An Overview of Stochastic Modeling of Groundwater Flow and Transport: From Theory to Applications

"Philosophers have only interpreted the world, in various ways; the point, however, is to change it."—Karl Marx, Theses on Feuerbach, XI

"The maxim 'Nothing but perfection' may be spelled "Paralysis."—Winston Churchill

The stochastic modeling of groundwater has developed considerably in the last twenty years and a large body of knowledge has accumulated. However, there is also room for criticism: the community has evolved into a specialized one and research has started to address esoteric topics. Furthermore, in spite of its expansion, the stochastic approach hasn't yet become a routine tool of hydrological modeling.

The main message of this article is an appeal for application of stochastic modeling to practical problems on a regular basis (see motto above). This is regarded as a task of utmost importance and high priority. Furthermore, in spite of outstanding problems, it is surmised that the accumulated knowledge makes it possible.

The stochastic approach in modeling groundwater flow and solute transport regards the aquifer properties and the parameters that influence flow and transport as random. The randomness reflects the uncertainty of their values: the most common example is the hydraulic conductivity K that varies in space by orders of magnitude in a seemingly erratic manner. The field data based on measurements are generally scarce and permit estimating K in statistical terms only. The same is true for many other properties of heterogeneous formations (e.g., storativity, pore-scale dispersivity, reactive properties, natural recharge, transport initial conditions, aquifer geometry, etc.).

The pdf (probability density functions) of properties and parameters serve as input to the quantitative modeling of flow and transport, resulting in stochastic differential equations for the dependent variables (pressure head H, water Darcian velocity **V**, solute concentration C). As a result, the latter can also be characterized only statistically by their pdf or in a more restricted manner by a few moments (mean, variance, etc.). Prediction is therefore subjected to uncertainty and aquifer management under risk is the appropriate approach. This is in contrast with the traditional deterministic modeling of groundwater flow and transport.

The emergence of stochastic modeling in a hydrological context can be traced back to *Freeze* [1975], though earlier theoretical developments were present in the monographs by *Shvidler* [1962] and *Matheron* [1967].

The research of the topic has known a tremendous impetus as manifested for instance in the number of articles published in the literature. Thus, in a review article [*Dagan*, 1986], the total number of WRR papers related to the subject in the years 1975–1984 was around 120. In 2000 solely, the number was around 60 and above 180 in all professional journals.

Another indication of the mature stage of development of the field is the publication of a few monographs that summarize the subject: e.g., *Dagan* [1989], *Gelhar* [1994], *Dagan and Neuman*, [1997], *Kitanidis* [1997], *Cushman* [1997], and *Rubin*, in press.

What are the causes of the flurry of research in this field?

First of all, the public concern about groundwater pollution (an indirect expression of that is the release of two Hollywood popular movies around the subject in the last few years). This concern was translated into increased research funding by various public sources. In this context, it was realized that aquifer heterogeneity has an important impact on spreading of pollutants in groundwater. Another factor that facilitated advanced modeling is the availability of numerical codes and large computation facilities. Those codes require detailed information as input and can handle it.

Last but not least, the subject has posed challenging theoretical problems that call for development of innovative mathematical and physical concepts and tools.

A few main achievements of the research so far are the development of new conceptual and modeling tools, a new perspective on aquifer characterization, design and analysis of elaborate field experiments, application to a few major projects (mainly nuclear repositories), and advancement of the scientific knowledge.

Essentials of the Stochastic Approach

Porous formations are heterogeneous at a variety of scales, the smallest one being the pore-scale ($10^{-4} - 10^{-3}$ meters). Averaging flow and transport variables over this scale leads to the macroscopic equations of flow (mass conservation, Darcy's Law) and transport (solute mass conservation), to be satisfied by the dependent variables H, V, and C. The macroscopic properties of the medium (porosity, permeability, pore-scale dispersivity, etc.) appear as coefficients in these equations.

Natural formations (aquifers, petroleum reservoirs) are heterogeneous at a much larger scale ($10^{-1} - 10^{4}$ meters). Large-scale modeling regards formations as continua and the macroscopic equations as valid pointwise. The aim of modeling is to solve the equations in order to obtain solutions for the upscaled dependent variables $\overline{H}, \overline{V}$, and $\overline{C}$ that are space averaged over different scales of interest, from pointwise to the entire formation.

Due to the availability of alternative models and the randomness of various parameters and properties, the solutions $\overline{H}, \overline{V}$, and $\overline{C}$ are subjected to uncertainty of a few types.

The first type is the conceptual one, which is difficult to quantify. Such uncertainty stems for instance from the freedom of choice of models; e.g., discrete versus continuous medium for fractured rocks, three-dimensional (local) versus two-dimensional (regional) for aquifers, parametric (boundary and initial conditions) versus distributed (spatially variable properties), wide unimodal distributions versus bimodal ones of aquifer properties, unconditional versus conditional (on data) probabilities, etc. The examination of different conceptual models and quantification of the ensuing uncertainty is a topic that has received little attention, but need be considered in applications.

The complex source of uncertainty, which has also received most attention, is spatial variability, mainly of the conductivity $K(\mathbf{x})$, that is modeled as a RSF (random space function). The conductivity of the actual aquifer is regarded as a realization that belongs to an ensemble that is compatible with observations. The aim of stochastic modeling is then to determine the statistical moments of the upscaled dependent variables $\overline{H}, \overline{C}$ for given statistical moments of $K(\mathbf{x})$, (direct problem) and similarly for the inverse problem.

If upscaling is at a much larger scale than the one characterizing heterogeneity, the only needed information consists of the ensemble means $\overline{H}, \overline{C}$ that are approximately equal to $\overline{H}, \overline{C}$ by ergodic considerations. Then, the solution is obtained by replacing the actual formation by a homogeneous one of effective properties. The derivation of the latter is one of the central problems of the theory of heterogeneous media. However, this is not the typical case in groundwater applications, and higher-order statistical moments of $\overline{H}, \overline{C}$ have to be determined to quantify uncertainty.

Parametric uncertainty, stemming for example from errors of estimation of recharge, mean conductivity, etc. are impacting the solution of the dependent variables as well. Its impact can be assessed in a simpler manner by using classical statistical tools.

Main Methods of Solution

Two main methods were used in the past.

Monte Carlo Simulations. Replicas of the formation are computer-generated and the flow and transport problem are solved repetitively to generate the ensemble of dependent variables values. Advantages: conceptual simplicity, generality, and comprehensive characterization of solutions. Limitations: the method is numerically intensive especially for three-dimensional simulations, upscaling is needed to increase elements size, the entire statistical structure (the joint pdf of K values at nodes) must be given (in practice, stationarity and multi-Gaussianity are assumed), conditioning on data is simple for K, but not so for H and C.

First-order approximation in logconductivity variance (weak heterogeneity). The equations of flow and transport are expended in a power series and a first-order approximation is sought. Advantages: equations are considerably simplified by linearization and analytical or semi-analytical solutions are feasible, the mean and second-order statistical moments of dependent variables depend only on the mean and autocorrelation of logconductivity, conditioning on data is easy, and understanding of the processes is gained in a simple manner. Limitations: in principle, the results are valid only for small variances, though it was found that some of them apply to relatively high values, analytical solutions are not feasible for complex boundaries, non-uniform mean flow and unsteady conditions.

A Few Results and Outstanding Issues

Monte Carlo simulations were used mainly as a numerical laboratory to validate approximate solutions and to grasp the behavior of the solution, especially for strong heterogeneity. The simulations were confined mainly to two-dimensional configurations.

First-order analytical or semi-analytical solutions were obtained for unbounded domains and mean uniform flow. They provided two-point covariances and cross-covariances of the dependent variables that can be used conveniently for conditioning on data by co-kriging, for instance. These results were generalized to account for the presence of simple trends or boundaries and for transport of reactive solutes.

In spite of the tremendous progress, there are still many outstanding issues. A few examples: the need to improve formation characterization such as to provide the modeler with the needed data, processes occurring in highly heterogeneous formations are not well understood, the investigation of the impact of wells (strong nonuniformity of the mean flow) is at its beginning, dependence of solute transverse spreading on heterogeneity is still under debate and the same is true for unsteady transport. Procedures to upscale flow and transport variables are still under development and there is only qualitative understanding of the influence of density, two-phase fluids, and complex reactive properties of solutes.

From Theory to Applications

In spite of past achievements and the change of outlook of subsurface hydrology, stochastic modeling is not yet applied on a regular basis to aquifer modeling and management.

The main theses of the talk underlying this article were: it is important to start applying the approach, even in an approximate manner (see epilogue at beginning); first steps shall be carried out with relatively simple tools (first-order approximations) to be improved later; the best strategy is to develop stochastic modules to be attached to codes of widespread use; a multidisciplinary effort involving geohydrologists, geochemists, geophysicists, and stochastic and numerical modelers is needed.

A tentative outline of implementation comprises a few first steps:

(i) Geological characterization: improved use of soft geological data and geophysical methods; building of a catalog of statistical properties of formations of different geological structure;

(ii) Regional flow and transport (two-dimensional): identification of transmissivity statistical moments by hard and soft data and inverse procedure; generation of maps of log-transmissivity conditioned on transmissivity and head measurements; first-order numerical solution of flow supplemented by Monte Carlo simulations when computational resources are available; solution of transport by particle tracking with velocity fields generated from the solution of flow.

Summary and Conclusions

Stochastic modeling of subsurface flow and transport has reached an advanced stage of development of concepts and tools. In parallel with theoretical advances, it has been applied to aquifer characterization, to design and analysis of elaborate field experiments, and to a few major projects. Still, stochastic modeling has not yet become a tool used by the hydrological community to aquifer management on a regular basis. Our main thesis is that this is a timely task of utmost importance and a blueprint of implementation was outlined.

Acknowledgments

This article summarizes the opening lecture presented by the author at the International Groundwater Symposium "Bridging the Gap between Measurement and Modeling in Heterogeneous Media," organized by IAHR and co-sponsored by IAHS, AGU, and ASCE/EWRI at Lawrence Berkeley National Laboratory, Berkeley, California, 25–28 March 2002. I am grateful to the organizers of the symposium, and primarily to Chin-Fu Tsang and Angelos Findikakis for inviting me to present the opening lecture.

References

Cushman, J. H., *The Physics of Fluids in Hierarchical Porous Media: Angstroms to Miles*, 490 pp., Kluwer Academic Press, Dordrecht, The Netherlands, 1997.

Dagan, G., *Flow and Transport in Porous Formations*, 465 pp., Springer-Verlag, Berlin, 1989.

Dagan, G., Statistical theory of groundwater flow and transport: pore to laboratory, laboratory to formation and formation to regional scale, *Water Resour. Res., Special Anniversary Edition, 22(9)*, 120S–135S, 1986.

Dagan, G. and S. P. Neuman (Eds), *Subsurface Flow and Transport: A Stochastic Approach*, International Hydrology Series, 241 pp., Cambridge University Press, Cambridge U.K., 1997.

Freeze, R. A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res., 11*, 725–741, 1975.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, 390 pp., Prentice Hall, Englewood Cliffs, N.J., 1994.

Kitanidis, P. K., *Introduction to Geostatistics*, pp. 249, Cambridge University Press, Cambridge U.K., 1997.

Matheron, G., *Elements pour une Théorie des Milieux Poreux*, Masson et Cie, Paris, 1967.

Rubin, Y., *Applied Stochastic Hydrogeology*, Oxford University Press, Oxford, U.K., (in press).

Shvidler, M. I., Flow in heterogeneous media (in Russian), Izv. Akad. Nauk USSR Mekh. Zhidk. Gaza, 3, 185, 1962.

Author

Gedeon Dagan
Faculty of Engineering, Tel Aviv University,
Ramat Aviv, Israel