Advances in Water Resources 51 (2013) 247–260

Contents lists available at SciVerse ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Assessment and management of risk in subsurface hydrology: A review and perspective

Daniel M. Tartakovsky

Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Available online 5 May 2012

Keywords:
Probabilistic risk analysis
Uncertainty quantification
Decision under uncertainty
Stochastic optimization

ABSTRACT

Uncertainty plagues every effort to model subsurface processes and every decision made on the basis of such models. Given this pervasive uncertainty, virtually all practical problems in hydrogeology can be formulated in terms of (ecologic, monetary, health, regulatory, etc.) risk. This review deals with hydrogeologic applications of recent advances in uncertainty quantification, probabilistic risk assessment (PRA), and decision-making under uncertainty. The subjects discussed include probabilistic analyses of exposure pathways, PRAs based on fault tree analyses and other systems-based approaches, PDF (probability density functions) methods for propagating parametric uncertainty through a modeling process, computational tools (e.g., random domain decompositions and transition probability based approaches) for quantification of geologic uncertainty, Bayesian algorithms for quantification of model (structural) uncertainty, and computational methods for decision-making under uncertainty (stochastic optimization and decision theory). The review is concluded with a brief discussion of ways to communicate results of uncertainty quantification and risk assessment.

1. Introduction

Uncertainty plagues every modeling effort in subsurface hydrology (and other environmental fields) and every decision made on the basis of such models. How much groundwater is stored in a given aquifer? Is a depleted oil reservoir safe to host significant amounts of sequestered carbon dioxide? Will natural attenuation degrade a contaminant to an acceptable level? How much groundwater can be extracted from a coastal aquifer before seawater intrusion becomes a problem? Mathematical models of subsurface flow and transport are routinely used to answer these and other questions. Just as ubiquitous are the lack of sufficient data to parameterize these models and disagreements over the degree to which a given model represents reality. This uncertainty makes a unique answer to such questions impossible.

Instead, multiple modeling predictions are plausible. Each prediction must be accompanied by an assessment of predictive uncertainty, i.e., by rigorous *uncertainty quantification* (UQ). A systematic UQ effort aims to estimate the effects of *structural uncertainty* (uncertainty about the validity of a particular mathematical model) and *parametric uncertainty* (uncertainty about parameters and driving forces in a particular model) on predictive uncertainty. These two sources of uncertainty are sometimes referred to as *epistemic uncertainty*, since they arise from incomplete knowledge and can be reduced by collecting more data. A substantial part of this review surveys approaches to quantification of epistemic uncertainty in subsurface modeling. (Quantification of *aleatory or irreducible uncertainty*, which refers to the limits of predictability of inherently random phenomena such as population dynamics of microbial colonies in the subsurface, lies outside the scope of this review.)

Uncertainty is intimately intertwined with the concepts of *risk* or *hazard*. The Merriam-Webster dictionary defines hazard and risk as "a source of danger" and "something that creates or suggests a hazard", respectively. Given pervasive uncertainty about the subsurface, these definitions suggest that most if not all practical problems in hydrogeology can be formulated in terms of (ecologic, monetary, health, regulatory, etc.) risk. Yet there is a surprising dearth of quantitative assessments of risk in subsurface hydrology, unless one counts as such a few Monte Carlo realizations. A rigorous assessment of risk addresses the following three questions: "What can happen? How likely is it to happen? Given that it occurs, what are the consequences?" [17]. The vast majority of applied geosciences (e.g., seismology [53], surface hydrology [6], and ecology [45] and geotechnics [40,57]) have adopted risk assessment as a standard practice. A typical example is the US National Research Council's report on seismic hazard analysis (SHA) [165], which concludes that any SHA must be probabilistic and that any probabilistic SHA must focus on uncertainty quantification.

The field of probabilistic risk assessment (PRA) is relatively new, owing its creation to the Three Mile Island nuclear power plant

E-mail address: dmt@ucsd.edu

0309-1708/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.advwatres.2012.04.007

accident (1979) and the space shuttle Challenger disaster (1986) [17]. (A much longer historic perspective can be found in [36].) Many government agencies, e.g., US Environmental Protection Agency (http://epa.gov/riskassessment), US Nuclear Regulatory Commission (http://www.nrc.gov/about-nrc/regulatory/risk-informed.html), US Army Corps of Engineers (Pub. Number EM200–1-4, http://140.194.76.129/publications/eng-manuals), and European Environment Agency (http://www.eea.europa.eu/publications/GH-07-97-595-EN-C2), have incorporated PRAs into their regulations and operation manuals. A large number of such documents and guidelines are reviewed in [102].

PRA has developed its own jargon (e.g., basic and initiating events) and methodology (e.g., fault trees and binary decision diagrams) that can be applied across many disciplines. This facilitates a system's approach to risk quantification and management in complex multi-component, multi-physics problems. For example, a PRA of the Yucca Mountain nuclear waste repository dealt with a wide range of hazards, including geological/geotechnical hazards, weather-related hazards, aircraft crash hazards, industrial and military-related activity hazards, etc. [167, Chapter 3]. Assessment of the risks posed by the waste's subsurface migration (e.g., [89]) is but a small component in the overall PRA. Reliance on the structured PRA approach discussed in Section 2 facilitates its incorporation into the overall risk assessment. This section also provides a brief overview of several alternatives to the structured PRA approach.

Being a quantitative tool, PRA is fundamentally distinct from both qualitative and pseudo-quantitative approaches to uncertainty and risk quantification. Examples of such approaches include the RISQUE method (e.g., [27]), which relies on a panel of experts to assign probabilities to various adverse events; the scenario approach [181] in which an expert assigns probabilities to deterministic computer simulations of alternative scenarios; and the screening and ranking framework [128] that relies on an expert to provide educated guesses for values (and corresponding measure of uncertainty) of "proxies for site characterization data and model analyses that may not be available". These and other similar approaches are invaluable for cursory risk assessments (used, for example, in site screening) and can be used to compute some of the probabilities in a rigorous PRA when either the required data are not available or numerical modeling is prohibitively expensive. They can also be used as prior distributions in Bayesian risk assessments [59].

Risk assessment is often accompanied by risk management. The latter can be defined as [164, p. 28] "the process by which risk assessment results are integrated with other information (e.g., technical feasibility, cost and offsetting benefits) to make decisions about the need for, method of, and extent of risk reduction". Risk management is the subject matter of optimization under uncertainty and, more generally, decision theory. Closely related to risk management is the challenge of communicating results of uncertainty quantification and probabilistic risk assessment to decision-makers, the general public, and other stake-holders.

This review addresses, with various degrees of completeness, all of these subjects. Section 2 provides a description of standard approaches to PRA and a survey of its applications in hydrogeology. Quantitative approaches to UQ, the mathematical and statistical foundation of PRA, are reviewed in Section 3. Section 4 provides a cursory survey of the use of optimization under uncertainty (Section 4.1) and decision theory (Section 4.2) in management of groundwater resources and subsurface decontamination. It also briefly discusses ways to communicate the outcomes of UQ and PRA (Section 4.3).

2. Structured PRA procedure

The structured PRA methodology consists of [17]

(i) defining a system failure,
(ii) identifying all basic events contributing to the system failure,
(iii) building a fault tree (or other modes of failure analysis some of which are discussed in Section 2.2.1) to relate the basic events to the system failure,
(iv) constructing minimal sets of the basic events whose joint occurrence causes the system failure,
(v) relating (joint) probabilities of occurrence of these basic events to the probability of the system failure, and
(vi) computing probabilities of the basic events.

Steps (i) and (ii) deal with the questions "What can happen?" and "Given that it occurs, what are the consequences?", while steps (iii)–(vi) address the question "How likely is it to happen?".

To the best of our knowledge, this methodology was introduced into subsurface hydrology in [157]. It has since been used to estimate the probability of failure of reactive barriers as a groundwater remediation strategy [26], to quantify groundwater-related risks at underground excavation sites [85], to determine the likelihood of success of NAPL remediation [59], and to construct probabilistic maps of well-head protection zones [144].

As an example, let us consider the EPA's definition of the Excess Lifetime Cancer Risk (ELCR) factor [163]

$$\mathrm{ELCR} = \alpha C, \quad \alpha = \frac{\mathrm{IR} \times \mathrm{EF}}{365\ \mathrm{days} \times \mathrm{BW}} \tag{1}$$

where $C(\mathbf{x},t)$ is the concentration of a carcinogen in groundwater at location $\mathbf{x}$ and time t; and α is the exposure factor defined in terms of the human ingestion rate (IR), the exposure frequency (EF), and the body weight (BW). Concentration of a carcinogen in groundwater is deemed acceptable if ELCF does not exceed a certain level. Both the exposure factor α and the contaminant concentration C are fundamentally uncertain and must be treated probabilistically. Probabilistic descriptions of α are obtained via statistical analyses of the population at risk, while those of C are a subject matter of hydrogeology. Since α and C are in most cases statistically independent, their probabilistic analyses can be carried out separately.

A brief overview of the approaches used to probabilistically determine exposure factors is provided in Appendix A. We focus instead on methods used to compute the probability of a contaminant's concentration exceeding a safe level. It is worthwhile emphasizing that the use of PRA is not limited to water quality problems wherein contaminant concentration is the key state variable. In other problems of interest, e.g., environmentally and/or legally sound groundwater exploitation and underground construction, hydraulic head is an uncertain state variable whose probabilistic descriptions serve as an input into PRA.

2.1. Fault tree analyses of state variables

The following two problems, the first dealing with contaminant transport [26] and the second with subsurface flow [85], demonstrate the versatility of the standardized PRA framework in dealing with dynamical distributed systems (i.e., in obtaining probabilistic descriptions of state variables, such as contaminant concentrations and hydraulic heads and fluxes). The second example also shows how PRA provides for a seamless integration of groundwater models into an overall engineering design.

2.1.1. Contaminant transport

PRA was used in [26] to estimate the probability of failure of a permeable reactive barrier to intercept and remediate a contaminant plume migrating from a point source towards a protected zone (a well field, a body of surface water, etc.). The structured PRA procedure outlined above proceeds as follows. We say that the "system fails" if contaminant concentration $C(\mathbf{x},t)$ within the protected zone exceeds a certain level $C^{\star}$ before a prescribed time T. Basic events that can contribute to system failure (SF) are a contaminant's spill occurs (SO); the plume travels through the reactive barrier (P3); the plume bypasses the reactive barrier, reaching the protected zone (P2); natural attenuation fails to reduce the contaminant concentration below $C^{\star}$ (NA); the remediation effort fails to do the same (RE).

Next, Boolean operators AND and OR are employed to relate the basic events to the system failure, forming a *fault tree* shown in Fig. 1. This fault tree reveals the existence of two *minimal cut sets*, the smallest collections of basic events that must occur *jointly* in order for the system to fail: {SO,P2,NA} and {SO,P3,NA,RE}. The statement that either of these cut sets causes system failure, is then written in terms of Boolean algebra as SF = (SO AND P2 AND NA) OR (SO AND P3 AND NA AND RE). Replacing the Boolean operators with their set counterparts ($X \cap Y \equiv X$ AND Y and $X \cup Y \equiv X$ OR Y) and accounting for the statistical independence of SO from the rest of the basic events, this Boolean expression allows one to express probability of system failure $P[\mathrm{SF}]$ in terms of probabilities of the (joint) occurrence of the relevant basic events as

$$\frac{P[\mathrm{SF}]}{P[\mathrm{SO}]} = P[\mathrm{P2} \cap \mathrm{NA}] + P[\mathrm{P3} \cap \mathrm{RE} \cap \mathrm{NA}]. \tag{2}$$

2.1.2. Subsurface flow

The analysis of groundwater-related hazards at subsurface excavation sites [85] provides another example of the use of structured PRAs. System failure (SF) is now defined as the occurrence of wet conditions during the construction of an underground station. The authors identify the following basic events leading to system failure: the failure of the excavation bottom (BEF) due to a combination of blow-in (BI), liquefaction (L), and bad drainage design (BD); the occurrence of lateral inflow (LI) caused by poor construction of the diaphragm walls; and drainage system failure (DSF) with the subsequent flooding of the excavation site due to a combination of the pumping well's clogging (C), heavy rainfall (R), and the pump failure (PDW). The latter can be caused by the pump's

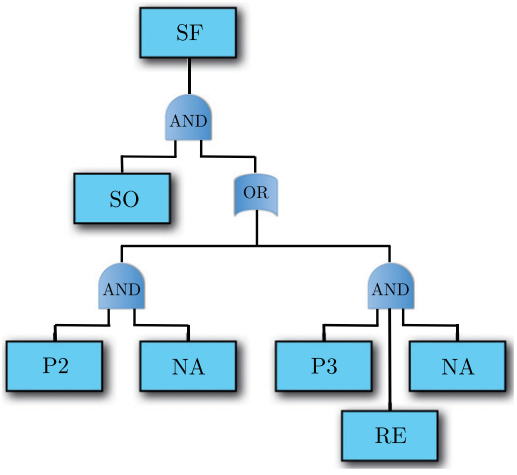

Fig. 1. The fault tree used in [26] to assess the probability of failure of a permeable reactive barrier to remediate a contaminant plume.

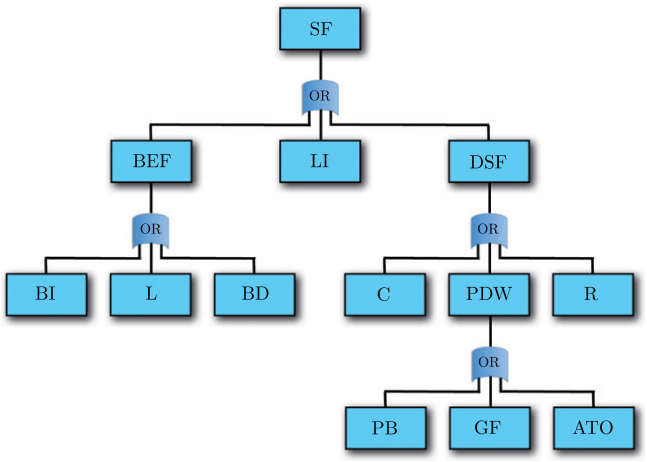

Fig. 2. The fault tree used in [85] to assess the probability of flooding of an excavation site.

breakdown (PB), an accidental turn-off (ATO), or the power generator failure (GF).

The fault tree in Fig. 2 links these events to the system failure and reveals that minimal cut sets consist of the single events {BI}, {L}, {BD}, {LI}, {C}, {R}, {PB}, {GF}, and {ATO}. Hence, the system failure can be expressed by the following Boolean expressions: SF = BEF OR LI OR DSF = (BI OR L OR BD) OR LI OR (C OR PB OR GF OR ATO OR R). Probability of the system failure is related to probabilities of the union of compound (or mega-basic) events BEF, LI, and BD by

$$P[\mathrm{SF}] = P[\mathrm{BEF} \cup \mathrm{LI} \cup \mathrm{DSF}]. \tag{3}$$

Next, probabilities of each of these mega events (but not of their joint occurrence) are expressed in terms of joint probabilities of the underlying basic events. For example, probability of the excavation-bottom failure is given by $P[\mathrm{BEF}] = P[\mathrm{BI} \cup \mathrm{L} \cup \mathrm{BD}]$ or

$$P[\mathrm{BEF}] = P[\mathrm{BI}] + P[\mathrm{L}] + P[\mathrm{BD}] - P[\mathrm{BI} \cap \mathrm{L}] - P[\mathrm{L} \cap \mathrm{BD}] - P[\mathrm{BD} \cap \mathrm{BI}] + P[\mathrm{BI} \cap \mathrm{L} \cap \mathrm{BD}]. \tag{4}$$

2.1.3. Computation of probabilities of basic events

The structured PRA described above replaces a formidable task of computing probability of system failure with a more tractable problem of computing probabilities of the occurrence of basic events. The former typically involves multiple physical and/or (bio-) geochemical processes, whose mathematical representations are often uncertain, and a prohibitively large number of uncertain parameters. The latter deals with subsets of these processes and correspondingly smaller numbers of uncertain models and parameters.

The ability to minimize the size of such subsets becomes paramount, if rigorous PRA and UQ in subsurface hydrology are to become practical. By themselves, neither (2) nor (4) accomplish this goal. Since they depend on the joint occurrence of basic events, the knowledge of probabilities of each basic event is not sufficient to compute the probability of system failure. Following [157], we advocate the use of the following approximations, which are routinely used in PRAs of engineered systems (e.g., [17]) but not in PRAs dealing with natural environments [26,85].

A *rare-event approximation* postulates that for basic events with low probabilities of occurrence, the probability of their joint occurrence is negligibly small. This approximation allows one, for example, to approximate (4) with

$$P[\mathrm{BEF}] \approx P[\mathrm{BI}] + P[\mathrm{L}] + P[\mathrm{BD}]. \tag{5}$$

Since probability is a non-negative quantity, the rare event approximation (5) provides a conservative estimate of probability $P[\mathrm{BEF}]$.

A *common cause approximation* attributes the joint occurrence of basic events to an underlying event that fully couples, but does not necessarily cause, them. In the example considered in Section 2.1.1, the existence of preferential flow paths might serve as a common cause of the simultaneous failures of the permeable reactive barrier to intercept the contaminant plume and of natural attenuation to degrade the contaminant. Let PF denote a basic event "occurrence of preferential flow path", and PF′ its complement (the absence of preferential flow paths). Then the common cause approximation replaces $P[\mathrm{P2} \cap \mathrm{NA}]$ in (2) with

$$P[\mathrm{NA} \cap \mathrm{RE}] \approx P[\mathrm{NA}|\mathrm{PF}]P[\mathrm{PF}] + P[\mathrm{NA}|\mathrm{PF}']P[\mathrm{RE}|\mathrm{PF}']P[\mathrm{PF}'], \tag{6}$$

where $P[\mathrm{PF}'] = 1 - P[\mathrm{PF}]$. This approximation can be viewed as a particular case of conditional independence in Bayesian systems, whose detailed analysis can be found in [134]. Here it suffices to say that the common cause approximation (6) provides a conservative estimate of probability $P[\mathrm{P2} \cap \mathrm{NA}]$ in (2).

Quantification of geologic uncertainty allows one to estimate probability of the occurrence of preferential flow paths, $P[\mathrm{PF}]$. Approaches for dealing with this type of uncertainty are surveyed in Section 3.1. Computation of the probabilities of the basic events in, e.g., (5) and (6) requires one to quantify model and parametric uncertainties. This can be accomplished by employing the methods discussed in Sections 3.2 and 3.3.

2.2. Alternatives to structured PRA

Throughout this review we use the term "structured PRA" to designate a procedure that follows the algorithm introduced in the beginning of Section 2. Some of the alternative modes of failure analysis, which do not use the fault tree analysis but otherwise preserve the structured PRA algorithm, are discussed in Section 2.2.1. Subsurface risk analyses that dispense with the structured PRA altogether are surveyed in Section 2.2.2.

2.2.1. Alternative modes of failure analysis

Fault trees, like those shown in Figs. 1 and 2, are but one example of the visual representations of the relationship between basic events and system failure. Such representations can be divided into two categories, deductive and inductive. Deductive approaches, such as fault tree analysis (FTA), use system failure as a starting point and then establish a chain of basic events that can cause it.

Inductive approaches start with a basic event and then establish a chain of events that can lead to system failure. Examples of inductive techniques in PRAs are logic tree analysis, failure mode and effect analysis (FMEA), influence diagrams, and event tree analysis. In the words of Vesely and Roberts [171], "inductive methods are applied to determine *what* system states (usually failed states) are possible, deductive methods are applied to determine *how* a given system state (usually a failed state) can occur".

In many problems of interest in subsurface hydrology, the distinction between these two paradigms is largely a matter of semantics, as long as the mode of system failure (as defined in the examples above) is unambiguous and unique. First, the number of basic events is small enough to avoid errors and omissions in the construction of trees of any kind. Second, when properly constructed, the inductive and deductive trees convey the same information [89]. Finally, the two paradigms (e.g., FTA and FMEA) can be combined as is done, for example, in the cause–consequence diagram method [3].

Bayesian belief nets (BBNs) [84] and state transition diagrams [186] provide an alternative representation of the relationship between basic events and system failure. BBNs can be constructed in deductive and inductive fashion. BBNs have been employed in subsurface hydrology as a tool of risk assessment [186] and risk management [58].

2.2.2. Unstructured PRAs

Monte Carlo simulations (MCS) are a tool that is routinely used in lieu of structured PRAs. (MCS are also employed within the structured PRA framework to compute both probabilistic exposure factors, e.g., the references in Appendix A, and probabilities of occurrence of basic events, e.g., [85]). Such approaches typically assume that risk is associated with a small number of uncertain hydraulic and/or transport parameters, which are treated as random fields. MCS are then used "to solve" corresponding stochastic flow and/or transport equations. In the context of risk assessment, "solving" means obtaining the full distribution of a state variable (e.g., hydraulic head or contaminant concentration) rather than its mean and variance, the two quantities typically computed in stochastic hydrogeology.

In such PRAs, hydraulic conductivity (or transmissivity T) is the main—in the references below the only—source of uncertainty in flow and transport simulations. MCS of the groundwater flow equation and (reaction-)advection–dispersion equations were used to assess the risk of human exposure to toxins in groundwater [20,44,90,97,111,113–115], to assess the risks to human health from CO_2 leakage into groundwater [152], and to delineate probabilistically well-head protection zones [74,56]. MCS of the coupled soil-mechanics/groundwater flow equations, with the shear modulus acting as the sole uncertain parameter, were used in PRA of seismically-induced soil liquefaction [121].

The computational burden associated with computation of full distributions of system states, i.e., their probability density functions (PDFs) or cumulative distribution functions (CDFs), explains why these and other MCS-based PRAs are limited to a single uncertain subsurface parameter. The importance of ensuring that MCS converged to the system state's PDF has been highlighted in [97]. The computational efficiency of MCS-based PRAs can be somewhat increased by employing accelerated (quasi) MCS approaches, e.g., Latin hypercube sampling [98].

Unstructured PRAs dealing with large numbers of uncertain subsurface parameters treat these parameters as random variables (constants) rather than random fields (functions). This allows one to reduce significantly the computational burden of MCS by relying on analytical solutions of (often zero- or one-dimensional) flow and transport equations. Examples of such PRAs include MCS on analytical solutions of either mass-balance (zero-dimensional) equations [78,161,196] or one-dimensional advection–dispersion-reaction equations [101], MCS coupled with the Analytical Contaminant Transport Analysis System (ACTS) [7, Chapter 9], MCS of combined numerical and analytical models [71], and FORM (first-order reliability method) [194] and SORM (second-order reliability method) [80] analyses. Apart from obvious limitations of treating subsurface parameters as (random) constants, such approaches have other drawbacks. For example, FORM might be poorly suited for computing concentration CDFs [154].

This review deals with *probabilistic* risk assessment, which employs probability theory to quantify various sources of uncertainty. Alternative approaches to risk assessment and uncertainty quantification rely on fuzzy logic, its generalization known as plausibility theory, and combined fuzzy-probabilistic approaches. The latter were used to quantify risks of groundwater contamination [13] and human exposure to toxins [88], as well as risks in petroleum exploration prospects [146]. In addition to suffering from an unfortunate association with the phrase "fuzzy math", fuzzy logic and its generalizations lack the scientific heft of probability theory. While probability theory is grounded in the seminal work of de Fermat, Pascal and Kolmogorov, fuzzy set theory dates back to the 1960s

[193]. The relationship between the two frameworks is subject to heated debate in the contemporary literature, which lies outside the scope of this review.

3. Probabilities of basic events

This section contains an overview of methods that enable one to calculate probabilities of occurrence of the basic events similar to those identified in Section 2.1.

3.1. Quantification of geologic uncertainty

Geologic uncertainty determines the probability of occurrence of interconnected high-conductivity structures (e.g., paleochannels) that may lead to preferential flow paths, the common cause of several basic events (see Section 2.1.3). The type of geological media and data availability determine which of the following approaches is appropriate to quantify this source of uncertainty, i.e., to compute the probability $P[\mathrm{PF}]$ in (6).

3.1.1. Data-driven quantification of geologic uncertainty

This class of methods is broadly applicable to sedimentary subsurface environments, which exhibit large-scale dominant preferential flow paths that are observable from available data (point measurements, outcrops, etc.). Fig. 3 provides a schematic representation of such an environment. Its key characteristics are statistical inhomogeneity (non-stationarity) and multi-modality of subsurface parameters, both of which can be inferred from measurements $K_i = K(\mathbf{x}_i)$ of hydraulic parameters $K(\mathbf{x})$ sampled at N locations $\{\mathbf{x}_i\}_{i=1}^N$ and/or from geophysical surveys.

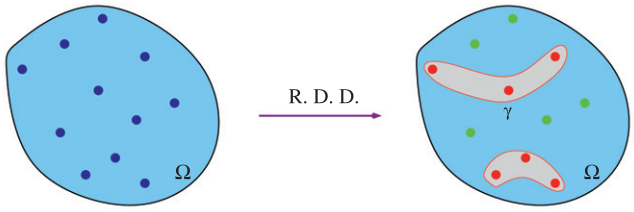

Fig. 3. A schematic representation of a subsurface environment Ω, whose property is sampled at the locations indicated by dots. Random domain decompositions (RDDs) [184,185] can be used to delineate probabilistically the boundaries γ of highly permeable sub-domains and to quantify predictive uncertainty associated with such delineation.

A plethora of approaches that can be used to delineate boundaries of (highly permeable) hydrofacies from a dataset $\{K_i\}_{i=1}^N$ include structure- and process-imitating methods [94], sequential Gaussian and transition-probability geostatistics [31], machine learning theory [189], Bayesian technics coupled with Markov chain Monte Carlo sampling [151], and deterministic classification techniques [188]. Lithologic data collected at outcrops can be used in lieu of, or in conjunction with, $\{K_i\}_{i=1}^N$ to delineate facies with transition-probability geostatistics [41]. Integration of geophysics and geostatistics (e.g., [55] and the references therein) and other "data fusion" techniques (e.g., [92] and the references therein) provide dramatic improvements in one's ability to delineate lithofacies and hydrofacies, including preferential flow paths.

These and other similar approaches to lithofacies and hydrofacies delineation are designed to provide *the best estimates* of the boundaries between geologic features, rather than their *probabilistic descriptions* that are required by PRA. There is a need to develop tools for probabilistic reconstruction of subsurface environments. It can be accomplished, for example, by modifying the existing techniques, such as indicator kriging [75]. This approach allows for aggregation of various data types, e.g., stratigraphic data and conductivity measurements, to probabilistically reconstruct a subsurface environment, as shown in Fig. 4.

Within the random domain decomposition (RDD) framework [184,185], geologic uncertainty (Fig. 4), coupled with parametric uncertainty about hydraulic conductivity and other hydraulic or transport properties of each hydrofacies, can be propagated through the modeling process by moment differential equations and/or Monte Carlo simulations [187], generalized polynomial chaos expansions (stochastic finite elements) [190], stochastic collocation on sparse grids [103], and other techniques some of which are described in Section 3.3.

3.1.2. Model-driven quantification of geologic uncertainty

If the amount of parameter data $\{K_i\}_{i=1}^N$ is not sufficient to probabilistically reconstruct high-permeability zones, one can employ concepts from percolation theory. According to this theory, a subsurface environment contains preferential flow paths (PFPs) or connected fracture networks (i.e., "percolation of highly conductive cells of a numerical grid occurs") if the relative volume of highly conductive material (e.g., sand) exceeds a certain level (the so-called percolation threshold). Percolation theory was used to determine the existence and probability of occurrence of PFPs in [77,81].

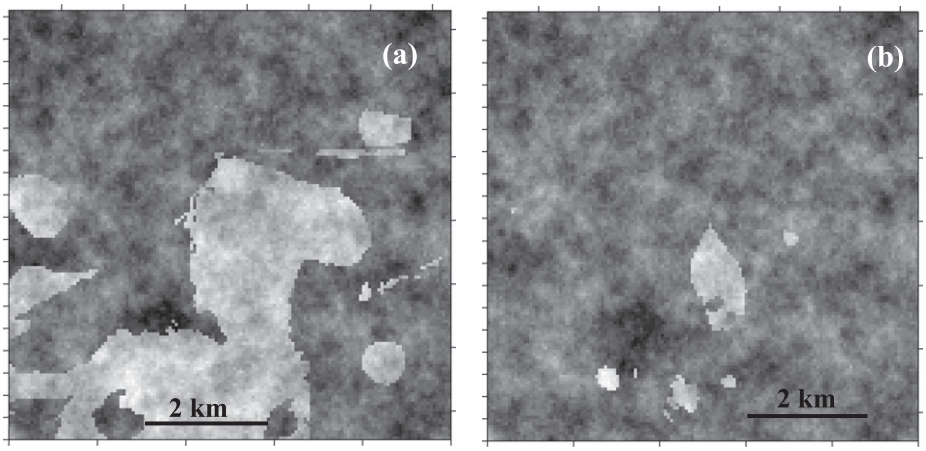

Fig. 4. A probabilistic reconstruction of the high-conductivity (light shades) and low-conductivity (dark shades) heterogeneous lithofacies in an aquitard [75]. Figures (a) and (b) display reconstructions with probability 74% and 87%, respectively.

Like all model-based identifications of PFPs, the percolation-based results should be used with caution. By treating the ambient geologic material as impervious, they neglect the possibility of flow and transport between permeable inclusions that do not touch. Moreover, transition probability-based indicator simulations of Markov chain models showed that PFPs can occur even if the relative volume of highly conductive material does not exceed the percolation threshold [139]. Such contradictory outcomes emphasize the need to determine not only the possibility but also the probability of encountering PFPs.

Various indicators of connectivity provide a conceptually attractive approach to identifying the presence of PFPs [23,49,93,160]. These approaches attempt to relate the presence of PFPs to such observable macroscopic characteristics as the relative volume of high-conductivity zones, the ratio between the effective and geometric mean of hydraulic conductivity, and hydraulic response to pumping. As of today, none of these indicators provides an unambiguous measure of the presence of PFPs, suffering from a large number of false positives and false negatives.

Percolation-based approaches to identification of PFPs were used in a number of risk assessments (not all of them probabilistic). The assessment of risks posed by geological sequestration of CO_2 [195] used percolation theory to estimate the probability of a fracture/fault network forming PFPs between a CO_2 repository and the ambient environmental receptors, and fuzzy logic to quantify the likelihood of the CO_2 plume reaching these PFPs. Percolation theory was also used in the PRA [156] to compute the probability of CO_2 injection causing the formation of interconnected fracture networks (PFPs) in the caprock.

3.2. Quantification of structural (model) uncertainty

To the extent that the probabilities of basic events in Section 2.1 are defined by solutions of mathematical models, they reflect structural (also known as model or conceptual) uncertainty. Reliance on wrong conceptual models introduces a systematic predictive bias regardless of whether their parameters are perfectly known or uncertain. While pore-scale models of subsurface flow and transport rely on the "first principles" (e.g., the Navier–Stokes equations of fluid flow and Fick's law of diffusion), they are not suitable for field-scale applications. The latter employ largely phenomenological descriptions (e.g., Richards' equation and advection–dispersion-reaction equations), whose validity is not universal and subject to ongoing debates [11,124,129]. Additionally, it is not uncommon in subsurface hydrology to treat geologic [140] and parametric [126] uncertainties, and uncertainty about driving forces (initial and boundary conditions, sources and sinks) [192], as model uncertainty.

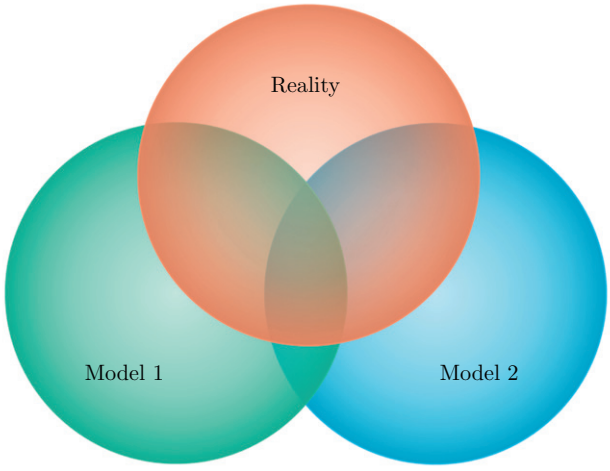

Fig. 5. A Venn diagram of the fidelity with which Models 1 and 2 represent "reality".

The importance of model uncertainty is problem-specific, as demonstrated by the Venn diagram in Fig. 5. Two different models, e.g., an advection–dispersion equation (ADE) and its nonlocal counterparts [124], might provide similar predictions of some aspects of "reality" (e.g., the dynamics of a contaminant plume's center of mass) but not others (e.g., tails of breakthrough curves). Moreover, even when competing models yield different predictions, these differences might not be large enough to alter probabilistic assessments of risk and corresponding decision-making.

At the same time, model uncertainty can, and often does, have a significant impact on predictability of system behavior and science-based predictions [70]. Two conceptual frameworks for quantification of structural uncertainty, *Bayesian model averaging* and *Bayesian model selection*, are discussed briefly below.

Bayesian model averaging (BMA) [123] seeks to combine predictions of competing models, each of which might reflect complementary aspects of reality (Fig. 5) with different degrees of fidelity. BMA starts with identification of competing models of reality and assignment of prior model probabilities. As with any Bayesian approach, these steps are highly subjective. Then data are used to update these probabilities, i.e., to compute posterior probabilities. Various computational implementations and hydrologic applications of BMA can be found in [52,125,140,147,162]. Comparative analyses of alternative model-averaging techniques, including BMA, are reported in [50,68,96,153]. BMA results depend critically on one's ability (expertise) to identify an exhaustive set of relevant models; otherwise a single model can outperform an average of multiple models [183].

If computational demands of running multiple conceptual models are prohibitive, one is forced to select the "best" among them. Within the Bayesian framework, it is common to define the best model as a model with the highest posterior probability, even though it is not always the case [9]. One can define the best model by relying on various (Bayesian and frequentist) criteria, such as the Bayesian information criterion (BIC), the Akaike information criterion (AIC), the Kashyap information criterion (KIC), etc. [29]. These and other information-theoretical criteria come with their limitations and caveats. For example, AIC cannot be used to rank models with different data sets [29], while BIC and KIC "assume that the true (or quasi-true) model exists in the set of candidate models" [137,29]. The latter assumption is not applicable to situations in which different models outperform their competitors in different regimes (Fig. 5). The relative performance of AIC and BIC is discussed in [65].

3.3. Quantification of parametric uncertainty

Probabilities of all the basic events in Section 2.1 are defined as solutions of subsurface flow and/or transport equations. Consider a typical set of such equations, defined on a domain Ω bounded by a surface Γ,

$$S_s \frac{\partial h}{\partial t} = -\nabla \cdot \mathbf{q} - f, \quad \mathbf{q} = -\mathbf{K}\nabla h \tag{7a}$$

and (for $n = 1,\ldots,N$)

$$\frac{\partial c_n}{\partial t} = -\nabla \cdot \mathbf{J}_n - R_n(c_1,\ldots,c_N), \quad \mathbf{J}_n = -\mathbf{D}\nabla c_n + \frac{\mathbf{q}}{\omega} c_n \tag{7b}$$

subject to initial and boundary conditions

$$h(\mathbf{x},0) = h_0; \quad h = H, \quad \mathbf{x} \in \Gamma_D; \quad \mathbf{n}\cdot\mathbf{q} = Q, \quad \mathbf{x} \in \Gamma_N \tag{7c}$$

and (for $n = 1,\ldots,N$)

$$c_n(\mathbf{x},0)=c_{n_0};\qquad c_n=C_n,\quad \mathbf{x}\in\Gamma_D;\quad \mathbf{n}\cdot\mathbf{J}_n=Q_n,\quad \mathbf{x}\in\Gamma_R. \tag{7d}$$

Here specific storage $S_s(\mathbf{x})$, hydraulic conductivity (tensor) $\mathbf{K}(\mathbf{x})$, and porosity $\omega(\mathbf{x})$ are the (uncertain) hydraulic properties of a heterogeneous subsurface environment Ω. Groundwater flow, i.e., changes in hydraulic head $h(\mathbf{x},t)$ and Darcy velocity $\mathbf{q}(\mathbf{x},t)$, are induced by (uncertain) sources/sinks $f(\mathbf{x},t)$; initial hydraulic head distribution $h_0(\mathbf{x})$; and hydraulic head $H(\mathbf{x})$ and flux $Q(\mathbf{x})$ prescribed along, respectively, the Dirichlet (Γ_D) and Neumann (Γ_N) segments of the boundary $\Gamma=\Gamma_D\cup\Gamma_N$ with the unit normal vector $\mathbf{n}(\mathbf{x})$. Migration and fate of the n-th reactive species ($n=1,\dots,N$) with concentration $c_n(\mathbf{x},t)$ is driven by advection in the velocity field $\mathbf{q}/\omega$, hydrodynamic dispersion with dispersion coefficient tensor $\mathbf{D}$, chemical reactions R_n with (uncertain) reaction rates $\boldsymbol{\kappa}=\{\kappa_1,\dots,\kappa_N\}$, and (uncertain) driving terms: initial concentration distribution $c_{n_0}(\mathbf{x})$, and concentration $C_n(\mathbf{x})$ and flux $Q_n(\mathbf{x})$ prescribed along, respectively, the Dirichlet (Γ_D) and Robin (Γ_R) segments of the boundary $\Gamma=\Gamma_D\cup\Gamma_R$.

Within a probabilistic framework, uncertainty in any or all of these parameters and driving terms is quantified by treating them as random fields. The corresponding boundary-value problems (7) become stochastic. Their solutions are given in terms of probability density functions (PDFs) of the system states such as hydraulic head $h(\mathbf{x},t)$, Darcy velocity $\mathbf{q}(\mathbf{x},t)$, concentrations $c_n(\mathbf{x},t)$ and mass fluxes $\mathbf{J}_n(\mathbf{x},t)$. By their very definition, these PDFs allow one to compute the probability of a system state exceeding a certain value, i.e., to compute probabilities of basic events.

It is worthwhile recognizing that for specified functional relations $\mathbf{D}(\mathbf{q})$ and $R_n(c_1,\dots,c_N)$ Eqs. (7) form *a single* model with uncertain parameters and driving terms, rather than a set of alternative models. (This is in contrast to many recent analyses that treat as competing models alternative spatial variability of hydraulic conductivity $\mathbf{K}$ due to geologic uncertainty [65,140], statistical parameterizations of $\mathbf{K}$ [126], recharge f [153,192] or boundary fluxes Q [70].) Consequently, these sources of uncertainty are amenable to various techniques for quantification of parametric uncertainty.

Examples of the latter approach include probabilistic analyses of uncertain (random) boundary fluxes, such as precipitation/infiltration and evapotranspiration [138,143,145,172]. Likewise, the random domain decompositions [184,185] discussed above treat geologic uncertainty as a subset of parametric (rather than structural) uncertainty by decomposing a flow domain $\Omega\cup_{m=1}^{M}\Omega_m$ into M non-overlapping subdomains (heterogeneous hydrofacies) $\{\Omega_m\}_{m=1}^{M}$ and representing an uncertain (random) parameter, e.g., hydraulic conductivity $\mathbf{K}(\mathbf{x})$, as (see Fig. 3 wherein $M=2$)

$$\mathbf{K}(\mathbf{x})=\begin{cases}\mathbf{K}_1(\mathbf{x}) & \mathbf{x}\in\Omega_1\\ \vdots & \vdots\\ \mathbf{K}_M(\mathbf{x}) & \mathbf{x}\in\Omega_M.\end{cases} \tag{8}$$

A general RDD formulation [187] allows for uncertainty in the spatial extent of subdomains Ω_m, their number M, and hydraulic/transport properties $K_m(\mathbf{x})$ of each subdomain. This uncertainty is propagated through the modeling process by solving (7) and (8) supplemented with the continuity conditions for h and c_n, and the normal components of their respective fluxes $\mathbf{q}$ and $\mathbf{J}_n$, along the (uncertain) interfaces between the adjacent subdomains.

Table 1
Computed (☆) and assumed (†) probability density functions (PDFs) of state variables, whose predictions are uncertain due to uncertainty in one or more hydraulic parameters. Saturated flow: h is the hydraulic head, and K is the saturated hydraulic conductivity. Unsaturated flow: ψ is the pressure head, i is the infiltration rate, z_f is the infiltration depth, and $\boldsymbol{\alpha}$ are the parameters in soil constitutive relations (relative conductivities and/or retention curves). Solute transport: $\mathbf{m}$ is the center of mass of a solute plume, $\mathbf{J}$ is the mean hydraulic head gradient, $\mathbf{v}$ is the flow velocity, $\mathbf{X}$ is the solute particle trajectory, τ is the particle travel time along $\mathbf{X}$, and c is the solute concentration described by the advective-dispersion equation (7b) (with $R_i\equiv 0$ for conservative solutes), κ is the reaction rate constant, α is the stoichiometric coefficient, and c_{eq} and c_{in} are the equilibrium and initial concentrations, respectively. For all random fields $\mathcal{A}$, $\langle\mathbf{A}\rangle$ and $\mathbf{A}'$ represent their ensemble means and zero-mean random fluctuations, such that $\mathbf{A}=\langle\mathbf{A}\rangle+\mathbf{A}'$; and $\sigma_{\mathcal{A}}^2$ and $l_{\mathcal{A}}$ denote the variance and correlation length of $\mathbf{A}$, respectively.

State variable	Sources of uncertainty	Modeling assumptions	PDF	Reference
Groundwater flow				
$h'(\mathbf{x})$, $h'(\mathbf{x},t)$	$Y=K(\mathbf{x})$	Stationary, normal Y; small σ_Y^2; $l_Y\ll\lvert\Omega\rvert$	Multivariate normal†	[122]
			Joint (with Y) multivariate normal†	[122]
$\lvert\nabla h(\mathbf{x})\rvert$, $\lvert\mathbf{q}(\mathbf{x})\rvert$	$Y(\mathbf{x})$	Stationary, normal Y; $l_Y\ll\lvert\Omega\rvert$;	Lognormal☆	[169]
Unsaturated flow				
$\psi'(\mathbf{x})$, $\psi'(\mathbf{x},t)$	$Y(\mathbf{x})$, $\mathbf{A}=\ln\boldsymbol{\alpha}(\mathbf{x})$	Stationary Y small σ_Y^2 and $\sigma_{\mathbf{A}}^2$	Multivariate normal†	[2]
			Joint (with Y and $\mathbf{A}$) multivariate normal†	[2]
$i(x_1,x_2,t)$, $z_f(x_1,x_2,t)$	$Y(\mathbf{x})$, $\mathbf{A}(\mathbf{x})$	$z_f\leqslant\min\{l_Y,l_{\mathbf{A}}\}$ Green–Ampt/Parlange models	Non-canonical (multivariate) PDFs☆	[177,178]
Conservative transport				
$\mathbf{m}(t)$	$Y(\mathbf{x})$, $\mathbf{J}(t)$	Stationary Y and $\mathbf{v}'$, small σ_Y^2 and $\sigma_{\mathbf{v}}^2$	Normal☆	[47,46]
			Normal☆	
$c(\mathbf{x},t)$	$\mathbf{v}(\mathbf{x},t)$	Stationary, multivariate normal $\mathbf{v}$ normal $\mathbf{m}(t)$	Non-canonical PDF☆	[48]
$\mathbf{X}(t)$	$Y(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$	Stationary normal Y, small σ_Y^2, $\langle\mathbf{v}\rangle=$ const, lognormal $c(\mathbf{x},t)$	Non-canonical PDF☆	[63]
$\tau(\mathbf{X})$	$Y(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$	Stationary normal Y, small σ_Y^2, $\langle\mathbf{v}\rangle=$ const, $\mathbf{D}\equiv\mathbf{0}$ in (7b),	Non-canonical PDF☆ lognormal☆ for large $\lvert\mathbf{X}\rvert$	[39]
$c(\mathbf{x},t)$	$\mathbf{v}(\mathbf{x})$	Infinite Ω, other assumptions are not explicitly stated	β-distribution†	[19,63]
Reactive transport				
Langmuir sorption: $\tau(\mathbf{X})$	$Y(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$	Stationary normal Y, small σ_Y^2, $\langle\mathbf{v}\rangle=$ const, $\mathbf{D}\equiv\mathbf{0}$ in (7b)	Non-canonical PDF☆, lognormal☆ for large $\lvert\mathbf{X}\rvert$	[39]
(7b) with $N=2$ and $R_n=\kappa c_1c_2$: $c_n(\mathbf{x},t)$	$Y(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$	Stationary normal Y, $\langle\mathbf{v}\rangle=$ const, infinite Ω	Map of β-distribution†	[18,33]
(7b) with $N=1$ and $R=\kappa\left(c^{\alpha}-c_{\text{eq}}^{\alpha}\right)$: $c(\mathbf{x},t)$	$\kappa(\mathbf{x})$, $c_{\text{in}}(\mathbf{x})$	Stationary κ, steady $\mathbf{v}$, $\mathbf{D}\equiv\mathbf{0}$ in (7b)	Equation for PDF☆	[28,159]
(7b) with $N=1$ and $R=\kappa\left(c^{\alpha}-c_{\text{eq}}^{\alpha}\right)$: $c(\mathbf{x},t)$	$\kappa(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$, $c_{\text{in}}(\mathbf{x})$	Stationary κ and $\mathbf{v}$, $\mathbf{D}\equiv\mathbf{0}$ in (7b)	Equation for PDF☆	[158]

Multivariate sensitivity analyses, such as analysis of variance (ANOVA), can be used to assess the uncertainty about a statistical parameterization of the parameters $\{\mathbf{K}_m(\mathbf{x})\}_{m=1}^{M}$, e.g., the existence and importance of cross-correlations between $\mathbf{K}_l(\mathbf{x})$ and $\mathbf{K}_n(\mathbf{x})$ for $l \neq n$ [182].

Since each basic event is affected by a relatively small number of uncertain parameters, the structured PRA approach described in Section 2.1 facilitates the numerical implementation of the RDD. It also makes possible the use of the basic event PDFs as building blocks to assess the overall probability of system failure. Examples of such PDFs, which can be used in lieu of computationally expensive Monte Carlo simulations, are provided in Table 1.

The use of the PDFs in Table 1 comes with the following caveats. First, the list of assumptions and corresponding limitations underlying these PDFs is not exhaustive, reflecting a certain degree of opacity in some analyses. Second, reduced complexity models (e.g., the Green–Ampt model used instead of the Richards equation or advection-reaction equations used instead of their advection–dispersion-reaction counterparts) provide a trade-off between modeling complexity and the comprehensive/rigorous treatment of parametric uncertainty. Third, these and other PDFs can be used as prior distributions that can be refined through Bayesian updating as relevant site-specific data become available. The PRAs conducted in [26,135] provide examples of the use of such PDFs.

4. Risk management: decisions under uncertainty

Once risk has been assessed, it has to be managed. A decision on how to mitigate this risk is inevitably made in the presence of uncertainty. This can be accomplished within one of the two conceptual frameworks, optimization under uncertainty (Section 4.1) and decision analysis (Section 4.2). While these two frameworks are sometimes viewed as opposite (e.g., [66,120]), they are closely related (see, for example, the title of [10]) rendering this subdivision somewhat artificial.

4.1. Optimization under uncertainty

Optimization under uncertainty is a rapidly developing field with applications in many fields of science and engineering. The state-of-the-art and comprehensive reviews of various approaches to optimization under uncertainty, including various flavors of stochastic programming and fuzzy programming, can be found in [148,82,170] and [51, Chapter 5]. We focus on applications of such techniques in subsurface hydrology.

A typical setting for optimization analyses of this type is the optimal design of pump-and-treat remediation strategies in heterogeneous subsurface environments with uncertain hydraulic and transport parameters. The problem can be formulated as follows. Given a probabilistic description of subsurface parameters, find the optimal number and locations of pumping wells and injection/extraction rates. Groundwater flow and contaminant transport are described by (7) wherein the source function $f(\mathbf{x},t) = F(\mathbf{x},t) + W(\mathbf{x},t)$ is a combination of (typically spatially-distributed) sources and sinks unrelated to the remediation effort (e.g., groundwater recharge), $F(\mathbf{x},t)$, and point sources/sinks representing pumping wells, $W(\mathbf{x},t) = \sum_{i=1}^{I} Q_i \delta(\mathbf{x} - \mathbf{x}_i)$. Here I is the number of wells, $\mathbf{x}_i$ ($i = 1,\dots,I$) are their locations, and $Q_i(t)$ are injection ($Q_i \leqslant 0$) and extraction ($Q_i > 0$) rates. The goal is to minimize the number of wells and their overall pumping rates, while reducing contaminant concentrations to acceptable levels. If all the wells are extracting ($Q_i > 0$ for all $i = 1,\dots,I$), this optimization problem can be written as

$$\min_{I,\{\mathbf{x}_i,Q_i\}_{i=1}^{I}} \sum_{i=1}^{I} Q_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{9a}$$

subject to possible hydraulic restrictions and (probabilistic) constraints

$$c_n(\mathbf{x},t) \leqslant c_n^{\star}, \quad \mathbf{x} \in \Omega, \quad n = 1,\dots,N \tag{9b}$$

where $c_n^{\star}$ ($n = 1,\dots,N$) are the contaminant concentrations consistent with water quality standards. For given I and $\{\mathbf{x}_i, Q_i\}_{i=1}^{I}$, the constraints (9b) are to be satisfied probabilistically since $c_n(\mathbf{x},t)$ are random fields whose PDFs satisfy the flow and transport Eqs. (7) with uncertain (random) coefficients. Let p denote an acceptable probability of the failure of a remediation effort. Then the proper formulation of the "chance constraints" (9b) is $\Pr[c_n(\mathbf{x},t) \leqslant c_n^{\star}] \geqslant p$.

Solving the nonlinear optimization problem (9) is computationally demanding even if all the parameters in the flow and transport equations (7) are known with certainty (deterministic). If one or more of these parameters are uncertain (random), this task might become computationally prohibitive. All solutions of the stochastic optimization problem (9) we were able to find are restricted to vertically-averaged flow equation (7a) with uncertain transmissivity T (the vertically averaged analog of hydraulic conductivity $\mathbf{K}$), and correspondingly averaged transport equation (7b) for a single (passive or reactive) species.

Monte Carlo simulations (MCS), which solve the deterministic optimization problem (9) for each realization of a random parameter and then average the results of multiple realizations, are a case in point. High computational costs of MCS limit their applicability to numerical models with few degrees of freedom and few decision variables. On a computer ca. 1987, such simulations of two-dimensional steady-state flow on a computational domain discretized into 336 finite elements and transport carried out over 80 time steps took 20 h of CPU time [72,174]. These simulations treated transmissivity $T(x_1,x_2)$ as the only source of uncertainty and had five decision variables. A decade later, a nearly identical MCS approach was able to handle 1440 elements [120]. The analysis dealt with two decision variables (constant Q_1 and Q_2 in two wells whose location is fixed), used a response matrix approach in lieu of solving the flow equation (7a), and relied on 200 realizations of the log-normal transmissivity field. A conceptual generalization of this MCS approach to stochastic optimization, which includes chance constraints on hydraulic head values, incorporates capital costs associated with construction and operation of wells, and accounts for structural (model) uncertainty, is described in [73]. Other uses of MCS to solve stochastic optimization problems similar to (9) can be found in [119,175]. MCS were also used as part of a genetic algorithm [104] and stochastic programming with recourse [176].

The computational efficiency of sampling-based solutions of the stochastic optimization problem (9) can be improved by employing accelerated or quasi-MCS, such as importance sampling and Latin hypercube sampling (LHS) of realizations of $T(x_1,x_2)$, or other approaches to identify "critical realizations" [14,15,95]. Additional computational savings can be obtained, at the cost of reduced accuracy, by replacing solutions of the flow and transport equations for each realization of an uncertain parameter with various surrogate techniques such as kriging [12] and the response matrix approach [120]. LHS and neural-network surrogates were used, within a noisy genetic algorithm (nGA) optimization framework, to solve (9) in which uncertain hydraulic conductivity values of 9 zones of a numerical mesh were treated as random variables [191]. The reliance on the surrogates was shown to reduce the computational cost by up to 90%. This study also contains an overview of applications of alternative nGA techniques to solving (9).

Stochastic differential dynamic programming (sDDP) is another approach to optimization under uncertainty. By employing the Bayesian framework or Kalman filter-based data assimilation techniques, sDDP allows one to adjust the optimal remediation

pump-and-treat strategy—primarily to select the optimal pumping rates $\{Q_i\}_{i=1}^I$ in (9)—as more data (e.g., measurements of hydraulic head and solute concentrations) become available, thus reducing parametric uncertainty. Like all methods of stochastic optimization, sDDP is computationally demanding. The first application of sDDP to solve (9) dates back to 1990, when it was limited to one-dimensional flow and transport models with uncertain transmissivity and dispersivity, and had only one decision variable (pumping rate Q at a well whose location was fixed) [5]. Its extension to two spatial dimensions [99] proved to be more efficient, in terms of both computational cost and accuracy, than a MCS-based feedback-control method. The sDDP approach to addressing various aspects of the pump-and-treat optimization (9) was further refined and generalized in [4,37,109,179].

Many of the above analyses formulated the pump-and-treat optimization problem (9) in terms of monetary costs associated with well construction (the number of wells I and their locations $\{\mathbf{x}_i\}_{i=1}^I$) and operation (pumping rates $\{Q_i\}_{i=1}^I$). Such formulations raise the question of data's monetary worth for uncertainty reduction and its impact on decision-making (e.g., optimal operation of wells). This question was addressed in several subsurface applications by means of either an extended Kalman filter built into sDDP [100] or Bayesian approaches that share many basic features but differ in names and implementation: a data-worth analysis [60,83], comparative information yield curves [43,44], and a risk-cost-benefit decision analysis [8,67].

In addition to pump-and-treat remediation, the stochastic optimization techniques discussed above (and several others whose descriptions can be found in [82]) were used to reduce fertilizer use to a level that minimizes agricultural income losses while ensuring that the maximum permissible concentration in groundwater is not exceeded [133], to optimize the exploitation of groundwater [4,86] and groundwater/surface water [136,141] resources, and to select an optimal DNAPL remediation strategy that minimizes remediation costs while not exceeding the probability of failure arising from parametric uncertainty [30,130]. These transport phenomena are described by (a system of) nonlinear differential equations with a large number of (typically uncertain) parameters. Yet the computational demands of stochastic optimization algorithms are such that they can handle at most two sources of uncertainty. With the exception of the studies of conjunctive groundwater/surface water supplies that treated inflow from streams [136] or river diversion [141] as the sole source of uncertainty, transmissivity (or hydraulic conductivity) is the common source of uncertainty in all of these studies. Recharge/leakage was considered as an additional source of uncertainty in [4,86].

Several computationally efficient approaches to stochastic optimization, e.g., the sDDP implementations in [4,5,99], achieve computational speed-up by replacing MCS with moment differential equations that propagate deterministically means and (co)variances of the system states. This necessitates the use of Taylor series and perturbation-based closures, which compromises the accuracy and range of applicability of such methods. Equally important, most state variables, e.g., solute concentrations $c_i(\mathbf{x},t)$ in (7) and (9), are highly non-Gaussian, so that their means and variances provide only partial information about their PDFs. This information is insufficient for risk analysis and risk management, both of which require the knowledge of full PDFs to estimate probabilities of rare events. It is also of limited use for data assimilation through either Bayesian updating or ensemble Kalman filtering. To obviate these limitations, we advocate the use of the PDF equations and closed-form PDFs (see Section 3.3) in lieu of both MCS and the moment equations.

As mentioned earlier, the present review deals primarily with the probabilistic treatment of risk assessment and management. A review of groundwater-related optimization under uncertainty, in which uncertain parameters are represented as fuzzy sets, can be found in [76].

4.2. Decision theory and analysis

Decision theory is an active field of research that is the subject matter of many textbooks and monographs [34,131,180], and scientific publications including the journal *Decision Theory*. Decision theory is commonly subdivided into two parts, *prescriptive* (or normative) and *descriptive* (or positive), with the former based on the notion that a decision maker will always make a rational decision and the latter disputing this notion. The normative approach to decision theory [34], which seems to be well-suited for engineering and natural sciences including hydrology, is based on the concepts of *rational behavior* as defined by the Savage axioms [150] and their subsequent extensions and generalizations (e.g., [54]).

At the heart of the Savage theory of rational behavior lies a postulate that alternative actions (e.g., different remediation strategies) can be arranged in order of *preference*. This is not always straightforward. For example, the relative efficiency of alternative remediation strategies must be reconciled with their costs, local conditions, tolerance to the risk of failure, etc. Virtually every decision a hydrogeologist (and relevant stakeholders) must take is made under uncertainty and leads to uncertain outcomes.

The normative approach to decision-making under uncertainty is typically based on the von Neumann–Morgenstern *utility function* [173], which was originally proposed as an alternative to the Bernoulli utility function [21] to represent preferences over lotteries. The utility function $u(\cdot)$ assigns numerical values to each alternative strategy, so that strategy S_1 (e.g., the pump-and-treat remediation considered in Section 4.1) is preferable to strategy S_2 (e.g., bioremediation) if and only if the utility of S_1, $u_1 = u(S_1)$, is larger than the utility of S_2, $u_2 = u(S_2)$.

Expected utility theory (EUT) posits that "the decision maker chooses between risky or uncertain prospects by comparing their expected utility values, i.e., the weighted sums obtained by adding the utility values of outcomes multiplied by their respective probabilities" [118]. Consider N remediation strategies $S_1, \ldots, S_N$, whose success is determined by their ability to meet safety standards for drinking water, e.g., the ELCR factor (1). If the probability of failure of the ith remediation strategy S_i is $P[S_i]$, then the probability of its success is $p_i = 1 - P[S_i]$. The expected utility of $u(S_i)$ can be found as [118]

$$E[u(S_i)] = \mu_{i1} p_i + \mu_{i2}(1 - p_i), \quad i = 1, \ldots, N, \tag{10}$$

where the coefficients μ_{ij} are utilities of choosing decision S_i and obtaining positive ($j = 1$) or negative ($j = 2$) outcomes. According to EUT, the optimal remediation strategy has the highest expected utility, $S^\star = \arg\max_i E[u(S_i)]$, where the maximum is computed over alternative strategies $i = 1, \ldots, N$.

Probabilistic risk analyses (Section 2) can provide the probabilities of failure $P[S_i]$ or, equivalently, the probabilities of success p_i. The utilities μ_{ij} must be supplied by the decision maker based on financial and other preferences as well as his/her tolerance to risk. In other words, EUT allows for a clear delineation of responsibilities between the experts (who compute probabilities) and the decision makers (who provide utilities). If the decision maker is not certain of which values to assign to each utility or several decision makers (stakeholders) are not able to agree on these values, the experts can help by providing a sensitivity analysis of $S^\star$ with respect to μ_{ij} while leaving the probability estimates intact [17].

Cost-benefit analysis (CBA) [25] replaces expected utility with expected monetary value. CBA has a number of appealing features, including its relative simplicity and flexibility, the ability to deal with multiple sources of uncertainty, and representation of risk and uncertainty in familiar sounding monetary terms [170]. Many

of the stochastic optimization studies reviewed in Section 4.1 are examples of CBA. Its other hydrogeologic applications include CBAs of waste management facilities [111], pump-and-treat remediation [66,112], and conjunctive water management [69].

Despite its widespread use, CBA has a number of drawbacks and limitations. These include

- its definitions of both risk and cost of risk, and "its poor ability to communicate risk-related information" [90];
- the inability to create "the basis for a consensus decision except when all participants agree with prices used, and when those benefiting agree to compensate those not benefiting" [17]; and
- high sensitivity of CBA-based decisions to cost of failure, which arises from the fact that "post-failure consequences, litigation costs, regulatory penalties, loss of opportunity or investment, and damage to public relations are hard to quantify or predict" [170, Section 8].

As alternatives to CBA, multi-attribute utility theory, analytical hierarchy process, outranking methods, etc. can be employed to determine the utilities μ_{ij} in (10). A comparative review of these approaches can be found in [24, p. 16]. Multiple criteria decision analysis (MCDA) [61] has been used for water resource planning [170] and management of subsurface contamination [90], and as an integral component of several decision support systems [110]. A key advantage of MCDA over CBA lies in its ability "to deal with noncommensurable aspects and to facilitate stakeholder's involvement for collaborative decision-making" [170].

4.3. Communication of risk and uncertainty

A recent report by the US Environmental Protection Agency (EPA) Science Advisory Board [42] concluded that ecological risk assessments "have been most effective when clear management goals were included in the problem formulation; translated into information needs; and developed in collaboration with decision makers, assessors, scientists, and stakeholders." This conclusion reaffirms EPA's "cardinal rules of risk communication" [35]:

Rule 1: Accept and involve the public as a legitimate partner.
Rule 2: Plan carefully and evaluate performance.
Rule 3: Listen to your audience.
Rule 4: Be honest, frank, and open.
Rule 5: Coordinate and collaborate with other credible sources.
Rule 6: Meet the needs of the media.
Rule 7: Speak clearly and with compassion.

A detailed survey of the research on risk and uncertainty communication lies outside the scope of the present review. The interested reader is referred to the comprehensive and entertaining assessments of the progress in this field [64,106,155]. The current state-of-the-art, best practices, and challenges in risk communications are discussed [166].

Of immediate relevance to communication of uncertainty quantification and risk assessment results in subsurface hydrology is the language recommended by the Intergovernmental Panel on Climate Change [22]. It is reproduced in Table 2.

Table 2
The likelihood ranges for expressing the probability of occurrence recommended by the Intergovernmental Panel on Climate Change [22].

Description	Probability of occurrence
Virtually certain	>99%
Extremely likely	>95%
Very likely	>90%
Likely	>66%
More likely than not	>50%
Unlikely	<33%
Very unlikely	<10%
Extremely unlikely	<5%

5. Summary

Uncertainty plagues every effort to model subsurface processes and every decision made on the basis of such models. Given this pervasive uncertainty, virtually all practical problems in hydrogeology can be formulated in terms of (ecologic, monetary, health, regulatory, etc.) risk. This review deals with applications to subsurface hydrology of recent advances in uncertainty quantification, probabilistic risk assessment, and decision-making under uncertainty.

Special attention is paid to the computationally efficient tools for uncertainty quantification and risk assessment that are still in early stages of their development. These include

- Systems-based approaches to probabilistic risk assessment (PRA) that replace a formidable task of computing probability of system failure with a more tractable problem of computing probabilities of the occurrence of basic events. The former involves multiple physical and/or (bio-) geochemical processes, whose mathematical representations are often uncertain, and a prohibitively large number of uncertain parameters. The latter deals with subsets of these processes and correspondingly smaller numbers of uncertain models and parameters.
- PDF (probability density function) methods to propagate parametric uncertainty through the modeling process. Such methods aim to provide a computationally efficient alternative to Monte Carlo simulations by deriving deterministic equations for PDFs of state variables (e.g., hydraulic heads or solute concentrations). By their very definition, these PDFs allow one to compute the probability of a system state exceeding a certain value, i.e., to compute probabilities of basic events.
- Random domain decompositions (RDDs) and other frameworks for quantification of geologic uncertainty. While many currently used approaches to lithofacies and hydrofacies delineation are designed to provide the best estimates of the boundaries between facies, PRAs require their probabilistic descriptions. RDDs and transition probability-based approaches can be used to probabilistically identify the existence and spatial extent of preferential flow paths that often serve as a common cause of system failure.
- Computationally efficient methods for optimization under uncertainty. These include (i) stochastic optimization via accelerated or quasi Monte Carlo simulations (e.g., importance sampling and Latin hypercube sampling) coupled with surrogate techniques (e.g., kriging, neural-network surrogates, and response matrix approaches), and (ii) stochastic differential dynamic programming enhanced with Kalman filter-based data assimilation techniques.
- Quantitative approaches for decision-making under uncertainty that are grounded in expected utility theory. A distinction is made between commonly used cost-benefit analysis (which reduces utility to monetary value) and its alternatives, such as multi-attribute utility theory, analytical hierarchy process, and outranking methods.

These and other approaches discussed above provide furtive areas of research in the rapidly growing fields of uncertainty quantification (UQ) and risk assessment (RA) in subsurface modeling.

Acknowledgment

The author thanks Peter Quinlan and the six anonymous reviewers whose insightful comments helped to improve the original version of this manuscript. This work was supported by Applied Mathematics Program of the US DOE Office of Advanced Scientific Computing Research.

Appendix A. Statistical characterization of exposure factor

In general, every parameter entering the expression for the exposure factor α in (1) is uncertain and subject to spatio-temporal variability. For example, a statistical analysis of the US Public Health Service data from 1976–1980 [62] led its authors to conclude that bivariate (between the height and the natural logarithm of weight) histograms of the body weight (BW) of men and women follow a normal distribution and two superposed normal distributions, respectively. It is worthwhile recognizing that such models of uncertainty are uncertain as well: while Finley et al. [62] postulate that "body weight will not vary significantly from setting to setting" which confers a degree of universality to their probabilistic models of BW, Riederer et al. [142] make a case for the importance of site-specific data.

The other two parameters defining α in (1), the human ingestion rate (IR) and exposure frequency (EF), are determined by the environmental pathways between a contaminant and population. These include dermal uptake (e.g., showering), inhalation of volatile chemicals in soils and groundwater, tap-water ingestion, soil ingestion, etc. Probabilistic distributions for these and other processes can be found in [1,16,62,108,149]. Depending on the problem, some or all of these pathways are to be accounted for. They can be either mutually independent or, as is the case with dermal uptake and body weight, strongly correlated [62].

A typical example of exposure-centered studies are the investigations of health risks posed by perchloroethylene (PCE) [116] and tetrachloroethylene (TCE) [117] in water supplies derived from groundwater. In both analyses, contaminant concentrations in water supplies were inferred from the water quality data and a set of assumptions (a so-called source-characterization step), and the main focus was placed on determining exposure pathways (tap-water ingestion, dermal uptake, etc.) and probabilities of their occurrence. Additional examples of probabilistic exposure-centered investigations include assessments of risks associated with (i) ingestion of drinking water containing infectious microorganisms at regulatory levels [79], (ii) consumption of agricultural products irrigated with reclaimed wastewater [168], and (iii) human arsenic exposure through a number of nodes in the food chain [91]. In-depth reviews of health-related PRAs that focus primarily on exposure can be found in [38,127,132].

Probabilistic analyses of exposure pathways are not limited to human health. They have been used in ecological risk assessment to estimate exposure of other end-point species, including mink and great blue herons [107], game birds and raptors [87], and raccoons [32], among many others. Salient differences in modeling wildlife and human exposure pathways are discussed in [105].

References

[1] Allan M, Richardson GM, Jones-Otazo H. Probability density functions describing 24-hour inhalation rates for use in human health risk assessments: an update and comparison. Human Ecol Risk Assess 2008;14(2):372–91.
[2] Amir O, Neuman SP. Gaussian closure of transient unsaturated flow in random soils. Transp Porous Media 2004;54(1):55–77.
[3] Andrews JD, Ridley LM. Reliability of sequential systems using the cause-consequence diagram method. Proc Instn Mech Engrs, Part E 2001;215:207–20.
[4] Andricevic R. Coupled withdrawal and sampling designs for groundwater supply models. Water Resour Res 1993;29(1):5–16.
[5] Andricevic R, Kitanidis PK. Optimization of the pumping schedule in aquifer remediation under uncertainty. Water Resour Res 1990;26(5):875–85.
[6] Apel H, Merz B, Thieken AH. Quantification of uncertainties in flood risk assessments. Int J River Basin Manage 2008;6(2):149–62.
[7] Aral MM. Environmental modeling and health risk analysis (ACTS/RISK). New York: Springer; 2010.
[8] Back P-E. A model for estimating the value of sampling programs and the optimal number of samples for contaminated soil. Environ Geol 2007;52:573–85.
[9] Barbieri MM, Berger JO. Optimal predictive model selection. Ann Statist 2004;32(3):870–97.
[10] Bather JA. Decision theory: an introduction to dynamic programming and sequential decisions. 1st ed. New York: Wiley; 2000.
[11] Battiato I, Tartakovsky DM. Applicability regimes for macroscopic models of reactive transport in porous media. J Contam Hydrol 2011;120-121:18–26.
[12] Bau D, Mayer AS. Stochastic management of pump-and-treat strategies using surrogate functions. Adv Water Resour 2006;29(12):1901–17.
[13] Baudrit C, Guyonnet D, Dubois D. Joint propagation of variability and imprecision in assessing the risk of groundwater contamination. J Contam Hydrol 2007;93(1–4):72–84.
[14] Bayer P, Burger CM, Finkel M. Computationally efficient stochastic optimization using multiple realizations. Adv Water Resour 2008;31:399–417.
[15] Bayer P, de Paly M, Bürger CM. Optimization of high-reliability-based hydrological design problems by robust automatic sampling of critical model realizations. Water Resour Res 2010;46:W05504. http://dx.doi.org/10.1029/2009WR008081.
[16] Beamer P, Canales RA, Leckie JO. Developing probability distributions for transfer efficiencies for dermal exposure. J Expo Sci Environ Epidemiol 2008;19(3):274–83.
[17] Bedford T, Cooke R. Probabilistic risk analysis: foundations and methods. Cambridge, U.K.: Cambridge Univ. Press; 2003.
[18] Bellin A, Severino G, Fiori A. On the local concentration probability density function of solutes reacting upon mixing. Water Resour Res 2011;47:W01514.
[19] Bellin A, Tonina D. Probability density function of non-reactive solute concentration in heterogeneous porous formations. J Contam Hydrol 2007;94:109–25.
[20] Benekos ID, Shoemaker CA, Stedinger JR. Probabilistic risk and uncertainty analysis for bioremediation of four chlorinated ethenes in groundwater. Stoch Environ Res Risk Assess 2007;21(4):375–90.
[21] Bernoulli D. Exposition of a new theory on the measurement of risk. Econometrica 1954;22(1):23–36.
[22] Bernstein L et al. (Core Writing Team), Climate change 2007: synthesis report (fourth assessment report). Technical report, The Intergovernmental Panel on Climate Change, IPCC Plenary XXVII, Valencia, Spain, 12–17 November;2007.
[23] Bianchi M, Zheng C, Wilson C, Tick GR, Liu G, Gorelick SM. Spatial connectivity in a highly heterogeneous aquifer: from cores to preferential flow paths. Water Resour Res 2011;47:W05524.
[24] Black P, Stockton T. Basic steps for the development of decision support systems. In: Marcomini A, Suter GW, Critto A, editors. Decision support systems for risk-based management of contaminated sites. New York: Springer; 2009. p. 1–27 [Chapter 1].
[25] Boardman A, Greenberg D, Vining A, Weimer D. Cost-benefit analysis: concepts and practice. 3rd ed. New York: Prentice Hall; 2005.
[26] Bolster D, Barahona M, Dentz M, Fernández-Garcia, Sanchez-Vila X, Trinchero P, et al. Probabilistic risk analysis of groundwater remediation strategies. Water Resour. Res. 2009;45:W06413.
[27] Bowden AR, Rigg AJ. Assessing risk in CO_2 storage projects. Aust Petrol Product Explor Assoc J 2004;44(1):677–702.
[28] Broyda S, Dentz M, Tartakovsky DM. Probability density functions for advective–reactive transport in radial flow. Stoch Environ Res Risk Assess 2010;24(7):985–92.
[29] Burnham KP, Anderson DR. Model selection and multi-model inference: a practical information-theoretic approach. New York: Springer-Verlag; 2002.
[30] Cardiff M, Liu X, Kitanidis PK, Parker J, Kim U. Cost optimization of DNAPL source and plume remediation under uncertainty using a semi-analytic model. J Contam Hydrol 2010;113(1–4):25–43.
[31] Lee S-Y, Carle SF, Fogg GE. Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. Adv Water Res 2007;30:1914–32.
[32] Chow TE, Gaines KF, Hodgsona ME, Wilson M. Habitat and exposure modelling for ecological risk assessment: a case study for the raccoon on the Savannah River site. Ecol Model 2005;189(1-2):151–67.
[33] Cirpka OA, Schwede RL, Luo J, Dentz M. Concentration statistics for mixing-controlled reactive transport in random heterogeneous media. J Contam Hydrol 2008;98(1-2):61–74.
[34] Clemen R. Making hard decisions: an introduction to decision analysis. 2nd ed. Belmont, CA: Duxbury Press; 1996.
[35] Cohrssen JJ, Covello VT. Risk analysis: a guide to principles and methods for analyzing health and environmental risks. Council on Environmental Quality, Washington, D.C., 1989.
[36] Covello VT, Mumpower J. Risk analysis and risk management: an historical perspective. Risk Anal 1985;5(2):103–20.

[37] Culver TB, Shoemaker CA. Optimal control for groundwater remediation by differential dynamic programming with Quasi-Newton Approximations. Water Resour Res 1993;29(4):823–31.
[38] Cushman DJ, Driver KS, Ball SD. Risk assessment for environmental contamination: an overview of the fundamentals and application of risk assessment at contaminated sites. Can J Civil Eng 2001;28(S1):155–62.
[39] Cvetkovic V, Dagan G. Reactive transport and immiscible flow in geological media. II. Applications. Proc R Soc Lond A 1996;452:303–28.
[40] Dai FC, Lee CF, Ngai YY. Landslide risk assessment and management: an overview. Eng Geol 2002;64(1):65–87.
[41] Dai Z, Ritzi RW, Dominic DF. Improving permeability semivariograms with transition probability models of hierarchical sedimentary architecture derived from outcropanalog studies. Water Resour Res 2005;41:W07032.
[42] Dale VH, Biddinger GR, Newman MC, Oris JT, Suter GW, Thompson T, et al. Enhancing the ecological risk assessment process. Integr. Environ. Assess. Manage. 2008;4(3):306–13.
[43] de Barros FPJ, Ezzedine S, Rubin Y. Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics. Adv Water Resour 2012;36:51–63.
[44] de Barros FPJ, Rubin Y, Maxwell R. The concept of comparative information yield curves and its application to risk-based site characterization. Water Resour Res 2009;45(6):W06401.
[45] De Lange HJ, Sala S, Vighi M, Faber JH. Ecological vulnerability in risk assessment – A review and perspectives. Sci Total Environ 2010;408(18): 3871–9.
[46] Dentz M. Concentration statistics for transport in heterogeneous media due to stochastic fluctuations of the center of mass velocity. Adv Water Resour 2012;36:11–22.
[47] Dentz M, Carrera J. Effective solute transport in temporally fluctuating flow through heterogeneous media. Water Resour Res 2005;41:W08414.
[48] Dentz M, Tartakovsky DM. Probability density functions for passive scalars dispersed in random velocity fields. Geophys Res Lett 2010;37:L24406. http://dx.doi.org/10.1029/2010GL045748.
[49] Deutsch CV. FORTRAN programs for calculating connectivity of three-dimensional numerical models and for ranking multiple realizations. Comput Geosci 1998;24(1):69–76.
[50] Diks CGH, Vrugt JA. Comparison of point forecast accuracy of model averaging methods in hydrologic applications. Stoch Environ Res Risk Asses 2010;24(6):809–20.
[51] Diwekar U. Introduction to applied optimization. 2nd ed. Springer; 2008.
[52] Duan Q, Ajami NK, Gao X, Sorooshian S. Multi-model ensemble hydrologic prediction using bayesian model averaging. Adv Water Resour 2007;30(5):1371–86.
[53] Ellingwood BR, Kinali K. Quantifying and communicating uncertainty in seismic risk assessment. Struct Saf 2009;31(2):179–87.
[54] Ellsberg D. Risk, ambiguity, and the Savage axioms. Quart J Econom 1961;75(4):643–69.
[55] Engdahl NB, Weissmann GS, Bonal ND. An integrated approach to shallow aquifer characterization: combining geophysics and geostatistics. Comput Geosci 2010;14:217–29.
[56] Enzenhöfer R, Nowak W, Helmig R. Probabilistic exposure risk assessment with advective–dispersive well vulnerability criteria. Adv Water Resour 2012;36:121–32.
[57] Eskesen SD, Tengborg P, Kampmann J, Veicherts TH. Guidelines for tunnelling risk management: International Tunnelling Association, Working Group No. 2. Tunnel Underground Space Technol 2004;19(3):217–37.
[58] Farmani R, Henriksen HJ, Savic D. An evolutionary Bayesian belief network methodology for optimum management of groundwater contamination. Environ Model Softw 2009;24(3):303–10.
[59] Fernández-Garcia D, Bolster D, Sanchez-Vila X, Tartakovsky DM. A Bayesian approach to integrate temporal data into probabilistic risk analysis of monitored NAPL remediation. Adv Water Resour 2012;36:108–20.
[60] Feyen L, Gorelick SM. Framework to evaluate the worth of hydraulic conductivity data for optimal groundwater resources management in ecologically sensitive areas. Water Resour Res 2005;41:W03019. http://dx.doi.org/10.1029/2003WR002901.
[61] Figueira J, Greco S, Ehrgott M, editors. Multiple criteria decision analysis: state of the art surveys. New York: Springer; 2005.
[62] Finley B, Proctor D, Scott P, Harrington N, Paustenbach D, Price P. Recommended distributions for exposure factors frequently used in health risk assessment. Risk Anal 1994;14(4):533–53.
[63] Fiori A. The Lagrangian concentration approach for determining dilution in aquifer transport: theoretical analysis and comparison with field experiments. Water Resour Res 2001;37(12):3105–14.
[64] Fischhoff B. Risk perception and communication unplugged: twenty years of process. Risk Anal 1995;15(2):137–45.
[65] Foglia L, Mehl SW, Hill MC, Perona P, Burlando P. Testing alternative ground water models using cross-validation and other methods. Ground Water 2007;45(5):627641.
[66] Freeze RA, Gorelick SM. Convergence of stochastic optimization and decision analysis in the engineering design of aquifer remediation. Ground Water 1999;37(6):934–54.
[67] Freeze RA, Bruce J, Massman J, Sperling T, Smith L. Hydrogeological decision analysis: 4. The concept of data worth and its use in the development of site investigation strategies. Ground Water 1992;30(4):574–88.
[68] Gaganis P, Smith L. Accounting for model error in risk assessments: alternatives to adopting a bias towards conservative risk estimates in decision models. Adv Water Resour 2008;31(8):1074–86.
[69] Gemma M, Tsur Y. The stabilization value of groundwater and conjunctive water management under uncertainty. Appl Econ Perspect Pol 2007;29(3):540–8.
[70] Gondwe BRN, Merediz-Alonso G, Bauer-Gottwein P. The influence of conceptual model uncertainty on management decisions for a groundwater-dependent ecosystem in karst. J Hydrol 2011;400(1–2):24–40.
[71] Goodrich MT, McCord JT. Quantification of uncertainty in exposure assessments at hazardous waste sites. Ground Water 1995;33(5):727–32.
[72] Gorelick SM. Large scale nonlinear deterministic and stochastic optimization: formulations involving simulation of subsurface contamination. Math Program 1990;48(1-3):19–39.
[73] Gorelick SM. Incorporating uncertainty into aquifer management models. In: Dagan G, Neuman SP, editors. Subsurface flow and transport: a stochastic approach. New York: Cambridge Univ. Press; 1997. p. 101–11.
[74] Guadagnini A, Franzetti S. Time-related capture zones for contaminants in randomly heterogeneous formations. Ground Water 1999;37(2):253–60.
[75] Guadagnini L, Guadagnini A, Tartakovsky DM. Probabilistic reconstruction of geologic facies. J Hydrol 2004;294:57–67.
[76] Guan J, Aral MM. Remediation system design with multiple uncertain parameters using fuzzy sets and genetic algorithm. J Hydrol Eng 2005;180(5):386–94.
[77] Guin A, Ritzi Jr RW. Studying the effect of correlation and finite-domain size on spatial continuity of permeable sediments. Geophys Res Lett 2008;35:L10402.
[78] Haas CN, Anotai J, Engelbrecht RS. Monte Carlo assessment of microbial risk associated with landfilling of fecal material. Water Environ Res 1996;68(7):1123–31.
[79] Haas CN, Rose JB, Gerba C, Regli S. Risk assessment of virus in drinking water. Risk Anal 1993;13(5):545–52.
[80] Hamed MM, Bedient PB. On the performance of computational methods for the assessment of risk from ground-water contamination. Ground Water 1997;35(4):638–46.
[81] Harter T. Finite-size scaling analysis of percolation in three-dimensional correlated binary Markov chain random fields. Phys Rev E 2005;72(2):026120.
[82] Infanger G, editor. Stochastic programming: the state of the art. Springer; 2010.
[83] James BR, Gorelick SM. When enough is enough: the worth of monitoring data in aquifer remediation design. Water Resour Res 1994;30(12):3499–513.
[84] Jensen FV. Bayesian networks and decision graphs. New York: Springer; 2002.
[85] Jurado A, De Gaspari F, Vilarrasa V, Bolster D, Sánchez-Vila X, Fernández-Garcia D, et al. Probabilistic analysis of groundwater-related risks at subsurface excavation sites. Eng Geol 2012;125:35–44.
[86] Katic P, Grafton RQ. Optimal groundwater extraction under uncertainty: resilience versus economic payoffs. J Hydrol 2011;406(3–4):215–24.
[87] Kendall RJ, Lacker Jr TE, Bunck C, Daniel B, Driver C, Grue CE, et al. An ecological risk assessment of lead shot exposure in non-waterfowl avian species: upland game birds and raptors. Environ Toxicol Chem 1996;15(1):4–20.
[88] Kentel E, Aral MM. Probabilistic-fuzzy health risk modeling. Stoch Environ Res Risk Asses 2004;18:324–38.
[89] Kessler JH, McGuire RK. Total system performance assessment for waste disposal using a logic tree approach. Risk Anal 1999;19(5):915–31.
[90] Khadam IM, Kaluarachchi JJ. Multi-criteria decision analysis with probabilistic risk assessment for the management of contaminated ground water. Environ Impact Assess Rev 2003;23(6):683–721.
[91] Khan NI, Owens G, Bruce D, Naidu R. Human arsenic exposure and risk assessment at the landscape level: a review. Environ Geochem Health 2009;31:143–66.
[92] Knödel K, Toloczyki M, Bohn A, Abel T, Lange G, Tejedo A. Data fusion. In: Knödel K, Lange G, Voigt H-J, editors. Environmental geology: handbook of field methods and case studies. Berlin: Springer; 2007. p. 1053–98 [Chapter 6].
[93] Knudby C, Carrera J. On the relationship between indicators of geostatistical, flow and transport connectivity. Adv Water Resour 2005;28(4):405–21.
[94] Koltermann CE, Gorelick SM. Heterogeneity in sedimentary deposits: a review of structure-imitating, process-imitating, and descriptive approaches. Water Resour Res 1996;32(9):2617–58.
[95] Kourakos G, Mantoglou A. Remediation of heterogeneous aquifers based on multiobjective optimization and adaptive determination of critical realizations. Water Resour Res 2008;44:W12408. http://dx.doi.org/10.1029/2008WR007108.
[96] Krueger T, Freer J, Quinton JN, Macleod CJA, Bilotta GS, Brazier RE, et al. Ensemble evaluation of hydrological model hypotheses. Water Resour Res 2010;46:W07516.
[97] Lahkim MB, Garcia LA. Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer. 1: Monte Carlo uncertainty analysis. Environ Eng Sci 1999;16(5):315–28.
[98] Lahkim MB, Garcia LA, Nuckols JR. Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer. 2: Application of Latin Hypercube sampling. Environ Eng Sci 1999;16(5):329–43.

[99] Lee S, Kitanidis PK. Optimal estimation and scheduling in aquifer remediation with incomplete information. Water Resour Res 1991;27(9):2203–17.
[100] Lee S, Kitanidis PK. Optimization of monitoring well installation time and location during aquifer decontamination. Water Resour Manage 1996;10(6):439–62.
[101] Lemke LD, Bahrou AS. Partitioned multiobjective risk modeling of carcinogenic compounds in groundwater. Stoch Environ Res Risk Assess 2009;23:27–39.
[102] Lester RR, Green LC, Linkov I. Site-specific applications of probabilistic health risk assessment: review of the literature since 2000. Risk Anal 2007;27(3):635–58.
[103] Lin G, Tartakovsky AM, Tartakovsky DM. Uncertainty quantification via random domain decomposition and probabilistic collocation on sparse grids. J Comput Phys 2010;229(19):6995–7012.
[104] Liu W, Medina Jr MA, Thomann W, Piver WT, Jacobs TL. Optimization of intermittent pumping schedules for aquifer remediation using a genetic algorithm. J Am Water Resour Assoc 2000;36(6):1335–48.
[105] Loos M, Schipper AM, Schlink U, Strebel K, Ragas AMJ. Receptor-oriented approaches in wildlife and human exposure modelling: a comparative study. Environ Model Softw 2010;25(4):369–82.
[106] Lundgren RE, McMakin AH. Risk communication: a handbook for communicating environmental, safety, and health risks. 4th ed. Hoboken, NJ: Wiley; 2009.
[107] Macintosh DL, Suter II GW, Hoffman FO. Uses of probabilistic exposure models in ecological risk assessments of contaminated sites. Risk Anal 1994;14(4):405–19.
[108] Maddalena RL, McKone TE, Sohn MD. Standardized approach for developing probabilistic exposure factor distributions. Risk Anal 2004;24(5):1185–99.
[109] Mansfield CM, Shoemaker CA, Liao L-Z. Utilizing sparsity in time-varying optimal control of aquifer cleanup. J Water Resour Plan Manage. 1998;124(1):15–21.
[110] Marcomini A, Suter GW, Critto A, editors. Decision support systems for risk-based management of contaminated sites. New York: Springer; 2009.
[111] Massmann J, Freeze RA. Groundwater contamination from waste management sites: the interaction between risk-based engineering design and regulatory policy, 1. Methodology. Water Resour Res 1987;23(2):351–67.
[112] Maxwell RM, Carle SF, Tompson AFB. Contamination, risk, and heterogeneity: on the effectiveness of aquifer remediation. Environ Geol 2008;54(8):1771–86.
[113] Maxwell RM, Kastenberg WE. Stochastic environmental risk analysis: an integrated methodology for predicting cancer risk from contaminated groundwater. Stoch Environ Res Risk Asses 1999;13:27–47.
[114] Maxwell RM, Kastenberg WE, Rubin Y. A methodology to integrate site characterization information into groundwater-driven health risk assessment. Water Resour Res 1999;35(9):2841–55.
[115] Maxwell RM, Pelmulder SD, Tompson AFB, Kastenberg WE. On the development of a new methodology for groundwater-driven health risk assessment. Water Resour Res 1998;34(4):833–47.
[116] McKone TE, Bogen KT. Predicting the uncertainties in risk assessment: a California groundwater case study. Environ Sci Technol 1991;25(10):1674–81.
[117] McKone TE, Bogen KT. Uncertainties in health-risk assessment: an integrated case study based on tetrachloroethylene in California groundwater. Regul Toxicol Pharmacol 1992;15(1):86–103.
[118] Mongin P. Expected utility theory. In: Davis J, Hands W, Maki U, editors. Handbook of economic methodology. London, UK: Edward Elgar; 1997. p. 342–50.
[119] Morgan DR, Eheart JW, Valocchi AJ. Aquifer remediation design under uncertainty using a new chance constrained programming technique. Water Resour Res 1993;29(3):551–62.
[120] Mylopoulos YA, Theodosiou N, Mylopoulos NA. A stochastic optimization approach in the design of an aquifer remediation under hydrogeologic uncertainty. Water Resour Manage 1999;13(5):335–51.
[121] Na UJ, Chaudhuri SR, Shinozuka M. Effects of spatial variation of soil properties on seismic performance of port structures. Soil Dynam Earthquake Eng 2009;29(3):537–45.
[122] Nachabe MH, Morel-Seytoux HJ. Perturbation and Gaussian methods for stochastic flow problems. Adv Water Resour 1995;18(1):1–8.
[123] Neuman SP. Maximum likelihood Bayesian averaging of uncertain model predictions. Stoch Environ Res Risk Asses 2003;17(5):291–305.
[124] Neuman SP, Tartakovsky DM. Perspective on theories of anomalous transport in heterogeneous media. Adv Water Resour 2009;32(5):670–80.
[125] Neuman SP, Xue L, Ye M, Lu D. Bayesian analysis of data-worth considering model and parameter uncertainties. Adv Water Resour 2011;36:75–85.
[126] Nowak W, de Barros FPJ, Rubin Y. Bayesian geostatistical design: task-driven optimal site investigation when the geostatistical model is uncertain. Water Resour Res 2010;46(W03535). http://dx.doi.org/10.1029/2009WR008312.
[127] Öberg T, Bergbäck B. A review of probabilistic risk assessment of contaminated land. J Soils Sediments 2005;5(4):213–24.
[128] Oldenburg CM. Screening and ranking framework for geologic CO_2 storage site selection on the basis of health, safety, and environmental risk. Environ Geol 2008;54:1687–94.
[129] Pachepsky Y, Timlin D, Rawls W. Generalized Richards' equation to simulate water transport in unsaturated soils. J Hydrol 2003;272(1-4):3–13.
[130] Parker J, Kim U, Kitanidis PK, Cardiff M, Liu X. Stochastic cost optimization of multistrategy DNAPL site remediation. Ground Water Monit Rem 2010;30(3):65–78.
[131] Parmigiani G, Inoue L. Wiley series in probability and statistics, decision theory: principles and approaches. New York: Wiley; 2009.
[132] Paustenbach DJ, Fehling K, Scott P, Harris M, Kerger BD. Identifying soil cleanup criteria for dioxins in urban residential soils: how have 20 years of research and risk assessment experience affected the analysis? J Toxicol Environ Health Part B 2006;9(2):87–145.
[133] Peña-Haro S, Pulido-Velazquez M, Llopis-Albert C. Stochastic hydro-economic modeling for optimal management of agricultural groundwater nitrate pollution under hydraulic conductivity uncertainty. Environ Model Softw 2011;26(8):999–1008.
[134] Pearl J. Causality: models, reasoning, and inference. New York: Cambridge Univ. Press; 2000.
[135] Persson K, Destouni G. Propagation of water pollution uncertainty and risk from the subsurface to the surface water system of a catchment. J Hydrol 2009;377(3–4):434–44.
[136] Philbrick CR, Kitanidis PK. Optimal conjunctive-use operations and plans. Water Resour Res 1998;34(5):1307–16.
[137] Poeter E, Anderson D. Multimodel ranking and inference in ground water modeling. Ground Water 2005;43(4):597–605.
[138] Porporato A, Vico G, Fay P. Interannual hydroclimatic variability and ecosystem superstatistics. Geophys Res Lett 2006;33:L5402.
[139] Proce CJ, Ritzi RW, Dominic DF, Dai Z. Modeling multiscale heterogeneity and aquifer interconnectivity. Ground Water 2004;42(5):658–70.
[140] Refsgaard JC, Christensen S, Sonnenborg TO, Seifert D, Hojberg AL, Troldborg L. Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. Adv Water Resour 2012;36:36–50.
[141] Reichard EG. Ground watersurface water management with stochastic surface water supplies: A simulation optimization approach. Water Resour Res 1995;31(11):2845–65.
[142] Riederer AM, Thompson KM, Fuentes JM, Ford TE. Body weight and water ingestion estimates for women in two communities in the Philippines: the importance of collecting site-specific data. Int J Hyg Environ Health 2006;209(1):69–80.
[143] Rigby JR, Porporato A. Spring frost risk in a changing climate. Geophys Res Lett 2008;35:L12703.
[144] Rodak C, Silliman S. Probabilistic risk analysis and fault trees: initial discussion of application to identification of risk at a wellhead. Adv Water Resour 2011;36:133–45.
[145] Rodriguez-Iturbe I, Porporato A, Ridolfi L, Cox D, Isham V. Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation. Proc Roy Soc A 1999;455:3789–805.
[146] Roisenberg M, Schoeninger C, Rodrigues da Silva R. A hybrid fuzzy-probabilistic system for risk analysis in petroleum exploration prospects. Expert Syst Appl 2009;36(3):6282–94.
[147] Rojas R, Kahunde S, Peeters L, Batelaan O, Feyen L, Dassargues A. Application of a multimodel approach to account for conceptual model and scenario uncertainties in groundwater modelling. J Hydrol 2010;394(3-4):416–35.
[148] Sahinidis NV. Optimization under uncertainty: state-of-the-art and opportunities. Comput Chem Eng 2004;28(6–7):971–83.
[149] Sander P, Bergbäck B, Öberg T. Uncertain numbers and uncertainty in the selection of input distributions consequences for a probabilistic risk assessment of contaminated land. Risk Anal 2006;26(5):1363–75.
[150] Savage LJ. The foundations of statistics. 2nd ed. New York: Dover; 1972.
[151] Short M, Higdon D, Guadagnini L, Guadagnini A, Tartakovsky DM. Predicting vertical connectivity within an aquifer system. Bayesian Anal 2010;5(3):557–82.
[152] Siirila ER, Navarre-Sitchler AK, Maxwell RM, McCray JE. A quantitative methodology to assess the risks to human health from CO_2 leakage into groundwater. Adv Water Resour 2012;36:146–64.
[153] Singh A, Mishra S, Ruskauff G. Model averaging techniques for quantifying conceptual model uncertainty. Ground Water 2010;48(5):701–15.
[154] Skaggs TH, Barry DA. The first-order reliability method of predicting cumulative mass flux in heterogeneous porous formations. Water Resour Res 1997;33:1485–94.
[155] Slovic P. Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. Risk Anal. 1999;19(4):689–701.
[156] Smith J, Durucan S, Korre A, Shi J-Q. Carbon dioxide storage risk assessment: analysis of caprock fracture network connectivity. Int J Greenhouse Gas Control 2011;5(2):226–40.
[157] Tartakovsky DM. Probabilistic risk analysis in subsurface hydrology. Geophys Res Lett 2007;34:L05404.
[158] Tartakovsky DM, Broyda S. PDF equations for advective-reactive transport in heterogeneous porous media with uncertain properties. J Contam Hydrol 2011;120–121:129–40.
[159] Tartakovsky DM, Dentz M, Lichtner PC. Probability density functions for advective–reactive transport in porous media with uncertain reaction rates. Water Resour Res 2009;45:W07414. http://dx.doi.org/10.1029/2008WR007383.
[160] Trinchero P, Sánchez-Vila X, Fernàndez-Garcia D. Point-to-point connectivity, an abstract concept or a key issue for risk assessment studies? Adv Water Resour 2008;31:1742–53.

[161] Troldborg M, Lemming G, Binning PJ, Tuxen N, Bjerg PL. Risk assessment and prioritisation of contaminated sites on the catchment scale. J Contam Hydrol 2008;101(1–4):14–28.
[162] Troldborg M, Nowak W, Tuxen N, Bjerg PL, Helmig R, Binning PJ. Uncertainty evaluation of mass discharge estimates from a contaminated site using a fully Bayesian framework. Water Resour Res 2010;46:W12552.
[163] US Environmental Protection Agency, Guidelines for Exposure Assessment (USEPA 600Z-92/001), Washington DC; 1992.
[164] US National Research Council, Science and Judgment in Risk Assessment, Taylor & Francis, New York; 1996.
[165] US National Research Council, Review of Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts, Natl. Academy Press, Washington, DC; 1997.
[166] US National Research Council, Complete the Forecast, US National Academies Press, Washington DC; 2006.
[167] US Nuclear Regulatory Commission, Safety Evaluation Report Related to Disposal of High-Level Radioactive Wastes in a Geologic Repository at Yucca Mountain, Nevada (NUREG-1949, Volume 2), Washington, DC, 2010.
[168] van Ginneken M, Oron G. Risk assessment of consuming agricultural products irrigated with reclaimed wastewater: an exposure model. Water Resour Res 2000;36(9):2691–9.
[169] Veneziano D, Tabaei A. Nonlinear spectral analysis of flow through porous media with isotropic lognormal hydraulic conductivity. J Hydrol 2004;294:4–17.
[170] Verderame PM, Elia JA, Li J, Floudas CA. Planning and scheduling under uncertainty: a review across multiple sectors. Ind Eng Chem Res 2010;49(9):3993–4017.
[171] Vesely WE, Roberts NH. (NUREG-0492), Fault Tree Handbook, US Nuclear Regulatory Commission.
[172] Vico G, Porporato A. Traditional and micro-irrigation with stochastic soil moisture. Water Resour Res 2010;46. http://dx.doi.org/10.1029/2009WR008130.
[173] von Neumann J, Morgenstern O. Theory of games and economic behavior. 2nd ed. Princeton, NJ: Princeton University Press; 1947.
[174] Wagner BJ, Gorelick SM. Optimal groundwater quality management under parameter uncertainty. Water Resour Res 1987;23(7):1162–74.
[175] Wagner BJ, Gorelick SM. Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. Water Resour Res 1989;25(10):2211–25.
[176] Wagner JM, Shamir U, Nemati HR. Groundwater quality management under uncertainty: stochastic programming approaches and the value of information. Water Resour Res 1992;28(5):1233–46.
[177] Wang P, Tartakovsky DM. Probabilistic predictions of infiltration into heterogeneous media with uncertain hydraulic parameters. Int J Uncert Quant 2011;1(1):35–47.
[178] Wang P, Tartakovsky DM. Reduced complexity models for probabilistic forecasting of infiltration rates. Adv Water Resour 2011;34:375–82.
[179] Whiffen GJ, Shoemaker CA. Nonlinear weighted feedback control of groundwater remediation under uncertainty. Water Resour Res 1993;29(9):3277–89.
[180] White DJ. Decision theory. Rutgers, NJ: Transaction Publishers; 2006.
[181] Wildenborg AFB, Leijnse AL, Kreft E, Nepveu MN, Obdam ANM, Orlic B. Risk assessment methodology for CO_2 storage: the scenario approach. In: Thomas DC, Benson SM, editors. Carbon dioxide capture for storage in deep geologic formations, vol. 2. Amsterdam: Elsevier; 2005. p. 1293–316.
[182] Winter CL, Guadagnini A, Nychka D, Tartakovsky DM. Multivariate sensitivity analysis of saturated flow through simulated highly heterogeneous groundwater aquifers. J Comput Phys 2006;217(1):166–75.
[183] Winter CL, Nychka D. Forecasting skill of model averages. Stoch Environ Res Risk Assess 2010;24:633–8.
[184] Winter CL, Tartakovsky DM. Mean flow in composite porous media. Geophys Res Lett 2000;27(12):1759–62.
[185] Winter CL, Tartakovsky DM. Groundwater flow in heterogeneous composite aquifers. Water Resour Res 2002;38(8). 10.1029/2001WR00045.
[186] Winter CL, Tartakovsky DM. A reduced complexity model for probabilistic risk assessment of groundwater contamination. Water Resour Res 2008;44:W06501. http://dx.doi.org/10.1029/2007WR006599.
[187] Winter CL, Tartakovsky DM, Guadagnini A. Moment equations for flow in highly heterogeneous porous media. Surv Geophys 2003;24(1):81–106.
[188] Wohlberg BE, Tartakovsky DM. Delineation of geological facies from poorly differentiated data. Adv Water Resour 2009;32(2):225–30.
[189] Wohlberg BE, Tartakovsky DM, Guadagnini A. Subsurface characterization with support vector machines. IEEE Trans Geosci Remote Sens 2006;44(1):47–57.
[190] Xiu D, Tartakovsky DM. A two-scale non-perturbative approach to uncertainty analysis of diffusion in random composites. J Multiscale Model Simul 2004;2(4):662–74.
[191] Yan S, Minsker B. Applying dynamic surrogate models in noisy genetic algorithms to optimize groundwater remediation designs. J Water Resour Plan Manage 2011;137(3):284–92.
[192] Ye M, Pohlmann KF, Chapman JB, Pohll GM, Reeves DM. A model-averaging method for assessing groundwater conceptual model uncertainty. Ground Water 2010;48(5):716–28.
[193] Zadeh LA. Fuzzy sets. Inform Control 1965;8(3):338–53.
[194] Zhang P, Aagaard P, Gottschalk L. Probability method used in predicting contaminant risk in groundwater adjacent to airport. Water Air Soil Pollut 2010;211:323–39.
[195] Zhang Y, Oldenburg CM, Finsterle S. Percolation-theory and fuzzy rule-based probability estimation of fault leakage at geologic carbon sequestration sites. Environ Earth Sci 2010;59(7):1447–59.
[196] Zhang Y, Vouzis P, Sahinidis NV. GPU simulations for risk assessment in CO_2 geologic sequestration. Comput Chem Eng 2011;35(10):1631–44.