D-1020

PDR-1
LPDR-
WM-10 (2)
WM-11 (2)
WM-16 (2)

WM Project 10,11,16
Docket No.
PDR ✓
LPDR B,N,S

Distribution:
Ross
Verma
Pohle
(Return to WM, 623-SS)

WM Record File D-1020 WM-RES WIA

STATISTICAL HYDROGEOLOGY--THEORY OR HYPOTHESIS:
A REVIEW OF "STATISTICAL THEORY OF GROUNDWATER FLOW
AND TRANSPORT: PORE TO LABORATORY, LABORATORY TO FORMATION,
AND FORMATION TO REGIONAL SCALE" BY GEDEON DAGAN,
WATER RESOURCES RESEARCH, 1986, VOL. 22, NO. 9, P. 120S TO 134S

by

Roy E. Williams
Professor of Hydrogeology
and Acting Head
Department of Geology and Geological Engineering
University of Idaho, Moscow, Idaho 83843

This paper constitutes my comments on Dagan, G., 1986, "Statistical Theory of Groundwater Flow and Transport: Pore to Laboratory, Laboratory to Formation, and Formation to Regional Scale", WRR, V. 22, No. 9, 120S-134S. My comments are presented from the point-of-view of a hydrogeologist who, along with colleagues and students, has spent four years working inside metamorphosed quartzite Precambrian rocks in which hydraulic conductivity is controlled primarily by fractures and bedding planes. Access to the inside of this population of rocks is via an underground, lead-zinc mine approximately 5,000 feet in depth. The drifts of the mine penetrate rocks at 200 foot intervals between the ground surface and the bottom of the mine. Please refer to Ralston and Williams (1985) for a presentation of some of the aspects of the subject study. Williams et al. (1986) present additional perspectives that are derived in part from this study. My comments follow:

1. Irregular Heterogeneity: Dr. Dagan states in the Abstract of his paper that the development of the statistical approach "has been motivated by the recognition of the fact that porous formations are heterogeneous, i.e., with properties which vary in an irregular manner in space." I suggest that this assertion is not a foregone conclusion. At best, we

3750

may be able to conclude that portions or subsets of populations of rock formations display hydraulic properties that vary in an irregular manner in space, particularly at small scales (pore scale). Freeze and Cherry (1979, p. 30) classify heterogeneities into three categories: layered heterogeneity, discontinuous heterogeneity and trending heterogeneity. Our work inside mine openings suggests that the true challenge to the hydrogeological scientist is to identify and classify the nature of the structural, sedimentary and stratigraphic geologic controls that have produced order among the apparent disorder. In many cases hydrogeologists have been forced to assume that properties vary in an irregular manner because we have been unable to identify the order in a manner that enables us to classify the populations of rocks into stratified sample sets or subsets for statistical analysis. This constraint is in large part a consequence of the fact that our scientific community has studied rocks through boreholes. Steinhorst and Williams (1985) present examples of two studies in which statistics were useful for identifying order within rocks when knowledge about site geology was used carefully in combination with statistical analysis. The utility of any statistical analysis of the hydrogeologic properties of rocks depends heavily on how the population of rocks is defined and on how that population is stratified for purposes of sampling. This procedure requires that the hydrogeologist be able to classify heterogeneities into a stratified sample set, some portions of which may constitute aquifers. The stratification usually precedes the sampling (testing) although this is not absolutely essential. Cochran (1977, Chapter 5) presents a discussion of stratified sampling procedures.

2. <u>Aquifers, Populations of Rocks, and the Classification of Heterogeneities</u>: Dr. Dagan does not provide a rigorous definition of his use of the work "aquifer;" concomitantly he uses the word loosely and inconsistently with the aforementioned concept of irregular heterogeneity. As explained above the delineation of an aquifer within a population of rocks requires the classification of heterogeneities in space. The creation of population substrata or subsets (aquifers, etc.) can be accomplished on the basis of professional judgement by hydrogeologists or it can be accomplished statistically by applying cluster analysis, MANOVA, canonical analysis and discriminant analysis to the entire sample population (see Steinhorst and Williams, 1985).

3. <u>Theory or Hypothesis</u>: The final sentence in Dr. Dagan's abstract states that, "In the concluding remarks it is submitted that the statistical approach to groundwater flow has become a comprehensive theory, beyond the stage of an Ad Hoc modeling technique." I submit that the statistical approach to groundwater flow may in fact have become comprehensive but it is not yet a theory. It is at best an hypothesis. The elevation of an hypothesis to the level of a valid theory that is applicable to rocks requires significantly more reliable hydrogeologic data than the data that are available currently in our scientific discipline. In many cases the data used in statistical analyses that have been conducted to date were collected by persons other than the author of the paper (see Hoeksema and Kitanidis, 1985, for example, in which <u>all</u> the data were collected by others). Under these conditions the reliability or applicability of test data is questionable at best. Dr. Dagan relies heavily on the paper by Hoeksema and Kitanidis (1985) throughout his development of statistical theory for groundwater flow.

This reliance is important; consequently portions of the data base for that paper are discussed more extensively under item 5 below. The study by Hoeksema and Kitanidis (1985) illustrates the dangers of data collected by one investigator for one purpose being used for another purpose by a second party.

4. Scale: Dr. Dagan discusses three levels of scale in his paper. But these levels of scale are defined very loosely. He implies (p. 122S) that pore scale data can be elevated to local and regional scale data simply by applying small scale test data to bigger and bigger chunks of rock. This transition may be valid for some rocks whose hydraulic conductivity is homogeneous and ubiquitously very low; but for most rocks tests conducted at different scales test different chunks of rock among which the characteristics of hydrogeologic coefficients do not have the same meaning. In all rocks except those with ubiquitously small and homogeneous hydraulic conductivity different scale tests reflect different degrees of connectedness of different openings. Large scale and small scale tests are not comparable except by subjective professional judgement about the geologic structure of the rock being tested. A scientist who is attempting to design test programs for characterizing so-called heterogeneous rock would have great difficulty in making the transition from Dr. Dagan's hypothetical scale concept to the experimental scale concept. Dr. Dagan's concept of scale appears to allow him to accept and compare whatever scale of test data that are available to him. I submit that this approach is a dangerous one. Our work inside the rocks of our research mine suggests that the type of test selected controls to a considerable extent the value of the hydraulic property that is measured for a given portion of

the rock. Stated in more direct terms, if we were to rely on laboratory tests of cores to obtain values of saturated hydraulic conductivity, our observations about geologic structure in the vicinity of the mine would suggest that we intentionally could obtain a substantially predetermined statistical distribution of values of saturated hydraulic conductivity for the rock by using our knowledge of the structural geology of the area to delineate subsets or substrata of the total rock population. If we were to employ larger scale tests and if we elected to locate those tests in selected portions of the rock that we have identified by structural geologic controls, we believe we could influence greatly the statistical distribution of the values of the measurements obtained. The concept of scale is very important to hydrogeologic analysis of rocks, but our discipline has not yet obtained sufficient data via carefully controlled experiments to understand completely the concept of scale. Williams (1985) and Moench (1984) have discussed the importance of scale in the interpretation of test data for purposes of rock characterization. The reader would be well advised to review Moench's comments about scale presented along with Williams (1985) in particular.

5. <u>Measurements</u>: The concept of scale is related directly to measurements. Dr. Dagan makes the following statement about stationarity and measurements.

> "We shall therefore assume that the random functions considered here enjoy some kind of stationarity, such that the statistical moments can be inferred from a single realization, by using measurements and some prior information derived from similar cases."

This assumption is critical to the validity of the application of Dr. Dagan's proposed statistical theory to rocks. The implications of this assumption are particularly important with regard to the manner in which the total population of rocks has been substratified (or not substratified) for purposes of sampling. Dr. Dagan refers the reader interested in the theoretical and practical aspects of inference of statistical moments of aquifer properties from field measurements to the aforementioned article by Hoeksema and Kitanidis (1985). When the reader studies the paper cited he learns that the measurements on which the assumption is based consist primarily of a combination of single well specific capacity tests and a few multiple well measurements of coefficients of transmissivity and storativity. Unfortunately Hoeksema and Kitanidis (1985) did not address the question of applicability of test design or substratification of the total population of rocks as it influences the statistical properties they investigated. They did not differentiate between injection tests, single well pumping tests, multiple well pumping tests and specific capacity tests. For purposes of illustration it is useful to investigate one of the principal studies from which Hoeksema and Kitanidis (1985) borrowed data (Kempton et al., 1982). That study reported transmissivity values from 288 tests on four different glacial drift aquifers in East Central Illinois. But when the reader researches that study he learns that only seven of the wells tested were designed or drilled by the investigators; they used existing water wells for the remainder of the study. Of these seven wells only three wells had "permeable material that was thick enough to set a screen and conduct a pumping test." This observation reflects the need for sample stratification among the data set if the study was intended for use beyond the reconnaissance level. Furthermore, specific capacity data were used to calculate transmissivity values for 194 of the tests (p. 26). Data for 107 of the wells were collected from the post 1948 files of the Illinois State Water Survey. Kempton et al. (1982) explain very clearly (p. 3) that

> "water well logs (drillers logs) provided the bulk of the information, supplemented by data from about 25 geophysical logs (not identified) of water wells."

Use of this type of data precludes assurance that the wells penetrated only one aquifer, that they fully penetrated that aquifer, that pretest water level trends were identified, and that a host of other factors prerequisite to characterizing the transmissivity of rocks beyond the reconnaissance level were evaluated. Indeed Kempton et al. (1982) never intended that their data base be used as a part of the basis for a major statistical theory of groundwater flow and transport. They intended their study to be a reconnaissance level study that would provide a general idea of the availability of water in a 50 by 80 mile area of the glacial drifts of East Central Illinois. Kempton et al. (1982) make the following statement about their project:

> "The second phase of the assessment attempted to locate suitable areas to explore for sources of supplementary water for each community. Because of the magnitude of the problem, however, it was decided that regional assessment rather than the detailed investigation of individual communities would be a more profitable procedure."

Nevertheless Hoeksema and Kitanidis (1985) state that

> "The best source of data (for their statistical analysis) was a series of Illinois Geological Survey and Illinois Water Survey Cooperative groundwater reports."

In accord with this precedent Dagan (1986) writes,

> "Fortunately Hoeksema and Kitanidis (1985) have undertaken the task of analyzing data from the literature for about 20 aquifers in the United States..."

Clearly the authors of neither paper bothered to check their data base for state of the art testing methodology that is internally consistent and compatible prior to using the data base to pursue their statistical hypotheses.

The purpose of this paper is not to review the well known difficulties and problems inherent in obtaining hydrogeologic coefficients of large volumes of rock from reconnaissance level single well tests. However a few comments appear to be in order. Specific capacity tests in particular constitute one of the most reconnaissance level tests among all field tests.

Walton (1970, p. 311) explains that the specific capacity of a well reflects some combination of aquifer loss, well loss due to turbulent flow, partial penetration loss, and recharge boundary gain. Walton states (p. 314), "The specific capacity of a well cannot be an exact criterion of the coefficient of transmissibility because ..." Consistent with the reconnaissance concept, Walton goes on to explain that

> "Because of the usefulness of <u>rough</u> <u>estimates</u> of transmissibility, an examination of the relation between the coefficient of transmissibility and specific capacity is useful",

which he provides. This author agrees with Walton's statement.

It is important to point out that the utilization of the combined results of a variety of different types and scales of tests for statistical analysis probably forces the distributions of the resulting hydrogeologic coefficients to be a log normal distribution, even though the population (the transmissivity of the rocks) may be distributed otherwise. The principal point to be made here is the fact that Dagan (1986) and Hoeksema and Kitanidis (1985) did not differentiate among measurements in terms of scale of test or applicability of test, even though the scale of the measurements can be expected to influence heavily the statistical moments of the hydraulic properties assigned to the rock being tested. It has long been accepted by experienced hydrogeologists that measurements on core for example are less desirable than single well test measurements and that single well test measurements are less desirable than large scale multiple well test measurements, if in fact multiple well tests are feasible under the hydrogeologic conditions at the site of the experiment. This acceptance is based on the widely held belief among hydrogeologists that larger scale measurements of the hydrogeologic coefficients of populations of rock are more valid at man's time scale of interest. Additional discussion about statistical substratification of rock populations and testing scale is beyond the scope of this review.

6. Average Data Values: Dr. Dagan states on page 122S of his paper that "The basic statistical operation is to take the ensemble average or the expected value which will be designated by the angle brackets." The key words in this sentence are "ensemble average" but other averaging concepts could be substituted. I submit that hydrogeologists have not demonstrated that the "ensemble average" (or any other average) of any hydraulic property is necessarily the critical element of interest in the theory of groundwater flow and transport at man's time scale of interest. For purposes of the prediction of groundwater flow velocities or of the prediction of transport characteristics we may very well be interested exclusively in some other value obtained from tests in which the scale is as large as the substrata or population of rocks will permit. This concept constitutes a very important subject on which much more research is needed prior to the acceptance of a proposed

comprehensive statistical theory for ground water flow and transport. To be more precise, highly conductive geologic features (faults, fold axes, buried stream channel deposits) as identified by single well tests (relatively small scale) may not be significant to the transport of dissolved solids if they have not been shown by large scale multiple well tests and careful geologic mapping to be connected into a hydraulically cohesive subset (substrata) of the total population of rocks. Small scale tests that in combination yield some type of average value of hydraulic coefficients may very well reflect only dead ended openings whose coefficient values are of little importance to contaminant transport at man's time scale of interest. The identification of the so-called fastest path is critical. It is unreasonable to bury the identification of the fastest path in statistical parameters of the total rock population. The importance of the identification of permeable pathways (fastest path) for contaminant transport is illustrated in Osiensky, Winter and Williams (1984) and in Molz et al. (1986).

7. <u>Deterministic Data for Use in Statistical Models</u>: A very strong argument that we should question a proposed comprehensive statistical theory of ground water flow and transport (as opposed to an hypothesis) centers on the nature of the test data on which any scientific theory is supposed to be built. All the field test data that exist about the hydrogeologic properties of rocks have been derived deterministically. Data on coefficients of hydraulic conductivity, transmissivity, effective porosity and storativity all must be derived from numerical or analytical solutions to partial differential equations. The solutions to these equations are defined as deterministic mathematical

models. The Theis equation, for example, is explicitly a deterministic model. No stochastic methods exist for measuring the hydraulic properties of rocks. Even the derivation of head data requires the development of a deterministic model of the rocks because one must ascertain that all the head data measurements intended for kriging, contouring, or any other type of comparison are derived from the same hydrostratigraphic unit (population subset or substrata). A variety of differential equations and a variety of solutions subject to various boundary conditions, internal conditions, and initial conditions are available for the designation of a deterministic model to fit the rocks being analyzed. But regardless of which solution is selected and applied to the acquisition of field data, certain well understood and well defined deterministic assumptions must be met (or at least assumed to be met) in order to derive values of hydrogeologic coefficients. These assumptions reflect the scientific investigator's professional perception of the rocks that he is testing. Some of the well known geologic conditions that are required for the acquisition of hydraulic property data are:

a. The rocks are homogeneous and isotropic (or anisotropic) with respect to transmissivity.
b. The aquifer being analyzed is infinite in areal extent.
c. The aquifer is not infinite in areal extent but is subject to known geologic boundary conditions (image well theory).
d. The well used for the test is fully penetrating (or partially penetrating).
e. The aquifer is leaky (or not leaky).
f. The aquifer is confined (or unconfined).

The point to be made with respect to the use of deterministic data for statistical purposes is that the deterministic assumptions and constraints have already been placed on the hydrostratigraphic unit (population subset) prior to the utilization of the data for statistical calculations. Consequently one must ask the question: What are the implications of transferring data acquired by deterministic methodologies (for which the assumptions about hydraulic properties have been fixed) to statistical models that by definition change the assumed characteristics of the rocks that were tested by the deterministic method(s)? Essentially all the analytical or numerical deterministic procedures that are required for the acquisition of field data bases require that the population or subset of rocks being tested be treated as homogeneous at the scale of the test. Moench discusses this issue at some length in Williams (1985). Heterogeneities must be defined (usually as some type of boundary) in order to analyze the deterministic drawdown curves or other types of curves that produce the data about the hydrogeologic coefficients of the rocks. Is it reasonable to ignore these required deterministic procedures and then use the resulting data bases to develop statistical models that characterize the rocks in terms of random spatial variability of hydrogeologic coefficients? This transition from deterministic data base to statistical treatment of the rocks clearly constitutes a giant step that should not be taken lightly. To my knowledge, the implications of this transition have not been addressed specifically in the literature. Dr. Dagan's paper does not address it. He assumes the issue away when he assumes that

> "The random functions considered here enjoy some kind of stationarity, such that the statistical moments can be

inferred from a single realization, by using measurements and some prior information derived from similar cases."

8. I suggest that the hydrogeologic scientific community be wary of falling into the trap of assuming that the statistical approach to characterizing hydraulic properties of populations of rocks is valid simply because the mathematics are valid when in fact the approach may not be applicable to the geologic structural, stratigraphic and sedimentary characteristics of rocks. We may be susceptible to this trap for two reasons. These reasons are:

    a. We have been hampered in our ability to conduct adequate scientific inquiry (experiments) about the hydrogeologic properties of rocks because our investigations have been restricted largely to the use of boreholes. In only a few instances have we conducted scientific experiments inside man-made openings in rocks where we can actually study geologic structural, stratigraphic and sedimentary heterogeneities. Such studies have proved to be very enlightening but we have only scratched the surface of their potential utility.

    b. Many investigators of the hydrogeologic properties of rocks have had no training in structural geology, sedimentology or stratigraphy. Training in structural geology in particular is essential to one's ability to conceptualize the hydrogeologic framework that constitutes the basis for defining populations and subsets of populations of rocks for purposes of testing. Heterogeneities usually are caused by structural geologic processes, by terrestrial sedimentary geologic processes, by geologic processes that produce marine stratigraphy or by combinations of all three processes. We need much more interdisciplinary cooperation in order to deal with these matters properly.

In summary I point out that the statistical approach to the analysis of hydrogeologic coefficients of rocks can be described as an effort to characterize the assumed disorder in the hydraulic properties of rocks. I suggest that heterogeneities within rocks usually are not random except possibly at the very smallest of scales (pore scale). I suggest that the nature of the test data that are available to us influences heavily the reported statistical characteristics of hydrogeologic coefficients. The true challenge to the science of hydrogeology is to identify the order among the apparent disorder. In exploration geology, for example, this order is worked out by the exploration geologist during the discovery of the ore body. The statistician enters the picture at the scale of the ore body after the large scale order has been worked out. If we are to do likewise with the purely deterministic testing methodologies currently available to us, it will be necessary to utilize the principles of structural geology, stratigraphy and sedimentology. Stratification of populations of rocks into analyzable homogeneous subsets requires the implementation of these principles. Concepts such as correlation structure, cannot be explored without these tools. Only preliminary field efforts have been directed at studying the relationship between geology and correlation structure to date (see for example Byers and Stephens, 1983). In addition it may be very difficult to identify the order among the disorder as long as we are restricted to the acquisition of data from boreholes at the earth's surface. More experiments need to be conducted inside man-made openings in rocks according to state of the art scientific methodologies, which requires a great deal of time and money. We ought to remember the advice of Goethe who said "To reason without (reproducible experimental) data is folly" (parentheses mine).

Finally I point out that all the above comments, with the possible exception of comment number 7, could be accommodated in Dr. Dagan's paper if it were presented in a different light. Such a revision could constitute significant guidance toward the collection of field data that could prove useful for purposes of the evaluation of his hypothesis.

References Cited

Byers, Elizabeth, and Stephens, D.B., 1983, Statistical and Stochastic Analyses of Hydraulic Conductivity and Particle-Size in a Fluvial Sand. Soil Sci. Soc. Am. J., vol. 47.

Cochran, W.G., 1977, Sampling Techniques. 3rd ed., Wiley, N.Y., N.Y.

Dagan, Gedeon, 1986, Statistical Theory of Groundwater Flow and Transport: Pore to Laboratory, Laboratory to Formation, and Formation to Regional Scale. Water Resources Research, vol. 22, no. 9, p. 120S-134S.

Freeze, R.A., and Cherry, J.A., 1979, Groundwater. Prentice-Hall, Inc., Englewood Cliffs, NJ, 604 p.

Hoeksema, R.J., and Kitanidis, P.K., 1985, Analyis of the Spatial Structure of Properties of Selected Aquifers. Water Resources Research, vol. 21, no. 4, p. 563-572.

Kempton, J.P., Morse, W.J., and Visocky, A.P., 1982, Hydrogeologic Evaluation of Sand and Gravel Aquifers for Municipal Groundwater Supplies in East-Central Illinois. Illinois State Geological Survey and Illinois State Water Survey, Champaign, IL, Cooperative Groundwater Report 8, 59 p.

Moench, A.F., 1984, Double-Porosity Models for a Fissured Groundwater Reservoir with Fracture Skin. Water Resources Research, vol. 20, no. 7, p. 831-846.

Molz, F.J., Oktay, G., Melville, J.G., Crocker, R.D., and Matteson, K.T., 1986, Performance, Analysis, and Simulation of a Two-Well Tracer Test at the Mobile Site. Water Resources Research, vol. 22, no. 7, p. 1031-1037.

Osiensky, J.L., Winter, G.V., and Williams, R.E., 1984, Monitoring and Mathematical Modeling of Contaminated Ground-Water Plumes in Fluvial Environments: Ground Water, vol. 22, no. 3.

Ralston, D.R., and Williams, R.E., 1985, Hydrogeologic Aspects of the Abandonment of An Underground Lead-Zinc Mine. Proceedings of the Second International Congress of the International Mine Water Association, September, 1985, Granada, Spain, p. 777-788.

Steinhorst, R.K., and Williams, R.E., 1985, Discrimination of Groundwater Sources Using Cluster Analysis, MANOVA, Canonical Analysis and Discriminant Analysis. Water Resources Research, vol. 21, no. 8, p. 1149-1156.

Walton, W.C., 1970, Groundwater Resource Evaluation. McGraw Hill, N.Y., N.Y., 664 p.

Williams, R.E., 1985, Comment on "Double-Porosity Models for a Fissured Groundwater Reservoir with Fracture Skin" by Allen F. Moench. Water Resources Research, vol. 21, no. 6, p. 889-891.

Williams, R.E., Winter, G.V., Bloomsburg, G.L., and Ralston, D.R., 1986, Mine Hydrology. Society of Mining Engineers, AIME, Caller No. D, Littleton, CO 80127, 169 p.