projected sea level and tide scenarios for the mouth of the Delaware Bay, along with the projected freshwater inflows to the Delaware Estuary and Bay, as input to multidimensional flow and salinity models. The primary models are 2- dimensional, vertically-integrated models for flow and salt. The models use spectral decomposition in time in order to separate low frequency motions from tidal motions and thus provide a much less CPU intensive approach to analyzing a large number of sea level rise/ freshwater inflow scenarios. The salt flux in the salt model depends upon both tidal and nontidal processes. Tidal processes are approximated directly with the salinity-velocity correlations in the transport terms. Nontidal processes are approximated with an empirically-derived relation for the diffusivity.

This analysis of the salt field is one component of a larger project that will examine runoff in the entire Delaware River Basin. The results from this component will be used to specify the boundary conditions for intrusion of salt water into the Coastal Plain aquifers.

Prediction Using Geostatistics and Mathematical Models

Peter K. Kitanidis[1]
Member ASCE

Abstract

This paper discusses a new interpretation of geostatistical estimation and presents some thoughts on the meaning of probabilities in geostatistics and the role of mathematical models of groundwater flow or solute transport.

Introduction

Consider that the piezometric head (or some other variable, such as the coefficient of transmissivity), is measured at a set of irregularly spaced locations. Some prediction or estimation problems are: (a) From measurements of the piezometric head at given locations, estimate the value of the head at another point. Estimation on all the nodes of a fine grid is often required before using contouring subroutines. (b) From measurements of the piezometric head at given locations and measurements of transmissivity also at given locations, estimate the value of the transmissivity and of the piezometric head on the nodes of a fine grid. (c) From measurements of permeability, head, and porosity, determine the seepage velocity at selected points.

All of these cases involve measurements and unknowns. The problem is how to use the measurements to determine the unknowns. Unfortunately, in practically all cases, there are many solutions which are consistent with the measurements. The information provided by the measurements is incomplete in the sense that based on this information alone, one cannot find a unique answer

---

[1]Associate Professor, Civil Engineering, Stanford University, Stanford, CA 94305-4020.

to the posed question. For example, in the problem of interpolating the hydraulic head, there is more than one plausible piezometric head surface which passes through the available measurements. Which one of them should one use for interpolation purposes?

Some of the earlier methods in geohydrology and petroleum engineering were based on deterministic "direct inversion" methods. For example, consider the problem of estimating transmissivity in an aquifer under steady-state conditions with no sources or sinks given measurements of the piezometric head. Transmissivity can be viewed as the dependent variable and head as the independent variable in the partial differential equation which describes flow in such an aquifer. To obtain a unique solution from head measurements the piezometric head must be assumed known everywhere and the transmissivity must be known at one point on each streamline [Nelson, 1962, Emsellem and DeMarsily, 1971, Frind and Pinder, 1973]. If the available information is less than that, the solution is nonunique and can be very unstable. If the available information is more than that, e.g. the transmissivity is known at many points along the same streamline, some of the information is redundant or inconsistent.

The study of such "improperly posed" problems arising in physics and engineering is an interesting topic in applied mathematics. See Laurentiev [1967]. From a practical viewpoint, however, it would appear that the most important lesson from the study of the mathematical requirements to make an inverse problem well-posed or precisely specified is that these requirements are seldom satisfied in practice. Usually only a modest number of irregularly distributed and error-affected measurements of head and transmissivity are available, a far cry from meeting the strict requirements for a well-posed problem. Assuming that inference must be made with these measurements, how can a reasonably stable and well-behaved solution be obtained?

In a deterministic framework, one could make the problem well-posed by representing the unknown (e.g., transmissivity) through parameters in a way that inversion is possible. For example, the piezometric head may be represented as a polynomial of degree such as the number of the uknown coefficients is exactly equal to the number of measuremnts. This fix-up may be referred to as <u>parameterization</u>. It is equivalent to using <u>additional information about the structure</u> of the spatially distributed variable which restricts the type

of possible solutions.

The most important among the desired features of parameterization is that it should be consistent with what is really known about the variable but not predetermine the solution by using information that is not really there. Deterministic representations of the spatial variability of hydrogeologic parameters are not well-suited to meet this criterion. Available information about the spatial structure of geologic formations is usually not complete enough to be represented though deterministic functions without unduly influencing the final result. For example, assume that the aquifer is divided into zones of constant transmissivity. If the zone boundaries are really known ("using available information") this method will work well. However, if the boundaries are selected arbitrarily for purposes of regularization ("using information which is not really there"), the solution will be largely predeteremined by the choice of zones.

Probabilistic methods, on the other hand, deal directly with the nonuniqueness problem. They are well suited to describe incomplete information and to incorporate qualitative information about spatial variability of a parameter, such as "the variable is continuous" or "the typical length of fluctuations is 100 meters". Furthermore, these methods can deal with measurement error and yield measures of estimation precision.

<u>A Probabilistic Approach</u>

For cases of incomplete information, the usual deterministic approach runs into difficulties. If the information is incomplete, any search for the unique solution appears to be in vain. A much more straightforward and honest approach is to accept that there is more than one solution to the given problem. However, some solutions may be more plausible than others. Therefore, each possible solutions should be given a weight signifying its degree of plausibility.

This approach provides answers which are useful in making decisions. For example, one may use the weighted average of all possible solutions of the piezometric head at a point as the "best" estimate of the value of the actual piezometric head at the same point. Also, one could use the variance of all possible solutions of the piezometric head as a measure of predictive uncertainty. The larger the difference among plausible solutions, the less the accuracy of the best estimate.

However, we are left with the thorny issue of how to assign the weights of plausibility. The answer is simple in cases of complete symmetry. If two solutions appear equally plausible, each is assigned weight 1/2. The justification for this weighting is the <u>principle of indifference</u>, according to which if there is no logical reason to give more weight to one solution over the other, weight the solutions equally. More generally, if there are n possible solutions and we have no reason to think that one solution is more likely than any other, we should adopt equal weights of 1/n for all of them.

The problem may appear more complicated when there are solutions which are not equally likely. The fact that the solutions are not equally likely results from additional information which needs to be taken into account. In this case, the principle of indifference says that we should assign weights as uniformly as possible while being consistent with the available information. Assigning probabilities as uniformly as possible is mathematically expressed as the maximization of the entropy, a mathematical measure of the multiplicity of possible solutions.

The roots of this idea are very deep. The entropy function has played a fundamental role in the statistical mechanics of Gibbs [1902] and the information theory of Shannon [1948]. Jaynes [1957a,b] presented the maximum-entropy method of selecting distributions clearly using a modern statistical language. It is beyond the scope of this paper to discuss its significance in greater depth. However, much has been written by others. [See Levine and Tribus, 1979, Smith and Grady, 1985, Justice, 1986.]

Returning to the problem of estimating an unknown and spatially variable quantity, the information which is usually available may be classified into the following categories:

1. Information about the spatial structure of the uknown function. For example, in the case of piezometric head at the regional scale, it is expected that it trends from areas of recharge to areas of discharge and varies smoothly.

2. The principles of flow and transport in porous media. The piezometric head, the hydraulic conductivity, etc must satisfy continuity and Darcy's law. That is, even though there are many possible solutions, they must all be physically plausible.

3. The site-specific measurements. Although the head is not known everywhere, it is known at the locations of observation wells. That is, all plausible solutions must honor the data within the range of measurement error.

Thus one can formulate the problem of choosing weights for the possible solutions as follows:

> Assign probabilities as uniformly as possible subject to constraints of spatial structure, physics, and site-specific measurements.

Linear Geostatistics [Matheron, 1971] may be interpreted in this framework with constraints on both the structure, given by variograms or generalized covariance functions, and the site specific measurements. A methodology which allows the use of the flow equations to further constrain the problem and improve the accuracy has been presented in Kitanidis and Vomvoris [1983] and Hoeksema and Kitanidis [1984].

In geostatistics, probability should be interpreted as degree of plausibility given available information. That is, the probability distribution of hydraulic conductivity expresses what is known about the value of a function at a specific point in the one and only geological formation that is of interest: The actual one. This state of knowledge changes as new information becomes available. It would be incorrect, however, to suggest that a "degree of plausibility" interpretation necessarily entails "subjectivity". Mathematical discipline characterizes probability theory, no matter how probabilities are interpreted.

APPENDIX I. REFERENCES

1. Emsellem, Y., and DeMarsily , G., An automatic solution for the inverse problem, <u>Water Resour. Res.</u>, Vol 7, No 5, pp. 1264-1283, 1971.

2. Frind, E. O., and Pinder G. F., Galerkin solution of the inverse problem for aquifer transmissivity, <u>Water Resour. Res.</u>, Vol 9, No 5, pp. 1397-1410, 1973.

3. Gibbs, J. W., <u>Elementary Principles in Statistical Mechanics</u>, 1902, as reprinted in the <u>Collected</u>

Works of J. W. Gibbs, Dover Publications, 1960.

4. Hoeksema, R. J., and Kitanidis, P. K., An application of the geostatistical approach to the inverse problem in two-dimensional groundwater Modelling, Water Resour. Res., Vol 20, No 7, 1003-1020, 1984.

5. Jaynes, E.T., Information theory and statistical mechanics, Vol 106, pp. 620, 1957

6. Jaynes, E.T., Information theory and statistical mechanics, Vol 106, pp. 620, 1957

7. Justice, J.H., (ed) Maximum Entropy and Bayesian Methods in Applied Statistics, Cambridge Univ. Press, 319 pp.1986.

8. Kitanidis, P. K., and Vomvoris, E.G., A geostatistical approach to the inverse problem in groundwater modelling (steady state) and one-dimensional simulations, Water Resour. Res., Vol 19, No 3, pp. 677-690, 1983.

9. Laurentiev, M. M., Some Improperly Posed Problems of Mathematical Statistics, Springer Verlang, 1967.

10. Levine, R. D., and Tribus, M., (eds) The Maximum Entropy Formalism, MIT Press, 498 pp., 1979

11. Matheron, G., The Theory of Regionalized Variables and Its Applications, Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau, NO 5, 1971.

12. Nelson, R. W., Conditions for determining areal permeability distribution by calculation, Soc. Petrol. Eng. J., pp. 223-224, Sep 1962.

13. Shannon, C. E., The mathematical theory of communication , Bell Sys. Tech. J., Vol 27,pp. 379-423 and 623-656, 1948. Reprinted in The Mathematical Theory of Communication by C. E. Shannon and Weaver, The U. of Ill. Press, 1959.

14. Smith, C. R., and Grady, W. T., (edts) Maximum-Entropy and Bayesian Methods in Inverse Problems, D. Reidel Pub., 492 pp., 1985