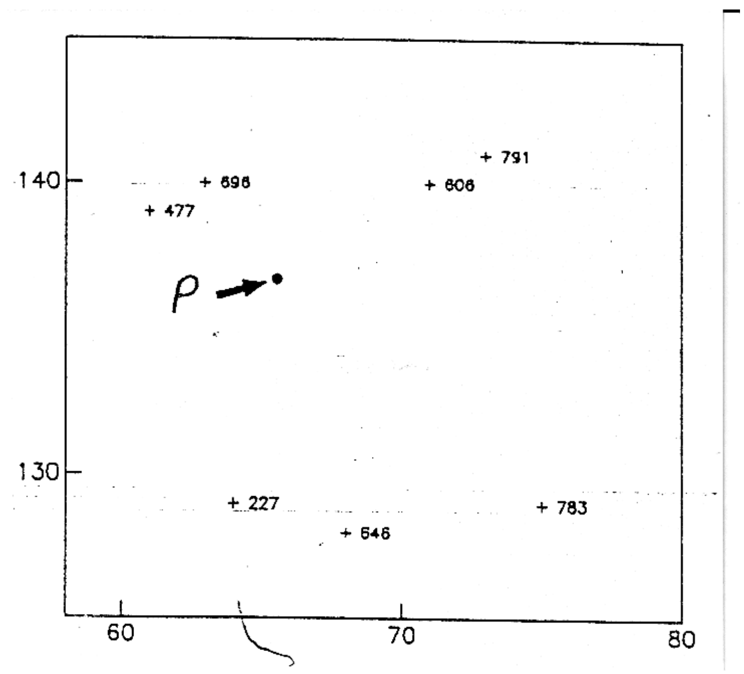
The Problem

- Estimation of unknown values (geological parameters) at specified locations (points or blocks).



- Classical Estimation Methods
  - Averaging
  - Geometrical
  - Distance weighting
  - Polynomial equations
- Geostatistical Techniques

Local Sample Mean
(Averaging Methods)

- A simple approach that weights all samples equally
- Arithmetic mean

$$\overline{x}_a = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Local Sample Mean
(Averaging Methods)

- Geometric mean

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

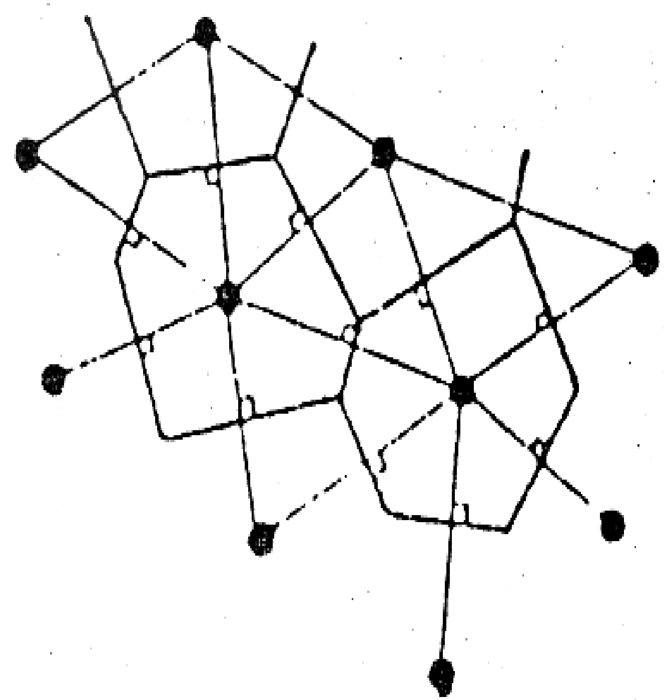
- Harmonic mean

$$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

$$\bar{x}_a > \bar{x}_g > \bar{x}_h$$

Polygons
(Geometrical Methods)

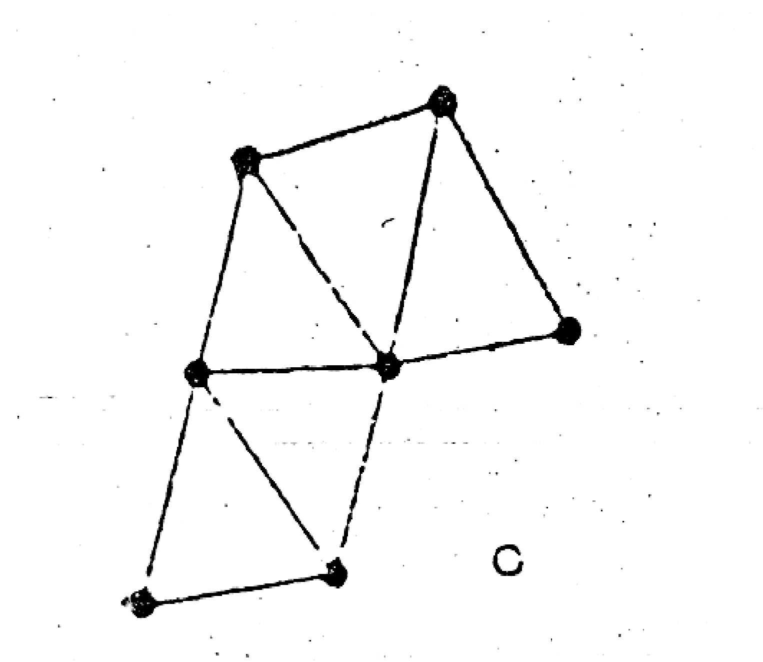
- If a measured data point is located inside a polygon, the whole polygon is assigned the value of that point.



A polygon is constructed by drawing perpendicular bisectors of lines joining neighboring sampled points.

Triangulation
(Geometrical Methods)

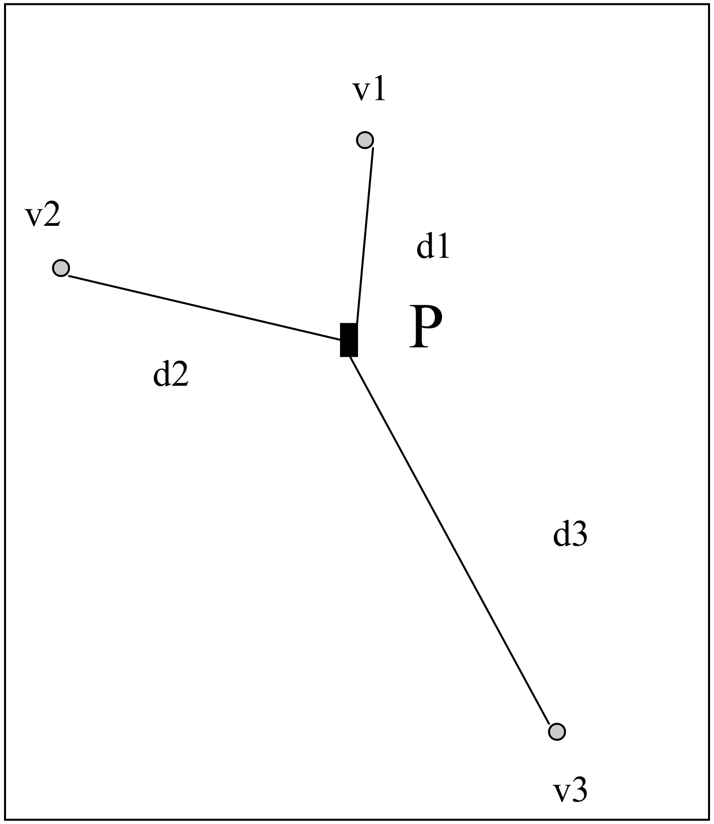
- The mean value of the 3 closest data points at the corners is assigned to the triangle.



It is an alternative method for polygonal technique.

Inverse Distance Method
(Distance Weighting Techniques)

- Estimating the value of a point (P) by assigning weights to known points (v's)



- Weights are functions of the distances between point P and v's.
- The influence of a known sample value (v) on the point to be estimated (P) decreases with the increase of the distance (d).

Inverse Distance Method
(Distance Weighting Techniques)

- The inverse of the "power" distance is used as a weight.
- Power "p" is chosen arbitrarily
- The most common and traditional choice of the power p is 2.

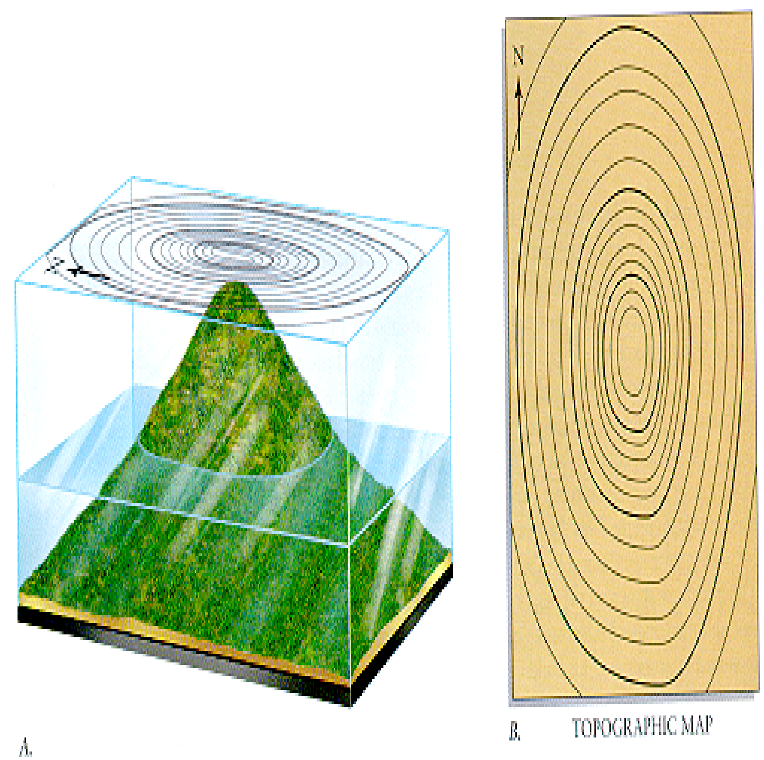
$$\widehat{P} = \frac{\sum_{i=1}^{n} v_i \frac{1}{d_i^p}}{\sum_{i=1}^{n} \frac{1}{d_i^p}}$$

Polynomial Equations
(Simple Plane Equation)

- Estimate the value of a point by polynomials.
- The simplest from is:

$$\hat{z} = ax + by + c$$

where x and y are samples' coordinates and a,b, and c are coefficients. Three sample points are needed to solve the equation



Disadvantages of Classical Methods

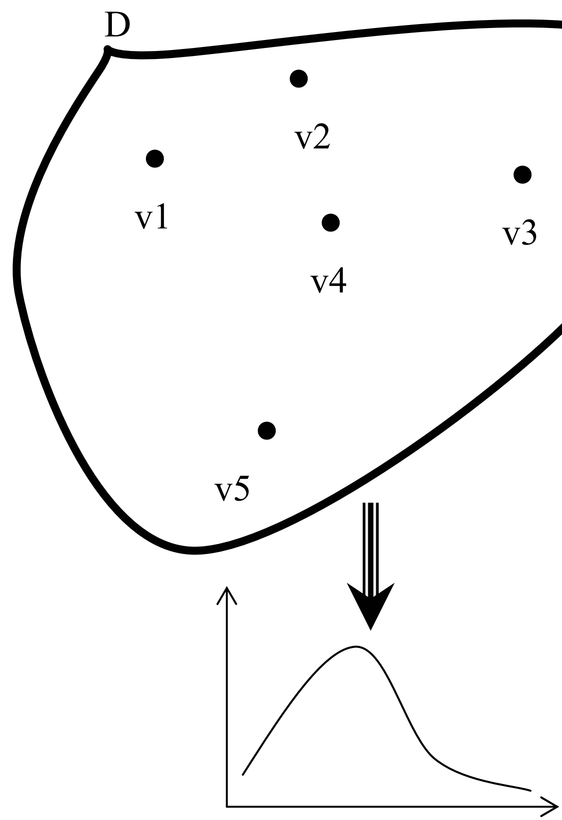
- Averaging is a very simple approach to deal with complex geological media because it smoothes all details.
- Discontinuous estimates from region to region if using polygonal technique.
- Triangulation is not unique.
- 3D implementation is a mess for both polygonal and triangulation techniques.
- If using the inverse distance method to estimate a block, the same estimate is assigned to different block sizes!

- How many samples should be used in the inverse distance method estimator?
- Polynomial equations estimate a value at a location without considering its geology.
- All classical methods do not provide a measure of uncertainty that is associated with the estimation procedure. Thus, reliable confidence intervals can not be assigned to the estimates.

Geostatistical methods overcome all previously mentioned disadvantages.

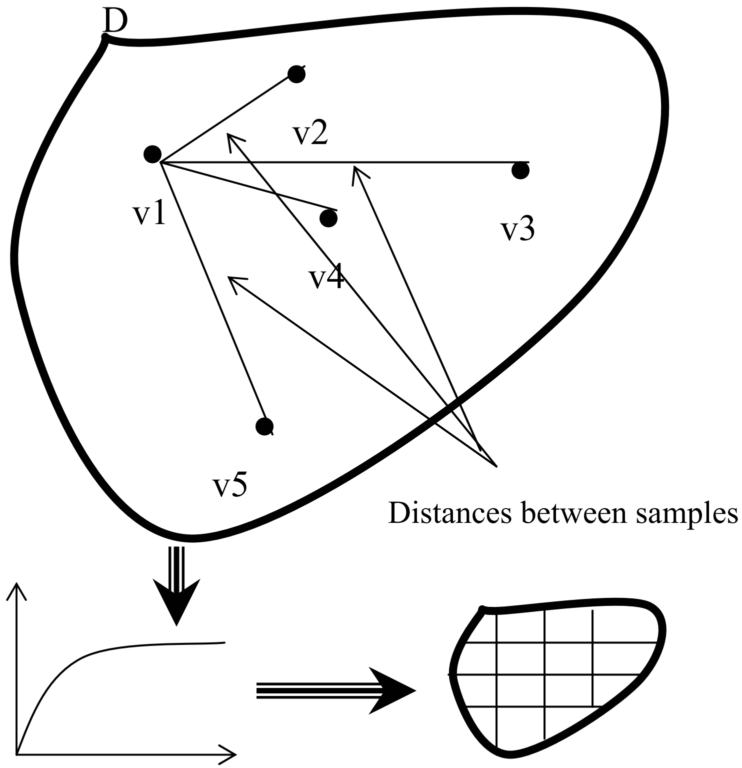
Difference between Statistics and Geostatistics

- Considers only values of the realizations of random variables in a domain D (e.g. porosity values in a reservoir)
- Assumes that the random variables are independent and uncorrelated.
- Assumes that the statistical inference and distribution of a limited data set are applicable to the whole population (or for the whole domain).
- Used for preliminary analysis of a data set and for final interpretations of the results.



Difference between Statistics and Geostatistics

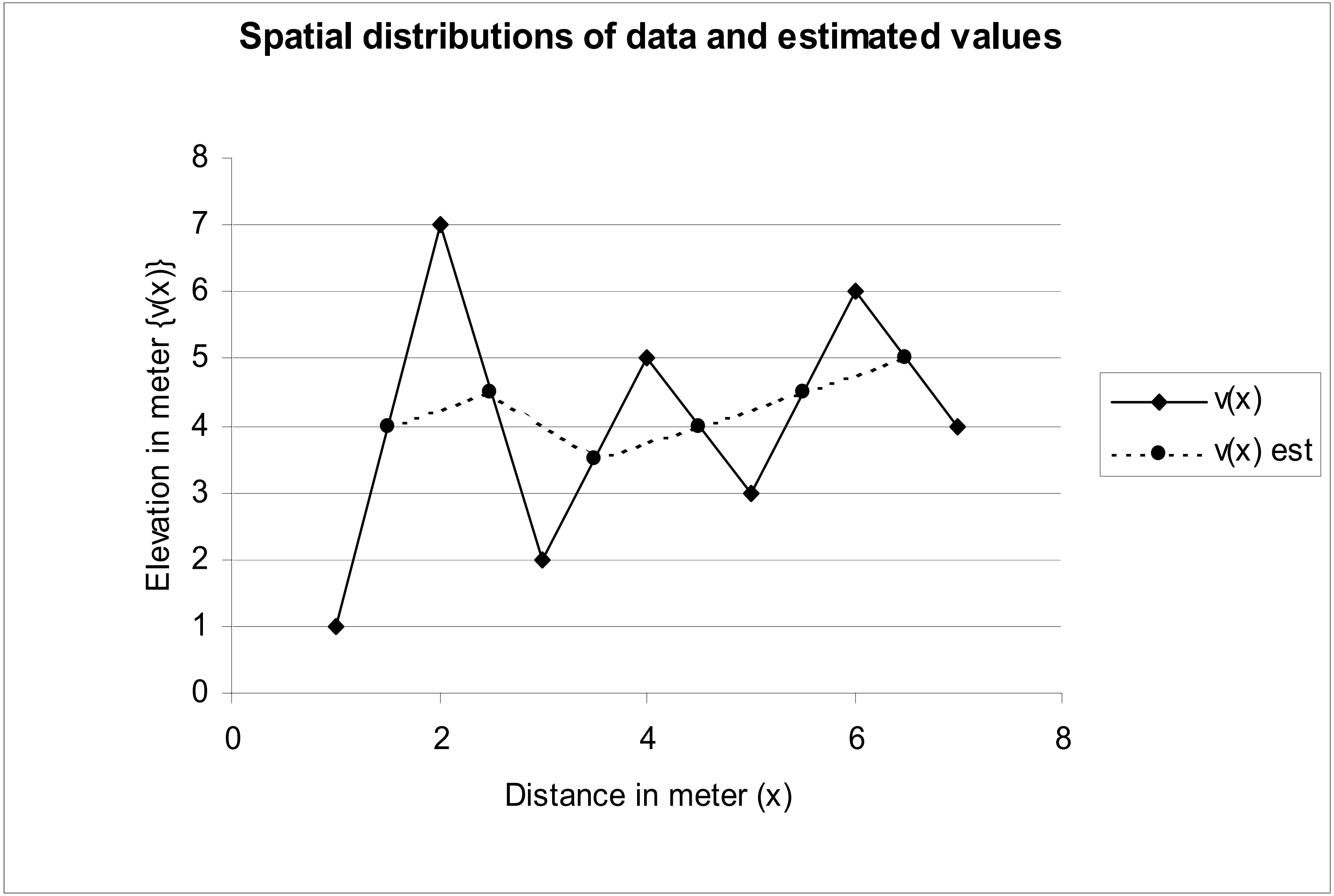
- Considers values of the realizations of random variables in a domain D as well as their locations, time series or both of them (e.g. porosity values and locations in a reservoir)
- Assumes that the random variables are spatially, temporally or spatially-temporally correlated.
- Assumes that the geostatistical interpretations of a limited data set are applicable to the whole domain.
- Used for estimating, simulating and mapping the unknowns.

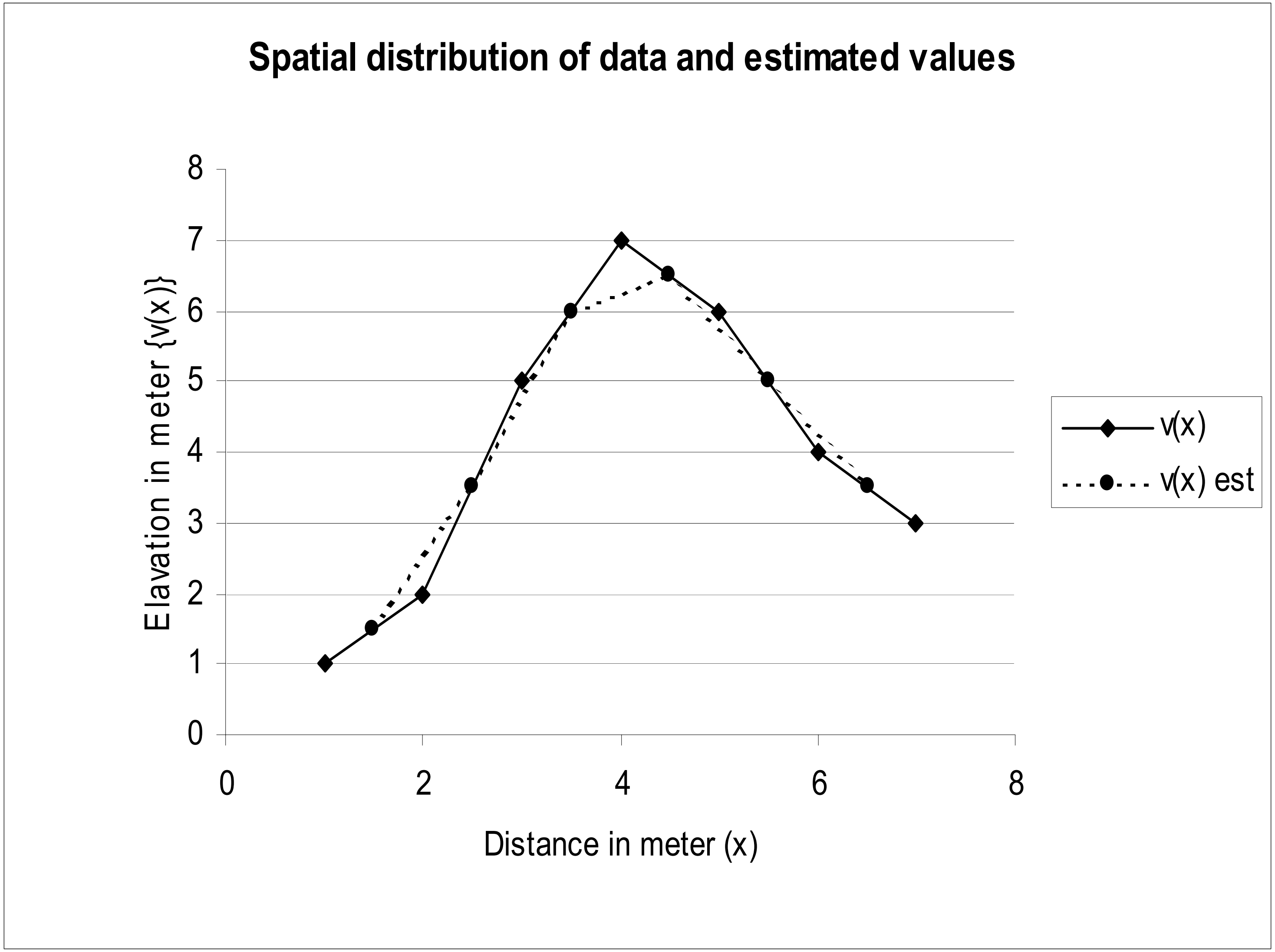


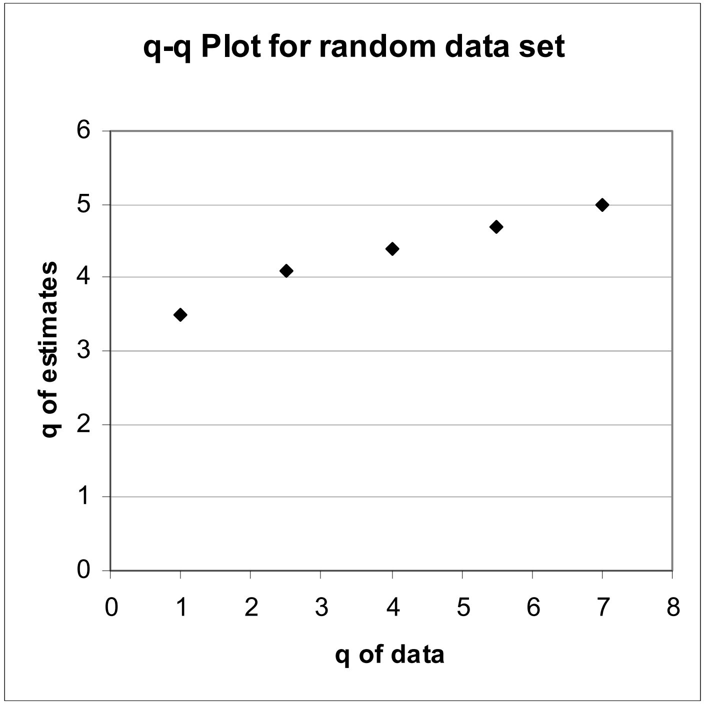
Value and Position

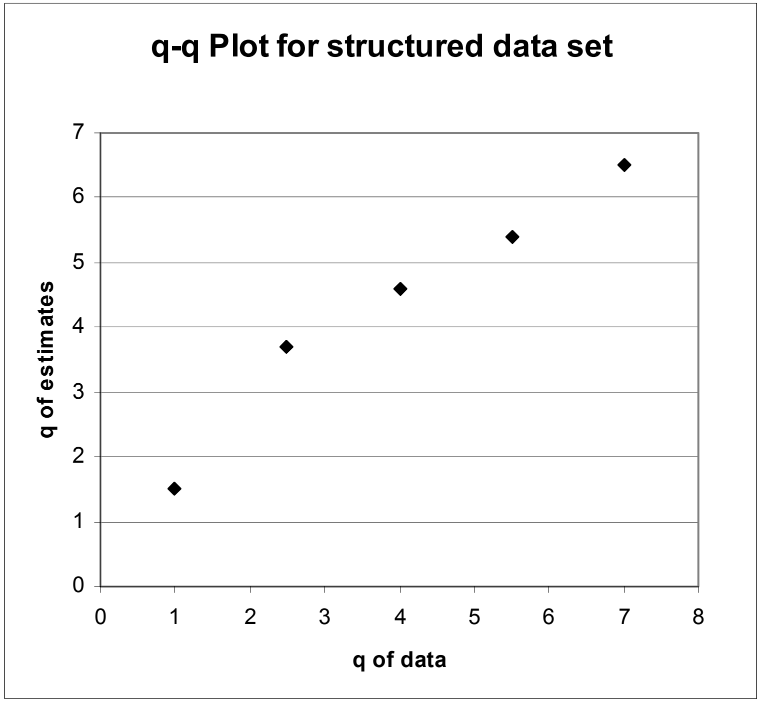
- Set # 1
- (1 7 2 5 3 6 4)
  - Same values as set# 2
  - Same mean (m=4)
  - Same variance(S=2.2)
  - Identical distribution
  - Different spatial arrangements
  - More random

- Set # 2
- (1 2 5 7 6 4 3)
  - Same values as set# 1
  - Same mean (m=4)
  - Same variance (S=2.2)
  - Identical distribution
  - Different spatial arrangements
  - Less random and more structured
  - Spatial arrangement or structure can be revealed by geostatistics (Theory of Regionalized Variables).


Spatial distributions of data and estimated values
Elevation in meter {v(x)}
Distance in meter (x)
0
1
2
3
4
5
6
7
8
0
2
4
6
8
v(x)
v(x) est


Spatial distribution of data and estimated values
Elavation in meter {v(x)}
0
1
2
3
4
5
6
7
8
0
2
4
6
8
Distance in meter (x)
v(x)
v(x) est


q-q Plot for random data set
q of estimates
q of data
0
1
2
3
4
5
6
7
8


q-q Plot for structured data set
q of estimates
q of data
0
1
2
3
4
5
6
7
8

The Theory of Regionalized Variables (or The Theory of Geostatistics)

- Regionalized Variable (ReV): A single-valued function defined over space and/or time. *Locally, a ReV is considered as a Random Variable (rv) but regionally it is linked with other rv's by a specific correlation function.*

  - Geostatistics models regionalized variables as realizations of random functions (RF).
  - The variation of a RF is not totally random in space. In several geological case studies some structure is imposed.
  - Examples of RF: porosity, permeability, elevation of top of a confining layer of an aquifer, thickness of a reservoir, mineral grade …. etc.

The Theory of Regionalized Variables

- ReV possesses two “contradictory” characteristics:
  - A local, random, and erratic aspect which calls to mind the concept of random variables.
  - A general (average) structured aspect that requires a certain functional representation.

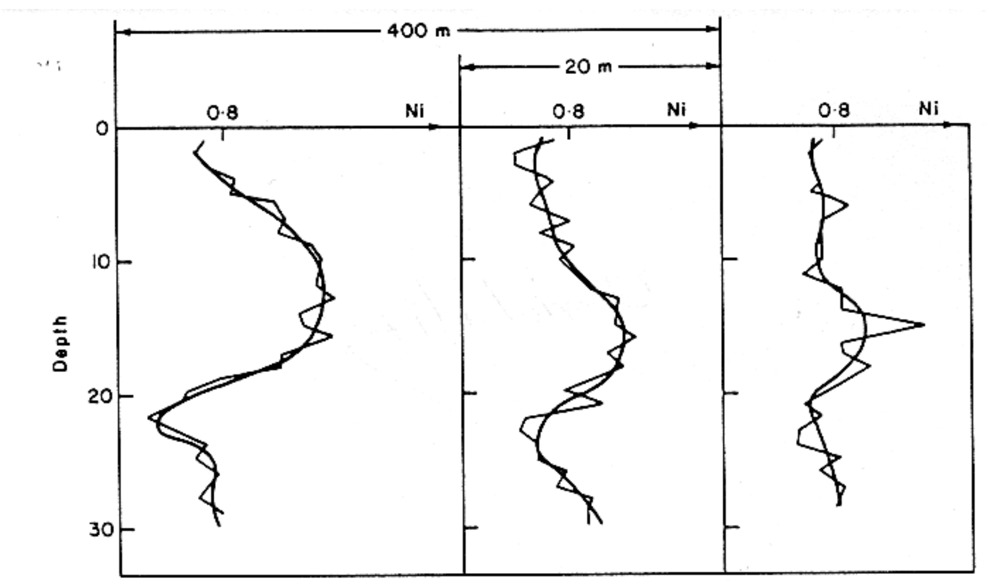


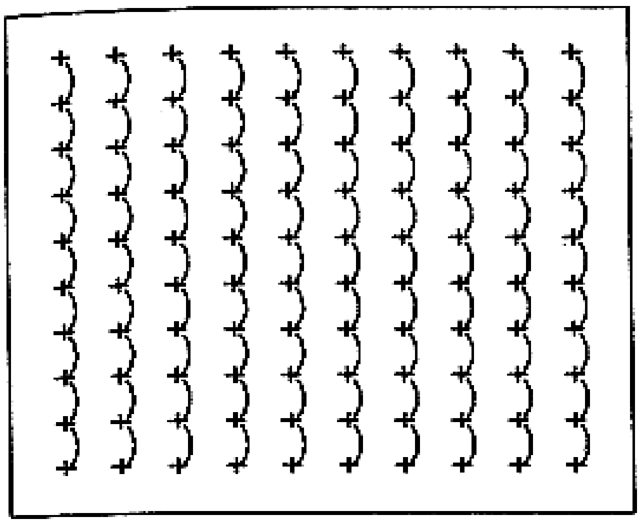
FIG. II.1. Vertical variability of Ni grades.

Spatial Continuity

- Spatial continuity exists in most geological parameters
- Spatial continuity can not be captured by univariate statistics
- Two data points close to each other are more likely to have similar values than two data points that are far apart.
- Tools used to describe spatial continuity:
    - Scatterplots
    - Correlation function
    - Covariance functions
    - Moment of inertia

h-Scatterplots

- An h-Scatterplot shows all possible pairs of data values whose locations are separated by a **VECTOR "h"** (i.e. certain distance in a particular direction).



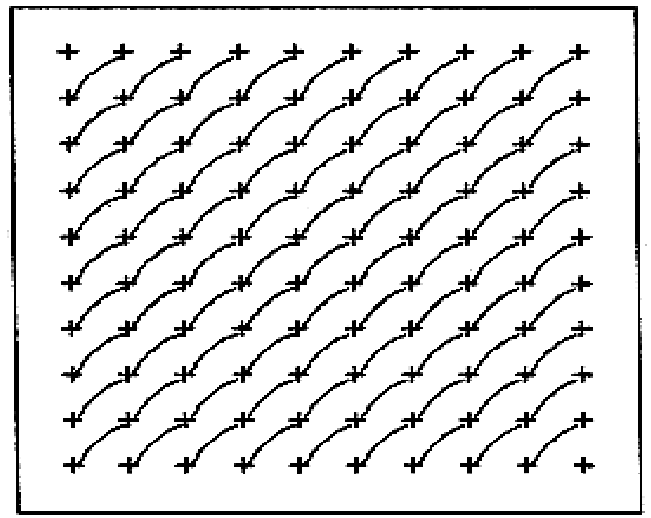


Figure 4.11 **Examples of how data can be paired to obtain an h-scatterplot.**

h-Scatterplots

for 4 separation distances

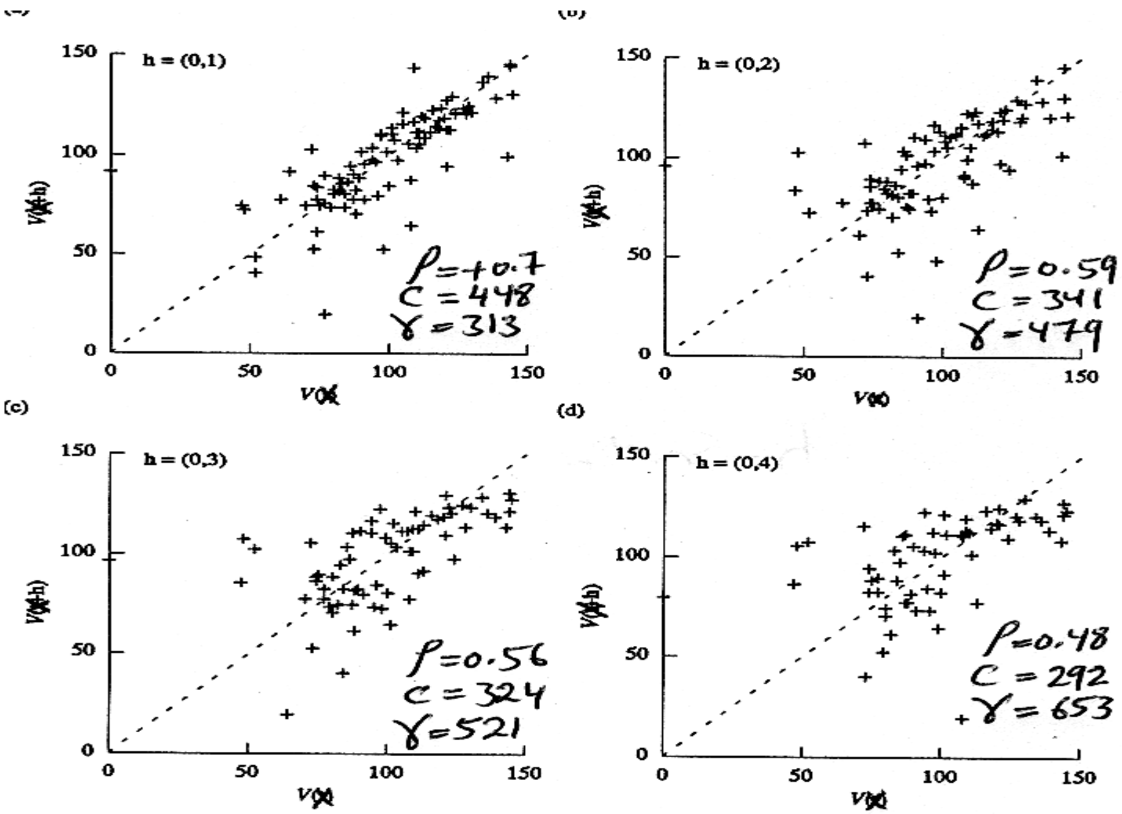


Figure 4.12 h-Scatterplots for four separation distances in a northerly direction between pairs of the 100 *V* values. As the separation distance increases, the similarity between pairs of values decreases and the points on the h-scatterplot spread out further from the diagonal line.

h-Scatterplots

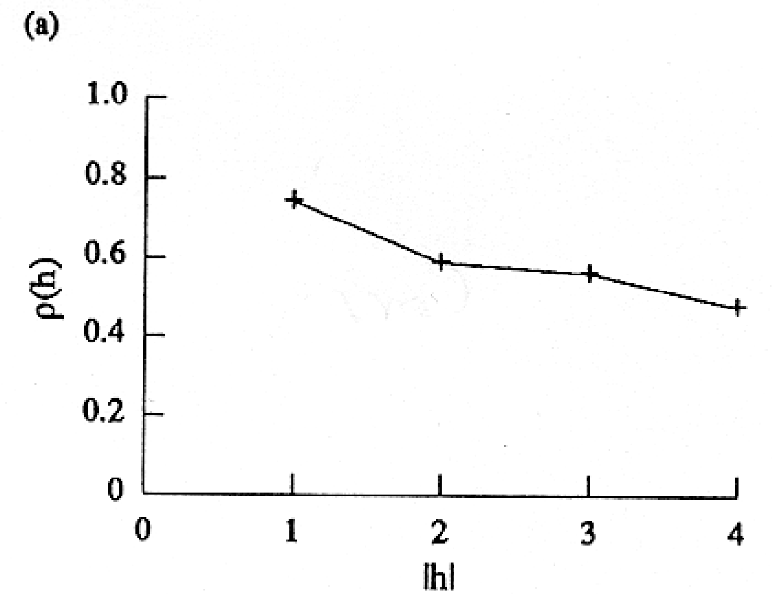
for 4 separation distances

Table 4.1 Statistics summarizing the fatness of the four h-scatterplots shown in Figures 4.12a-d.

h	Correlation Coefficient	Covariance (ppm^2)	Moment of Inertia (ppm^2)
(0,1)	0.742	448.8	312.8
(0,2)	0.590	341.0	479.2
(0,3)	0.560	323.8	521.4
(0,4)	0.478	291.5	652.9

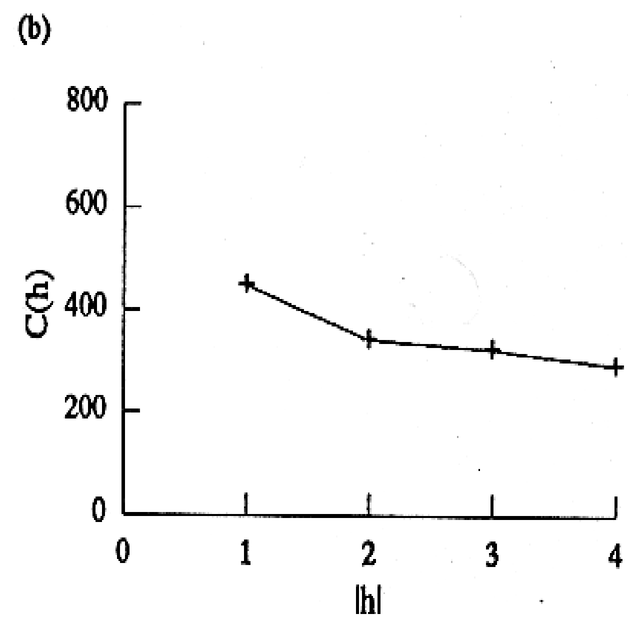
Correlation Coefficient and Correlation Function

- Correlation coefficient (ρ) decreases as vector **h** increases.
- The relationship between correlation coefficient of an h-Scatterplot and vector **h** is called correlation function or ***CORRE-OGRAM*** ($\rho(h)$).



Covariance and Covariance Function

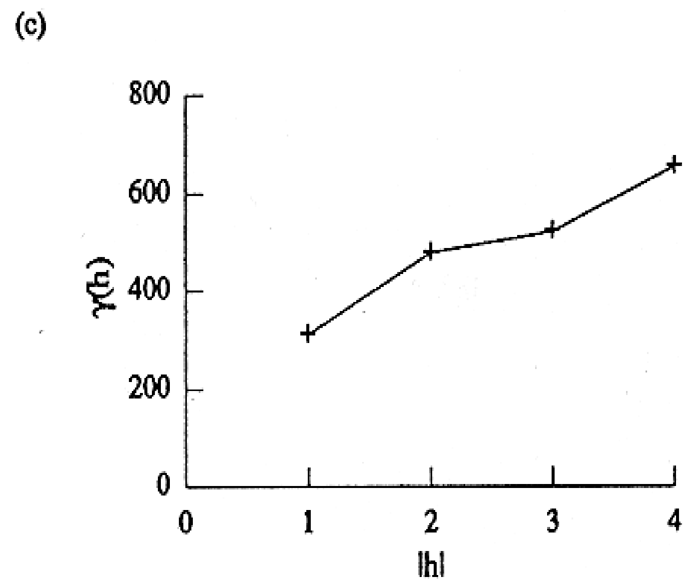
- Covariance decreases as vector **h** increases.
- The relationship between covariance of an h-Scatterplot and vector **h** is called covariance function or simply:



COVARIANCE, C(h)

Moment of Inertia or Semivariogram

- Moment of inertia increases as vector **h** increases.
- The relationship between moment of inertia of an h-Scatterplot and vector **h** is called the ***SEMIVARIOGRAM*** or simply ***VARIOGRAM***, $\gamma(h)$



Relationship between Semivariogram and Covariance Functions

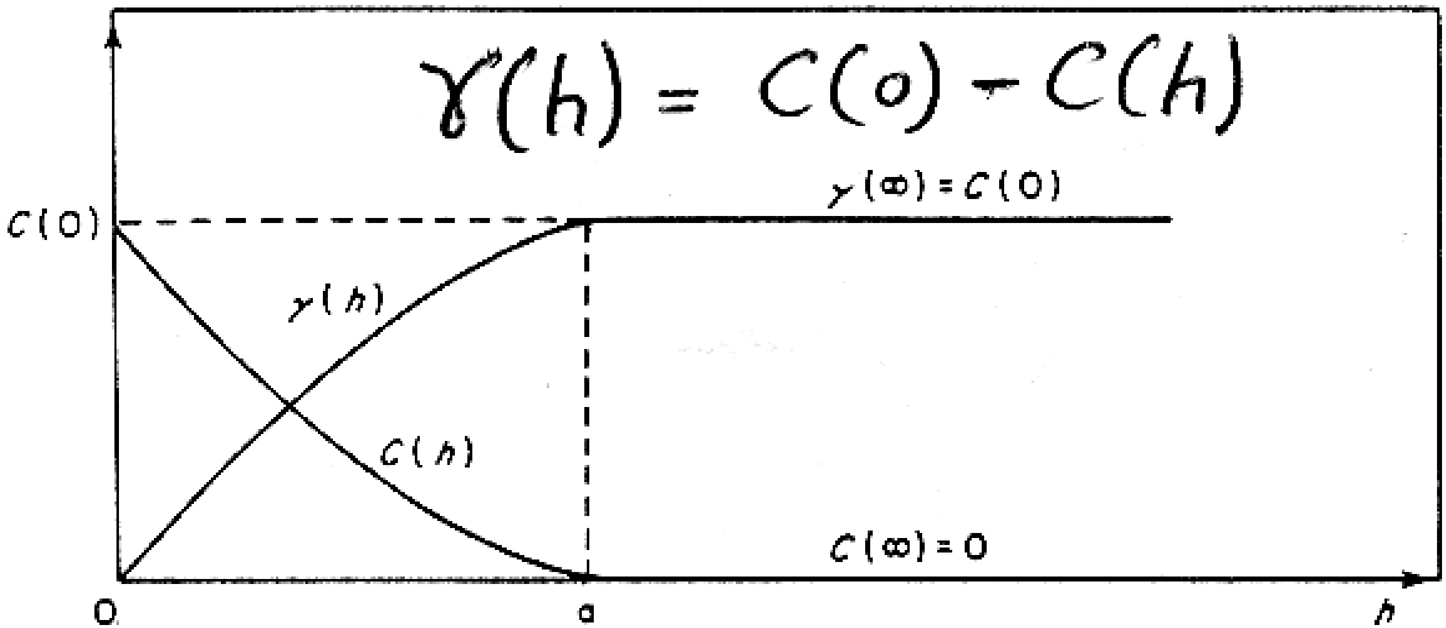


FIG. II.4. Covariance and semi-variogram.