SPECIAL ISSUE

Conditional Probability Density Functions of Concentrations for Mixing-Controlled Reactive Transport in Heterogeneous Aquifers

**X. Sanchez-Vila · A. Guadagnini ·
D. Fernàndez-Garcia**

Abstract This paper presents an approach conducive to an evaluation of the probability density function (pdf) of spatio-temporal distributions of concentrations of reactive solutes (and associated reaction rates) evolving in a randomly heterogeneous aquifer. Most existing approaches to solute transport in heterogeneous media focus on providing expressions for space–time moments of concentrations. In general, only low order moments (unconditional or conditional mean and covariance) are computed. In some cases, this allows for obtaining a confidence interval associated with predictions of local concentrations. Common applications, such as risk assessment and vulnerability practices, require the assessment of extreme (low or high) concentration values. We start from the well-known approach of deconstructing the reactive transport problem into the analysis of a conservative transport process followed by speciation to (a) provide a partial differential equation (PDE) for the (conditional) pdf of conservative aqueous species, and (b) derive expressions for the pdf of reactive species and the associated reaction rate. When transport at the local scale is described by an Advection Dispersion Equation (ADE), the equation satisfied by the pdf of conservative species is non-local in space and time. It is similar to an ADE and includes an additional source term. The latter involves the contribution of dilution effects that counteract dispersive fluxes. In general, the PDE we provide must be solved numerically, in a Monte Carlo framework. In some cases, an approximation can be obtained through suitable localization of the governing equation. We illustrate the methodology to depict key features of transport in randomly stratified media in

X. Sanchez-Vila (✉) · D. Fernàndez-Garcia
Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia,
08034 Barcelona, Spain
e-mail: xavier.sanchez-vila@upc.edu

A. Guadagnini
Dipartimento di Ingegneria Idraulica, Ambientale, Infrastrutture Viarie, Rilevamento (DIIAR),
Politecnico di Milano, Piazza L. Da Vinci, 32, 20133, Milano, Italy

� Springer

the absence of transverse dispersion effects. In this case, all the pdfs can be explicitly obtained, and their evolution with space and time is discussed.

Keywords Reactive transport · Heterogeneous porous media

1 Introduction

The complete analysis of a reactive transport problem typically involves specification of a large number of aqueous and non-aqueous species. A common method to estimate spatial distributions of chemical species concentrations and associated reaction rates is to employ numerical modeling. A series of mathematical formulations are available in the literature, which are included in a variety of codes (Rubin 1990; Yeh and Tripathi 1991; Friedly and Rubin 1992; Lichtner 1996; Steefel and Mac-Quarrie 1996; Tebes-Stevens et al. 1998; Clement et al. 1998; Saaltink et al. 1998; Parkhurst and Appelo 1999; Robinson et al. 2000; Molins et al. 2004). All these methodologies are based on the idea that reactive transport problems can be reformulated mathematically in terms of chemical components, defined as linear combinations of reactive species concentrations.

The methodology typically consists of the following steps: (a) decoupling the solution of the equations associated with the chemical reactions from the mass balance equations governing chemical components; (b) solving these conservative transport equations by means of either Eulerian or Lagrangian methods; (c) performing speciation calculations to obtain the concentration of aqueous species from the space–time distribution of components; and (d) substituting the latter into the transport equations to evaluate numerically reaction rates. This approach can be applied when the concentrations of the reacting species stand in algebraic relationships, and the coefficients describing physical mixing in the system coincide for all compounds. The former requirements are met by systems either in local chemical equilibrium or instantaneous, complete, irreversible reactions (Ham et al. 2004; Liedl et al. 2005) and can also be used for specific cases of kinetic reactions (Molins et al. 2004; Cirpka and Valocchi 2007). A methodology to compute directly homogeneous and heterogeneous reaction rates under instantaneous equilibrium has been presented recently by De Simoni et al. (2005, 2007). Their general expressions illustrate that mixing processes control equilibrium reaction rates, and includes Phillips's model (1991) as a particular case. The approach has been used by Guadagnini et al. (2008), in conjunction with laboratory-scale $CaCO_3$ dissolution experiments reported in Singurindy et al. (2004) to describe spatial distribution of reactions rates in a homogeneous flow cell.

Natural heterogeneity constitutes a key challenge to modeling flow and transport in the subsurface. Adopting a stochastic framework has become a convenient and popular way to address the key dynamics of flow and transport processes evolving in natural porous aquifers. In this context, solute transport in randomly heterogeneous media is described by stochastic transport equations. The latter can be solved according to a variety of formulations. The most intuitive amongst these is the Monte Carlo method. Notwithstanding the limitations of the method, it is largely used either in

practical applications or to support analytical/numerical results based on some approximation of the governing equations. The impact of heterogeneity in transport of conservative species is often described by adapting observations to a particular shape of the transport equation. The resulting upscaled parameters, describing dispersion, are generally larger than those describing transport at the local scale, giving raise to the concept of macrodispersion. Different approximations adopted in the mathematical treatment of the stochastic transport equation often lead to different results for the macrodispersion or effective dispersion tensor (Gelhar and Axness 1983; Neuman and Zhang 1990; Dagan 1994; Andricevic and Cvetkovic 1998; Dentz et al. 2000; Fiori 2001) and typically describes the spreading of a plume within a randomly heterogeneous aquifer.

[Besides this type of description, several works are concerned with providing leading order statistical space–time moments of concentrations distributions.] Starting from the work of Dagan (1989), several analytical solutions have been proposed rendering space–time distributions of (ensemble) mean and variance–covariance of concentrations or trajectories and time-of-residence of conservative solutes in multidimensional porous systems (Cvetkovic et al. 1992; Kapoor and Gelhar 1994; Kapoor and Kitanidis 1998; Fiori and Dagan 2000; Vanderborght 2001; Guadagnini et al. 2003; Sanchez-Vila and Guadagnini 2005; Riva et al. 2006). All of these solutions rely on (different flavors of) the perturbation theory to provide approximations of governing equations and associated solutions. These types of analytical solutions are of general interest, as they provide insights on the nature of the solution and allow evaluating the relative importance of the different parameters and processes involved.

A key feature of all these solutions is the observation that the coefficient of variation of a target variable evolves in time and space and can be very large, being linked to the descriptors of the physical heterogeneity of the domain. Therefore, the assessment of the state of an aquifer, based only on low order (statistical) moments, provides a picture which is at best incomplete in the context of risk and/or vulnerability assessment. In particular, Kapoor and Kitanidis (1998) and Fiori and Dagan (2000) showed that in unbounded formations under uniform mean flow conditions the coefficient of variation of the solute concentration is a non-monotonic function of travel time and reaches a maximum at a time linked with the time scale of processes characterizing pore-scale dispersion. In this sense, the interest lays in the knowledge of the extreme values where quantities such as concentration, travel time and/or trajectories can attain in a region of investigation. While this can be accomplished by numerical Monte Carlo simulations (see Riva et al. 2008a for a recent field application related to the probabilistic interpretation of a conservative tracer test), methods based on the assessment of the probability density function (pdf) of solutes are also being proposed as powerful tools. While these methods are well established and developed in the turbulence and combustion literature (Pope 1981; Chen et al. 1989; Girimaji 1991), their application to transport in randomly heterogeneous aquifers has emerged only recently. A recent hydrological application concerning the probabilistic assessment of dynamics of soil nitrate is presented by Botter et al. (2008).

Shvidler and Karasaki (2003) presented a method to compute the pdf, $p(x, t, c)$, of the concentration $c(x, t)$ of nonreactive solutes transported by random purely advective processes in groundwater. In this case the equation satisfied by $p(x, t, c)$ can

be determined by (ensemble) averaging of a Liouville stochastic equation. They illustrate a method of computing the pdf and the mean and variance of concentration for the case of one-dimensional transport upon assuming that the flow velocity can be described by either a Gaussian or a telegraph random process. Fiorotto and Caroni (2002) and Caroni and Fiorotto (2005) used a numerical Lagrangian approach to analyze the combined effect of pore-scale dispersion and sampling volume on the solute concentration moments. Fiorotto and Caroni (2002) simulated a synthetic tracer test in an unconditional heterogeneous formation and obtained a good fit of the simulated observations with a Beta distribution. Caroni and Fiorotto (2005) proposed an integral expression of the concentration pdf and showed numerically that the Beta distribution provides a good approximation for moderate values of the variance of the log-transmissivity ($\sigma_Y^2 \leq 2.0$). It is noted that the Beta distribution is very similar to a Gaussian distribution when the variance is low and tends to be bimodal in the case of a high coefficient of variation. In this sense, these results might shed new insights on the findings of Bellin et al. (1994), who showed that although the cumulative distribution of concentration is close to bimodality for small sampling volumes, it approaches the Normal distribution as the sampling volume increases.

While Fiorotto and Caroni (2002) adopted the Beta distribution as a fitting functional format, Bellin and Tonina (2007) developed a theoretical model for the pdf of solute concentration in heterogeneous formations. Their model was based on the assumption that a solute plume is composed by a large number of independent particles whose motion can be described on the basis of an Ito stochastic differential equation. Their findings suggested that under ergodic conditions, the concentration can be interpreted as a Beta distribution. The parameters of the latter depends on the first two (ensemble) moments of the solute concentration. These authors found a good correspondence between their model and bromide concentrations from the first Cape Cod tracer test (Hess et al. 1992). A number of stochastic analyses of reactive species transport in heterogeneous formations are mainly concerned with (or limited to) the prediction of spatial moments of plumes composed by (eventually kinetically) absorbing solutes within spatially varying velocity fields (Quinodoz and Valocchi 1993; Hu et al. 1995; Reichle et al. 1998) and mainly targeted at deriving key quantities describing the spreading and shape of an evolving contaminant plume. Other works consider the combined effects of physical and chemical heterogeneity on solutes residence times in the context of sorption–desorption processes (Lawrence et al. 2002; Riva et al. 2008b and others). In all these studies, either analytical or Monte Carlo based, the groundwater velocity field can be conveniently represented in terms of rescaled (or retarded) velocity.

Adopting the conceptual framework of De Simoni et al. (2005) allows taking advantage of works related to stochastic analysis of conservative solutes and extending these to deal with reactive transport in randomly heterogeneous aquifers to include different types of reactions. Two recent examples are Luo et al. (2008), who presented a macroscopic formulation of reaction parameters using perturbation theory, and Fernandez-Garcia et al. (2008), who obtain exact analytical/numerical predictions (and associated uncertainty) of reaction rates in randomly stratified heterogeneous media. The latter work provides a solution directly relating the statistical and chemical properties of the formation to the spatial distribution of the reaction rate.

Regarding reactive solutes concentrations, Lichtner and Tartakovsky (2003) apply a pdf formulation to a heterogeneous chemical reaction involving an aqueous solution reacting with a solid phase in a batch. Upon considering both linear and nonlinear kinetic rate laws, they found an approximation for an effective rate constant describing the change in mean concentration with time.

Cirpka et al. (2008) start from recognizing that knowledge of the spatial distribution of mixing ratios of waters of different chemical composition can be unambiguously mapped onto reactive species concentrations (Ham et al. 2004; Liedl et al. 2005; Cirpka et al. 2006; De Simoni et al. 2007). In their study, Cirpka et al. (2008) assumed the pdf of mixing ratios to be known and adopted the Beta distribution. Then they related it to the pdf of concentrations of reactive species concentrations calculated by means of Taylor-series truncations of expressions relating them to mixing ratios expanded around (ensemble) mean mixing ratios. They showed that the low order moments thus computed can lead to nonphysical results, since such expansions rely on the smallness of the variance of mixing ratios, which is not necessarily the case in heterogeneous systems.

The distinctive aim of this work is to present a methodology conducive to the evaluation of the complete pdf of reaction rates and concentrations of reactive species within randomly heterogeneous aquifers. The starting point of our methodology is similar to the one adopted by Shvidler and Karasaki (2003). We consider that the space–time distribution of solutes can be described by means of a local-scale ADE with appropriate initial and boundary conditions. Since our interest is in illustrating the methodology, we choose a simple transport setup: a bi-molecular system where two aqueous solutes are in chemical equilibrium with an immobile solid mineral that undergoes instantaneous precipitation. We then derive the exact equations satisfied by the pdf of both the conservative component and the reactive species. Whenever the pdf for a conservative quantity can be obtained (analytically or numerically), it is then possible to map onto the pdf of reaction rates and concentrations of reactive species, thus avoiding the need for heavy Monte Carlo computations of multicomponent reactive transport problems. The expressions we offer hold for a general distribution of physical heterogeneity of the aquifer. For illustration purposes, we then provide explicit solutions of the pdf and associated moments for reaction rates in the case of a stratified formation.

2 Equation Governing the Evolution of the PDF of Concentration of a Conservative Species

2.1 Problem Statement

We denote by $u(\mathbf{x}, t)$ the concentration of a conservative species (or a suitable conservative linear combination of otherwise reactive species). In principle, we assume that the local scale u is governed by an advection–diffusion equation in a D-dimensional medium

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{v}(\mathbf{x})\nabla u(\mathbf{x}, t) - \nabla \cdot \left[\mathbf{D}_d \nabla u(\mathbf{x}, t)\right] = 0, \quad \mathbf{x} \in \Omega\left(\subset \mathbb{R}^D\right), \qquad (1)$$

subject to initial and boundary conditions

$$u(\mathbf{x}, 0) = U_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{2}$$

$$u(\mathbf{x}, t) = U_D(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_1 \ (\subset \partial \Omega), \tag{3a}$$

$$-\mathbf{D}_d \nabla u(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = W(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_2 \ (\subset \partial \Omega), \tag{3b}$$

where $\mathbf{D}_d$ is a constant (local scale) dispersion tensor, U_D a prescribed value of u on a surface boundary Γ_1, W a dispersive flux of u prescribed normal to the boundary surface Γ_2, and $\mathbf{n}$ an outward unit normal to Γ_2. All quantities are defined at some scale ω. Without loss of generality, flux (mixed) boundaries are not considered in this study.

Whenever either $\mathbf{v}, U_0, U_D, W$ (or any combination of these spatio-temporal functions) are not deterministically specified, (1)–(3) constitutes a stochastic partial differential equation (SPDE) and $u(\mathbf{x}, t)$ is viewed as a random function (stochastic process). In this work we consider all forcing terms U_0, U_D, W to be deterministic. Hence, the randomness of the system is embedded in that of $\mathbf{v}$. The latter stems directly from the variability of hydraulic parameters, such as hydraulic conductivity, $K(\mathbf{x})$.

2.2 Equation Satisfied by the PDF of Conservative Components, $p_u(\mathbf{x}, t)$

For any given realization of the random system we can write an auxiliary function, Π, to indicate that the concentration at a given space–time location take a specified value U, i.e.,

$$\Pi(u, U; \boldsymbol{x}, t) = \delta\big(u(\boldsymbol{x}, t) - U\big). \tag{4}$$

Randomness stems from the different $u(\mathbf{x}, t)$ values at each space–time location in each individual realization. The Π function is termed Density Distribution Function (Shvidler and Karasaki 2003). The ensemble of Π renders the conditional pdf of U, p_u, according to

$$\langle \Pi \rangle = \int \delta\big(u(\boldsymbol{x}, t) - U\big) p_u \, du = p_u(U; \boldsymbol{x}, t), \tag{5}$$

where $\langle \cdot \rangle$ stands for ensemble averaging. For illustration purposes, a conceptualization of the relationship between u and p_u is presented in Fig. 1. Figure 1(a) corresponds to 100 Monte Carlo realizations of concentrations as a function of time. Each curve represents the simulated breakthrough for a conservative solute at the pumping location of a tracer test performed at the Lauweisen site (Riva et al. 2008b) near Tübingen (Germany), under convergent flow conditions. Figures 1(b) and 1(c), respectively, correspond to the numerically evaluated pdf of concentrations at two given times, and can be constructed by taking the 100 Monte Carlo generated concentration values (each one corresponding to one particular realization) and constructing the corresponding frequency distribution. We observe that the associated coefficients
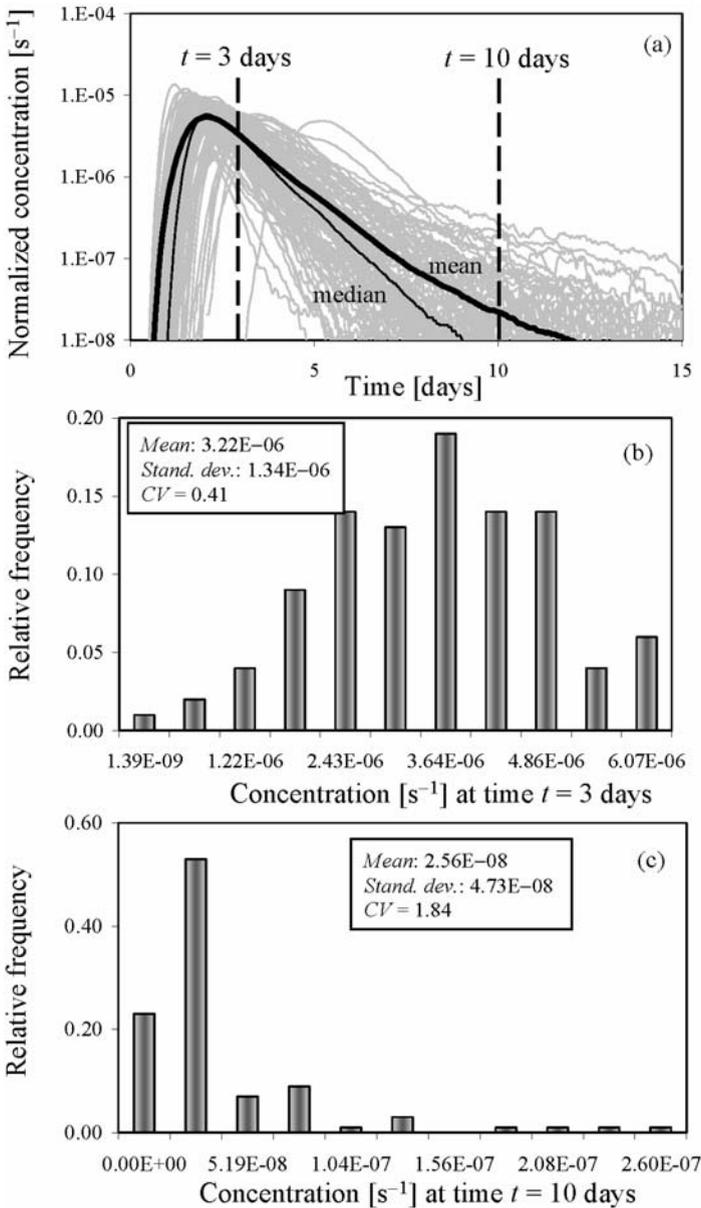
Fig. 1 Schematic representation of the relationship between u and p_u. The value of u (here being mass flux over total mass) at a point in space and time and for a single realization drawn from the ensemble is a given value, U. The ensuing pdf is related to the distribution of the relative frequencies of all the values obtained (at a given observation point in space-time) amongst the ensemble of realizations. The resulting pdf, p_u, varies in space and time

of variation can be quite large and tend to increase with time, indicating that the average value is not necessarily a good predictor for concentrations. The very same effect will be observed in Sect. 5 in the case of uniform flow in stratified aquifers.

We are now in a position to write the equation satisfied by p_u. We start by pre-multiplying (1) by $-\frac{\partial \Pi}{\partial U}$, thus obtaining

$$-\frac{\partial \Pi}{\partial U}\frac{\partial u(\mathbf{x},t)}{\partial t} - \frac{\partial \Pi}{\partial U}\mathbf{v}(\mathbf{x})\nabla u(\mathbf{x},t) + \frac{\partial \Pi}{\partial U}\nabla\cdot\left[\mathbf{D}_d\nabla u(\mathbf{x},t)\right] = 0. \qquad (6)$$

Noticing that,

$$\frac{\partial \Pi}{\partial t} = \frac{\partial \Pi}{\partial u}\frac{\partial u}{\partial t} = -\frac{\partial \Pi}{\partial U}\frac{\partial u}{\partial t}, \qquad (7)$$

$$\nabla \Pi = \frac{\partial \Pi}{\partial u}\nabla u = -\frac{\partial \Pi}{\partial U}\nabla u, \qquad (8)$$

$$\nabla\cdot[\mathbf{D}_d\nabla \Pi] = \nabla\cdot\left[\mathbf{D}_d\frac{\partial \Pi}{\partial u}\nabla u\right] = \frac{\partial \Pi}{\partial u}\nabla\cdot[\mathbf{D}_d\nabla u] + \nabla\frac{\partial \Pi}{\partial u}\cdot[\mathbf{D}_d\nabla u]. \qquad (9)$$

Also, taking advantage of the properties of the derivatives of the delta function $(\delta^{(n)}(z) = \frac{(-1)^n}{z^n}n!\delta(z))$, leads to

$$\frac{\partial \Pi}{\partial u} = -\frac{\Pi}{u-U}, \qquad (10)$$

$$\begin{aligned}\nabla\frac{\partial \Pi}{\partial u} &= -\nabla\left(\frac{\Pi}{u-U}\right) = -\frac{\nabla \Pi}{u-U} - \Pi\nabla\left(\frac{1}{u-U}\right) \\ &= -\frac{1}{u-U}\frac{\partial \Pi}{\partial u}\nabla u + \Pi\left(\frac{1}{u-U}\right)^2\nabla u \\ &= 2\Pi\left(\frac{1}{u-U}\right)^2\nabla u \equiv \frac{\partial^2 \Pi}{\partial u^2}\nabla u.\end{aligned} \qquad (11)$$

By combining (7) to (11) and substituting into (6), we obtain the following format for the equation satisfied by $\Pi(\mathbf{x},t)$

$$\frac{\partial \Pi}{\partial t} + \mathbf{v}(\mathbf{x})\cdot\nabla \Pi(\mathbf{x},t) - \nabla\cdot\left[\mathbf{D}_d\nabla \Pi(\mathbf{x},t)\right] = -\frac{\partial^2 \Pi}{\partial u^2}\nabla u^T\mathbf{D}_d\nabla u. \qquad (12)$$

This result indicates that $\Pi(\mathbf{x},t)$ satisfies an equation, which is formally identical to (1) with an additional sink/source term. In the particular case of non-diffusive transport, we recover the equation already presented by Shvidler and Karasaki (2003). Their equation includes an additional term to account for external sources which we disregard here.

Taking ensemble mean of (12) leads to the following integro-differential equation for p_u

$$\frac{\partial p_u}{\partial t} + \langle\mathbf{v}(\mathbf{x})\rangle\nabla p_u - \nabla\cdot[\mathbf{D}_d\nabla p_u] = -\langle\mathbf{v}'(\mathbf{x})\nabla \Pi\rangle - \int\delta^{(2)}(u-U)p_u\nabla u^T\mathbf{D}_d\nabla u\,du. \qquad (13)$$

Defining an auxiliary function, $f_M(u) = \nabla u^T \mathbf{D}_d \nabla u$, and using integration by parts, the last term on the right-hand side of (13) becomes

$$-\int \delta^{(2)}(u-U) f_M(u) p_u \, du = -\int \delta(u-U) \frac{\partial^2}{\partial u^2} \left[f_M(u) p_u \right] du$$

$$\equiv -\sum \frac{\partial^2}{\partial u^2} \left(p_u \nabla u^T \mathbf{D}_d \nabla u \right) \Big|_{u_i = U}. \qquad (14)$$

Here, the sum stems from the fact that, since u is not necessarily a monotonic function, $u = U$ can take several values at different points in space–time.

A similar approach can be used to assess the first term on the right-hand side of (13). We start by noting that

$$-\langle \mathbf{v}'(\mathbf{x}) \nabla \Pi \rangle = -\left\langle \mathbf{v}'(\mathbf{x}) \frac{\partial \Pi}{\partial u} \nabla u \right\rangle$$

$$= -\int \delta^{(1)}(u-U) \mathbf{v}' \nabla u p_u \, du = \int \delta(u-U) \frac{\partial}{\partial u} [\mathbf{v}' \nabla u p_u] \, du. \, (15)$$

Defining an auxiliary function, $f_D = \mathbf{v}' \nabla u$, and incorporating expressions (14) and (15) into (13) leads to the following equation for p_u

$$\frac{\partial p_u}{\partial t} + \langle \mathbf{v}(\mathbf{x}) \rangle \nabla p_u - \nabla \cdot [\mathbf{D}_d \nabla p_u]$$

$$= \int \delta(u-U) \frac{\partial}{\partial u} \left[f_D(u) p_u \right] du - \int \delta(u-U) \frac{\partial^2}{\partial u^2} \left[f_M(u) p_u \right] du. \quad (16)$$

The initial condition for p_u, $p_u(\mathbf{x}, t=0) = p_u(U; \mathbf{x}, t=0)$, is determined by the initial distribution of u in the system. For instance, if $u(\mathbf{x}, 0) = U_0(\mathbf{x})$ is deterministic, then

$$p_u(U; \mathbf{x}, t=0) = \delta(U - U_0(\mathbf{x})). \qquad (17)$$

Similarly, if the boundary condition along Γ_1 is deterministic, then

$$p_u(U; \mathbf{x}, t) = \delta(U - U_D(\mathbf{x})), \quad \mathbf{x} \in \Gamma_1. \qquad (18)$$

Along Γ_2 we assume continuity of u, so that $W = 0$. Thus, from (8) we can write

$$\mathbf{D}_d \nabla u(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = 0 \Rightarrow [\mathbf{D}_d \nabla \Pi] \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \Gamma_2. \qquad (19a)$$

Taking expected value leads to

$$[\mathbf{D}_d \nabla p_u] \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \Gamma_2. \qquad (19b)$$

We note that the range of the values of u associated with a nonzero p_u value is limited. For example, if the aquifer has a constant concentration U_0 and a pollution process takes place, the valid u values would be in the range $u \in [U_0, U_{max}]$, where U_{max} is a random spatio-temporal function.

2.3 Closure

The governing equation for p_u, (16), contains both the pdf of u and the variable u itself. In a heterogeneous medium (the only case where the pdf of u is non-trivial), the problem cannot be solved in general. The full pdf then could only be exactly computed through a full suite of Monte Carlo numerical simulations. A notable exception is the case of transport evolution within a stratified aquifer, where all the terms in (16) can be computed explicitly. This approach is carried out in Sect. 5. The terms appearing in the right-hand side of (16) deserve some additional comments. The first one of the two terms (given also by (15)) incorporates the effects of the variability of the velocity field. It contains the ensemble mean of the variations of the Π function, caused by variations in velocity. In this sense, its effects might be similar to the macrodispersive term appearing in conservative transport equations, which characterizes spreading of solute. Thus, we envision that this term in (16) causes spreading in the distribution of p_u, and we term it "spreading term".

   The last term in the right hand side of (16), given also by (14), incorporates the quantity $f_M = \nabla u^T \mathbf{D}_d \nabla u$. The latter is directly related to the concept of dilution index (Kitanidis 1994). The interplay between spreading (as represented by f_D) and mixing (as represented by f_M) in heterogeneous media is well known in stochastic hydrology literature and suggested as crucial for providing realistic modeling predictions. While spreading, f_D increases the production of the variance of concentrations at a given space–time location, mixing, f_M counteracts this effect, as it is the mechanism responsible for decreasing the maximum concentration values in the system (Kapoor and Gelhar 1994). The role of mixing is even more important when considering reactive transport. For instance, reaction rates at a given space–time location under local equilibrium conditions directly depend on mixing rather than spreading (De Simoni et al. 2007). Here, we investigate the effects of spreading and mixing on the complete pdf of the conservative and reactive species concentration. As we will show for the particular case of the stratified media, this term actually counteracts the effect of spreading (i.e., subtracts the effects of spreading from those caused by mixing). We refer to the last term in (16) as the 'dilution term'. This term can be computed in different ways, depending on which expression is adopted in (14). Our preliminary analysis of the problem suggests that the most efficient (from a numerical standpoint) approach is to evaluate numerically this term as

$$-\left\langle \frac{\partial^2 \Pi}{\partial u^2} \nabla u^T \mathbf{D}_d \nabla u \right\rangle = -2\left\langle \frac{\Pi}{(u-U)^2} \nabla u^T \mathbf{D}_d \nabla u \right\rangle. \tag{20}$$

3 Concentration PDFs of Reactive Species

3.1 Problem Statement

We consider a reaction of pure dissolution/precipitation at equilibrium where an immobile solid mineral S_{3s} dissolves reversibly to yield ions B_1 and B_2

$$B_1 + B_2 \leftrightarrow S_{3s}. \tag{21}$$

We further assume that the mineral, S_{3s}, is a pure phase, so that its activity equals 1. We define $m_1 = \phi c_1$ and $m_2 = \phi c_2$ as the mass per unit volume of medium of the aqueous species B_1 and B_2, respectively (ϕ and c_i, respectively, being porosity, which we assume as a deterministic constant, and concentration of aqueous species B_i, $i = 1, 2$), and m_3 as the mass per unit volume of medium of the solid mineral, S_{3s}.

The mass action law for the considered system (assumed diluted) is expressed as

$$\log c_1 + \log c_2 = \log K_{eq}. \tag{22}$$

The equilibrium constant, K_{eq}, is strictly related to solubility of the solid phase, S_{3s}, and usually depends on temperature, pressure and ionic strength of the solution. The mass balance equations for the three species are

$$\frac{\partial c_i}{\partial t} + \mathbf{v} \nabla c_i - \nabla \cdot (\mathbf{D}_d \nabla c_i) = -r, \quad i = 1, 2, \tag{23}$$

$$\frac{\partial (m_3)}{\partial t} = r, \tag{24}$$

where r is the reaction rate. The system can be fully defined by means of a single component

$$u = c_1 - c_2 \tag{25}$$

whose governing mass balance equation, obtained by substituting (25) into (23), has the format of an advection–dispersion equation without a reaction source term. Since u is a conservative quantity, it satisfies (1). Once u is solved for using (1), from (22) and (25), and assuming that K_{eq} is independent of c_1 and c_2, we can write explicit expressions for c_i

$$c_1 = \frac{1}{2}\left(u + \sqrt{u^2 + 4K_{eq}}\right) = f_1(u), \qquad c_2 = \frac{1}{2}\left(-u + \sqrt{u^2 + 4K_{eq}}\right) = f_2(u). \tag{26}$$

The work can then be extended to multicomponent reactive transport. In such a case, it is possible to write the system of equations in terms of n conservative components. Linear combinations of aqueous species, u_1 to u_n, and then species concentrations c_i ($i = 1, \ldots, m$) are written as implicit expressions of the type

$$f_i(c_i, u_1, u_2, \ldots u_n) = 0, \quad i = 1, \ldots, m. \tag{27}$$

3.2 Calculation of the PDF and Statistical Moments of Reactive Solutes

Whenever p_u is fully known, either analytically or numerically, it is possible to write the low order power moments of the concentrations of the reactive species. From (26), the central moments of species concentrations can be written as

$$\langle c_i^n \rangle = \int f_i^n(u) p_u \, du, \quad i = 1, 2. \tag{28}$$

Alternatively, it is also possible to derive the complete pdf of c_i ($i = 1, 2$), p_{ci}, by taking advantage of the rule of a univariate change-of-variable transformation for continuous random variables. Knowing that the relationship between u and c_i is monotone (and therefore invertible),

$$p_{ci} = \left| \frac{\partial u}{\partial c_i} \right| p_u. \tag{29}$$

Since

$$u = c_1 - \frac{K_e}{c_1} = \frac{K_e}{c_2} - c_2, \tag{30}$$

it follows that

$$\frac{\partial u}{\partial c_i} = \pm \left(1 + \frac{K_e}{c_i^2} \right). \tag{31}$$

This implies that $|\frac{\partial u}{\partial c_1}| = |\frac{\partial u}{\partial c_2}|$, and thus the pdf of the reactive concentrations can be formally written (recall (29)) as

$$p_{c_i} = \left(1 + \frac{K_e}{c_i^2} \right) p_u \left(\pm \left(c_i - \frac{K_e}{c_i} \right) \right). \tag{32}$$

3.3 The Governing Equation for the PDF of the Species Concentrations

The pdf p_{c_i} can be written explicitly in terms of u, directly from (32). Multiplying (16) by $(1 + \frac{K_e}{f_i^2(U)})$ leads to

$$\frac{\partial p_{ci}}{\partial t} + \langle \mathbf{v}(\mathbf{x}) \rangle \nabla p_{ci} - \nabla \cdot [\mathbf{D}_d \nabla p_{ci}]$$
$$= \left(1 + \frac{K_e}{f_i^2(U)} \right) \left(\int \delta(u - U) \frac{\partial}{\partial u} \left[f_D(u) p_u \right] du \right.$$
$$\left. - \int \delta(u - U) \frac{\partial^2}{\partial u^2} \left[f_M(u) p_u \right] du \right). \tag{33}$$

Here,

$$f_D(u) p_u = \mathbf{v}' \nabla u p_u = \pm \mathbf{v}' \nabla c_i W(c_i) p_u = \pm f_D(c_i) p_{ci}, \tag{34a}$$

$$f_M(u) p_u = \nabla^t u D \nabla u p_u = W^2(c_i) \nabla^t c_i D \nabla c_i p_u = W(c_i) f_M(c_i) p_{ci}, \tag{34b}$$

$$W(c_i) = 1 + \frac{K_e}{c_i^2}. \tag{34c}$$

Finally, the PDE satisfied by the pdf of the concentrations of the reactive species becomes

$$\frac{\partial p_{ci}}{\partial t} + \langle \mathbf{v}(\mathbf{x}) \rangle \nabla p_{ci} - \nabla \cdot [\mathbf{D}_d \nabla p_{ci}]$$

$$= W(C_i) \left(\int \delta(c_i - C_i) \frac{\partial}{\partial u} \left[f_D(c_i) p_{ci} \right] dc_i \right.$$

$$\left. - \int \delta(c_i - C_i) \frac{1}{W(c_i)} \frac{\partial^2}{\partial c_i^2} \left[W(c_i) f_M(c_i) p_{c_i} \right] dc_i \right). \tag{35}$$

4 Direct Calculation of the (Ensemble) Moments and PDF of the Reaction Rates

When K_{eq} is a constant, the reaction rate can be expressed as (De Simoni et al. 2005)

$$\frac{r}{\phi} = \frac{\partial^2 c_2}{\partial u^2} \left(\nabla^T u \mathbf{D}_d \nabla u \right), \tag{36a}$$

$$\frac{\partial^2 c_2}{\partial u^2} = \frac{2 K_{\text{eq}}}{(u^2 + 4 K_{\text{eq}})^{3/2}}, \tag{36b}$$

and r can be formally written in terms of u, $r = G(u)$. Therefore, the (ensemble) mean value of r can be written as

$$\langle r \rangle = \int G(u) p_u \, du. \tag{37}$$

Considering that $r^2 = \Psi(u) (= G(u)^2)$, the expression for the variance of r, VAR(r) is then derived as

$$\text{VAR}(r) = \int \Psi(u) p_u \, du - \left[\int G(u) p_u \, du \right]^2. \tag{38}$$

Knowledge that $r = G(u)$ allows employing a univariate change-of-variable transformation of a continuous random variable to completely determine the pdf of r as a function of the pdf of u. However, in this case, reactions rates are not related monotonically to the conservative concentrations and a more general transformation technique is necessary. In general, the $G(u)$ function consists of a finite sequence of monotonic functions, g_i. Each function is defined over an open interval $I_j = (u_i, u_{i+1})$, within which the derivative of g_i is nonzero at each point:

$$G(u) = \sum_{i \in I_j} g_i(u) \big(H(u - u_i) - H(u - u_{i+1}) \big), \tag{39}$$

$H(\cdot)$ being the Heaviside function. The generalized univariate change-of-variable transformation then reads

$$p_r(r) = \sum_{i \in I_j} \left| \frac{dg_i^{-1}(r)}{dr} \right| p_u \big(g_i^{-1}(r) \big). \tag{40}$$

Remarkably, we note from (40) that when r approaches the peak value (where the first derivative vanishes) while p_u is finite, the probability density function of r will in turn exhibit an asymptote.

5 Stratified Media

An expression for the pdf of conservative and reactive species can be obtained explicitly for a perfectly stratified aquifer whenever transverse dispersion can be neglected. This scenario has been previously analyzed by Fernandez-Garcia et al. (2008) and allows for exploring the relevant features associated with the different terms appearing in (16).

In the stochastic framework, a perfectly stratified aquifer can be conceptualized as a realization of a heterogeneous medium whose hydraulic conductivity is characterized by a very large horizontal correlation length. As such, the adopted model can be applied for small travel distances compared to the correlation length characterizing the heterogeneity of the medium. Velocities and concentrations in such an aquifer are governed by the variations of hydraulic conductivity, K, along the coordinate perpendicular to flow (see Fig. 2 for an illustrative sketch). We further note that random K-fields with well connected extreme values and large variances of the natural logarithm of K (large degrees of heterogeneity) behave similarly to a layered medium.

We analyze a particular reactive transport scenario. For example, we consider an aquifer, r, which is initially under conditions of homogeneous geochemical equilib-
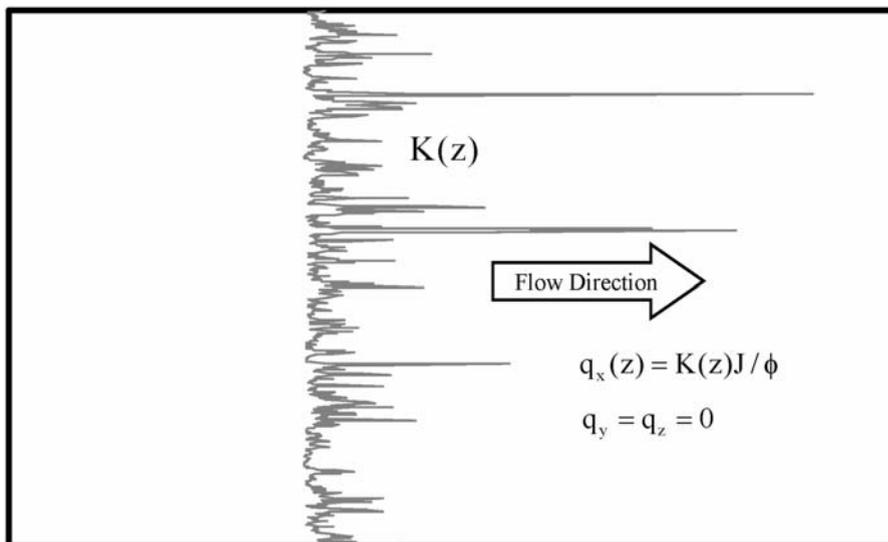

Fig. 2 Representation of the conceptual model of a perfectly stratified medium. Vertical correlation scale is taken to be much smaller than its horizontal counterpart. Hydraulic conductivity, K, and seepage velocity are only functions of the vertical coordinate, z

rium at all points,

$$c_i(x, z, t = 0) = c_{i,0}, \quad i = 1, 2, \tag{41}$$

where the $c_{i,0}$ values satisfy (22). We then inject instantaneously, along a given line perpendicular to the flow direction, water with a different chemical signature, but also in chemical equilibrium with the mineral. Writing the initial conditions directly in terms of u leads to

$$u(x, z, t = 0^+) = u_0 + \Delta u_0 \delta(x), \tag{42}$$

with $u_0 \equiv u(x, z, t = 0) = c_{1,0} - c_{2,0}$, and $\delta(x)$ being the Dirac delta function. The presence of the two end-members produces a chemical disequilibrium through mixing. The input water is advected through the system and, in the absence of transverse diffusion, the only mixing mechanism is that of longitudinal diffusion/dispersion, taking part at the moving boundary (initially sharp) between these two waters. It is precisely the process of mixing that drives mineral precipitation.

5.1 PDF of the Concentration of the Conservative Component

In a perfectly stratified medium, the distribution of the conservative component concentrations are explicitly written as a function of the hydraulic conductivity under the assumption of negligible transverse dispersion (Fernandez-Garcia et al. 2008)

$$u(K) = u_0 + \frac{\Delta u_0}{\sqrt{4\pi D_d t}} \exp\left(-\left(x - \frac{KJ}{\phi}t\right)^2 / 4D_d t\right). \tag{43}$$

The dependence of u on K is then described by a Gaussian-type functional format and is not a monotonic function of K. Yet, one can express $u(K)$ as a sequence of two invertible monotonic functions. When D_d is constant, these functions are

$$K_1(u) = \frac{\phi x}{Jt} + \frac{\phi}{Jt} \sqrt{4D_d t \ln\left(\frac{(u - u_0)\sqrt{4\pi D_d t}}{\Delta u_0}\right)}, \quad K \in \left(0, \frac{\phi x}{Jt}\right), \tag{44}$$

$$K_2(u) = \frac{\phi x}{Jt} - \frac{\phi}{Jt} \sqrt{4D_d t \ln\left(\frac{(u - u_0)\sqrt{4\pi D_d t}}{\Delta u_0}\right)}, \quad K \in \left(\frac{\phi x}{Jt}, +\infty\right). \tag{45}$$

It then follows that p_u can be directly written in terms of the pdf of K, p_K, by simply applying the generalized change-of-variable transformation of a continuous random variable

$$p_u(u) = \left|\frac{dK_1(u)}{du}\right| p_K(K_1(u)) + \left|\frac{dK_2(u)}{du}\right| p_K(K_2(u)). \tag{46}$$

The derivatives involved in (46) can be calculated as

$$\frac{dK_i(u)}{du} = \left(\frac{du}{dK}(K_i(u))\right)^{-1}, \tag{47}$$

$$\frac{du}{dK}(u) = \frac{\Delta u_0}{\sqrt{4\pi D_d t}} \exp\left(-\left(x - \frac{KJ}{\phi}t\right)^2 \Big/ 4D_d t\right) \frac{1}{2D_d \phi}\left(x - \frac{KJ}{\phi}t\right)J. \quad (48)$$

This formulation holds for any given distribution of the hydraulic conductivity K. Assuming that the natural logarithm of K ($Y = \ln K$) can be described by a Normal distribution with mean μ and variance σ_Y^2 yields

$$p_K(K) = \frac{1}{K\sqrt{2\pi\sigma_Y^2}} \exp\left(-\frac{(\ln K - \mu)^2}{2\sigma_Y^2}\right). \quad (49)$$

Figures 3(a) and 3(b) depict, respectively, the spatial distribution of the pdf of U for a given observation time and the temporal dependence of p_u for fixed observation location, both for $\sigma_Y^2 = 1$. All numerical simulations were performed using the following fixed parameters: $\phi = 0.3$, $K_G = 1$ m/day, $J = 0.05$, $D_d = 0.016$ m^2/d, $u_0 = 0$, $\Delta u_0 = 9 \cdot 10^{-4}$ Kg/m^3, $K_{eq} = 10^{-7}$. Figure 3(c) illustrates the dependence of p_u on σ_Y^2 for a given point in space–time. Our results illustrate that for a fixed time, the pdf tends to become U-shaped as the distance from the injection point increases. At a given time, the values of U lie within the range $[U_0, U_{max}]$ when $\Delta u_0 > 0$ (or, conversely, in the range $[U_{min}, U_0]$ for $\Delta u_0 < 0$). The value of U_{max} is independent from distance, because our set-up assumes D_d to be constant (i.e., independent of velocity). Since all values of K can be encountered in the system (albeit associated with different probabilities of occurrence), the pdf function displays a singularity at $U = U_{max}$. On the contrary, a large probability is associated with $U = U_0$. This leads to the asymmetric shape of the pdf when approaching the two limits of the interval of variability. Figure 3(b) further suggests that U_{max} decreases with time, regardless of the distance from the injection location. It is also clear from our results that the pdf can attain a U-shape that closely resembles a Beta distribution with parameters less than unity. The latter result is consistent with findings of Fiorotto and Caroni (2002) and Caroni and Fiorotto (2005), despite the fact that their works consider a different flow and transport configuration.

Figure 3(c) shows the variability of the pdf of U as a function of the variance of Y. The effect of the degree of heterogeneity of the system, as expressed by σ_Y^2, is to spread pdf values towards the extreme (largest and minimum) admissible concentration values, with a small probability of obtaining intermediate values. This renders predictions based only on (ensemble) mean values less significant as σ_Y^2 increases, because the associated coefficient of variance is much larger than one in many situations. We will find later that the pdfs of reactive species display similar patterns.

5.2 Direct Evaluation of the Dilution and Spreading Terms

Evaluation of the dilution and spreading terms (right-hand side of (16)) requires knowing the dependence of f_M and f_D on u. Taking the second partial derivative of (43) with respect to x, and expressing the resulting expression in terms of u leads to

$$f_M(u; x, t) = -\frac{(u - u_0)^2}{t} \ln\left(\frac{u - u_0}{\Delta u_0}\sqrt{4\pi D_d t}\right). \quad (50)$$

Fig. 3 Examples of the dependence of the probability density function for the concentration of a conservative species on: (**a**) space (for a given time and $\sigma_Y^2 = 1$); (**b**) time (for a given location and $\sigma_Y^2 = 1$); and (**c**) σ_Y^2 (for a fixed space–time location, $t = 7.3$ days and $x = 2$ m)
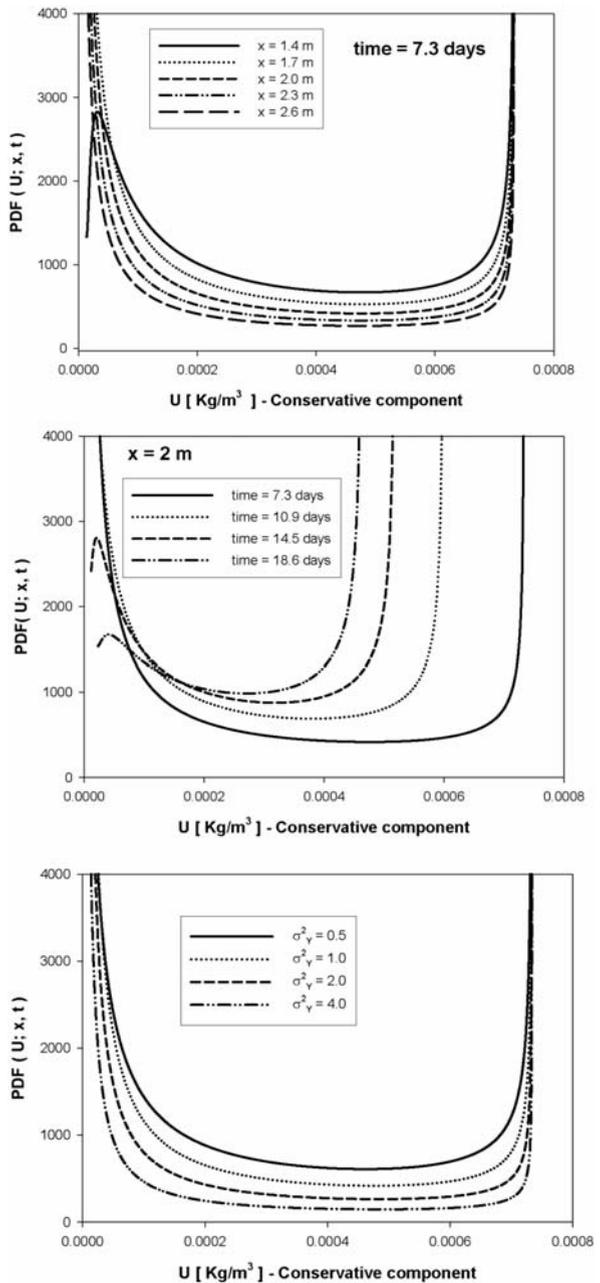
Since p_u is actually fully known explicitly from (44)–(49), one can analytically evaluate the product $(f_M p_u)$ and its derivatives. It can be shown that the second partial derivative of $(f_M p_u)$ with respect to u is invertible. This, in turn, implies that there is only one value at which $u = U$ for the function inside the integral involved in the

dilution term in (16). This dilution term is then written as

$$-2\int \delta(u-U)\frac{\partial^2}{\partial u^2}[f_M p_u]\,du = -2\frac{\partial^2}{\partial u^2}[f_M p_u]\bigg|_{u=U}. \tag{51}$$

The evaluation of the spreading term is performed by means of a similar analysis. In this case, f_D is not invertible but needs to be separated into two monotonic functions. For example, each one of these provides a distinct value of the integral involved in the spreading term in (16),

$$\int \delta(u-U)\frac{\partial}{\partial u}[f_D p_u]\,du = \sum_{i=1}^{2}\frac{\partial}{\partial u}[f_D p_u]\bigg|_{u_i=U}. \tag{52}$$

Knowing that $f_D = v'\frac{\partial u}{\partial x}$, where

$$v'(u;x,t) = \left(x+\sqrt{-4D_d t\ln\left(\frac{u-u_0}{\Delta u_0}\sqrt{4\pi D_d t}\right)}\right)\cdot\frac{1}{t} - \frac{\langle K\rangle J}{\phi}, \quad K \in \left(0,\frac{\phi x}{Jt}\right), \tag{53}$$

$$v'(u;x,t) = \left(x-\sqrt{-4D_d t\ln\left(\frac{u-u_0}{\Delta u_0}\sqrt{4\pi D_d t}\right)}\right)\cdot\frac{1}{t} - \frac{\langle K\rangle J}{\phi}, \quad K \in \left(\frac{\phi x}{Jt},\infty\right), \tag{54}$$

one obtains

$$\sum_{i=1}^{2}\frac{\partial}{\partial u}[f_D p_u]\bigg|_{u_i=U} = -4\frac{\partial}{\partial u}\left[\frac{f_M p_u}{u-u_0}\right]. \tag{55}$$

When dealing with reactive species, the dilution and spreading terms are simply obtained multiplying (51) and (55) by $W(f_i(U))$.

Figure 4(a) illustrates the relative importance of the spreading and dilution terms in (16) as a function of travel distance from the source (for a fixed time) and for a mildly heterogeneous medium ($\sigma_Y^2 = 0.1$). The corresponding depiction for $\sigma_Y^2 = 1.0$ is presented in Fig. 4(b). For both variances, it is found that the spreading term displays two negative peaks, with an intermediately located positive peak. As variance increases, this effect is more visible at short distances and can become difficult to appreciate in the plots. A negative sign of this term is representative of some sort of "macrodispersive effect", in the sense that the spread of the pdf of u (for a given location in space–time) tends to be visibly enhanced. The effect of the dilution term is somehow in the opposite sense, in that it somehow tends to compensate the influence of the spreading term. The net effect of the sum of the two terms can be either positive or negative in space (or time). The sign of the global sink/source term in (16) displays a clear transition in space when $\sigma_Y^2 = 1$ (Fig. 4(b)). For very small distances from the injection location, the spreading and dilution terms are both negative, indicating the tendency to increase the spreading in p_u. Both the spreading and dilution terms are positive for large distances, indicating the tendency to dampen the dispersion of p_u over different values. The immediate consequence is that p_u can attain
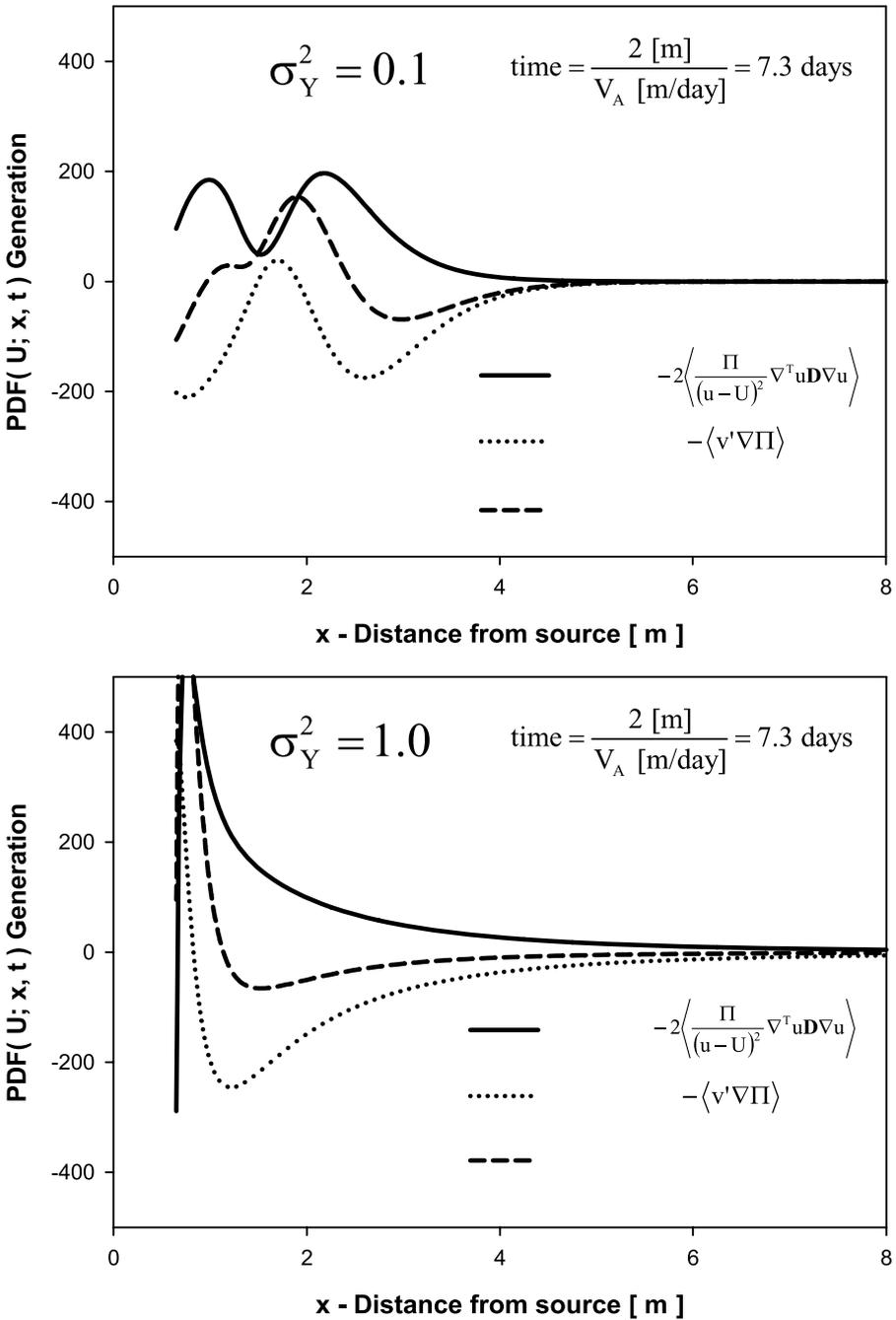
Fig. 4 Contributions of mixing and spreading processes (dilution and spreading terms) to the probability density function of U, $p_u(U; x, t)$, as a function of the degree of heterogeneity of the system: (**a**) $\sigma_Y^2 = 0.1$; (**b**) $\sigma_Y^2 = 1.0$

a nearly uniform shape in most of its domain of existence in the u-space, thus resembling a Beta-type distribution. For intermediate distances, the net generation of the pdf becomes negative, indicating enhanced global dispersion. It follows that the shape of p_u is subject to rapid changes in space and/or time. For very large distances, the net source term vanishes, indicating that p_u is essentially driven by an ADE-type equation. This, in turn, implies that the pdf of u tends to become Gaussian.

Figure 5 presents the net contribution of the spreading and dilution terms for $\sigma_Y^2 = 1$ and selected observation times and values of U. For a given value of U (i.e., $U = U_{\max}/2$), the transition of the sink/source term from positive values at short distances to negative values (not reached for the largest times in the plots) is noted for some values of the time of observation, with eventually leveling out toward a zero value (Fig. 5(b)). Figure 5(a) focuses on the effect of the spreading and dilution terms as a function of U. For a fixed location and time the effect of the spreading and dilution terms is very different depending on the particular value of U selected in the space of definition of u. First, it should be noted that, since $p_u(u = U_{\max}/2) < p_u(u = 0.1U_{\max})$, and also $p_u(u = U_{\max}/2) < p_u(u = 0.9U_{\max})$ (see Fig. 3), the contribution of the source term is also less significant for $u = U_{\max}/2$ than for $u = 0.1U_{\max}$ or $u = 0.9U_{\max}$. Our results indicate that the shape of the curves and the location of the associated largest values are very different for values of u in the proximity of the extremes of the admissible range (i.e., $u = 0.1U_{\max}$ and $u = 0.9U_{\max}$). The location of the maximum and minimum pdf generation term is clearly related to the choice of the target U value. Two important features should be emphasized: (i) the characteristic time and distance, at which the pdf of U approaches a Gaussian-like behavior, appears to be independent of the actual target value of U; and (ii) the largest U values are linked to the overall largest net production term for the pdf, dampening the dissipation of the maximum concentration values within the system.

5.3 PDF of the Concentration of the Reactive Species

Once the pdf of the conservative component concentration is calculated (see Sect. 5.1), the pdf of the concentrations of the associated reactive species can be directly estimated using (32). Figure 6(a) depicts the probability density function of u, c_1 and c_2 for a given space–time location and $\sigma_Y^2 = 1$. Overall, the pdf of c_1 and c_2 inherits the same main features previously observed for the pdf of u, i.e., two asymptotes at the extreme values (Fig. 6(b)). Furthermore, the pdf of the reactive species attains values which are consistently larger than those of the pdf of u because $p_{ci} = (1 + K_e/C_i^2)p_u$. As a consequence, small concentrations of reactive species are associated with very large values of their corresponding pdf (Fig. 6(a)). In the particular set-up of the transport problem used, c_2 decreases as chemical reactions take place and thereby the corresponding pdf describing the distribution of its values increases substantially. On the contrary, the distribution of c_2 is similar to that of u and therefore its corresponding pdf has been only slightly amplified. For $u = 0$, c_1 is minimum while c_2 is maximum, as $c_1 = c_2 = \sqrt{K_e}$ (Fig. 6(b)). This causes the pdf of c_1 and c_2 to be characterized by the same asymptote (Fig. 6(a)).

Figure 7(a) illustrates the relative importance of the spreading and dilution source terms appearing in the integro-differential equation governing p_{c1} (right-hand side
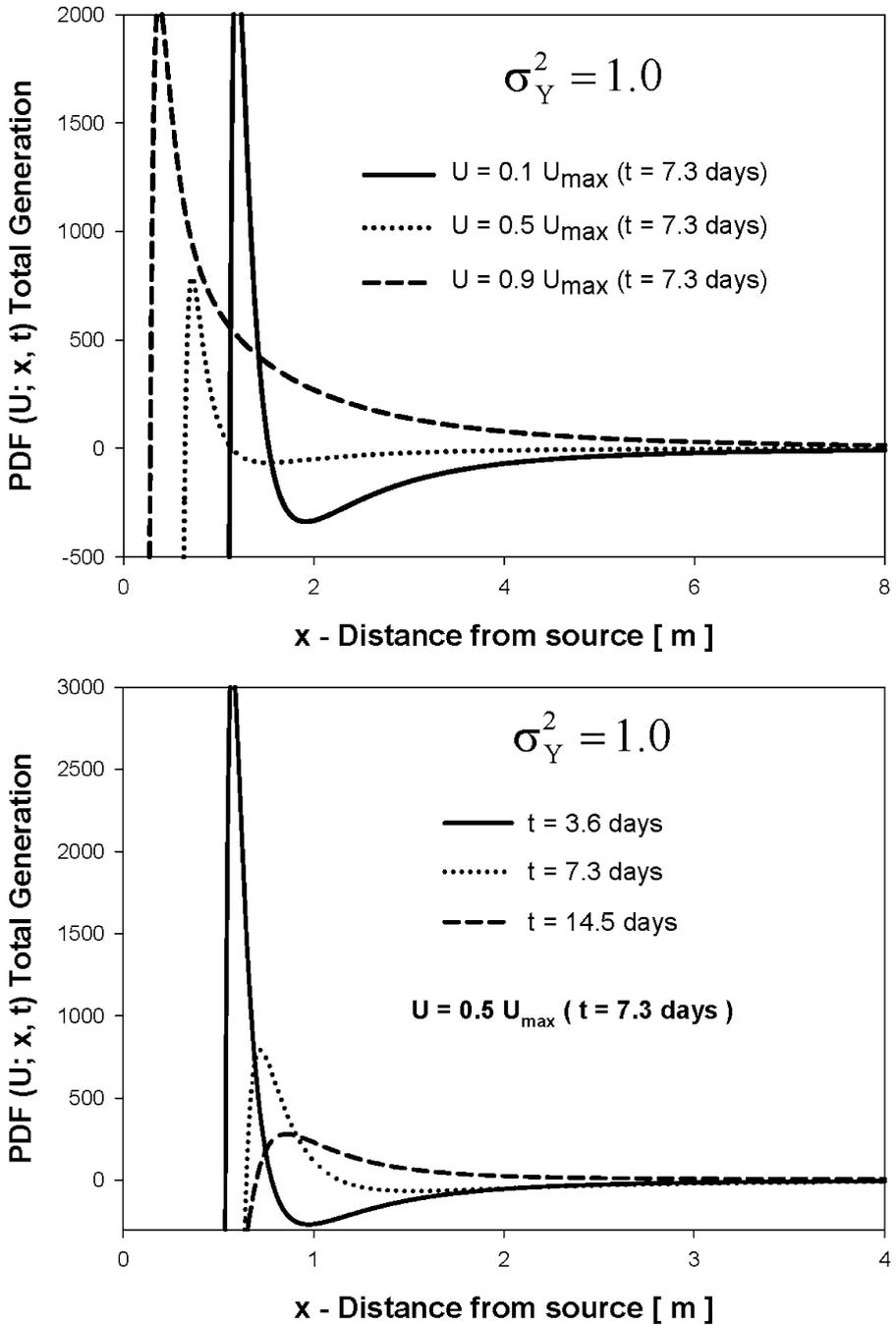
Fig. 5 Total generation of the probability density function of U, $p_u(U; x, t)$ as a function of distance, x, from the injection for a fixed variance of the natural log of hydraulic conductivity $\sigma_Y^2 = 1.0$: (**a**) dependence on given values of U at a given time; (**b**) dependence on observation times at a given space location. The total generation term is the sum of the p_u source terms due to spreading and dilution appearing in the right-hand side of (16)
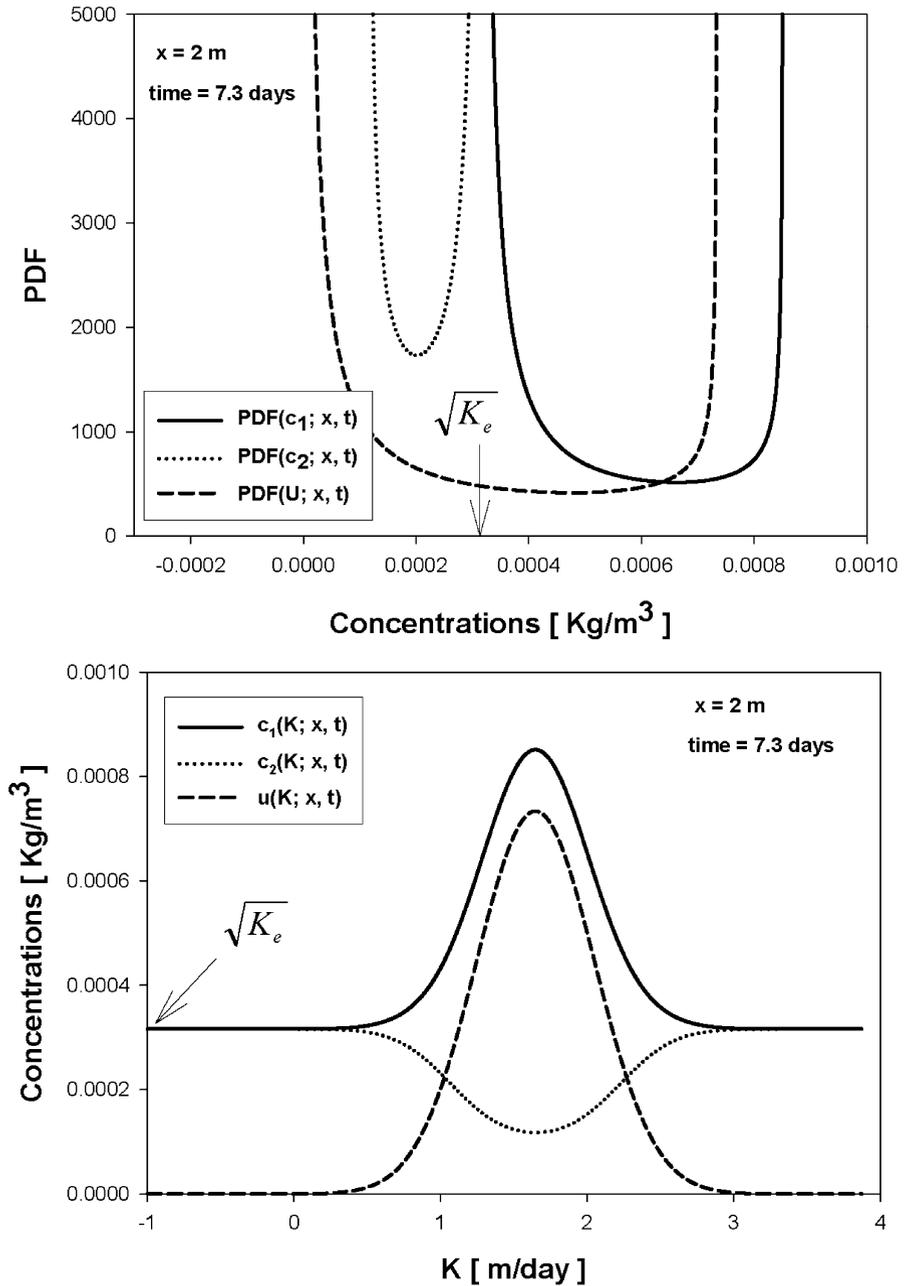
Fig. 6 (**a**) Probability density functions of the conservative component and reactive species for a given space–time location and $\sigma_Y^2 = 1.0$; (**b**) Corresponding distributions of conservative and reactive concentrations as a function of hydraulic conductivity, K
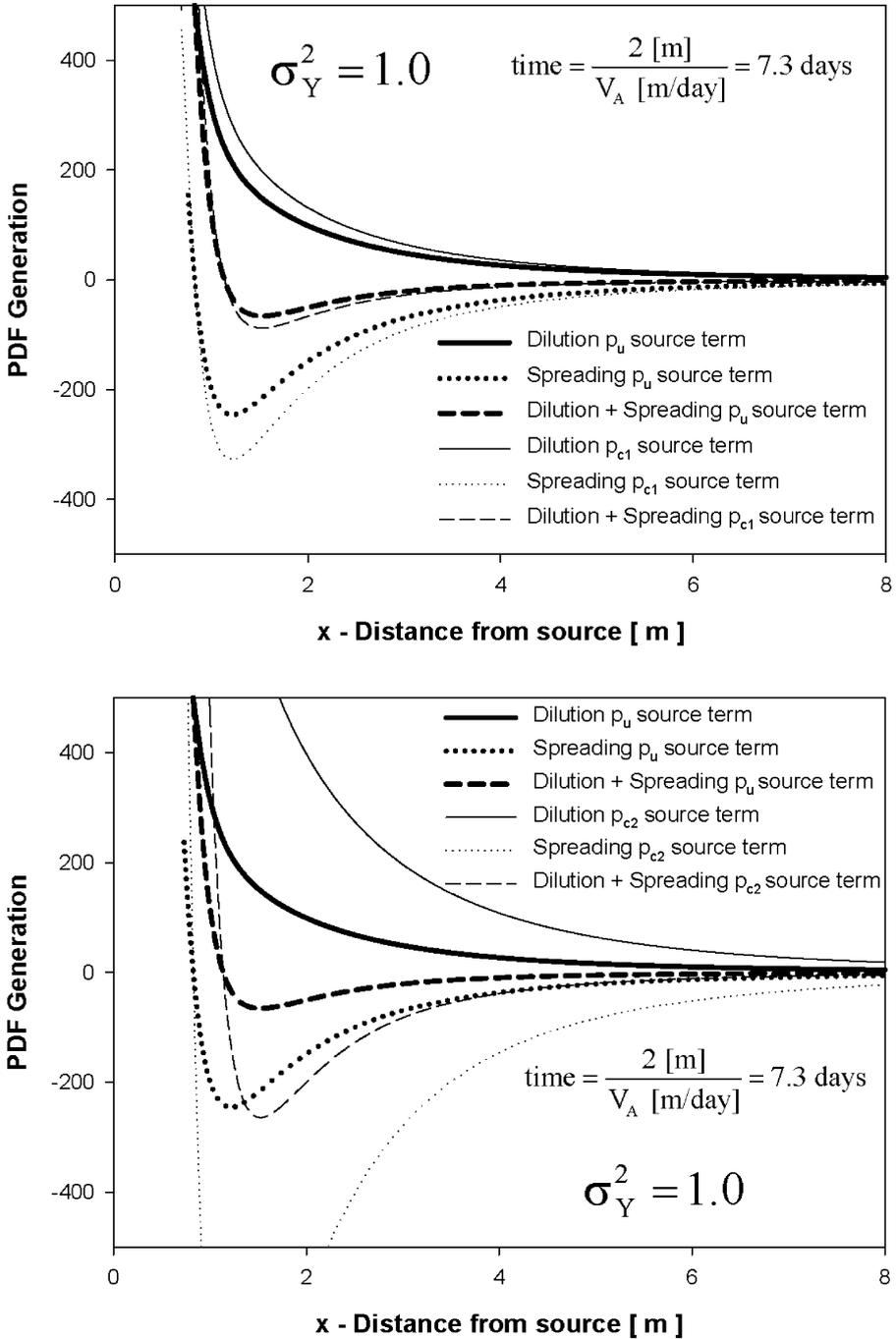
$$\sigma_Y^2 = 1.0 \qquad time = \frac{2\ [m]}{V_A\ [m/day]} = 7.3\ days$$

Legend (a):
— Dilution p_u source term
‥‥‥ Spreading p_u source term
– – – Dilution + Spreading p_u source term
— Dilution p_{c1} source term
······ Spreading p_{c1} source term
– – – Dilution + Spreading p_{c1} source term

x - Distance from source [m]

Legend (b):
— Dilution p_u source term
‥‥‥ Spreading p_u source term
– – – Dilution + Spreading p_u source term
— Dilution p_{c2} source term
······ Spreading p_{c2} source term
– – – Dilution + Spreading p_{c2} source term

$$time = \frac{2\ [m]}{V_A\ [m/day]} = 7.3\ days$$

$$\sigma_Y^2 = 1.0$$

x - Distance from source [m]

Fig. 7 Comparison of the effects of spreading and dilution on the generation terms associated with (**a**) p_{c1} and (**b**) p_{c2} (for a given $\sigma_Y^2 = 1.0$) and p_u

of (35)) contrasting them against the corresponding terms appearing in (16) (thus, governing the evolution of p_u). A companion depiction for the terms associated with p_{c2} is offered in Fig. 7(b). Similarly to what was previously observed for the probability density function of $p_{ci}(i = 1, 2)$, the source terms are amplified by a factor of $(1 + K_e/C_i^2)$, yielding the largest pdf source values for small concentration of the reactive species.

5.4 PDF of Reaction Rates

In the stratified scenario examined, reaction rates can be directly related to hydraulic conductivity as

$$r = G(K) = \phi \frac{\partial^2 c_2}{\partial u^2} D \left(\frac{\partial u}{\partial x} \right)^2$$

$$= \phi \frac{\partial^2 c_2}{\partial u^2} \frac{\Delta u_0^2}{16\pi D^2 t^3} \exp \left(- \left(x - \frac{KJ}{\phi} t \right)^2 \Big/ 2Dt \right) \left(x - \frac{KJ}{\phi} t \right)^2. \quad (56)$$

From (56), it follows that the reaction rate displays a dependence on K which is characterized by a bimodal functional format. It can then be expressed as a sequence of four monotonic invertible functions, denoted in the following as $g_i(r)(i = 1, 2, 3, 4)$, depicted in Fig. 8(a) for ease of illustration. This leads to

$$p_r(r) = \sum_{i=1}^{4} \left| \frac{dg_i^{-1}(r)}{dr} \right| p_K \left(g_i^{-1}(r) \right). \quad (57)$$

The derivatives involved in (47) can be computed as

$$\frac{dg_i^{-1}(r)}{dr} = \left(\frac{dG(g_i^{-1}(r))}{dK} \right)^{-1}, \quad (58)$$

where

$$\frac{dG(K)}{dK} = \phi \frac{d}{dK} \left(\frac{\partial^2 c_1}{\partial u^2} \right) D \left(\frac{\partial u}{\partial x} \right)^2 + \phi \left(\frac{\partial^2 c_1}{\partial u^2} \right) D \frac{d}{dK} \left(\frac{\partial u}{\partial x} \right)^2. \quad (59)$$

The derivatives involved in (59) can be analytically calculated from (43) and (36b) and thereby are explicitly known as a function of (K, x, t). The first term on the right-hand side of (59) includes the derivative (with respect to K) of the speciation term, whereas the last term is the derivative of the dilution term, f_M.

From (57), the pdf of the reaction rates exhibits two asymptotic values when the maximum and minimum reaction rates occur. This is shown in Fig. 8(b), which displays the pdf of the reaction rates at a fixed location ($x = 2$ m) and for different observation times. For all times the shape of the pdf resembles that of a Beta distribution with parameters less than unity, with large values at $r = 0$ and $r = r_{\max}$. The particular $r_{\max}$ value is a monotonically decreasing function of time, as it is controlled by the mixing process that causes the gradients of concentrations to reduce with time (this needs not to be valid for other transport setups).

Fig. 8 (**a**) Illustration of the separation of the reaction rate distribution as a function of K in a sequence of four monotonic invertible functions denoted as $g_i(r)$, in the case of a perfectly stratified medium in the absence of transverse dispersion; (**b**) Probability density function of the reaction rate as a function of r for different times at a given location; (**c**) The effect of neglecting the spatial variability of the speciation term in the estimate of the probability density function of the reaction rate
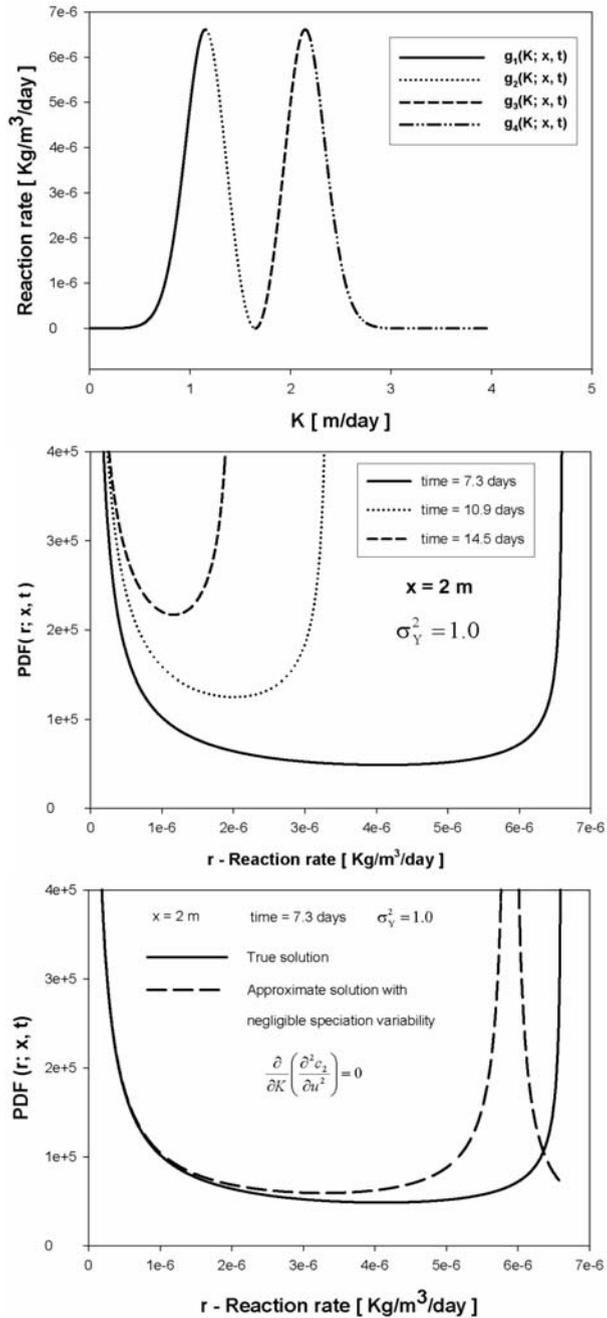
Figure 8(c) shows the effect of neglecting the first of the two terms in (59), as assumed by Luo et al. (2008) in their analysis. A key conclusion is that the pdf of the reaction rate obtained upon disregarding variability of the speciation term is not rep-

resented correctly for large r values (which are usually the most interesting values). The most relevant discrepancy is noted in the value of r displaying the asymptote, which is wrongly assessed if the heterogeneity of the speciation term is wiped out from the analysis. On the contrary, disregarding heterogeneity of the speciation term does not introduce significant errors in the estimation of the pdf of the reaction rate for the lowest r values. This is due to the location associated with zero reaction rate not being influenced by the speciation term, as the derivative of r with respect to K is always zero, regardless the variability of the speciation term.

6 Conclusions

Solute transport in heterogeneous media is conveniently studied in a stochastic framework. Species concentrations at any location and time are prone to uncertainty, and can be fully defined only in terms of probability density functions. As we reviewed in our Introduction, most of the existing literature concentrates on finding either the (ensemble, statistical) moments of concentrations or describing the evolution of a contaminant plume in terms of its spatial or temporal moments. These approaches have lead to a very large number of works that have contributed to increasing our understanding of the effects of underlying processes upon predictors of a contamination state and their associated uncertainty.

In line with recent developments in the hydrologic literature, we start from a different perspective and aim at providing a direct evaluation of the probability density function (pdf) of reaction rates and concentrations of reactive species. The pdf and the associated cumulative distribution function (cdf) are needed in risk analysis and vulnerability practices, where the probability of exceeding regulatory limits and the degree of geochemical activity of the system should be assessed. As such, our work opens novel and interesting perspectives in stochastic modeling and risk analysis.

We analyze a bi-molecular system where two aqueous solutes are in chemical equilibrium with an immobile solid mineral and undergo instantaneous precipitation. Transport is modeled by means of an Advection Dispersion Equation (ADE). We deconstruct the reactive transport problem into (a) the solution of a conservative transport problem, written in terms of nonreactive components, and (b) the evaluation of a speciation term. In this context, the rate of the reaction is only a function of quantities, such as the conservative components, the equilibrium constants, and the dispersion coefficients, and can be calculated without the need to actually evaluate the concentration of the dissolved (reactive) species (De Simoni et al. 2005, 2007).

Our work leads to the following major findings:

1. We provide a partial differential equation governing the spatial and temporal variation of the pdf, p_u, of the conservative component describing the evolution of the transport process in a generally heterogeneous system. This equation is an ADE-type with two sink/source terms. Both terms are integrals in the probability space of the conservative component. The first one is related to the variations of velocity with space (which are unknown). Thus, we call it the "spreading term". The second term is associated with spatial gradients of the conservative component itself. The resulting PDE equation for the pdf of conservative species in a

general heterogeneous medium can be fully solved only using a Monte Carlo approach. Actually computing the full pdf of concentrations is a direct outcome from Monte Carlo methods, and thus it does not introduce a new degree of complexity.

2. Once the pdf of the conservative components is obtained, it is possible to directly map it into reactive species and reaction rates, without the need to perform all the computations of reactive transport within a Monte Carlo framework. This is in line with the main finding of Cirpka et al. (2008).

3. Alternatively, it is also possible to write the governing partial differential equation for the pdf of the aqueous species concentrations.

4. We provide an explicit expression for the pdf of conservative and reactive species and associated reaction rates for transport within a perfectly stratified aquifer, whenever the effects of transverse dispersion can be neglected and under the assumption that the vertical distribution of hydraulic conductivity is Gaussian. This particular set-up has been studied by Fernandez-Garcia et al. (2008), who provided exact expressions for the low order moments of the reaction rate.

5. For a (statistically) stratified aquifer, our results indicate that in some cases p_u can display a U-shape, resembling that of a Beta distribution with parameters less than unity. However, this is not always the case because: (a) at a given time p_u tends to become bimodal at locations close to the injection; and (b) at a given location, p_u tends to be come bimodal for large times.

6. The effect of the degree of heterogeneity of the system, as expressed by the log-hydraulic conductivity variance, σ_Y^2, is to spread the pdf values towards the extreme (largest and minimum) admissible concentration values with a small probability of obtaining intermediate values. This renders predictions based only on (ensemble) mean values less significant as σ_Y^2 increases, because the associated coefficient of variance is much larger than one in many situations. The pdfs of reactive species display similar patterns.

7. The key ingredients in the pdf of the reaction rate are the derivatives of the speciation term and of the dilution term. The analysis of the relative influence of these terms clearly shows that neglecting the spatial variability of the former, as in the work of Luo et al. (2008) can produce severe modifications in the actual shape of the pdf, leading to underestimate the importance of the largest admissible rate values and to overestimate the relative weight of intermediate values of the rate.

References

Andricevic R, Cvetkovic V (1998) Relative dispersion for solute flux in aquifers. J Fluid Mech 361:145–174

Bellin A, Tonina D (2007) Probability density function of non-reactive solute concentration in heterogeneous porous formations. J Contam Hydrol 94(1–2):109–125

Bellin A, Rubin Y, Rinaldo A (1994) Eulerian–Lagrangian approach for modeling of flow and transport in heterogeneous geological formations. Water Resour Res 30(11):2913–2924

Botter GE, Daly A, Porporato I, Rodriguez-Iturbe I, Rinaldo A (2008) Probabilistic dynamics of soil nitrate: coupling of ecohydrological and biogeochemical processes. Water Resour Res 44:W03416. doi:10.129/2007WR006108

Caroni E, Fiorotto V (2005) Analysis of concentration as sampled in natural aquifers. Transp Porous Media 59(1):19–45

Chen HD, Chen SY, Kraichnan RH (1989) Probability distribution of a stochastically advected scalar field. Phys Rev Lett 63(24):2657–2660

Cirpka OA, Valocchi AJ (2007) Two-dimensional concentration distribution for mixing-controlled bioreactive transport in steady state. Adv Water Resour 30(6–7):1668–1679

Cirpka OA, Olsson Å, Ju Q, Rahman MA, Grathwohl P (2006) Determination of transverse dispersion coefficients from reactive plume lengths. Ground Water 44(2):212–221

Cirpka OA, Schwede RL, Luo J, Dentz M (2008) Concentration statistics for mixing-controlled reactive transport in random heterogeneous media. J Contam Hydrol 98(1–2):61–74

Clement TP, Sun Y, Hooker BS, Petersen JN (1998) Modeling multispecies reactive transport in ground water. Ground Water Monit Rem 18(2):79–92

Cvetkovic VD, Shapiro AM, Dagan G (1992) A solute flux approach in transport in heterogeneous formations: 2. Uncertainty analysis. Water Resour Res 28(5):1377–1388

Dagan G (1989) Flow and transport in porous formations. Springer, New York

Dagan G (1994) Upscaling of dispersion coefficients in transport through heterogeneous porous formations. In: Computational methods in water resources X. Kluwer Academic, Norwell

Dentz M, Kinzelbach H, Attinger S, Kinzelbach W (2000) Temporal behavior of a solute cloud in a heterogeneous porous medium 2. Spatially extended injection. Water Resour Res 36(12):3605–3614

De Simoni M, Carrera J, Sanchez-Vila X, Guadagnini A (2005) A procedure for the solution of multicomponent reactive transport problems. Water Resour Res 41:W11410. doi:10.1029/2005WR004056

De Simoni M, Sanchez-Vila X, Carrera J, Saaltink MW (2007) A mixing ratios-based formulation for multicomponent reactive transport. Water Resour Res 43:W07419. doi:10.1029/2006WR005256

Fernandez-Garcia D, Sanchez-Vila X, Guadagnini A (2008) Reaction rates and effective parameters in stratified aquifers. Adv Water Res 31(10):1364–1376

Fiori A (2001) The relative dispersion and mixing of passive solutes in transport in geologic media. Transp Porous Media 42(1–2):69–83

Fiori A, Dagan G (2000) Concentration fluctuations in aquifer transport: a rigorous first-order solution and applications. J Contam Hydrol 45(1–2):139–163

Fiorotto V, Caroni E (2002) Solute concentration statistics in heterogeneous aquifers for finite Peclet values. Transp Porous Media 48(3):331–351

Friedly JC, Rubin J (1992) Solute transport with multiple equilibrium-controlled or kinetically controlled chemical reactions. Water Resour Res 28(6):1935–1953

Gelhar LW, Axness CL (1983) Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resour Res 19(1):161–180

Girimaji SS (1991) Assumed beta-pdf model for turbulent mixing—validation and extension to multiple scalar mixing. Combust Sci Technol 78(4–6):177–196

Guadagnini A, Sanchez-Vila X, Riva M, De Simoni M (2003) Mean travel time of conservative solutes in randomly heterogeneous unbounded domains under mean uniform flow. Water Resour Res 39(3):1050. doi:10.1029/2002WR001811

Guadagnini A, Sanchez-Vila X, Saaltink MW, Bussini M, Berkowitz B (2008) Application of a mixing-ratios based formulation to model mixing-driven dissolution laboratory experiments. Adv Water Resour. doi:10.1016/j.advwaters.2008.07.005

Ham PAS, Schotting RJ, Prommer H, Davis GB (2004) Effects of hydrodynamic dispersion on plume lengths for instantaneous bimolecular reactions. Adv Water Resour 27(8):803–813

Hess KM, Wolf SH, Celia MA (1992) Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 3. Hydraulic conductivity variability and calculated macrodispersivities. Water Resour Res 28(8):2011–2017

Hu BX, Deng F-W, Cushman JH (1995) Nonlocal reactive transport with physical and chemical heterogeneity: linear nonequilibrium sorption with random Kd. Water Resour Res 31(9):2239–2252

Kapoor V, Gelhar LW (1994) Transport in three-dimensionally heterogeneous aquifers. 1. Dynamics of concentration fluctuations. Water Resour Res 30(6):1775–1788

Kapoor V, Kitanidis PK (1998) Concentration fluctuations and dilution in aquifers. Water Resour Res 34(5):1181–1193

Kitanidis PK (1994) The concept of the dilution index. Water Resour Res 30(7):2011–2026

Lawrence AE, Sanchez-Vila X, Rubin Y (2002) Conditional moments of the breakthrough curves of kinetically sorbing solute in heterogeneous porous media using multirate mass transfer models for sorption and desorption. Water Resour Res 38(11), Article Number 1248

Lichtner PC (1996) Continuous formulation of multicomponent-multiphase reactive transport. In: Lichtner PC, Steefel CI, Oeklers EH (eds) Reactive transport in porous media. Reviews in mineralogy, vol 34. Miner Soc Amer, Washington, pp 1–81

Lichtner PC, Tartakovsky DM (2003) Stochastic analysis of effective rate constant for heterogeneous reactions. Stoch Env Res Risk Assess 17(6):419–429

Liedl R, Valocchi AJ, Dietrich P, Grathwohl P (2005) Finiteness of steady state plumes. Water Resour Res 31(12):W12501. doi:10.1029/2005WR004000

Luo J, Dentz M, Carrera J, Kitanidis PK (2008) Effective reaction parameters for mixing controlled reactions in heterogeneous media. Water Resour Res 44(2):W02416. doi:10.1029/2006WR005658

Molins S, Carrera J, Ayora C, Saaltink MW (2004) A formulation for decoupling components in reactive transport problems. Water Resour Res 40(10):W10301. doi:10.1029/2003WR002970

Neuman SP, Zhang Y-K (1990) A quasi-linear theory of non-Fickian subsurface dispersion, 1. Theoretical analysis with application to isotropic media. Water Resour Res 26(5):887–902

Parkhurst DL, Appelo CAJ (1999) User's guide to PHREEQC (Version2)—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. US Geological Survey water-resources investigations report 99-4259

Phillips OM (1991) Flow and reactions in permeable rocks. Cambridge Univ Press, New York

Pope SB (1981) Transport equation for the joint probability density function of velocity and scalars in turbulent flow. Phys Fluids 24(4):588–596

Quinodoz HAM, Valocchi AJ (1993) Stochastic analysis of the transport of kinetically sorbing solutes in aquifers with randomly heterogeneous hydraulic conductivity. Water Resour Res 29(9):3227–3240

Reichle R, Kinzelbach W, Kinzelbach H (1998) Effective parameters in heterogeneous and homogeneous transport models with kinetic sorption. Water Resour Res 34(4):583–594

Riva M, Sanchez-Vila X, Guadagnini A, De Simoni M, Willmann M (2006) Travel time and trajectory moments of conservative solutes in two-dimensional convergent flows. J Contam Hydrol 82:23–43

Riva M, Guadagnini A, Fernandez-Garcia D, Sanchez-Vila X, Ptak T (2008a) Relative importance of geostatistical and transport models in describing heavily tailed breakthrough curves at the Lauswiesen site. J Contam Hydrol 101(1–4):1–13

Riva M, Guadagnini A, Sanchez-Vila X (2008b) Effect of sorption heterogeneity on moments of solute residence time in convergent flows. Comput Geosci, submitted

Robinson BA, Viswanathan HS, Valocchi AJ (2000) Efficient numerical techniques for modeling multicomponent ground-water transport based upon simultaneous solution of strongly coupled subsets of chemical components. Adv Water Resour 23(4):307–324

Rubin J (1990) Solute transport with multisegment, equilibrium-controlled reactions: a feed forward simulation method. Water Resour Res 26(9):2029–2055

Saaltink MW, Ayora C, Carrera J (1998) A mathematical formulation for reactive transport that eliminates mineral concentrations. Water Resour Res 34(7):1649–1656

Sanchez-Vila X, Guadagnini A (2005) Travel time and trajectory moments of conservative solutes in three dimensional heterogeneous porous media under mean uniform flow. Adv Water Res 28:429–439

Shvidler M, Karasaki K (2003) Probability density functions for solute transport in random field. Transp Porous Media 50:243–266

Singurindy O, Berkowitz B, Lowell RP (2004) Carbonate dissolution and precipitation in coastal environments: laboratory analysis and theoretical consideration. Water Resour Res 40:W04401. doi:10.1029/2003WR002651

Steefel CI, MacQuarrie KTB (1996) Approaches to modelling reactive transport. In: Reactive transport in porous media. Reviews in mineralogy, vol 34. Miner Soc Amer, Washington, pp 83–129

Tebes-Stevens C, Valocchi AJ, VanBriesen JM, Rittmann BE (1998) Multicomponent transport with coupled geochemical and microbiological reactions: model description and example simulations. J Hydrol 209(1–4):8–26

Vanderborght J (2001) Concentration variance and spatial covariance in second-order stationary heterogeneous conductivity fields. Water Resour Res 37(7):1893–1912

Yeh GT, Tripathi VS (1991) A model for simulating transport of reactive multispecies components: model development and demonstration. Water Resour Res 27(12):3075–3094