
19	1.536	.429	.5285714	.6659252	.137	.166	.166
20	1.704	.531	.5571429	.7024247	.145	.174	.174
21	1.712	.537	.5857143	.7042528	.119	.147	.147
22	1.743	.556	.6142857	.7107422	.096	.125	.125
23	1.916	.661	.6428571	.7456942	.103	.131	.131
24	1.972	.695	.6714286	.7565202	.085	.114	.114
25	1.974	.697	.7000000	.7569965	.057	.086	.086
26	2.016	.722	.7285714	.7649221	.036	.065	.065
27	2.060	.749	.7571429	.7730932	.016	.045	.045
28	2.157	.809	.7857143	.7906138	.005	.033	.033
29	2.211	.842	.8142857	.8000214	.014	.014	.014
30	2.239	.859	.8428571	.8047884	.038	.009	.038
31	2.298	.894	.8714286	.8144588	.057	.028	.057
32	2.328	.913	.9000000	.8193256	.081	.052	.081
33	2.450	.988	.9285714	.8383119	.090	.062	.090
34	2.563	1.057	.9571429	.8546852	.102	.074	.102
35	2.689	1.133	.9857143	.8714897	.114	.086	.114

DMAX = .219

THE NORMALITY OF THE OBSERVATIONS IS ACCEPTED
FOR SIGNIFICANCE LEVEL 5%
SINCE DMAX= .219 AND THE LIMIT IS .230

THE NORMALITY OF THE OBSERVATIONS IS REJECTED
FOR SIGNIFICANCE LEVEL 10%
SINCE DMAX= .219 AND THE LIMIT IS .206

THE MEAN IS .834177 AND STD. IS 1.636142

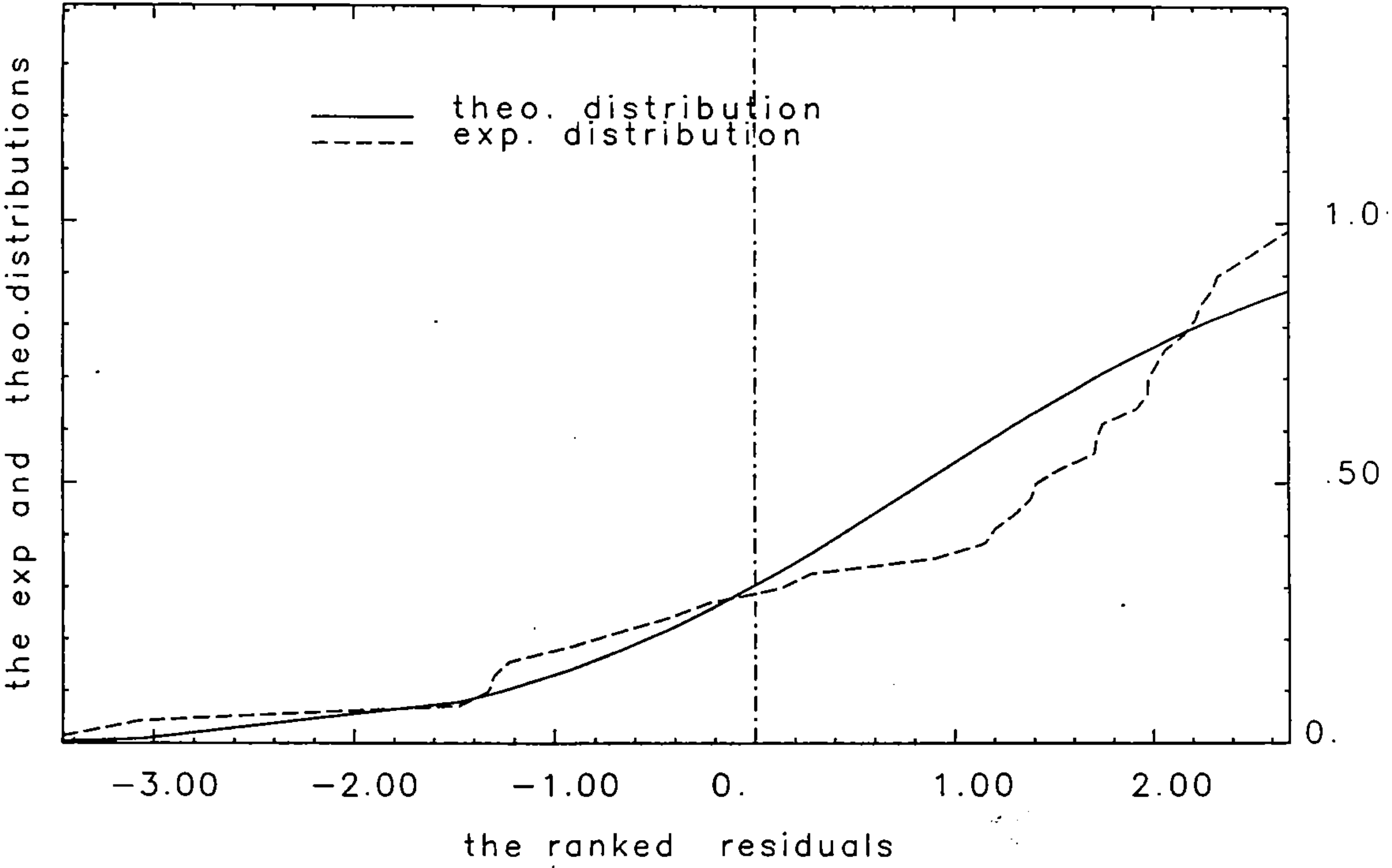


Fig. 28. Distribution curves for log of K values.

table 18. Kolmogorov-Smirnov test for K values

NUMBER OF SAMPLES= 35

RANK	Z VAL	NOR. VAL	EXPT. PROB	THEO. PROB	D1	D2	D
1	.000	-.703	.0142857	.2409369	.227	.000	.227
2	.001	-.703	.0428571	.2409382	.198	.227	.227
3	.033	-.703	.0714286	.2410267	.170	.198	.198
4	.047	-.703	.1000000	.2410638	.141	.170	.170
5	.050	-.703	.1285714	.2410709	.112	.141	.141
6	.060	-.703	.1571429	.2410981	.084	.113	.113
7	.121	-.702	.1857143	.2412644	.056	.084	.084
8	.209	-.701	.2142857	.2415033	.027	.056	.056
9	.373	-.700	.2428571	.2419525	.001	.028	.028
10	.590	-.698	.2714286	.2425421	.029	.000	.029
11	1.323	-.692	.3000000	.2445475	.055	.027	.055
12	1.897	-.687	.3285714	.2461260	.082	.054	.082
13	7.859	-.635	.3571429	.2628162	.094	.066	.094
14	14.122	-.580	.3857143	.2809511	.105	.076	.105
15	15.919	-.564	.4142857	.2862641	.128	.099	.128
16	20.206	-.527	.4428571	.2991280	.144	.115	.144
17	24.248	-.492	.4714286	.3114929	.160	.131	.160
18	25.642	-.479	.5000000	.3158058	.184	.156	.184
19	34.322	-.404	.5285714	.3432227	.185	.157	.185
20	50.536	-.262	.5571429	.3966052	.161	.132	.161
21	51.552	-.253	.5857143	.4000286	.186	.157	.186
22	55.353	-.220	.6142857	.4128992	.201	.173	.201
23	82.351	.016	.6428571	.5062157	.137	.108	.137
24	93.658	.114	.6714286	.5454947	.126	.097	.126
25	94.197	.119	.7000000	.5473564	.153	.124	.153
26	103.715	.202	.7285714	.5800710	.149	.120	.149
27	114.754	.298	.7571429	.6173151	.140	.111	.140
28	143.577	.550	.7857143	.7088569	.077	.048	.077
29	162.673	.717	.8142857	.7632335	.051	.022	.051
30	173.532	.812	.8428571	.7914703	.051	.023	.051
31	198.436	1.029	.8714286	.8482453	.023	.005	.023
32	212.640	1.153	.9000000	.8755299	.024	.004	.024
33	281.799	1.757	.9285714	.9605107	.032	.061	.061
34	365.727	2.489	.9571429	.9936000	.036	.065	.065
35	488.294	3.559	.9857143	.9998140	.014	.043	.043

DMAX = .227

THE NORMALITY OF THE OBSERVATIONS IS ACCEPTED
FOR SIGNIFICANCE LEVEL 5%
SINCE DMAX= .227 AND THE LIMIT IS .230

THE NORMALITY OF THE OBSERVATIONS IS REJECTED
FOR SIGNIFICANCE LEVEL 10%
SINCE DMAX= .227 AND THE LIMIT IS .206

THE MEAN IS 80.566149 AND STD. IS 114.555251

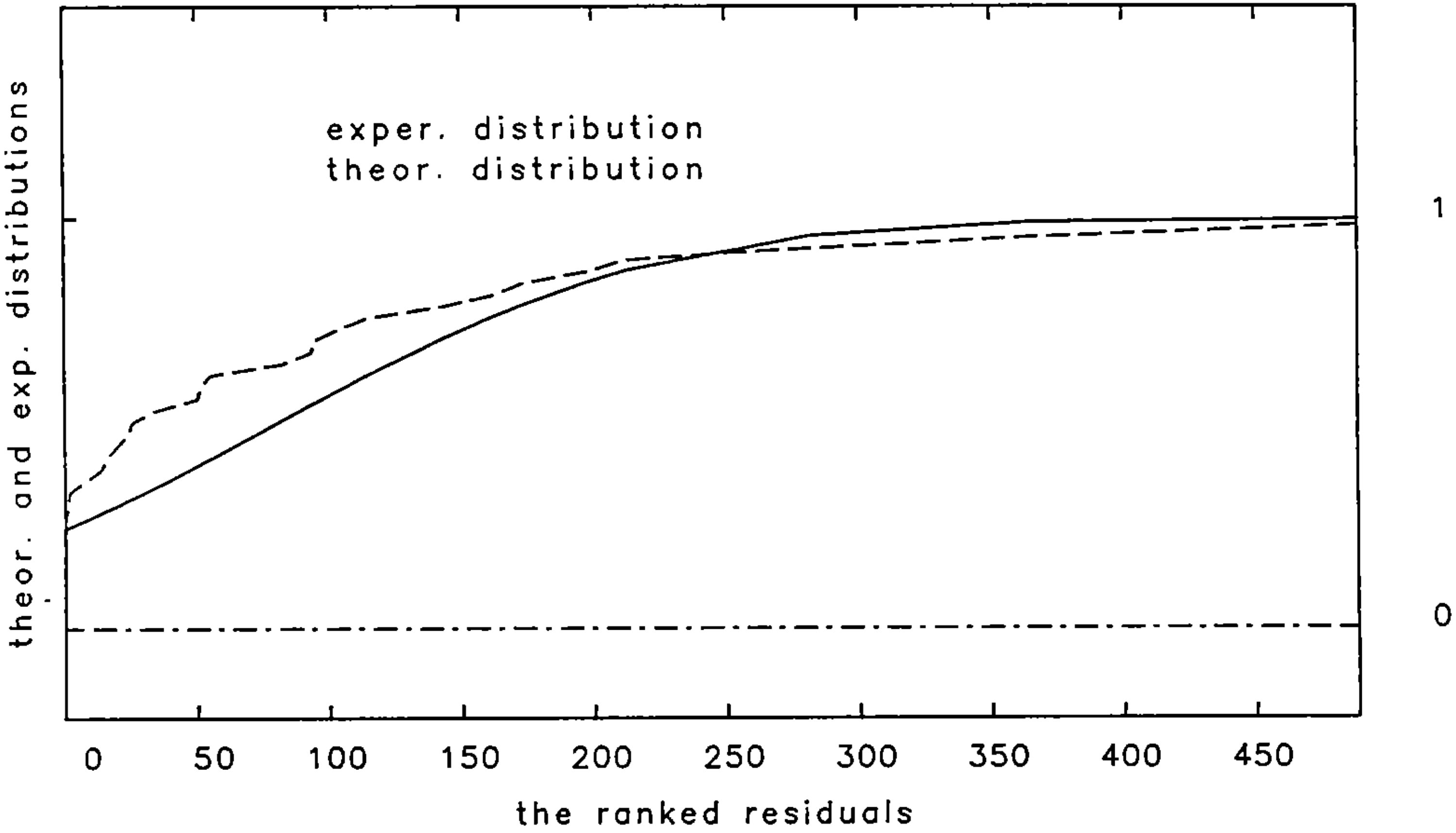


Fig. 29. Distribution curves for K values