ICIWRM – 2000, Proceedings of International Conference on Integrated Water Resources
Management for Sustainable Development, 19 – 21 December, 2000, New Delhi, India

Geostatistics: A brief look at its application in drainage management

MAHMOUD M. MOUSTAFA
Egyptian Public Authority for Drainage Projects, Ministry of Water Resources and Irrigation, 13 Giza Street, Giza, Egypt

Abstract

Many statistical tools are useful in developing qualitative insights into a wide variety of natural phenomena; many others can be used to develop quantitative answers to specific questions. Unfortunately, most classical statistical methods make no use of the spatial information in soil-water properties data sets. Geostatistics, as one tool of Spatial Information Technology (SIT), offers a way of describing the spatial continuity that is an essential feature of many natural phenomena and provides adaptations of classical regression techniques to take advantage of this continuity.

In view of the essential nature of subsurface drainage as a factor in the yield of farmlands in our struggle to produce food for our increasing population, it is important that any information about the theory and practice of drainage for agriculture be accurate. The management and design of drainage systems depend upon the reliability of the field measurements of soil-water properties that display a wide range of variability and classical statistical methods produce an incomplete description of this variability. Geostatistics recognizes these difficulties and provides the statistical tools for calculating the most accurate predictions of soil-water properties by making use of, yet powerful, tools for quantifying the accuracy of these predictions. It utilizes the fact that variations in soil-water properties are not always random, but have some spatial structures that may be expressed in a powerful mathematical form.

This article shows, in a brief look, how geostatistics could be one of the most effective tools for subsurface drainage management.

INTRODUCTION

Soil scientists who make and interpret soil maps deal with soil variability as a routine part of their jobs. In fact, if it were not true that soils vary from place to place on the landscape, there would be no need for soil scientists! Soils of course do vary from place to place, and it is the job of the field soil scientist to make some sense of the variation that occurs. How do we do this? To answer this question we must consider the nature of soil variability.

To the layman, the fact that the soil in one farm field is different from the soil in another field is often a mystery. To the soil scientist, these soil differences are understandable. This is because the soil scientist views soil variability as a function of the interaction of the five factors of soil formation. We consider the combined effects of climate, parent material, vegetation, landscape position and time and are able to understand and even predict (within limits) the nature of the soil at a given location.

The process is one in which we begin with an aerial photograph covering some geographic area which contains considerable variation in soil-water properties (e.g., hydraulic conductivity, soil salinity, water table depth, groundwater salinity, etc.). Based on an understanding of soil genesis for the area, the soil scientist partitions the variability of the entire area into a collection of delineations on a map where the variation within a delineation is less than the variation of the area as a whole.

The variation that can be understood and predicted by our knowledge of the factors of soil formation can be thought of as systematic variation (Wilding, 1985; Upchurch and Edmonds, 1991). The soil survey program has made extensive use of our knowledge of systematic variation to make reliable soil maps (Hudson, 1992). Even after partitioning the variability of soil properties into map units, however, we are still faced with variability within each of the map unit delineations. While we are doing field investigation for measuring soil-water properties necessary for subsurface drainage design, we are not able to explain why a soil property at one location is different than at another location a few meters away. This type of variability we would attribute to random variation (Wilding, 1985; Upchurch and Edmonds, 1991).

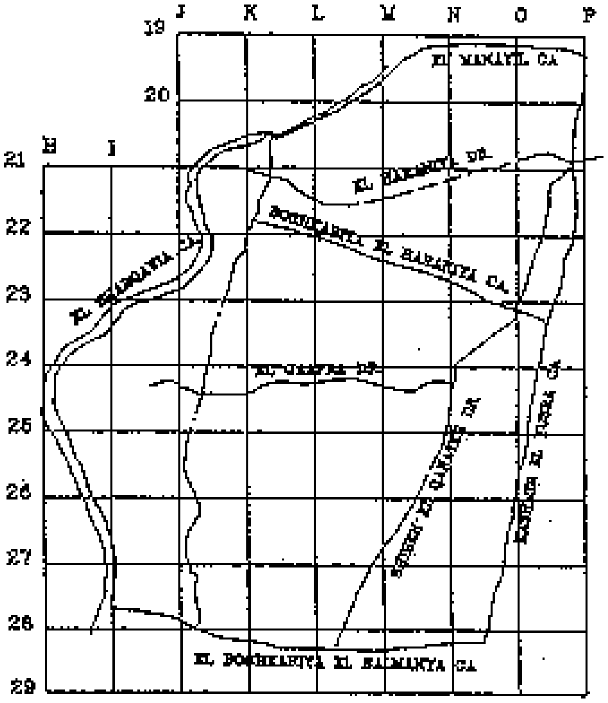

Figure 1. The grid system for field measurements of a drainage area.

GEOSTATISTICS AS A TOOL FOR DESCRIBING RANDOM VARIATION

The field of geostatistics originated during the 1950s in the South African mining industry. It was developed as a tool for estimation of gold content in ore bodies for potential future mining activities (Trangmar et al., 1985). What the South African gold miners knew (as do drainage specialists) is that while we may not be able to go to a location and predict precisely what the hydraulic conductivity or soil salinity will be at that spot (or the gold content in the case of mining), we intuitively know that once we learn what the value is at one location, a second observation very close by is likely to yield a similar result. As you go further away from your sampled location, the less able you are to predict what the value at the new location will be. Simply put, geostatistics is a technique for describing how a property varies with distance, how rapidly does the value change as you move across the landscape and how much total variability is there in the landscape.

The basic idea behind geostatistical analysis is as follows. A number of regularly spaced observations, N(h), separated by a lag distance h are made. These could be arranged as one or more transects, or more commonly, as a grid. Figure 1 shows an example of measurements locations, spaced every 500 m in a drainage area. In the geostatistics, the relation between variability and distance is quantified by the spatial correlation function, i.e., variogram, (Moustafa and Yomota, 1998; Moustafa, 2000) which might be estimated by:

$$\gamma(h) = C(0) - C(h) \tag{1}$$

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} z(x_i + h).z(x_i) - m_1.m_2 \tag{2}$$

where γ(h) and C(h) are the variogram and covariance function, respectively at a separation distance h, C(0) is the finite variance of the measured values and is assumed to be constant under the assumption of second-order stationarity and N(h) is the number of pairs of observations [z(xi), z(xi+h)] separated by the distance h. m1 and m2 are the means of z(xi+h) and z(xi) data values, respectively, and they are equal under the assumption of second-order stationarity. Second-order stationarity for the spatial structures of soil properties might be assumed in the analysis since it is all that is usually required in geostatistics (Olea, 1975). This assumption implies that the mean is the same everywhere and the covariance exists and is a unique function of separation distance h.

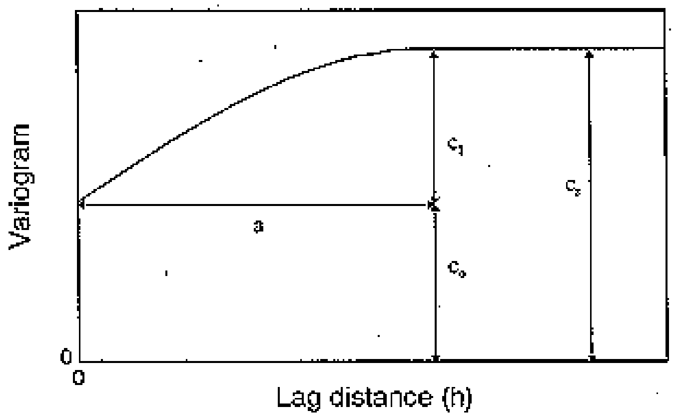

Figure 2. Spatial parameters of variogram.

After calculating variogram for all pairs of observations, the estimates are plotted on *Y*-axis of a graph with lag distance on its *X*-axis. Figure 2 shows a hypothetical variogram revealing spatial structure for a soil property with distance. The point where the curve intersects the *Y*-axis is referred to as the *nugget* (c_o). Intuition would suggest that the value of variogram at distance zero should be zero because this represents sampling the soil repeatedly at the same location. In practice, however, the plot of variogram often intersects the *Y*-axis at a point greater than zero, indicating that there is some inherent variability at distances less than the shortest distance used in the study area. Even if repeated observations could be made for the same location, the variogram is likely to be greater than zero and this estimate would be due to sampling measurement error.

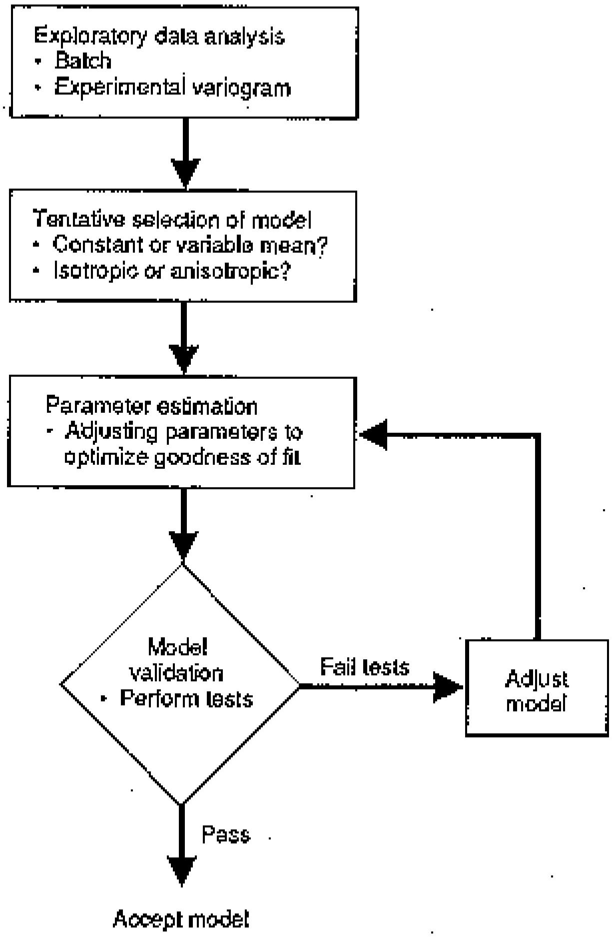

Figure 3. Flowchart of structural analysis for geostatistical model development.

The point at which the curve levels off is referred to as the *sill* (c_s). The value of variogram corresponding to the sill is a measure of the total variability for the property measured in the study area. As the distance between observations is increased beyond the distance where the sill begins, no further increase in variation is observed. The distance to the point where the sill is reached is referred to as the *range* (a). For distances less than the range it can be clearly seen that as points become further and further apart, more total variation is encountered, as indicated by the increasing estimates for variogram. At distances beyond the range, no additional variation is encountered.

APPLICATION TO DRAINAGE MANAGEMENT

In most geostatistical studies the variogram is modeled with a mathematical equation which in turn is used in a technique called *kriging*. This is a technique for calculating optimal-unbiased linear estimation of soil properties at unsampled locations with minimum estimation error variance. (Journel and Huijbregts, 1978). In addition, geostatistical analysis can be used to determine whether variation is different in one direction versus another (a condition known as anisotropy). Figure 3 shows a flowchart of structural analysis for geostatistical model development.

Many attempts have been made to infer spatial variability of soil-water properties by applying geostatistical technique for different practical applications (e.g., Moustafa and Yomota, 1998; Bracq and Delay, 1997; Fonteh and Podmore, 1994; Hosseini et al., 1994; Yost et al., 1982). In drainage management, the selection of the areas to be provided with subsurface drainage and the optimum drainage plans depend upon the reliability of the field observations of some soil-water properties. The decision of when and where to implement a drainage system is a very complex management process. For example, in Egypt, water table depth and soil salinity, depending on threshold values, play an important role in selecting areas that need subsurface drainage and determining their implementation priorities. Therefore, these two soil-water properties, which are inherently variable in both space and time, should be estimated as accurate as possible during feasibility stage and prior to design of the drainage system since a mistake made in subsurface drainage design is quite difficult to correct after its installation. In this case, an accurate estimation of soil properties based on limited field observations can be obtained with the geostatistics technique, and local estimates at unsampled points can be made from sampled data. The results are mainly presented as contour maps.

Table 1. Characteristics of water table depth for measured and kriged values.

Water table characteristics	Measured	Kriged
% of area where water table depth < 1.0 m	45	73
% of area where water table depth > 1.0 m	55	27

To illustrate this application, these two soil-water properties (i.e., water table depth and soil salinity) were sampled at 61 locations on a regular 500 m square grid for an area of 1533 ha located in West Delta of Egypt. Soil salinity samples were taken from depths of 0-25 (topsoil) and 25-50 cm (subsoil). Structural analysis was done to determine the spatial structures of these properties and the most appropriate spatial models were fitted to their variograms. Then, kriging technique was used to interpolate the original grid at a

100 m regular square grid covering the entire field using the spatial model and the neighboring observations. The contour maps of original measured and kriged values for water table depth and soil salinity were drawn and used to interpret characteristics of these properties as shown in Tables 1 and 2.

Table 2. Characteristics of soil salinity for measured and kriged values.

Soil salinity characteristics	Measured	Kriged
	Topsoil Subsoil	Topsoil Subsoil
% of area where soil salinity < 10 dS/m	58 52	85 77
% of area where soil salinity 10-12 dS/m	32 26	15 18
% of area where soil salinity > 12 dS/m	10 22	- 5

The primary criterion for selecting areas to be provided with subsurface drainage in Egypt is the proximity of the water table to the soil surface, where at least 75% of the area the water table depth is less than 1.0 m; this causes the saturation of a part of the root zone which decreases crop yields. High priority is given to areas where the soil salinity over a depth of 0-50 cm is above 4 dS/m at 25 ^{o}C. Following these criteria, areas that have drainage priority are ranked according to their soil salinity status.

Comparing this criterion with the characteristics of water table depth and soil salinity for both original observations and their kriged values as shown in Tables 1 and 2, we may conclude that the kriged estimates of water table depth almost met the criterion for providing the area with a drainage system, whereas the measured values underestimate the criterion by about 40% (Table 1). On the other hand, the measured values of soil salinity represent an overestimate of soil salinity levels of about 59% compared with their kriged values (Table 2).

These results reveal that based on the kriged estimates the study area is suffering from waterlogging and salinity problems, while based on the measured values the salinity problem is only the prevailing problem in the area. Since the kriged estimates are the optimal estimates and any necessary action should depend upon them (Moustafa, 2000), a subsurface drainage system should be implemented with high priority in the study area. Moreover, the results show how the geostatistics technique can be a very useful tool to determine the extent and severity of drainage problems over an area, and hence for proper decision regarding necessity and implementation priority of subsurface drainage in a region.

CONCLUDING REMARKS

Soil scientists have done a very effective job of understanding and mapping variation in soils arising from systematic variation across the landscape. Our understanding of the random variation found within our map unit delineations is not as good however, and our techniques for describing this variation are only moderately effective. Geostatistics technique can provide a tool for describing this variability. The use of geostatistics can help us to relate variation of important soil properties to the distance over which this variation occurs. This kind of information is potentially very useful to drainage professionals. It helps to maximize the accuracy of the management decisions and address the variability aspects of soil-water properties at a minimal cost owing to the need to visit only a few locations in the field and the ability to estimate the properties at unvisited locations accurately with minimal error.

References

Bracq, P. and Delay, F. (1997), "Transmissivity and morphological features in a chalk aquifer: a geastatistical approach of their relationship", J. of Hydrology, 191: 139-160.

Fonteh, M.F. and Podmore, T. (1994), "Application of geostatistics to characterize spatial variability of infiltration in furrow irrigation", Agric. Water Manage., 25: 153-165.

Hosseini, E., Gallichand, J. and Marcotte, D. (1994), "Theoretical and experimental performance of spatial interpolation methods for soil salinity analysis", Trans. ASAE, 37(6): 1799-1807.

Journel, A.G. and Huijbregts, C.J. (1978), "Mining Geostatistics", Academic Press, London, U.K.

Moustafa, M.M. (2000), "A geostatistical approach to optimize the determination of saturated hydraulic conductivity for large-scale subsurface drainage design in Egypt", Agric. Water Manage., 42: 291-312.

Moustafa, M.M. and Yomota, A. (1998), "Use of a covariance variogram to investigate influence of subsurface drainage on spatial variability of soil-water properties", Agric. Water Manage., 37: 1-19.

Olea, R.A. (1975), "Optimum mapping techniques", Kansas Geological Survey Series on Spatial Analysis, Kansas Geological Survey, Lawrence, Kansas.

Trangmar, B.B., Yost, R.S. and Uehara, G. (1985), "Application of geostatistics to spatial studies of soil properties", Adv. Agron., 38: 45-95.

Upchurch, D.R. and Edmonds, W.J. (1991), "Statistical procedures for specific objectives", in Mausbach M.J. and Wilding L.P. (ed.), Spatial variabilities in soils and landscapes, SSSA Spec. Publ. 28, SSSA, Madison, WI, pp. 49-71.

Wilding, L.P. (1985), "Spatial variability: Its documentation, accommodation and implications to soil surveys", in Nielsen D.R. and Bouma J. (ed.), Soil spatial variability, Proceedings, ISSS and the SSSA, Las Vegas, Produc, Wageningen, The Netherlands, pp. 166-189.

Yost, R.S., Uehara, G. and Fox, R.L. (1982), "Geostatistical analysis of soil chemical properties of large land areas. I. Semi-variograms", Soil Sci. Soc. Am. J., 46: 1028-1032.