Published in Ground Water 45, issue 5, 531-541, 2007
which should be used for any reference to this work

Stochastic Hydrogeology: What Professionals Really Need?

Philippe Renard

Abstract

Quantitative hydrogeology celebrated its 150th anniversary in 2006. Geostatistics is younger but has had a very large impact in hydrogeology. Today, geostatistics is used routinely to interpolate deterministically most of the parameters that are required to analyze a problem or make a quantitative analysis. In a small number of cases, geostatistics is combined with deterministic approaches to forecast uncertainty. At a more academic level, geostatistics is used extensively to study physical processes in heterogeneous aquifers. Yet, there is an important gap between the academic use and the routine applications of geostatistics. The reasons for this gap are diverse. These include aspects related to the hydrogeology consulting market, technical reasons such as the lack of widely available software, but also a number of misconceptions. A change in this situation requires acting at different levels. First, regulators must be convinced of the benefit of using geostatistics. Second, the economic potential of the approach must be emphasized to customers. Third, the relevance of the theories needs to be increased. Last, but not least, software, data sets, and computing infrastructure such as grid computing need to be widely available.

Introduction

Among all the environmental applications of geostatistics, stochastic hydrogeology has emerged as an almost independent field of research during the past 30 years. Stochastic hydrogeology is the part of hydrogeology that deals with stochastic methods to describe and analyze ground water processes. A large part of stochastic hydrogeology consists of solving the stochastic partial differential equations describing these processes in order to estimate the joint probability density functions (pdfs) of the parameters (e.g., transmissivity, storativity, thermal conductivity) and state variables (e.g., ground water levels, concentrations, temperature) involved in those equations.

From a practical point of view, the main advantage of stochastic techniques is their ability to quantify the uncertainty inherent to any underground study (Winter 2004). It allows evaluating risks resulting from heterogeneity and lack of information on design and management. From a scientific point of view, stochastic hydrogeology is used to understand the impact of heterogeneity on processes and models. For example, it allowed deriving expressions to estimate effective governing laws for composite media and effective properties (Sanchez-Vila et al. 2006). Of course, the scientific and engineering sides of stochastic hydrogeology are tightly linked. For example, improvements in the understanding of effective behaviors had a large impact on simulation techniques such as the multiscale approach (Lunati and Jenny 2006) later used by practitioners.

Despite the large number of publications and textbooks (Dagan 1989; Gelhar 1993; Kitanidis 1997; Zhang 2002; Rubin 2003), stochastic hydrogeology is still not used routinely (Dagan 2002). To make things worse, Christakos (2004) emphasizes that this situation is very different from common practices in other fields such as petroleum engineering, fluid dynamics, meteorology, or surface hydrology. Thus, why stochastic hydrogeology is not used in practice has been a subject of debate, as illustrated by the forum sponsored by the journal *Stochastic Environmental Research and Risk Assessment* in 2004 (Zhang and Zhang 2004). More recently, Pappenberger and Beven (2006) discussed what they consider to be the

Centre of Hydrogeology, University of Neuchâtel, 11 Rue Emile Argand, CP 158, CH-2000 Neuchâtel, Switzerland; (41) 32 718 26 90; philippe.renard@unine.ch

seven main reasons for not using uncertainty analysis in hydrological modeling. They conclude that none of these reasons is tenable and propose a common code of practice.

The aim of this paper is to analyze the foundations of this debate and to discuss the following questions: Is it true that stochastic hydrogeology is a very active field of research? Is it true that stochastic methods are not used in practice? If yes, what are the reasons for this situation? Is there any possibility to change the situation?

The Emergence of Stochastic Hydrogeology

The birth of quantitative hydrogeology is generally considered to be the year 1856 when Henry Darcy published *The Fountains of the City of Dijon* (Darcy 1856), which contained the first description of the law governing the flux of water through porous media. Interestingly, the same year Paramelle published a book entitled *The Art of Discovering Ground Water* (Paramelle 1856). Paramelle's book was descriptive, analyzing ground water occurrence in thousands of places in France and developing empirical rules to infer the presence of ground water from geological and geomorphological observations. Paramelle's book was the best seller at that time (not Darcy's book).

During the first 50 years, quantitative hydrogeology was mainly dealing with the derivation of basic laws and analytical solutions. Most of the work was focused on ground water flow in porous medium assuming steady-state flow. This assumption started to be relaxed by the pioneering work of Theis (1935). Looking back at the reports of the 100th anniversary of the publication of Darcy's book (Association Internationale d'Hydrologie Scientifique 1956), one can see that during the years prior to 1956, most research in quantitative hydrogeology was focused on problems related to the analysis and modeling of ground water flow mainly with analytical and simplified expressions. The subsurface was assumed to be homogeneous as it was the only way to develop mathematical models that could be solved analytically and used in practice. During the last 50 years (between 1956 and 2006), hydrogeology has evolved in an amazing fashion. First, theoretical problems related to ground water flow became somewhat marginal, and the most important theoretical developments concerned solute transport and ground water pollution. In the early 1960s, numerical techniques handling heterogeneity started to be applied to model flow and transport (Warren and Price 1961). New concepts had to be developed to model, first, the behavior of conservative solutes. Then these models evolved rapidly to include sorption, decay, and geochemical reactions among other transport processes. In parallel, while heterogeneity was not too problematic for flow problems (in terms of overall fluxes), it appeared to exert a major control over flow paths and transport problems. A mathematical tool was then required to model the heterogeneity, map it, and investigate its impact on flow and transport. Theis (1967) commented, "I consider it certain that we need a new conceptual model, containing the known heterogeneities of the natural aquifer to explain the phenomenon of transport in ground water." Petroleum engineers also faced the same problem in the context of the simulation of multiphase flow.

Roughly in the same period, geostatistics started to become a well-established theory through the work of Matheron (1962, 1965). It is important to remember that Matheron started to develop the theory of regionalized variables in the mid-1950s mainly in the framework of mining, but he showed very early an interest in the application of stochastic theories to investigate the impact of heterogeneity on flow and transport through porous media (Matheron 1966, 1967; Matheron and de Marsily 1980; Delhomme and de Marsily 2006).

In 1975, Freeze published the first paper that analyzed one-dimensional (1D) flow in porous medium in a stochastic manner (Freeze 1975). His model was simple; permeabilities were piecewise constant values taken randomly (in a lognormal distribution) and independently in each grid cell. Freeze applied the Monte Carlo method to analyze the first moments of heads. He extended his study to a 1D transient consolidation problem. Although this is not the first contribution dealing with stochastic treatment of flow in porous medium (Warren and Price 1961; Matheron 1966), it is indeed the first publication of this type in the hydrogeology literature and it had a very large impact. A debate started soon after its publication (Dagan 1976; Freeze 1977; Gelhar et al. 1977), and it was also the beginning of the generalization of the stochastic concepts to flow in more than one dimension (Bakr et al. 1978), stochastic analysis of the forcing terms (Sagar 1978), first conditional simulations (Delhomme 1979), etc. It would be too long to detail all the history of stochastic hydrogeology here. Instead, the reader is referred to a series of review papers that cover most of the accomplishments made in this field (Neuman 1984; Dagan 1986, 2002; Gelhar 1986; Gutjahr and Bras 1993; Chilès 2001; Zhang and Zhang 2004; Carrera et al. 2005; de Marsily et al. 2005; Delhomme and de Marsily 2006). To conclude this section, it is important to remind the reader that the ground water community has been and is still leading the theoretical research on stochastic partial differential equations for flow and transport in porous media.

Why Care about Stochastic Hydrogeology?

Practical applications of hydrogeology are diverse: identification of potential locations for new water supply, estimation of water resources, design of water supply wells, aquifer protection, identification of contamination sources, design of remediation systems, aquifer management, design of dewatering schemes, etc. In all these projects, the approach used by the professionals is the same. They collect information that allows characterizing the geometry and the properties of the subsurface. The values of the state variables are measured and analyzed to understand the behavior of the system. On the basis of this understanding, conceptual and numerical models are built and calibrated. In the best cases, the values of the parameters are optimized through an inverse model in such a way that the model fits the observations and the estimated parameters are plausible (Alcolea et al. 2006).

In the end, the models are used to forecast and to optimize (in a broad sense) the decision variables (position of the wells, pumping rates, quantity of solvents, etc.) with respect to a given management-oriented objective function (maximum yield, minimum contamination, best water quality, etc.) and within a prescribed regulatory framework that constrains the optimization problem.

In all such projects, there is a considerable amount of uncertainty. The geometry of the limits between aquifers and aquitards is sampled at only a few boreholes. The rock and soil properties are measured in only a few points. The temporal evolution of source terms such as recharge or a contaminant spill is only indirectly estimated. A major fracture is observed only within a tunnel, and its extension is not known. A review of uncertainties affecting solute transport modeling is addressed by Carrera (1993). Thus, it is clear that in most situations, professionals have to deal with important uncertainties that are due to an incomplete sampling of a complex reality. According to Winter (2004), "all significant applications of hydrogeology are intrinsically uncertain … the only settings in which deterministic models are actually appropriate are primarily academic."

From an engineering standpoint, the quantification of uncertainty is extremely important not only because it allows estimating risk but mostly because it allows taking optimal decisions in an uncertain framework. The nonintuitive point is that optimal decisions often include nonsymmetric economical loss functions; in this case, the optimum decision is almost never the one that corresponds to an average estimation of the variable of interest. The important idea is that the complete pdf of the variable of interest must be known when the economical losses due to an overestimation or an underestimation of this quantity are not identical (Srivastava 1990). For example, the complete pdf of the possible amount of contaminated ground water to be remediated must be known to minimize the economical losses when deciding what amount the treatment plant should be designed for.

For all of these reasons, stochastic hydrogeology has an important role to play in the hydrogeological practice.

What Is the Situation Today?

The Research Perspective

It is generally accepted that stochastic hydrogeology is very active with a large quantity of accumulated knowledge as demonstrated by the large number of publications in the field (Dagan 2002). However, I argue that the situation is slightly different and, to be provocative, I even argue that stochastic hydrogeology is a minor part of the overall ground water research activities. Indeed, if one makes a search in one of the most recognized databases of scientific publications (ISI Web of Knowledge: http://isiknowledge.com), selecting articles which contain words related both to stochastic techniques and hydrogeology within their title, abstract, or keywords—the exact keywords used for the search were the following: ([ground water or hydrogeol* or aquifer] and [stoch* or statist* or geostat* or random])—and compares it with the total number of publications that relate only to the hydrogeological keywords, one finds that 5% of the publications relate to stochastic techniques (2800 out of 55,700 between 1980 and 2006). By comparing these numbers with the total number of papers including in addition references to modeling—the keyword used for the search was (model*)—one finds that 16% of all the modeling articles relate to stochastic techniques (2800 out of 17,800). While the total number of publications has dramatically increased over the past 15 years (Figure 1a) in all fields of research, it is surprising to observe that the proportions of stochastic hydrogeology papers remained constant over time (Figure 1b) both with respect to the total number of publications in hydrogeology and with respect to the modeling articles. This shows that there is not an increasing interest in these techniques within the research community. A more detailed analysis of those results shows a high variability of the proportions of

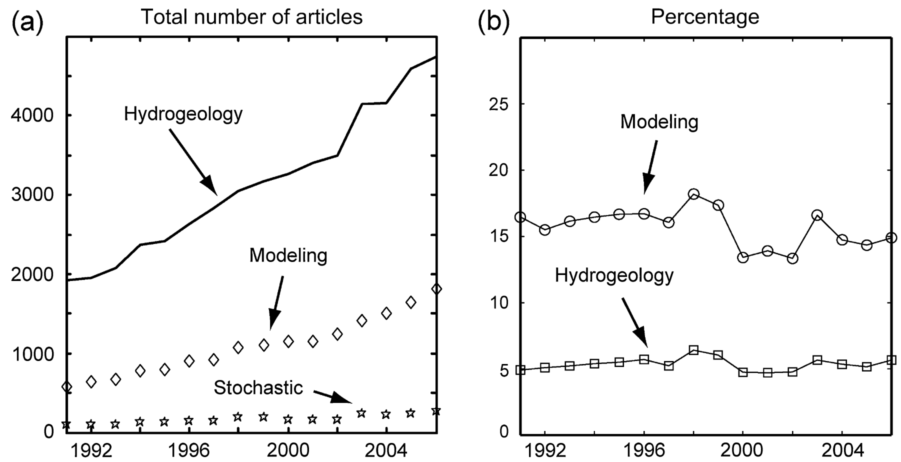

Figure 1. Time evolution of (a) the number of articles in international journals that relates to hydrogeology, hydrogeological modeling, and stochastic techniques; and (b) the proportion (in percent) of stochastic hydrogeology articles with respect to all hydrogeology articles or with respect to hydrogeological modeling articles (Source: ISI Web of Knowledge).

stochastic hydrogeology articles if we consider different journals.

Of course, these statistics are only rough indicators since they are just based on searches within a database. These words may not be sufficient to capture all the related papers and may capture papers that are not related to the topic. Nevertheless, I argue that these statistics indicate that stochastic hydrogeology is still a minor branch of research when compared on one hand with the total amount of ground water research and if we consider on the other hand that most real ground water problems have to deal with heterogeneity and uncertainty.

The Professional Perspective

The second point that must be discussed is the professional perspective. By professional, I mean an individual who is either a consultant, a hydrogeologist working for a government agency, or a scientist analyzing ground water systems but who is not a specialist in stochastic techniques.

In order to investigate if stochastic techniques are used by the professionals, a series of interviews was performed with colleagues working in 14 companies (16 persons) in Australia, Canada, France, Germany, Ireland, Italy, the United Kingdom, the United States, and Switzerland. The answers were more diverse than expected. They were different depending on the company and the country but also within different branches of the same company. The statement of Dagan (2002) is confirmed with 10 answers declaring that stochastic techniques are not used in practice. One of the answers provides a good feeling of what many professionals think:

> I would go as far as to say that stochastic hydrogeology has had no impact whatsoever in the practice of the great majority of professionals.

However, there were six answers declaring that stochastic techniques are used in 10% to 50% of the studies. These companies replied that they use stochastic techniques on a regular basis (not on the majority of their contracts but regularly). Among them, the pattern was quite clear. Some were located in some specific environmental sectors (e.g., waste disposal management) and in countries where regulations impose stochastic risk analysis. Again, a good illustration of that type of situation comes from one of the answers:

> Most hydrogeologists in the UK in waste management and to a lesser extent contaminated land are using a stochastic approach.

The same situation occurs in Ireland, and in Italy where regulations impose a minimum standard for risk assessment, which includes a Monte Carlo analysis. Under this pressure, software is developed, professionals are getting trained, and the whole profession is changing its standards. The other typical situation in which stochastic techniques are applied is when the clients have an important risk of economic losses. In order to minimize the risk, the client wants to evaluate the risks and is ready to invest in data acquisition and stochastic methods. The use of Monte Carlo techniques with either a numerical model or an analytical solution appears to be the standard methodology. First-order second-moment techniques were also mentioned by one of the interviewed teams. The applications are quite diverse, ranging from water supply to large-scale civil engineering projects such as horizontal well design for aquifer storage and recovery, tunneling, or contaminant migration problems.

Possible Reasons for the Lack of Application

Let us now try to understand why stochastic theories are not applied routinely. Obviously, there is not a single explanation but rather a set of reasons that act together. These reasons are organized in four main groups: structural reasons, invalid reasons, fundamental reasons, and technical reasons. But before going further, let us first discuss the question of time lag. Clearly, it makes sense to consider that the theories are available, already published in the scientific literature, and they will diffuse progressively into the practice. For example, the pilot points method was proposed in 1978 (de Marsily 1978; de Marsily et al. 1984), developed extensively in the 1990s (Lavenue et al. 1995; RamaRao et al. 1995), and is only available now, 20 years later, in the popular and widely distributed code PEST (Dougherty 2004). A time lag of more than 20 years was required for the application of a robust and successful method. Such a time lag is long as compared with the time life of a patent that is usually 20 years. Technology transfer for the same type of methods occurs much faster in petroleum engineering (5 to 10 years) or meteorology. Therefore, we have to look for other reasons to explain the lack of applications of stochastic hydrogeology.

The Structural Reasons

Economy and regulation are the two factors that structure the hydrogeology market and the work of professionals. Consultants must make profits to survive, and they must do activities that are either required by the regulations or that help their clients in solving practical problems. When the regulations do not demand the use of stochastic techniques or when the regulations impose single values as a base to make decisions (e.g., drinking water standards, travel time for protection zones), the majority of the professionals work in a deterministic manner. Similarly, under this type of regulations, when professionals have to provide expert opinions in courts, a probabilistic answer to a question is not acceptable. This argument is discussed in detail by Freeze (2004), Sudicky (2004), or Pappenberger and Beven (2006). But the situation is changing progressively (Rubin 2004; Winter 2004). This is the case today in the United Kingdom where the British Environment Agency is recommending the use of stochastic techniques for risk analysis at contaminated sites or in the design of new landfills. Driven by the regulations, consultants adapt their methods; they learn and apply stochastic techniques.

From an economic perspective, many companies describe a systematic decrease of available budgets for hydrogeological studies. The companies fight to propose

the smallest prices to their clients in order to survive in a very competitive market. Thus, the trend is to make studies faster, simpler, and cheaper. In this framework, most companies do not take the risk of investing in stochastic techniques that very few companies offer on the market.

Having still to deal with uncertainty, and while full uncertainty propagation with stochastic techniques is not feasible, professionals use alternative and cheaper techniques such as best-case/worst-case scenarios. Winter (2004) argues that, in many cases, overdesign (e.g., for a wellfield) is cheaper than a complete stochastic study. Overdesign may be an appropriate and economic solution to some well-known problems, but it is not always applicable. Winter (2004) mentions the limitations for mapping of contaminants. Another type of situation in which overdesign may be inappropriate is when there is not enough experience of failures to define accurately the safety factors. One issue of this type is the sequestration of CO_2 in deep aquifers. Other alternative methods to the stochastic ones exist, such as fuzzy logics (Guyonnet et al. 1999). However, these techniques are not used routinely either.

Economical constraints often hinder the applications of stochastic techniques, but economy can also be the driving force for the use of stochastic methods. This is the case in the petroleum industry but also in hydrogeology. The stochastic studies reported by consulting companies (in response to the e-mail interviews), which were not related to a regulatory pressure, were all conducted for clients that were concerned by important potential economic losses due to a lack of information. Often, the lack of application of stochastic techniques is certainly due to the fact that managers do not realize which are the economical risks related to a deterministic investigation or believe that uncertainty does not really matter in making the final decision as discussed by Pappenber and Beven (2006). To improve this situation, one simple thing that the academic world must do, and is already partly doing, is to publish stochastic studies including economical implications (Freeze et al. 1990; Srivastava 1990; Demougeot-Renard et al. 2004; Bayer et al. 2005; Bierkens 2006). Such studies can be used by professionals to demonstrate the economic interest of the stochastic approach to their clients.

Some Invalid Reasons

"God Does Not Play Dice with the Universe"

This famous quote from Albert Einstein illustrates a classical misunderstanding of stochastic hydrogeology concepts. For many professionals, the fact that there is only one single reality underground implies that it does not make sense to model it with random processes. This was a topic of hot debates in the development of geostatistics and is still one reason many professionals do not feel comfortable with stochastic approaches. The point here is a confusion between the objectivist and subjectivist interpretations of probabilities. In the subjectivist interpretation, the probability model is used to represent the degree of knowledge. In other words, stochastic hydrogeology does not claim that the world is random. It just claims that the world is partly unknown and that the lack of data requires the use of probabilistic techniques to estimate the corresponding uncertainty. In the objectivist interpretation, probabilities represent a truly random process such as a Brownian movement or a dice game and then probabilities can be calculated as frequencies of occurrence of some events. I argue that one important source of the origin of this misunderstanding is the fact that probabilities are taught in undergraduate classes essentially from an objectivist point of view.

Furthermore, hydrogeologists themselves are not sufficiently trained in the application of stochastic theories to be ready to apply them or to recommend them (Sudicky 2004). To the best of my knowledge, there are only few universities that provide stochastic hydrogeology classes in their programs. An Internet survey revealed that only 6% of the hydrogeology courses described online include a stochastic part.

Thus, education of future professionals must be one of the first priorities. Basic concepts such as the subjectivist interpretation of probability must be introduced in the early stage of the programs. Later, education should be focused on the engineering side of stochastic hydrogeology. Young professionals should be trained on how to apply those techniques to real problems. Teaching the theoretical and analytical tools is necessary for those who will pursue a career in research. However, these students are not the majority. Focusing on practical cases is, indeed, difficult as there are not so many available examples and not so many (friendly) tools allowing the students to work by themselves. The challenge is to make stochastic techniques understandable and usable by a vast majority of the students. Stochastic hydrogeology should not be a kind of obscure alchemy understandable only by the initiated.

Lack of Data

Most often, professionals must provide an answer without having the budget to collect a large data set. This motivates the paradox that when few data are available and the uncertainty is maximum, professionals tend to use a deterministic model (Neuman 2004; Rubin 2004). It does not mean that professionals do not care about uncertainty. Most often, they are cautious with their conclusions and point out that their answers are the best they can provide with the quantity/quality of available data.

While the lack of data is often real, it is also forgotten that data exist in other similar sites around the world. The idea of the "wwhypda" project (Comunian and Renard in press) is to facilitate the access to these data. This is not a novelty; Newell et al. (1990) built a database of typical values of hydraulic conductivities, seepage velocities, hydraulic gradients, etc., from a survey conducted at 400 sites in the United States. The aim was to provide basic data for Monte Carlo simulations in order to investigate the safety of waste disposals. Authors like Dagan (2002, 2004) or de Marsily et al. (2005) have also emphasized the need for such a catalog of statistical properties of earth materials to facilitate the application of stochastic

methods. The main difference between the previous projects and "wwhypda" is that it will be a collaborative environment open to user contributions. In the words of Diderot (speaking of the encyclopedia):

> This is a work that cannot be completed except by a society of men of letters and skilled workmen, each working separately on his own part, but all bound together solely by their zeal for the best interests of the human race and a feeling of mutual goodwill.

We argue that this is the only way for the hydrogeological community to build a sufficient knowledge base. Of course, "wwhypda" will not allow solving all statistical inference problems (we do not intend to replace the local data), but it should be a step forward to, at least, improve site characterization.

Some Fundamental Reasons

Irrelevance of the Theory

A very good reason for not applying stochastic theories is to consider them as irrelevant. For many professionals, the standard multi-Gaussian model is simply not adequate to represent the geological variability. The gap between the geostatistical models and the mental vision that the professionals have of the subsurface variability may prevent them from using these theories. Actually, from the early beginning of hydrogeology, there has been always two poles: on the one hand, engineers following the track opened by Henry Darcy, working mainly with equations and models, and on the other hand, geologists, following Paramelle, developing an in-depth (and conceptual) knowledge of the structures of the aquifers and their behavior. Up to now, these two approaches were very difficult to reconcile. For geologists, the most important features of an aquifer are the main channels, faults, boundaries, etc., and they concentrate most of their efforts in the field or with geophysical techniques to locate these structures. For stochastic modelers, efforts are devoted to finding the most accurate and most general relations between the various parameters appearing in the stochastic partial differential equations, but most often, the fields are represented by a simple multi-Gaussian random function. Such a representation allows finding (analytically or numerically) important relations and understanding the impact of heterogeneity. This is necessary but not sufficient to deal with real cases. The limits of the multi-Gaussian models have been frequently highlighted in the literature, and we will not revisit them here (Journel and Alabert 1990; Gómez-Hernández and Wen 1998; Zinn and Harvey 2003). The situation is changing today as new tools become available that can both be conditioned to local data and account for some general and geological constraints. At least three alternative techniques are following this track: the truncated plurigaussian method, the continuous-lag Markov chain, and the multiple-point approach. Most often, these methods decompose the simulation into two steps: first simulate the lithofacies and second simulate the property values within the facies.

In the author's opinion, the most appealing of these methods is the multiple-point statistics approach (Guardiano and Srivastava 1993; Strebelle 2002; Caers et al. 2003; Journel and Zhang 2006). I argue that it will radically change the application of stochastic hydrogeology; it has the potential to reconcile Darcy and Paramelle. The concept is simple and extremely flexible. The geologist provides a training image (can be three-dimensional [3D]) of the subsurface, which is a conceptual quantitative geological model. From this image, multiple-point statistics are calculated and used for simulation. The advantage of the multiple-point statistics as compared to traditional variograms or transition probabilities is that they integrate the possibility of modeling complex spatial relations between the facies. To illustrate its application in this paper, two stochastic simulations were generated with this technique (Figures 2b and 2d), with exactly the same conditioning data, but with two different conceptual models (Figures 2a and 2c). The crucial point here is that the point data do not contain enough

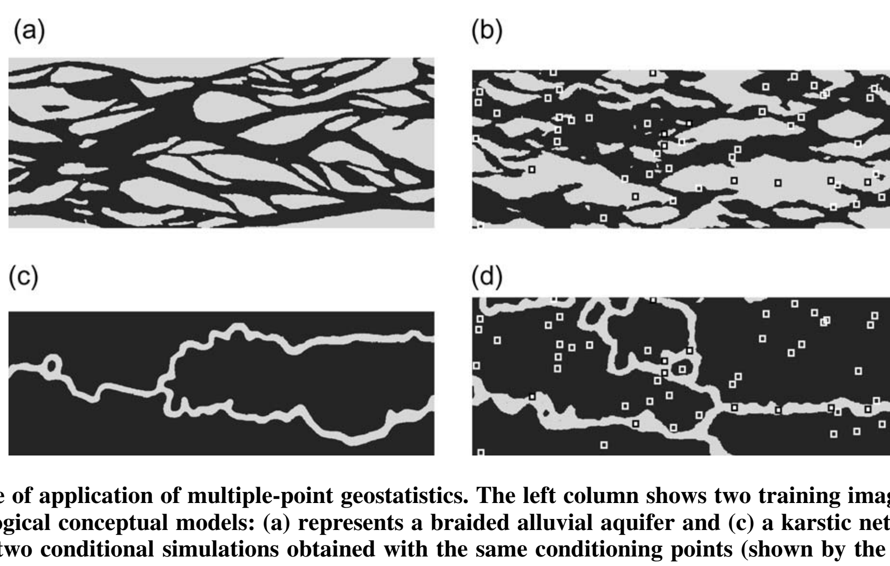

Figure 2. Example of application of multiple-point geostatistics. The left column shows two training images corresponding to two different geological conceptual models: (a) represents a braided alluvial aquifer and (c) a karstic network. The right column (b, d) shows two conditional simulations obtained with the same conditioning points (shown by the points in the figure) but with the two different training images (a, c), respectively.

information to distinguish if they have been sampled from one conceptual model or another. Assuming that the information necessary to distinguish between the different conceptual models can only be obtained from a variographic study is incorrect. This information must be provided by a geological analysis and transferred to the stochastic model. The multiple-point method allows doing that. To date, it has been applied in hydrogeology only for a small number of synthetic cases (Feyen and Caers 2003). A reformulation of the method handles both categorical and continuous variables (Zhang et al. 2006).

Another powerful approach is truncated plurigaussian simulation (Le Loc'h et al. 1994; Armstrong et al. 2003). Compared to the previous technique, this one is based on two-point statistics (rather than multiple points). The main advantage is that it allows geologists to describe qualitatively the relations between the lithofacies. For example, in the case of a fluvial deposit, one can impose that the levee will always be located between the channels and the floodplain but that the crevasse splay may cut all these lithotypes. Mariethoz et al. (2006) have used such a technique to build high-resolution 3D permeability and porosity fields for a contaminant migration problem in Switzerland (Figure 3).

The list of techniques that allow simulating realistic facies distribution could be easily extended here (see, for example, de Marsily et al. 2005), but this is not the objective of this paper. The important point is to show that there are available tools, which allow building stochastic models capable of representing the most relevant features of the internal architecture of aquifers. This represents a major step toward a better acceptance of stochastic models by the professionals.

Connectivity

Usually, geostatistics does not account specifically for connectivity, while it has been shown that it controls very strongly flow and transport (Zinn and Harvey 2003; Knudby and Carrera 2005). Multiple-point geostatistics is one technique capable of reproducing a global connectivity as shown in Figure 2: two training images with different types of connectivity (disconnected lenses or continuous channels) allow simulating fields that reproduce these types of connectivity. However, there is another aspect of connectivity that is not yet accounted for: the conditional connectivity. Tracer tests (conducted routinely) provide evidence that some points are connected by a continuous path of high conductivity values. This type of information is usually reproduced in stochastic ground water models by inverse techniques (Medina and Carrera 1996). The geostatistical permeability field is perturbed iteratively in order that the flow and transport model based on this permeability field reproduces the tracer test data. This is a widely accepted approach, but I argue that geostatistical simulation algorithms should allow imposing the connectivity between two or more points directly when these data are available. The geostatistical simulation technique should honor both the structural model (variogram or training image) and connectivity information. Up to now, only one technique (Allard 1994) has been proposed. It uses a Gibbs sampler and a truncated Gaussian model.

I proposed an alternative algorithm to impose connectivity (Renard 2006). The method is still in a preliminary stage. The main idea is to use a training image that can be either the training image used in a multiple-point framework or an unconditional simulation if the modeling technique is not based on the multiple-point principle (sequential indicator simulations or truncated plurigaussian simulations, for example). From the training image, a number of replicates of events that connect the points of interest are extracted. The simulation process consists of taking randomly one of these replicates and pasting it as conditioning data in the simulation grid before running the simulation algorithm. The method can be applied with any type of simulation method and is faster than the Gibbs sampler. A basic example of the results of using this methodology is shown in Figure 4.

Technical Reasons

Software

There is plenty of software for modeling ground water flow and transport in a deterministic fashion. The equivalent for stochastic models does not exist (Dagan 2004; Neuman 2004; Winter 2004). Nevertheless, this situation is slowly evolving (at least for Monte Carlo simulations).

MODFLOW2000 within GMS v5.0 (http://www.scisoft-gms.com/) allows running Monte Carlo simulations

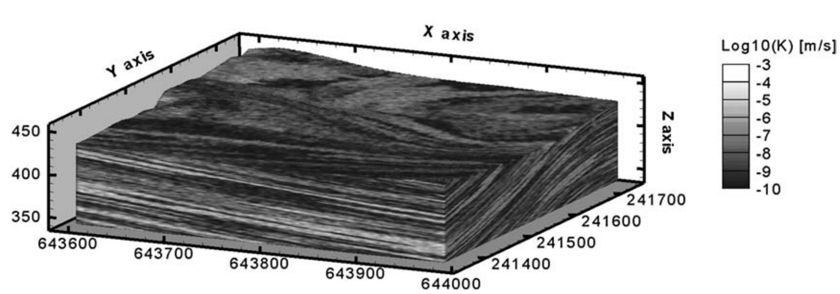

Figure 3. One stochastic simulation of a permeability field (3 million cells) for a contaminated site in Switzerland. The simulation represents a portion of a fluvial deposit containing five main lithofacies. The lithofacies simulations are then assigned physical properties that vary with different histograms and variograms within the facies. Scale is in meters (Mariethoz et al. 2006).

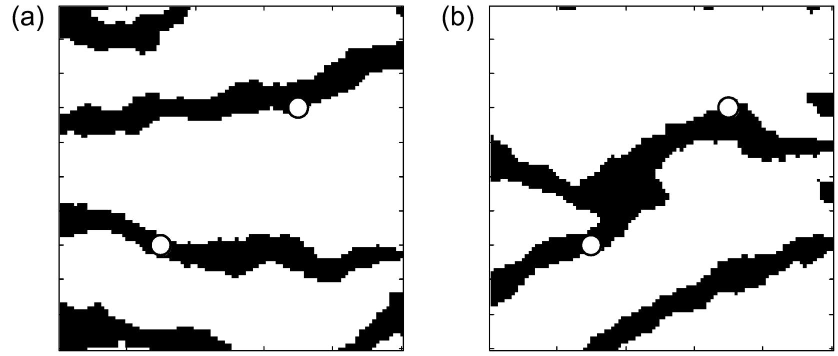

Figure 4. Example of a multiple-point simulation conditional to well data in two points (the circles): (a) simulation conditional to well data only, (b) simulation conditional to well data and a connectivity constraint between the two wells (Renard 2006).

and stochastic analysis as does Groundwater Vistas v4.0 (http://www.groundwatermodels.com/). There are commercial codes to run Monte Carlo simulations in a generic fashion such as GoldSim (http://www.goldsim.com/). There is software designed to conduct stochastic risk assessment for some specific problems such as landfills or waste disposal and even related to some specific regulations (http://www.landsim.co.uk/). There are also some semianalytical tools available. For example, PPpath allows postprocessing the results of a ground water flow model to estimate the uncertainty related to path lines or capture zones (http://www.ifu.ethz.ch/publications/software/pppath, Stauffer 2005). This is still very small compared to the number of deterministic software or purely geostatistical software available. In addition, Pappenberger and Beven (2006) point out the lack of information related to limitations and advantages of each of these codes and approaches. Overall, despite the current evolution, there is still a clear need to provide stochastic hydrogeology software for professionals.

Computing Resources

As revealed by the practitioners' interview, Monte Carlo simulation is the most general and most often applied technique to estimate uncertainties. While applying such a technique, a number of major difficulties arise, such as upscaling (Renard and de Marsily 1997) or downscaling of the geostatistical simulations onto an unstructured grid, or defining how many Monte Carlo simulations are required to obtain stable statistics (Ballio and Guadagnini 2004), but let us forget those difficulties for a moment. If the principle of the Monte Carlo procedure is straightforward, it requires a considerable amount of computing power to run a (serious) set of simulations. Often, professionals do not have or do not think they have access to the resources they need. This is not necessarily true. Over the past 5 years, the concept of grid computing has become a reality (Foster et al. 2001). All over the world, computer scientists developed software (middleware) and procedures that allow the user to send jobs in a transparent manner on a set of remote machines. Very large dedicated grids have been built in Europe, United States, and Japan (e.g., the EGEE project: Enabling Grid Computing in Europe for E-Science: http://www.eu-egee.org/). These grids are already used for research, and they are extremely well adapted to Monte Carlo simulations because every job can run independently on a computing node. Once all the jobs are launched, they do not communicate. The results have just to be collected, in the end, to extract meaningful statistics. An example of application in hydrogeology is shown in Kerrou et al. (2006). This technology is not just used in universities. Biomedical companies such as Novartis started to realize that the most important computing resources available in house are the PCs of the employees (from the general director to the secretary), and they set up an internal and private computing grid. To this end, open source codes, such as Condor (http://www.cs.wisc.edu/condor/), and commercial solutions are available.

Thus, we can safely state that currently, this technology can be used by professionals to bundle the already existing PCs in their company and make them widely available for Monte Carlo simulations. We cannot say that such a technology will solve all computing problems but it will be a step forward. Another possibility for small- or medium-size companies is to buy computing time within a computing grid when there is a need. This is a service provided by some companies and is based on grid technology (http://www.sun.com/service/sungrid/). At the academic level, or at a software development level, the target should be to ensure that the future stochastic hydrogeology codes will be as much as possible grid compatible.

Conclusions

Stochastic hydrogeology is still in its infancy as shown by the small number of published papers compared to the total number of publications in hydrogeology. Most of the research is hidden to the practitioners, and there is an important gap between theory and practice. Thus, there is still ample room for theoretical, software, and practical developments. Techniques such as multiple-point statistics may reconciliate the approaches of Darcy and Paramelle, i.e., quantitative and descriptive hydrogeology, but to bridge the gap between theoreticians and practitioners, some efforts still have to be made.

On the one hand, the academic community has to start by adapting teaching programs in order to ensure

that the basic concepts of stochastic hydrogeology are well assimilated at the undergraduate level. They must also offer the opportunity for professionals to learn how to use these techniques to solve real problems. By providing demonstrative examples of a wide range of applications, including cost/benefit analyses, they can participate with the professionals in the lobbying action that need to be conducted toward administration, regulators, and customers. Together with software companies, researchers have to promote a wider distribution of user-friendly software that allow running of stochastic analysis. Last, but not least, the research community can provide data that will allow practitioners to feed stochastic models even when they do not have sufficient data for a given site. This is possible to do if there is a concerted and collaborative effort to place the data already available in a centralized system such as "wwhypda" (http://wwhypda.org).

On the other hand, professionals have to become more familiar with the philosophy of stochastic theories. Stochastic research must leave the status of an obscure science (some kind of alchemy) to go to practitioners who need to appreciate the benefits of using such tools. Consequently, practitioners have an important role to play in convincing their administration and clients that stochastic techniques can help them to solve their practical problems in an efficient manner, better than the traditional deterministic methods. Overall, practitioners must realize that theoreticians are not "God" but, indeed, that both theoreticians and practitioners have to "play dice."

Acknowledgments

Funding for this work was provided by the Swiss National Science Foundation (contract PP002-106557) and the European Commission (degree project, contract 034619). I am very grateful to the generous feedback received from colleagues working in consulting companies and universities: F. Bertone, J.-P. Delhomme, H. Demougeot-Renard, H. Diersch, G.-P. Droli, J.-M. Côme, O. Jacquet, I. Krusic-Hrustanpasic, M.A. Lavenue, C. Neville, U. Ofterdinger, R. Pecci, V. Rabemanana, S. Rogers, M.-T. Schafmeister, J. Vermaak, and J. Watherhouse as well as to A. Alcolea for his early review of the manuscript and C. Simmons for providing the impact factors. I further would like to thank M. P. Anderson, G. Dagan, G. de Marsily, and D. Walker for their constructive reviews.

References

Alcolea, A., J. Carrera, and A. Medina. 2006. Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Advances in Water Resources* 29, no. 11: 1678–1689.

Allard, D. 1994. Simulating a geological lithofacies with respect to connectivity information using the truncated Gaussian model. In *Geostatistical Simulations: Proceedings of the geostatistical simulation workshop, Fontainebleau, France, 27-28 May 1993*, ed. M. Armstrong and P.A. Dowd, 197–211. Norwell, Massachusetts: Kluwer.

Armstrong, M., A. Galli, G. Le Loc'h, F. Geffroy, and R. Eschard. 2003. *Plurigaussian Simulations in Geosciences*. Berlin, Germany: Springer.

Association Internationale d'Hydrologie Scientifique. 1956. *Symposia Darcy (Dijon, 20-26 septembre 1956). Tome II.* Dijon, France: Eaux souterraines.

Bakr, A.A., L.W. Gelhar, A.L. Gutjahr, and J.R. MacMillan. 1978. Stochastic analysis of spatial variability in subsurface flows—1. Comparisons of one and three-dimensional flows. *Water Resources Research* 14, no. 2: 263–271.

Ballio, F., and A. Guadagnini. 2004. Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology. *Water Resources Research* 40, no. W04603: doi: 10.1029/2003WR002876.

Bayer, P., M. Finkel, and G. Teutsch. 2005. Cost-optimal contaminant plume management with a combination of pump-and-treat and physical barrier systems. *Ground Water Monitoring and Remediation* 25, no. 2: 96–106.

Bierkens, M.F.P. 2006. Designing a monitoring network for detecting groundwater pollution with stochastic simulation and a cost model. *Stochastic Environmental Research and Risk Assessment* 20, no. 5: 335–351.

Caers, J., S. Strebelle, and K. Payrazyan. 2003. Stochastic integration of seismic data and geologic scenarios: A West Africa submarine channel saga. *The Leading Edge* 22, no. 3: 192–196.

Carrera, J. 1993. An overview of uncertainties in modeling groundwater solute transport. *Journal of Contaminant Hydrology* 13, no. 1–4: 23–48.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L.J. Slooten. 2005. Inverse problem in hydrogeology. *Hydrogeology Journal* 13, no. 1: 206–222.

Chilès, J.-P. 2001. On the contribution of hydrogeology to advances in geostatistics. In *geoENV III—Geostatistics for Environmental Applications*, ed. P. Monestiez, D. Allard, and R. Froidevaux, 1–16. Dordrecht, The Netherlands: Kluwer Academic Publisher.

Christakos, G. 2004. A sociological approach to the state of stochastic hydrogeology. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 274–277.

Comunian, A., and P. Renard. wwhypda: A world wide hydrogeologic parameters database. *Hydrogeology Journal*. In press.

Dagan, G. 2004. On application of stochastic modeling of groundwater flow and transport. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 266–267.

Dagan, G. 2002. An overview of stochastic modeling of groundwater flow and transport: From theory to applications. *EOS Transactions AGU* 83, no. 53: 621.

Dagan, G. 1989. *Flow and Transport in Porous Formations*. New York: Springer-Verlag.

Dagan, G. 1986. Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale. *Water Resources Research* 22, no. 9: 120S–134S.

Dagan, G. 1976. Stochastic conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media—Comment. *Water Resources Research* 12, no. 3: 567–567.

Darcy, H. 1856. *Les Fontaines Publiques de la Ville de Dijon*. Paris, France: Victor Dalmont.

de Marsily, G. 1978. De l'identification des systèmes hydrogéologiques. Doctorat d'état thesis, Ecole des Mines de Paris, Centre d'Informatique Géologique, Paris.

de Marsily, G., F. Delay, J. Gonçalvès, P. Renard, V. Teles, and S. Violette. 2005. Dealing with spatial heterogeneity. *Hydrogeology Journal* 13, no. 1: 161–183.

de Marsily, G., C. Lavedan, M. Boucher, and G. Fasanino. 1984. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for Natural Resources Characterization*. NATO ASI Ser. C, 182, ed. G. Verly, M. David, A.G. Journel, and A. Marechal, 831–849. Dordrecht, The Netherlands: D. Reidel Publishing Co.

Delhomme, J.P. 1979. Spatial variability and uncertainty in groundwater flow parameters. *Water Resources Research* 15, no. 2: 281–290.

Delhomme, J.P., and G. de Marsily. 2006. Flow in porous media: An attempt to outline Georges Matheron's contributions. In *Space, Structure and Randomness: Contributions in Honor of Georges Matheron in the Fields of Geostatistics, Random Sets and Mathematical Morphology. Lecture Notes in Statistics*, ed. M. Bilodeau, F. Meyer, and M. Schmitt, 69–88. Berlin, Germany: Springer.

Demougeot-Renard, H., C. de Fouquet, and P. Renard. 2004. Forecasting the number of soil samples required to reduce remediation cost uncertainty. *Journal of Environmental Quality* 33, no. 5: Q03–0354.

Dougherty, J. 2004. *PEST Model Independent Parameter Estimation. User Manual*, 5th ed. Australia: Watermark Numerical Computing.

Feyen, L., and J. Caers. 2006. Quantifying geological uncertainty for flow and transport modelling in multi-modal heterogeneous formations. *Advances in Water Resources* 29, no. 6: 912–929.

Foster, I., C. Kesselman, and S. Tuecke. 2001. The anatomy of the grid. Enabling scalable virtual organization. *International Journal of High Performance Computing Applications* 15, no. 3: 200–222.

Freeze, R.A. 2004. The role of stochastic hydrogeological modeling in real-world engineering applications. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 286–289.

Freeze, R.A. 1977. A stochastic-conceptual analysis of one dimensional groundwater flow in non uniform homogeneous media—Reply. *Water Resources Research* 13, no. 2: 480.

Freeze, R.A. 1975. A stochastic-conceptual analysis of one dimensional groundwater flow in non uniform homogeneous media. *Water Resources Research* 11, no. 5: 725–741.

Freeze, R.A., J. Massman, L. Smith, T. Sperling, and B. James. 1990. Hydrogeological decision analysis: 1. A framework. *Ground Water* 28, no. 5: 738–766.

Gelhar, L.W. 1993. *Stochastic Subsurface Hydrology.* Engelwood Cliffs, New Jersey: Prentice-Hall.

Gelhar, L.W. 1986. Stochastic subsurface hydrology from theory to applications. *Water Resources Research* 22, no. 9: 135S–145S.

Gelhar, L.W., A.A. Bakr, A.L. Gutjahr, and J.R. MacMillan. 1977. Stochastic-conceptual analysis of 1-dimensional groundwater flow in nonuniform homogeneous media—Comment. *Water Resources Research* 13, no. 2: 477–479.

Gómez-Hernández, J.J., and X.-H. Wen. 1998. To be or not to be multi-gaussian? A reflection on stochastic hydrogeology. *Advances in Water Resources* 21, no. 1: 47–61.

Guardiano, F., and R.M. Srivastava. 1993. Multivariate geostatistics: Beyond bivariate moments. In *Geostatistics-Troia*, ed. A. Soares, 113–114. Dordrecht, The Netherlands: Kluwer Academic Publications.

Gutjahr, A.L., and R.L. Bras. 1993. Spatial variability in subsurface flow and transport—A review. *Reliability Engineering and System Safety* 42, no. 2–3: 293–316.

Guyonnet, D., B. Côme, P. Perrochet, and A. Parriaux. 1999. Comparing two methods for addressing uncertainty in risk assessments. *Journal of Ennvironmental Engineering* 125, no. 7: 660–666.

Journel, A., and T. Zhang. 2006. The necessity of a multiple-point prior model. *Mathematical Geology* 38, no. 5: 591–610.

Journel, A.G., and F.G. Alabert. 1990. New method for reservoir mapping. *Journal of Petroleum Technology* 42, no. 2: 212–218.

Kerrou, J., G. Lecca, P. Renard, and J. Tarhouni. 2006. Uncertainty in exploitation rates and numerical modeling of its impact on seawater intrusion in the Korba aquifer (TUNISIA). *First International Joint Salt Water Intrusion Conference, Cagliari—Chia Laguna, Italy.*

Kitanidis, P.K. 1997. *Introduction to Geostatistics: Applications in Hydrogeology.* Cambridge, U.K.: Cambridge University Press.

Knudby, C., and J. Carrera. 2005. On the relationship between indicators of geostatistical, flow and transport connectivity. *Advances in Water Resources* 28, no. 4: 405–421.

Lavenue, M., B.S. RamaRao, G. de Marsily, and M.G. Marietta. 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields 2. Application. *Water Resources Research* 31, no. 3: 495–516.

Le Loc'h, G., H. Beucher, A. Gali, B. Doligez, and H. Group. 1994. Improvement in the truncated gaussian method: Combining several gaussian functions, Ecmor 4. *4th European Conference on the Mathematics of Oil Recovery, Roros, Norway.*

Lunati, I., and P. Jenny. 2006. Multiscale finite-volume method for compressible multiphase flow in porous media. *Journal of Computational Physics* 216, no. 2: 616–636.

Mariethoz, G., P. Renard, F. Cornaton, and O. Jaquet. 2006. High resolution stochastic modelling of aquifers, example for a contamination migration problem. *Proceedings of the Swiss Geoscience Meeting, Bern, Switzerland, November 24–25,* 127–128. Bern, Switzerland: GEOForumCH.

Matheron, G. 1967. *Eléments Pour Une Théorie des Milieux Poreux*. Paris, France: Masson.

Matheron, G. 1966. Structure et composition des perméabilités. *Revue de l'IFP* 21, no. 4: 564–580.

Matheron, G. 1965. *Les Variables Régionalisées et Leur Estimation: Une Application de la Théorie des Fonctions Aléatoires aux Sciences de la Nature*. Paris, France: Masson.

Matheron, G. 1962. *Traité de Géostatistique Appliquée*. Paris, France: Technip.

Matheron, G., and G. de Marsily. 1980. Is transport in porous media always diffusive? A counterexample. *Water Resources Research* 16, no. 5: 901–917.

Medina, A., and J. Carrera. 1996. Coupled estimation of flow and solute transport parameters. *Water Resources Research* 32, no. 10: 3063–3076.

Neuman, S.P. 2004. Stochastic groundwater models in practice. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 268–270.

Neuman, S.P. 1984. Role of geostatistics in subsurface hydrology. In *Geostatistics for Natural Resources Characterization. Part 2*, ed. G. Verly, M. David, A.G. Journel, and A. Marechal, 787–816. The Netherlands: D. Reidel Publishing Co.

Newell, C.J., L.P. Hopkins, and P.B. Bedient. 1990. A hydrogeologic database for groundwater modeling. *Ground Water* 28, no. 5: 703–714.

Pappenberger, F., and K.J. Beven. 2006. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research* 42, no. W05302: doi: 10.1029/2005WR004820.

Paramelle, L'abé. 1856. *L'art de Découvrir les Sources*. Paris, France: Librairie Polytechnique.

RamaRao, B.S., M. Lavenue, G. de Marsily, and M.G. Marietta. 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. Theory and computational experiments. *Water Resources Research* 31, no. 3: 475–493.

Renard, P. 2006. Connectivity of binary media, measures, and simulation. Technical report. Neuchâtel, Switzerland: Centre of Hydrogeology of the University of Neuchâtel.

Renard, P., and G. de Marsily. 1997. Calculating equivalent permeability: A review. *Advances in Water Resources* 20, no. 5–6: 253–278.

Rubin, D.B. 2004. Stochastic hydrogeology—Challenges and misconceptions. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 280–281.

Rubin, Y. 2003. *Applied Stochastic Hydrogeology.* New York: Oxford University Press.

Sagar, B. 1978. Analysis of dynamic aquifers with stochastic forcing function. *Water Resources Research* 14, no. 2: 207–216.

Sanchez-Vila, X., A. Guadagnini, and J. Carrera. 2006. Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics* 44, no. 3: Art. No. RG3002.

Srivastava, R.M. 1990. An application of geostatistical methods for risk analysis in reservoir management. *66th Annual*

Technical Conference and Exhibition of the Society of Petroleum Engineers held in New Orleans, *Louisiana*, *September 23–26*, 825–834.

Stauffer, F. 2005. Uncertainty estimation of pathlines in ground water models. *Ground Water* 43, no. 6: 843–849.

Strebelle, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* 34, no. 1: 1–21.

Sudicky, E.A. 2004. On certain stochastic hydrology issues. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 185.

Theis, C.V. 1967. Aquifers and models. In *Proceedings of the Symposium on Ground-Water Hydrology*, ed. M.A. Marino, 138. San Francisco, California: American Water Resources Association.

Theis, C.V. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Transactions of the American Geophysical Union* 2, August: 519–524.

Warren, J.E., and H.S. Price. 1961. Flow in heterogeneous porous media. *Society of Petroleum Engineers Journal* 1, no. 3: 153–169.

Winter, C.L. 2004. Stochastic hydrology: Practical alternatives exist. *Stochastic Environmental Research and Risk Assessment* 14, no. 8: 271–273.

Zhang, D. 2002. *Stochastic Methods for Flow in Porous Media*. San Diego, California: Academic Press.

Zhang, T., P. Switzer, and A.G. Journel. 2006. Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology* 38, no. 1: 63–80.

Zhang, Y.-K., and D. Zhang. 2004. Forum: The state of stochastic hydrology. *Stochastic Environmental Research and Risk Assessment* 18, no. 4: 265.

Zinn, B., and C.F. Harvey. 2003. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. *Water Resources Research* 39, no. 3: 1051, doi: 10.1029/2001WR001146.