

Water Resources Research

RESEARCH ARTICLE

10.1029/2020WR027867

Key Points:

- We develop semianalytical solutions for the low-order moments of solute discharge in groundwater
- We investigate relative importance of source release rate, aquifer recharge, and log-conductivity variance on the moments of solute discharge
- Global sensitivity analysis is employed with surrogate models to characterize predictive uncertainty

Correspondence to:

V. Ciriello,
v.ciriello@unibo.it

Citation:



Ciriello, V., & de Barros, F. P. J. (2020). Characterizing the influence of multiple uncertainties on predictions of contaminant discharge in groundwater within a Lagrangian stochastic formulation. *Water Resources Research*, 56, e2020WR027867. <https://doi.org/10.1029/2020WR027867>

Received 2 MAY 2020

Accepted 23 SEP 2020

Accepted article online 25 SEP 2020

Characterizing the Influence of Multiple Uncertainties on Predictions of Contaminant Discharge in Groundwater Within a Lagrangian Stochastic Formulation

Valentina Ciriello¹  and Felipe P. J. de Barros² 

¹Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali (DICAM), Università di Bologna, Bologna, Italy, ²Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA, USA

Abstract Assessing the risks associated with transport of contaminants in hydrogeological systems requires the characterization of multiple sources of uncertainty. This paper examines the impact of the uncertainty in the source zone mass release rate, aquifer recharge, and the spatial structure of the hydraulic conductivity on transport predictions. Through the use of the Lagrangian framework, we develop semianalytical solutions for the first two moments of the total solute discharge through a control plane while accounting for source zone release conditions and recharge. We employ global sensitivity analysis (GSA) to investigate how the predictive uncertainty of the mass discharge is affected by uncertainty in source zone mass release rate, recharge, and the variance of the log-conductivity field. The semianalytical solutions are employed with the polynomial chaos expansion technique to perform a GSA. Our results reveal the relative influence of each source of uncertainty on the robustness of model predictions, which is critical for site managers to allocate resources and design mitigation strategies.

1. Introduction

Mathematical models are essential to compute the risks posed by the presence of contaminants in the subsurface environment. Yet it is well-established that quantitative predictions of contaminant transport in aquifers are fundamentally uncertain. Hydrogeological model input parameters (such as the hydraulic conductivity) are heterogeneous, and their complete characterization is not feasible due to high costs of data acquisition. Other key sources of uncertainty typically include the contaminant source loading history and aquifer recharge rates. Understanding the impact of each of these factors on the robustness of model predictions and lastly on the endpoint risk metric is important for site managers to better allocate resources (e.g., to address additional data collection) in order to narrow the uncertainty in model outputs, which is critical to design effective mitigation strategies in turn.

It is well-known that the spatial variability of the hydraulic conductivity impacts mixing and dispersion of a solute plume (see Dentz et al., 2011, and references therein). Several works studied the effects of the spatial random structure of the conductivity field in the statistics of the solute concentration (e.g., Ciriello, Di Federico, et al., 2013; Fiori & Dagan, 2000; Kapoor & Kitanidis, 1998; Tonina & Bellin, 2008) and travel times (e.g., Bellin et al., 1994; Gotovac et al., 2009; Sanchez-Vila & Guadagnini, 2005; Shapiro & Cvetkovic, 1988). Works on the full probabilistic distribution of the concentration in the subsurface environment are also reported in the literature (e.g., Bellin & Tonina, 2007; Boso et al., 2014; Boso & Tartakovsky, 2016; de Barros & Fiori, 2014; Dentz & Tartakovsky, 2010; Meyer et al., 2013). The spatial structure of the conductivity field controls first arrival times (Rizzo & de Barros, 2017) and the probability of concentration exceedance (Henri et al., 2015), two quantities which are critical for human health risk estimation and decision making. It also controls the uncertainty of the resilience loss of our subsurface resources in the presence of contaminants (Im et al., 2020).

Although hydrogeological heterogeneity has been recognized as a key control on the transport, there is a need to investigate the combined effect of physical heterogeneity, aquifer recharge, and source zone mass release rates on the probabilistic description of solute transport. For instance, aquifer recharge can impact

plume mobility and remediation strategies (Libera et al., 2019). Recent works (e.g., Libera et al., 2019; O'Connell & Hou, 2015) raised concerns regarding the risks to site remediation due climate change since precipitation regimes are expected to change. Recharge rates impact solute travel times and the macrodispersive behavior of the plume (e.g., Butera & Tanda, 1999; Destouni et al., 2001; Ezzedine & Rubin, 1997; Li & Graham, 1999; MacFarlane et al., 1983; Rubin & Bellin, 1994). The effects of the temporal recharge patterns on the uncertainty on transport have also been subject of investigation. Foussereau et al. (2001) showed that temporally random rainfall produced only slightly larger prediction uncertainty on transport when opposed to the uniform rainfall case. Furthermore, like the hydraulic conductivity, recharge rates are also uncertain. In the context of inverse modeling, Carrera et al. (2005) emphasized the need to account for uncertainty quantification in recharge rates (in its average value as well as its temporal patterns). Rubin and Dagan (1987) provided uncertainty estimates associated with aquifer recharge. Scanlon et al. (2002) discussed how the reliability of recharge rate estimation varies according to the measurement technique utilized.

Most studies investigating the impact of conductivity heterogeneity and recharge on transport in heterogeneous aquifers neglect the effects of mass release rates in the source zone. The works of Andričević and Cvetković (1996) and Selroos (1997a) provided a framework that allows to quantify the uncertainty of the risks associated with nuclear waste disposal while considering a generic source zone mass release function. Kokkinaki et al. (2014) developed an upscaled model that captures the spatial and temporal variability of DNAPL source zones. Still within the context of DNAPL source zone depletion, Falta et al. (2005) highlighted the importance of accounting for temporal dynamics for the mass release at the source zone. Soga et al. (2004) provided a comprehensive discussion on the relationship between mass fluxes at the source zone and aquifer remediation. Through the use of the Lagrangian framework, de Barros (2018) accounted for the joint impact of spatial heterogeneity of the conductivity field and solute mass loading on the probabilistic description of solute mass fluxes. It was shown that the mass release rate at the source zone had a profound impact on both risk uncertainty and the estimation of the safe source-receptor setback distance that complies with regulatory agencies (de Barros, 2018). The lack of knowledge on the release history of contaminants in aquifers has also led to the development of techniques to identify source zone characteristics given its importance on predicting the spread of pollution and improving the efficiency of remediation campaigns (e.g., Alapati & Kabala, 2000; Butera & Tanda, 2003; Michalak & Kitanidis, 2003; Neupauer et al., 2000). The fundamental question we wish to address is, What is the relative contribution of these parameters in the final uncertainty estimate of the solute mass flux at an environmentally sensitive target? Addressing this question can aid contaminant site managers to better allocate their resources to minimize the uncertainty in model predictions and define aquifer clean-up strategies.

In this paper we evaluate the relative impact of each of these factors on the low-order moments of the solute mass discharge at an environmentally sensitive location. We focus in understanding how multiple sources of uncertainty interact in the statistical characterization of the solute discharge at a control plane. To achieve this goal, we expand the solution derived in de Barros (2018) to incorporate the effects of aquifer recharge and its uncertainty. We develop semianalytical solutions for the mean and variance of the solute discharge as a function of the parameters characterizing aquifer heterogeneity, recharge, and solute mass release rates at the source zone. To quantify how the uncertainty in the mean and variance of the solute discharge can be apportioned to the above-mentioned uncertain parameters, we perform a global sensitivity analysis (GSA) (e.g., Cvetkovic et al., 2015; Sobol', 1993; Sudret, 2008; Wainwright et al., 2014). With the goal of alleviating the computational costs associated with GSA, we utilize polynomial chaos expansion (PCE) (Wiener, 1938; Ghanem & Spanos, 1991). The key idea behind PCE is to approximate the response surface through an orthonormal polynomial basis in the parameter space to represent the model output to change in input parameters. PCE has been employed for GSA (e.g., Crestaux et al., 2009) and has been applied to several hydrogeological problems (e.g., Ciriello et al., 2017, 2019; Oladyshkin et al., 2012). Specific applications of PCEs to address groundwater quality are also reported in the literature (Ciriello, Di Federico, et al., 2013; Oladyshkin et al., 2012; Riva et al., 2015).

This paper is structured as follows: section 2 provides the general description of the physical problem. The stochastic model for flow and transport is explained in section 3. Details regarding the methodology employed to perform the GSA is presented in section 4. Computational results are illustrated in section 5. Finally, section 6 provides a summary of our findings.

2. Problem Description

In many hydrogeological applications, decision makers are interested in quantifying the temporal evolution of the concentration signal of a given pollutant at an environmentally sensitive location. Let $c(\mathbf{x}_s, t)$ [MT^{-1}] denote the contaminant concentration at a given location \mathbf{x}_s as a function of time t . As a consequence of our inability to fully characterize the subsurface's intrinsic heterogeneity, model estimates of c are subject to uncertainty. In addition to the physical heterogeneity of the subsurface, other sources of uncertainty exist such as the spatial distribution of source zone, mass release rates, aquifer recharge rates, and biogeochemical processes. Therefore, risk managers are tasked to evaluate the probability that the concentration at an environmental sensitive location will exceed a regulatory threshold value denoted by c^* . This probability will be a function of the parameters characterizing the physics of flow and transport in the subsurface environment and is equal to the complementary cumulative distribution function

$$\mathbb{P}[c(\mathbf{x}_s, t) \geq c^*] = f(\boldsymbol{\theta}_s, \boldsymbol{\theta}_a, \boldsymbol{\theta}_e), \quad (1)$$

where $\boldsymbol{\theta}_s$ contains parameters that characterize the contaminant source zone such as its spatial extension, mass release rates into the aquifer and the injected mass. The vector $\boldsymbol{\theta}_a$ incorporates variables that describes the variability of the hydraulic conductivity (such as geostatistical parameters) as well as porosity and other parameters describing flow and transport physics in the subsurface environment. Finally, the $\boldsymbol{\theta}_e$ contains external parameters such as aquifer recharge and well pumping rates.

In this work, we will focus on the impact of three key parameters in controlling the uncertainty of the exposure concentration at an environmentally sensitive target. These parameters are the level of heterogeneity of the subsurface, aquifer recharge rates, and the contaminant mass release rate from the source zone. The concentration of exposure will be represented by the flux-averaged concentration, that is, $c = Q/Q_w$, where Q [MT^{-1}] is the solute discharge crossing an environmentally sensitive target and Q_w [L^3T^{-1}] is the volumetric water discharge of the aquifer. Given that the solute discharge Q is fundamental to estimate the flux-averaged concentration, we will consider the statistics of Q to be the quantities of interest (QoI) of our study. In particular, we will focus on estimating on the mean and variance of Q . Computing the complementary cumulative distribution in Equation 1 is outside the scope of this paper. However, several works have provided numerical evidence that Equation 1 can be approximated by a lognormal distribution for aquifers displaying low to mild levels of heterogeneity (e.g., Moslehi & de Barros, 2017; Schwede & Cirpka, 2010). Under these conditions, Equation 1 can be parameterized by the mean and variance of Q .

3. Stochastic Representation of Solute Transport in a Spatially Heterogeneous Flow Field

3.1. General Formulation and Background

We start by considering an incompressible and steady-state flow through a spatially heterogeneous aquifer in the absence of boundary effects in the presence of uniform recharge R [LT^{-1}]. The heterogeneity considered in this work originates from the spatially variable hydraulic conductivity $K(\mathbf{x})$ [LT^{-1}] with $\mathbf{x} = (x_1, \dots, x_d)$ denoting the Cartesian coordinate vector and the subscript d indicates the dimensionality of the flow domain. The flow field is assumed to be nonuniform and unidirectional (i.e., uniform-in-the-mean along the x_1 -direction). Nonuniformity stems from the presence of groundwater recharge (i.e., the mean hydraulic gradient is a function of x_1). Due to the spatial heterogeneity of the aquifer, we model the log-conductivity $Y = \ln[K]$ as a statistically stationary multivariate normal random space function (RSF) (see Dagan, 1989; Rubin, 2003, for details) characterized by its mean value $\langle Y \rangle$, variance σ_Y^2 , correlation scale λ_i (with $i = 1, \dots, d$), and spatial covariance function C_Y . In this work the angled brackets, that is, $\langle \cdot \rangle$, correspond to the expected value.

At time $t = 0$, a dissolved and nonreactive solute is injected at a given rate into the aquifer over a source domain Ω . For $t > 0$, the solute plume will be advected downstream toward a control plane (normal to the mean flow direction) located at a distance L from the source zone. Due to spatial variations in the Y field, the solute body will be distorted in a random manner. In this work, we neglect the effects of local-scale dispersion and chemical reactions. Through the use of the Lagrangian framework, the total solute mass discharge, that is, Q [MT^{-1}], crossing the control plane is given by (e.g., Cvetković et al., 1992; Dagan et al., 1992; Fiori et al., 2002; Rubin, 2003)

$$Q(t, \tau | L, R) = \int_0^t \int_{\Omega} \dot{m}(\tilde{t}, \mathbf{r}) \delta(t - \tilde{t} - \tau) d\tilde{t} d\mathbf{r}, \quad (2)$$

where τ is the solute travel time from the source to the control plane located at $x_1 = L$, \dot{m} [ML⁻³T⁻¹] is the mass release strength and δ is Dirac's Delta function. Following de Barros (2018), we re-write the mass release strength function for a point-like injection, for example, $\dot{m}(t, \mathbf{r}) = M_o \delta(\mathbf{r} - \mathbf{r}_o) \phi(t)$ where M_o is the solute mass injected at location \mathbf{r}_o and finally ϕ characterizes the mass release rate [T⁻¹]. The mass release function ϕ employed for our work is a variation of the model reported in the works of (Selroos, 1997a, 1997b):

$$\phi(t) = \kappa_r e^{-\kappa_r t}, \quad (3)$$

with κ_r representing the release rate constant [T⁻¹]. Other functional forms for ϕ can be adopted in our formulation.

3.2. Mean and Variance of the Solute Discharge

Equation 2 represents the total mass discharge for one possible outcome of the spatially variable Y field. In order to account for the randomness of the Y field, we model τ as a random variable characterized by its PDF, namely g_{τ} . The PDF g_{τ} incorporates all the spatial randomness originating from the conductivity field. The expected value and variance of Q , over all possible values of τ , at the control plane is as follows:

$$\langle Q(t, \tau | L, R) \rangle = \int_{\tau} Q(t, \tau | L, R) g_{\tau}(\tau | \mathbf{r}, L, R) d\tau, \quad (4)$$

$$\sigma_Q^2(t | L, R) \equiv \langle Q(t, \tau | L, R)^2 \rangle - \langle Q(t, \tau | L, R) \rangle^2, \quad (5)$$

where the second moment of Q is given by

$$\begin{aligned} \langle Q(t, \tau | L, R)^2 \rangle &= \int_{\tau'} \int_{\tau''} Q(t, \tau' | L, R) Q(t, \tau'' | L, R) \\ &\times g_{\tau\tau}(\tau', \tau'' | \mathbf{r}', \mathbf{r}'', L, R) d\tau' d\tau''. \end{aligned} \quad (6)$$

In the following, we provide semianalytical solutions for the first two moments of Q for an injected point source and the mass release function (3). Details pertaining the derivation of the moments of Q can be found in de Barros (2018).

By making use of Equations 2 and 4, the first and second moment of the solute discharge for a point-like injection are

$$\langle Q(t, \tau | L, R) \rangle = M_o \kappa_r \int_0^t e^{-\kappa_r \tilde{t}} g_{\tau}(t - \tilde{t} | \mathbf{r}_o, L, R) d\tilde{t}, \quad (7)$$

$$\langle Q(t, \tau | L, R)^2 \rangle = M_o^2 \kappa_r^2 \int_0^t e^{-2\kappa_r \tilde{t}} g_{\tau}(t - \tilde{t}' | \mathbf{r}_o, L, R) d\tilde{t}'. \quad (8)$$

The variance of Q at the control plane is evaluated by inserting Equations 7 and 8 into (5)

$$\sigma_Q^2(t | L, R) = M_o^2 \kappa_r^2 \Phi(t), \quad (9)$$

where

$$\Phi(t) = \int_0^t e^{-2\kappa_r \tilde{t}} g_{\tau}(t - \tilde{t} | \mathbf{r}_o, L, R) d\tilde{t} - \left[\int_0^t e^{-\kappa_r \tilde{t}} g_{\tau}(t - \tilde{t} | \mathbf{r}_o, L, R) d\tilde{t} \right]^2. \quad (10)$$

Therefore, in order to estimate the first two moments of Q , we need provide a PDF model for g_{τ} .

3.3. Travel Time Statistics for Planar Flow

Next, we compute the statistics of travel times which is needed for estimation of the mean and variance of Q at a control plane. The statistics of travel time are obtained for the case of a nonuniform, unidirectional flow. As previously mentioned, nonuniformity in the flow field stems from the presence of groundwater recharge which is assumed to be constant and uniformly distributed throughout the flow domain. Furthermore, we assume aquifers displaying low-to-mild levels of heterogeneity (i.e., $\sigma_Y^2 \lesssim 1$) and that the flow field is essentially horizontal. The latter implies that flow can be approximated as two-dimensional (i.e., $d = 2$) in the $x_1 - x_2$ plane. This assumption also implies that we neglect the vertical velocity component induced by recharge (i.e., the Dupuit-Forschheimer approximation is applicable). The logconductivity field is characterized by an isotropic exponential covariance function with correlation scale λ .

The above-mentioned nonuniform flow conditions are identical to the ones investigated by Rubin and Bellin (1994), Butera and Tanda (1999), and Destouni et al. (2001). These works showed that the average longitudinal velocity varies according to the linear relationship

$$\langle U(x_1) \rangle = U_o \left[1 + \frac{\beta x_1}{\lambda} \right], \quad (11)$$

with U_o denoting the mean velocity at $x_1 = 0$. The dimensionless parameter $\beta = R\lambda / (T_G J_o)$ captures the effects of recharge where T_G corresponds to the geometric mean of the transmissivity field and J_o is the mean hydraulic gradient at $x_1 = 0$. The parameter β quantifies the level of nonuniformity and if $\beta = 0$, the flow field is uniform.

In order to compute the travel time PDF (which is needed to estimate the uncertainty in Q), the particle trajectory covariance has to be defined. For a planar nonuniform flow field, Rubin and Bellin (1994) showed that the longitudinal one-particle displacement covariance function, in the absence of local-scale dispersion, is given by

$$X_{11}^{(R)}(t) = \sigma_Y^2 \lambda^2 \left[2 \left(\frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right) - 3 \ln \left[\frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right] + \frac{3}{2} - 3E + 3\Theta(t) \right], \quad (12)$$

with

$$\Theta(t) = \text{Ei} \left[- \left(\frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right) \right] + \frac{\exp \left[- \left(\frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right) \right] \left(1 + \frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right) - 1}{\left(\frac{\exp[\beta t U_o / \lambda] - 1}{\beta} \right)^2}. \quad (13)$$

The superscript “(R)” is used to emphasize the presence of groundwater recharge, E corresponds to Euler’s constant (equal to 0.577) and $\text{Ei}[\cdot]$ is the exponential integral function. For uniform flows (i.e., $\beta = 0$), Equation 12 reduces to the one-particle displacement covariance function derived by Dagan (1984):

$$X_{11}(t) = \sigma_Y^2 \lambda^2 \left\{ 2 \frac{t U_o}{\lambda} - 3 \ln \left[\frac{t U_o}{\lambda} \right] + \frac{3}{2} - 3E + 3 \left[\text{Ei} \left[- \frac{t U_o}{\lambda} \right] + \frac{\exp \left[- \frac{t U_o}{\lambda} \right] \left(1 + \frac{t U_o}{\lambda} \right) - 1}{\left(\frac{t U_o}{\lambda} \right)^2} \right] \right\}. \quad (14)$$

Equation 12 is necessary in order to compute the travel time statistics.

As shown in section 3.1, the travel time PDF g_r is a key component in order to compute the first two moments of Q at the control plane. The travel time PDF will be a function of $X_{11}^{(R)}$ (12) (Dagan, 1989; Cvetković et al., 1992; Rubin, 2003). In our work, we opt to adopt the approach described in Cvetković et al. (1992) and Destouni et al. (2001) in which a lognormal PDF model is used for the travel time PDF (see also Andričević & Cvetković, 1996; Andričević et al., 1994; de Barros, 2018; Gotovac et al., 2009; Persson & Destouni, 2009; Selroos, 1997a, 1997b). The lognormal approximation for g_r is appropriate for low-to-mild heterogeneous formations, that is, $\sigma_Y^2 \lesssim 1$ (pp. 254–256 of Rubin, 2003) and has been verified numerically (e.g., Bellin et al., 1994; Gotovac et al., 2009). Other models such as the inverse Gaussian and Gamma PDFs have also

been adopted to capture travel time statistics (e.g., de Barros, 2018; Fiori & Russo, 2008; Fiori et al., 2017; Kreft & Zuber, 1978). Destouni et al. (2001) showed that the lognormal travel time PDF can be parameterized by the following expressions for the mean (μ_τ) and variance (σ_τ^2) at the control plane located at a distance L from the source zone:

$$g_\tau(\tau) = \frac{1}{\tau \sigma_\tau \sqrt{2\pi}} \exp \left[-\frac{(\ln \tau - \mu_\tau)^2}{2\sigma_\tau^2} \right], \quad (15)$$

with

$$\mu_\tau = \begin{cases} \frac{\lambda}{U_o} \ln \left[1 + \frac{\beta L}{\lambda} \right], & \text{for } \beta \neq 0 \\ \frac{L}{U_o}, & \text{for } \beta = 0, \end{cases} \quad (16)$$

$$\sigma_\tau^2 = \begin{cases} \frac{1}{U_o^2} X_{11}^{(R)} \left(t = \frac{L}{U_o} \right), & \text{for } \beta \neq 0 \\ \frac{1}{U_o^2} X_{11} \left(t = \frac{L}{U_o} \right), & \text{for } \beta = 0. \end{cases} \quad (17)$$

In summary, g_τ will account for the statistical properties of the solute trajectory. This is achieved by inserting Equation 12 into (16) and (17) which will be used to parametrize g_τ (15). With the travel time PDF (15), we can then evaluate the moments of Q given in Equations 7 and 9. We point out that the Lagrangian formulation has been used to derive expressions for key statistics of the solute mass flux as a function of physical and chemical parameters in multiple hydrological settings (Andrićević & Cvetković, 1996; Botter et al., 2005; Cvetković et al., 1992; de Barros, 2018; Fiori et al., 2002). Using the Lagrangian framework to perform both parametric uncertainty and data worth analysis on environmental performance metrics can be found in Rubin and Dagan (1992), de Barros and Rubin (2008), and (de Barros et al., 2012). Severino et al. (2012) also utilized the Lagrangian framework to study the effects of parameter uncertainty (in reaction rates) on the mean flux-averaged concentration. Here, we employ the Lagrangian framework to develop *application oriented* semianalytical solutions for the mean and variance of the solute discharge at a control plane that accounts for a generic function of the mass release rate at the source zone and aquifer recharge rates, while considering nonergodic effects (i.e., small contaminant sources).

4. Global Sensitivity Analysis

4.1. Polynomial Chaos Expansion (PCE)

Let ω denote a quantity of interest (QoI) depending on a set of M parameters, through a model $f(\cdot)$, that is, $\omega = f(\mathbf{p})$ where $\mathbf{p} = \{p_1, \dots, p_M\}$ is the vector collecting the input parameters. The QoI can represent the solute arrival time, peak concentration, or the mass recovery at a remediation site. If the input parameters are affected by uncertainty, ω becomes random in turn.

When the variance of ω is finite, the approximation of the model $\omega = f(\mathbf{p})$, namely $\hat{\omega}$, via a polynomial chaos expansion (PCE) is admitted (e.g., Ghanem & Spanos, 1991; Wiener, 1938; Xiu & Karniadakis, 2002)

$$\hat{\omega} = \sum_{\mathbf{a} \in \mathbb{N}^M} s_{\mathbf{a}} \Psi_{\mathbf{a}}(\mathbf{p}). \quad (18)$$

In (18), multi-indices $\mathbf{a} = \{a_1, \dots, a_M\} \in \mathbb{N}^M$ are associated with multivariate polynomials $\Psi_{\mathbf{a}}$ of degree $|\mathbf{a}| = \sum_{i=1}^M a_i$, which constitute an orthonormal basis with respect to the joint PDF of \mathbf{p} ; coefficients $s_{\mathbf{a}}$ are deterministic coordinates of the spectral decomposition (Xiu & Karniadakis, 2002).

In practical applications, the PCE (18) is opportunely truncated as

$$\hat{\omega} = \sum_{i=0}^{P-1} s_i \Psi_i(\mathbf{p}), \quad P = \frac{(M+q)!}{M!q!}, \quad (19)$$

where q is the maximum degree of the expansion, i.e., $|\mathbf{a}| \leq q$ for all $\mathbf{a} \in \mathbb{N}^M$.

PCE coefficients can be conveniently computed by means of nonintrusive methods like a regression-based approach consisting in the minimization of the variance of a suitable residual, $\varepsilon = |\hat{\omega} - \omega|$, with respect

to the vector of the PCE coefficients s_i (Sudret, 2008). An optimal set of regression points in the random parameter space is returned by the probabilistic collocation method (Webster et al., 1996).

The zero-order coefficient s_0 corresponds to the mean of ω . If ω is also function of time and/or space (or eventually of other parameters not affected by uncertainty, that is, deterministic quantities treated as single-point values) this dependence is included into the PCE coefficients; this reflects on the need to compute the PCE coefficients for each space-time location at which the model has to be approximated.

Note that in case of correlated input parameters one has to preliminary transform \mathbf{p} onto a set of independent random variables to derive the PCE approximation (e.g., Section 4.1 in Um et al., 2019).

4.2. Analytic Computation of Sensitivity Indices

Global sensitivity analysis (GSA) allows exploring simultaneously the full range of variation of random input parameters. As a consequence, GSA (i) handles interaction effects among input uncertainties and (ii) identifies how uncertainty in the QoI can be apportioned to the different inputs.

Variance-based metrics, such as the Sobol (1993, 2001) indices can be derived with an analytical postprocessing of PCE coefficients, drastically reducing the computational cost associated with GSA (Ciriello, Di Federico, et al., 2013; Oladyshkin et al., 2012; Sudret, 2008). Specifically, the contribution of the i th parameter, p_i , to the total variance of ω , that is, σ_ω^2 , is quantified by the principle sensitivity index of Sobol, S_i , which is computed as (Sudret, 2008)

$$S_i = \frac{\sum_{\gamma \in \Gamma_i} s_\gamma^2 \langle \Psi_\gamma^2(p_i) \rangle}{\sigma_\omega^2}, \quad \sigma_\omega^2 = \sum_{i=1}^{P-1} s_i^2 \langle \Psi_i^2(\mathbf{p}) \rangle, \quad (20)$$

where $\Gamma_i = \{\gamma \in (1, \dots, P-1) : \Psi_\gamma(p_i)\}$ is the set including the polynomials of the expansion depending only on p_i , that is, $\Psi_\gamma(p_i)$. The definition provided by (20) may be extended to compute the joint influence of a subset of inputs when the variance of ω cannot be fully explained by the principle sensitivity indices. As an example, when computing S_{ij} , the numerator becomes $\sum_{\gamma \in \Gamma_{ij}} s_\gamma^2 \langle \Psi_\gamma^2(p_{ij}) \rangle$, where $\Gamma_{ij} = \{\gamma \in (1, \dots, P-1) : \Psi_\gamma(p_{ij})\}$.

It is also possible to derive the total sensitivity index of the i th parameter, ST_i , which is given by the sum of all the indices computed by (20) and associated with p_i . While S_i represents the reduction in σ_ω^2 if p_i is not uncertain, ST_i is related to the expected residual variance of the response if only p_i is uncertain. If S_i and ST_i return similar values, then sensitivity indices accounting for the joint influence of subsets of parameters including p_i are negligible (Sobol', 1993).

In case of correlated inputs, one has to move toward distribution-based GSA metrics whose computation can be also handled resorting to PCE as shown in Ciriello et al. (2019).

5. Illustration and Discussion

For our computational illustrations, all variables are dimensionless. The dimensionless groups for time, distance, and mass release rate are as follows: tU_o/λ , x_1/λ and $\kappa_r \lambda/U_o$. As for the mean and standard deviations of Q , the dimensionless groups adopted are $\langle Q \rangle \lambda / (M_o U_o)$ and $\sigma_Q \lambda / (M_o U_o)$. This is equivalent to setting $M_o = 1$, $U_o = 1$, $\lambda = 1$ in the expressions for the solute discharge moments. Therefore, all variables displayed in the upcoming figures are in dimensionless form.

This section is divided into two part: section 5.1 provides a systematic analysis of the physics involving the effects of recharge and mass release rates on the solute discharge uncertainty. Then, in section 5.2 we assess parameter uncertainty (i.e., GSA) on the moments of the solute discharge .

5.1. Effects of Aquifer Recharge and Mass Release Rates on the Solute Discharge Moments

Figure 1 shows the travel time CDFs for different values of β for a fixed level of heterogeneity ($\sigma_\gamma^2 = 1$) and two control planes ($x_1/\lambda = 5$ and 15). The travel time CDF is given by $G_\tau(t) = \int_0^t g_\tau(\tau) d\tau$. As depicted in Figure 1, an increase in groundwater recharge (i.e., $\beta > 0$) leads to a larger probability of observing earlier breakthrough at the control plane. This is aligned with the results reported in Bellin et al. (1994) and Destouni et al. (2001). By comparing Figures 1a and 1b, the effects of groundwater recharge become more evident with travel distance.

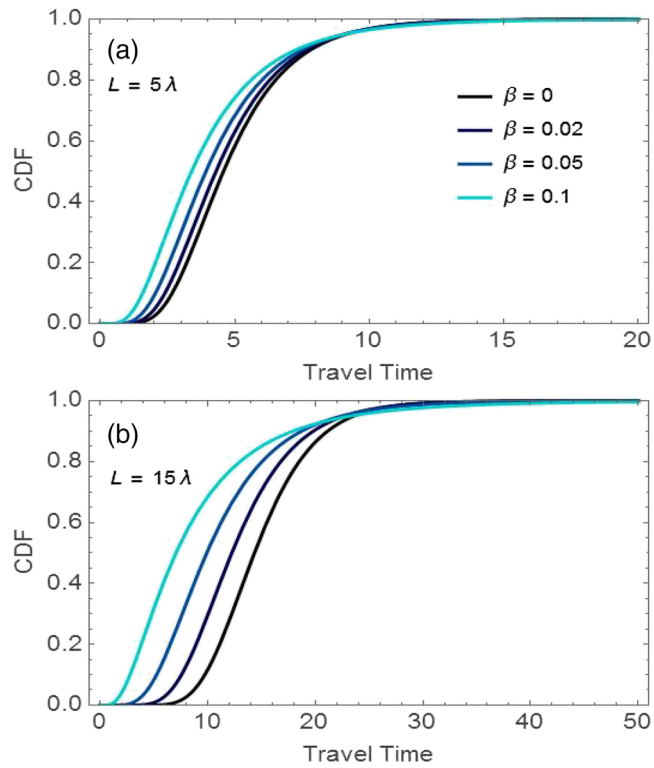


Figure 1. Travel time CDF at two distinct control plane locations (a) $x_1 = 5\lambda$ and (b) $x_1 = 15\lambda$ for different values of β and $\sigma_Y^2 = 1$. All results displayed are dimensionless.

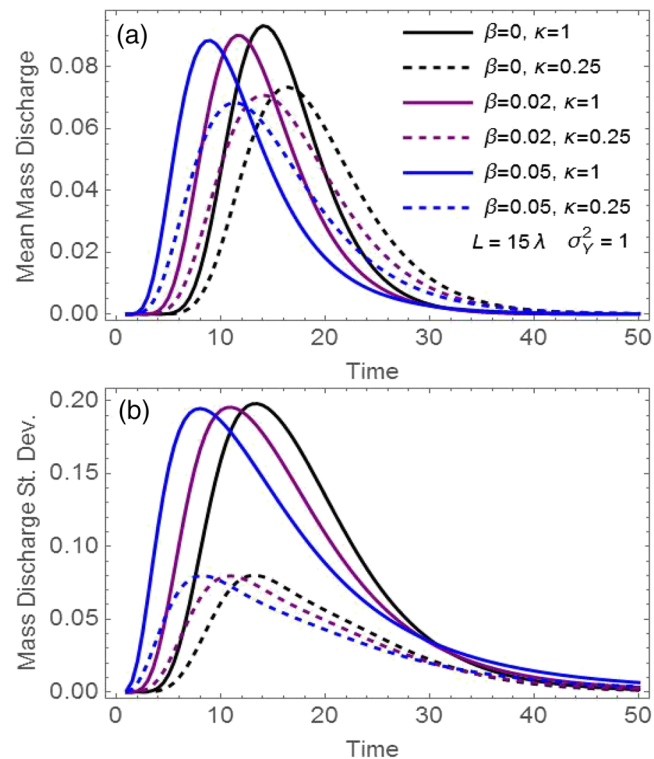


Figure 2. Temporal evolution of the (a) mean and (b) standard deviation of the solute mass discharge for $\kappa = 0.25$ and 1 and different values of $\beta = [0, 0.02, 0.05]$ at $x_1/\lambda = 15$ with $\kappa = \kappa_r \lambda / U_o$. All results displayed are dimensionless.

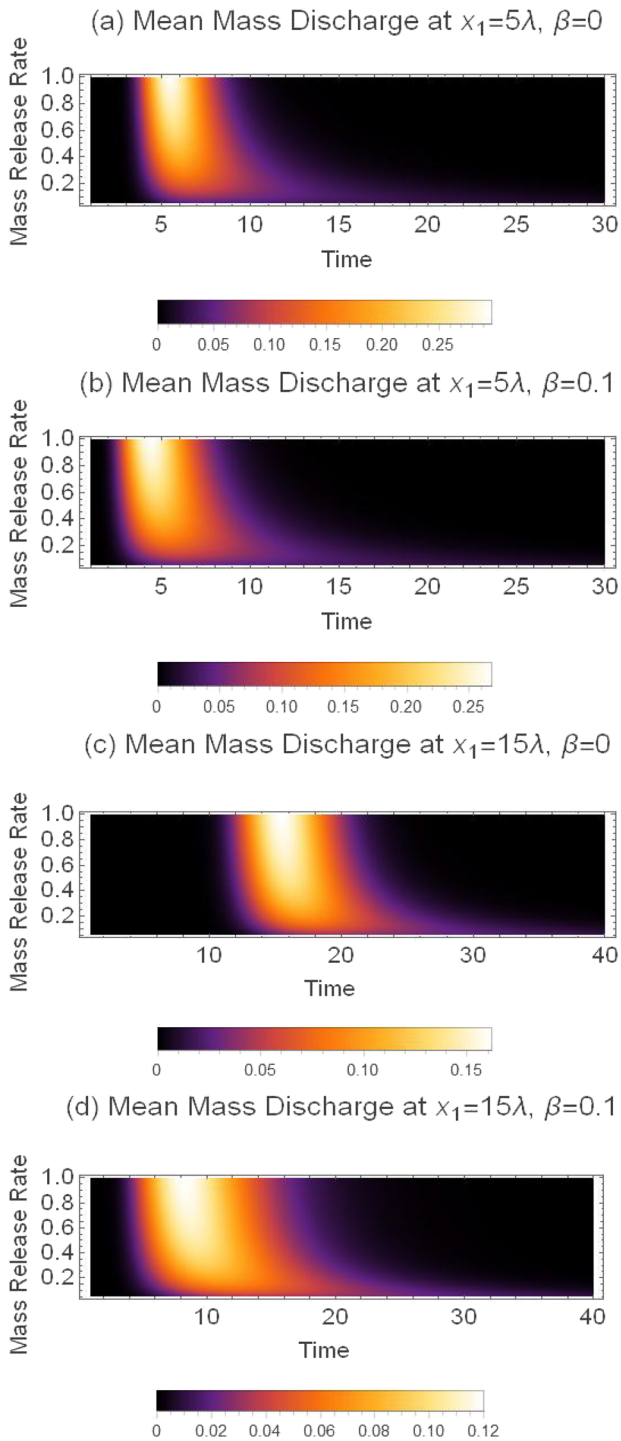


Figure 3. Contour maps of the mean solute discharge for $\sigma_Y^2 = 0.25$. Results obtained for $x_1/\lambda = [5, 15]$ and $\beta = [0, 0.1]$. All results displayed are dimensionless.

mean and standard deviation of the total solute mass discharge, Q . We do so by modeling σ_Y^2 , β and κ as independent random variables. The uncertainties in σ_Y^2 , β , and κ will propagate uncertainty to the first two moments of Q .

Based on the data reported in the literature (e.g., Murakami et al., 2010; Rubin, 2003) we select a truncated lognormal PDF to represent uncertainty associated with σ_Y^2 . The PDF is characterized by an arithmetic mean

Next, we analyze the effects of mass release rates and recharge on the solute discharge uncertainty. Figure 2 shows how the low-order moments of Q varies in time for distinct values of β and κ (where $\beta = R\lambda / (T_G J_o)$ and $\kappa = \kappa_r \lambda / U_o$). The results shown in Figure 2 were computed for $x_1/\lambda = 15$ and $\sigma_Y^2 = 1$. The simulation results depicted in Figure 2a show that the breakthrough time at the control plane for the expected value of the solute discharge decreases with an increase of β . The impact of recharge on the standard deviation is also illustrated (Figure 2b). Close inspection of Figures 2a and 2b reveals that aquifer recharge has a minor impact on the magnitude of the peak of the low-order moments. This effect was observed in Bellin et al. (1994). As depicted in Figure 2, aquifer recharge mainly shifts the peak of the low-order moments of the solute discharge at the control plane. On the other hand, the peaks of the solute discharge mean and standard deviation are sensitive to changes in the source zone mass release rate. We observe lower peaks in solute discharge mean and standard deviation when κ decreases from 1 to 0.25. This indicates the importance on the source zone characteristics in controlling the uncertainty in solute mass fluxes.

We further explore the effects of aquifer recharge and mass release rate on the low-order moments of Q in Figures 3 and 4. The results shown in Figures 3 and 4 were computed for $\sigma_Y^2 = 0.25$. Figures 3a–3d show how the temporal evolution of the mean mass discharge is affected by changes in aquifer recharge and control plane distance. Similar to Figure 2, increasing β from 0 to 0.1 leads to earlier breakthrough of the mean solute discharge at both the control planes (i.e., $x_1/\lambda = 5$ and 15). At the same time, the contour plots displayed in Figures 3a–3d show that the temporal spreading behavior of the mean Q is quite insensitive to variations in the mass release rate within the range $\kappa \gtrsim 0.4$. For $\kappa < 0.4$, the temporal spreading of the mean Q is significantly affected by changes in κ .

Similarly, Figures 4a–4d show the standard deviation of the mass discharge as a function of both time and mass release rate. We observe that the highest uncertainty in Q corresponds to higher values of κ . Decreasing the mass release rate, that is, slow injections of the solute into the aquifer, diminishes the standard deviation of the solute mass discharge. As for the effect of aquifer recharge, the peak of the standard deviation occurs at an earlier time for higher β values (compare Figure 4a with 4b and Figure 4c with 4d).

Similar to Figures 3 and 4, Figure 5 depicts the mean and standard deviation of the solute discharge at $x_1/\lambda = 15$ but for a higher level of heterogeneity, that is, $\sigma_Y^2 = 1$. The effects of σ_Y^2 on the moments of Q are elucidated when comparing Figures 3c, 3d, 4c, and 4d with Figures 5a–5d. The level of aquifer heterogeneity contributes significantly to the temporal spreading of the moments of Q . In particular we point out that the asymmetry of the mass discharge standard deviation becomes more evident for $\beta = 0.1$ when compared to $\beta = 0$ (compare Figure 5c with 5d).

5.2. Application of Global Sensitivity Analysis

We apply the approach described in section 4 to investigate the influence of the uncertainty in the parameters σ_Y^2 , β and κ on the QoI, that is, the

mean and standard deviation of the total solute mass discharge, Q . We do so by modeling σ_Y^2 , β and κ as independent random variables. The uncertainties in σ_Y^2 , β , and κ will propagate uncertainty to the first two

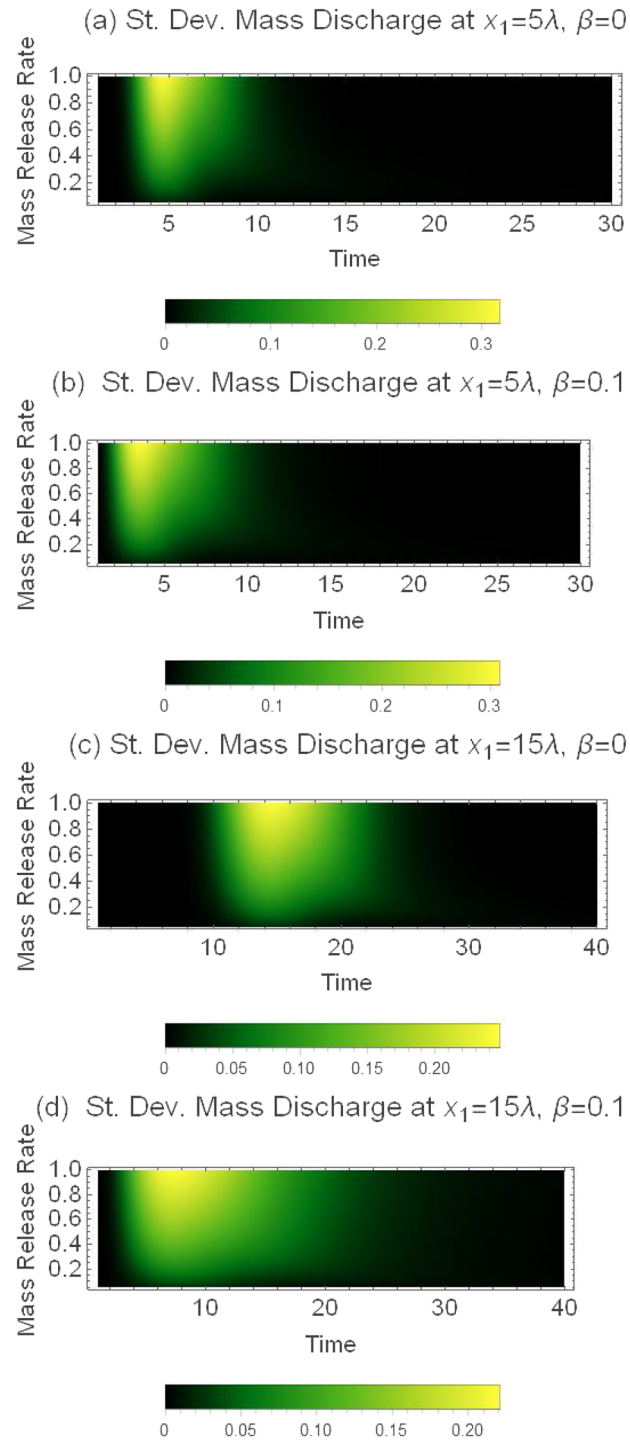


Figure 4. Contour maps of the standard deviation of the solute discharge for $\sigma_Y^2 = 0.25$. Results are obtained for $x_1/\lambda = [5, 15]$ and $\beta = [0, 0.1]$. All results displayed are dimensionless.

of 0.2 and variance of 0.025 and is truncated in the range $[0.005, 1]$. Following the observations reported in Xie et al. (2018, 2019), we model β as a normal random variable with mean 0.01, variance of 0.000064 and truncated in the range $[0, 0.075]$. As for κ , we consider a lognormal PDF with arithmetic mean 0.7 and variance 0.5. Note that these PDFs are selected for the purpose of illustration. Other models and ranges of values can be adopted within the methodological framework.

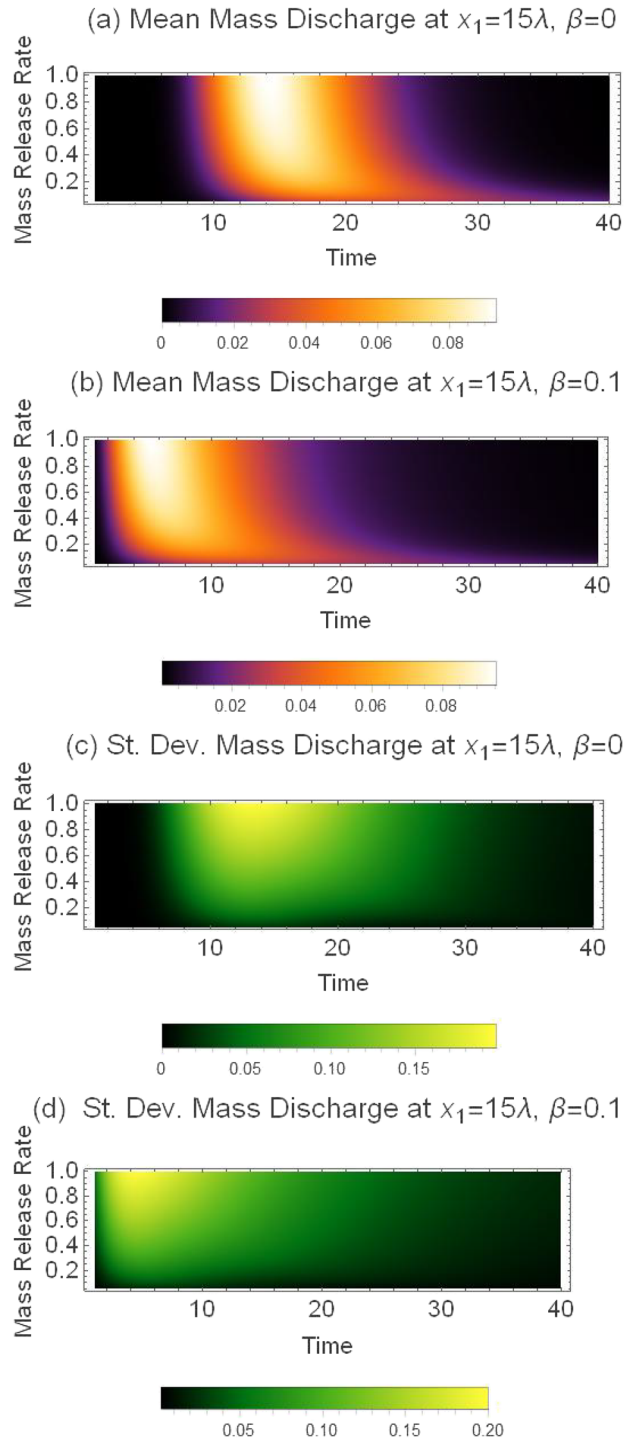


Figure 5. Contour maps of the mean and standard deviation of the solute discharge for $\sigma_Y^2 = 1$. Results are obtained for $x_1/\lambda = 15$ and for two distinct values of $\beta = [0, 0.1]$. All results displayed are dimensionless.

Based on the evidences reported in previous studies (e.g., Ciriello et al., 2015; Xiu & Tartakovsky, 2006), we select second-order PCEs to approximate the response surfaces of $\langle Q \rangle$ and σ_Q as follows:

$$\widehat{\langle Q(t) \rangle} = \sum_{i=0}^{P-1} a_i(t) \Psi_i(\xi_1, \xi_2, \xi_3), \quad (21)$$

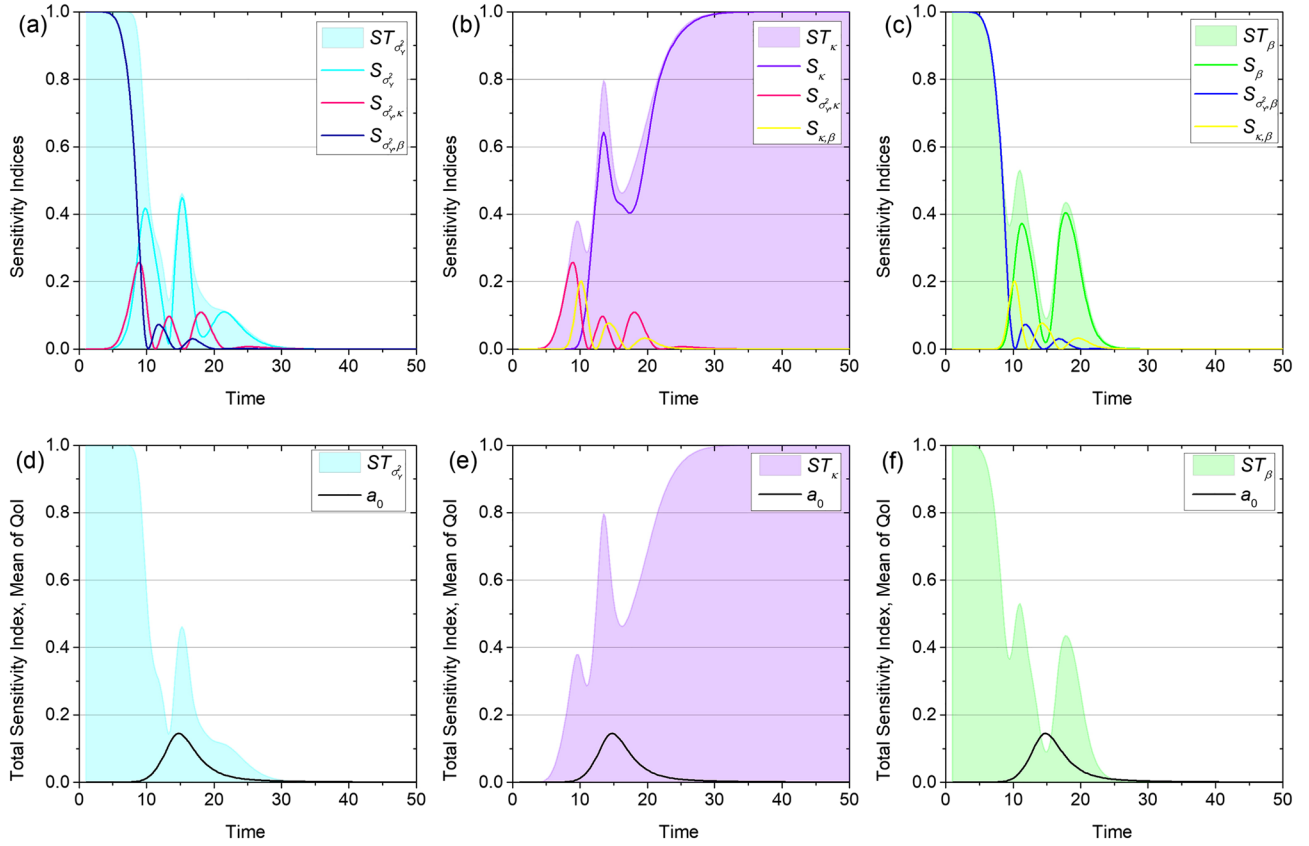


Figure 6. With reference to $\widehat{\langle Q \rangle}$: (a) $ST_{\sigma_Y^2}$ in time and its components $S_{\sigma_Y^2}$, $S_{\sigma_Y^2, \kappa}$, $S_{\sigma_Y^2, \beta}$; (b) ST_{κ} in time and its components S_{κ} , $S_{\sigma_Y^2, \kappa}$, $S_{\kappa, \beta}$; (c) ST_{β} in time and its components S_{β} , $S_{\sigma_Y^2, \beta}$, $S_{\kappa, \beta}$; (d) $ST_{\sigma_Y^2}$ in time and the mean of the QoI $\widehat{\langle Q \rangle}$ ($a_0(t)$ in (21)); (e) ST_{κ} in time and the mean of the QoI $\widehat{\langle Q \rangle}$ (a_0 in (21)); (f) ST_{β} in time and the mean of the QoI $\widehat{\langle Q \rangle}$ (a_0 in (21)).

$$\widehat{\sigma_Q}(t) = \sum_{i=0}^{P-1} b_i(t) \Psi_i(\xi_1, \xi_2, \xi_3), \quad (22)$$

where $P = 10$, Ψ_i denotes multivariate Hermite polynomials as the parameters are associated with normal/lognormal distributions (e.g., Xiu & Karniadakis, 2002), and ξ_1, ξ_2, ξ_3 are standard Gaussian variables related to σ_Y^2 , β , κ , respectively, by means of an isoprobabilistic transform (Sudret, 2008).

Following the approach described in section 4, the coefficients a_i and b_i are computed for dimensionless time $\mathcal{T} \in (1, \dots, 50)$ where $\mathcal{T} = tU_o/\lambda$ with a time step of $\Delta\mathcal{T} = 0.25$. Once the PCE coefficients are available, the sensitivity indices of Sobol are analytically computed for each uncertain parameter with respect to the two QoI approximated by Equations 21–22.

In the following, we first analyze GSA results related to $\widehat{\langle Q \rangle}$, Equation 21, at a control plane location $x_1/\lambda = 15$. Figure 6a shows the time evolution of the total sensitivity index, principal sensitivity index, and second-order sensitivity indices associated with σ_Y^2 . Figures 6b and 6c do the same for the sensitivity indices associated with κ and β , respectively. Figures 6d and 6f overlap the total sensitivity index of σ_Y^2 , κ , and β respectively, with the mean (in the random parameter space) of $\widehat{\langle Q \rangle}$, provided by coefficient $a_0(t)$ in (21).

We observe that the total sensitivity index of σ_Y^2 rapidly decreases after early times, with the presence of a second peak. At approximately $\mathcal{T} = 30$, the total sensitivity index reaches zero. This temporal pattern is mainly attributed to the combined effect of the uncertainty in σ_Y^2 and β at early times, while the principal sensitivity index associated with σ_Y^2 becomes predominant at about $\mathcal{T} = 10$ (Figure 6a). Figure 6c reveals a very similar behavior, that is, the total sensitivity index of β rapidly decreases and after a couple of residual peaks reaches zero at $\mathcal{T} \approx 25$. The combined effect of the uncertainty in σ_Y^2 and β is maximum at the very

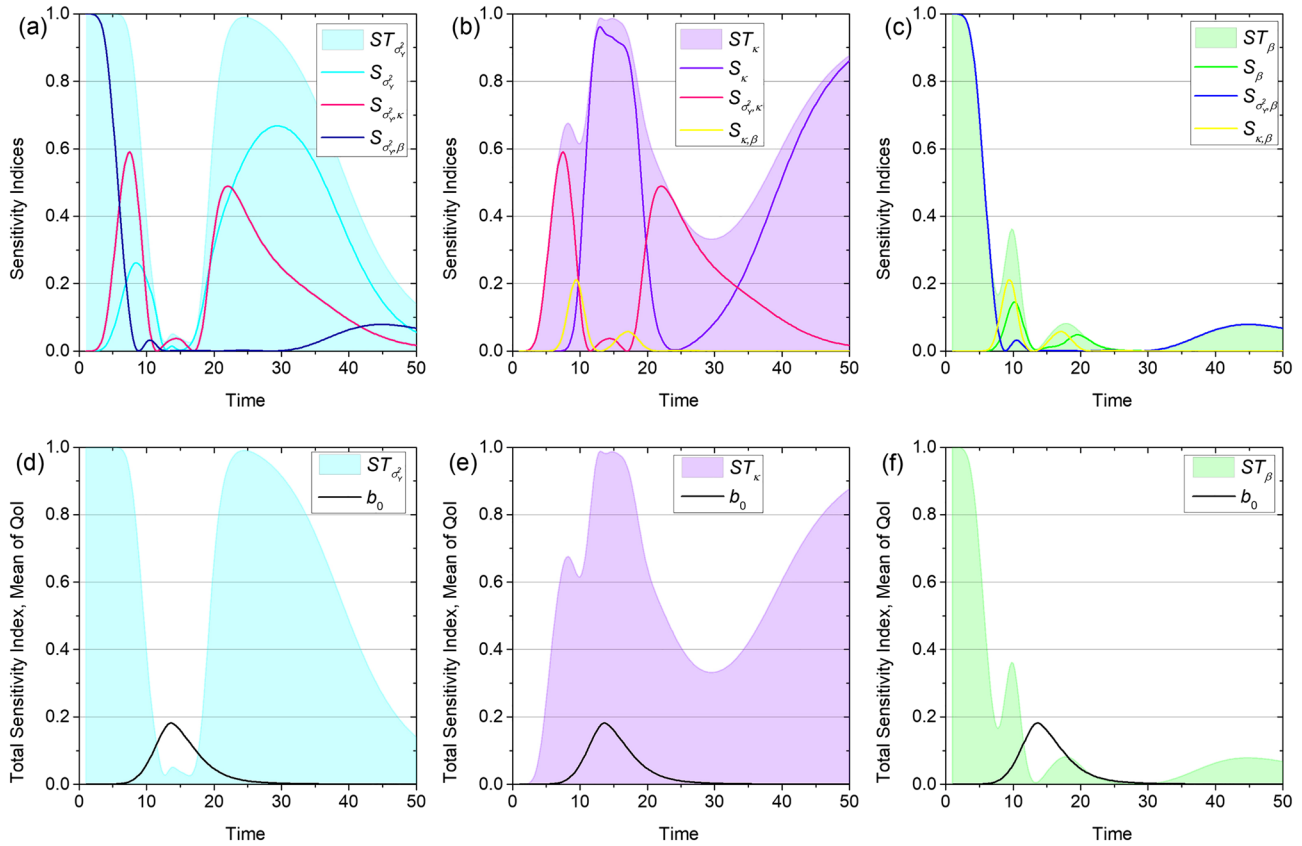


Figure 7. With reference to $\widehat{\sigma_Q}$: (a) $ST_{\sigma_Y^2}$ in time and its components $S_{\sigma_Y^2}$, $S_{\sigma_Y^2, \kappa}$, $S_{\sigma_Y^2, \beta}$; (b) ST_{κ} in time and its components S_{κ} , $S_{\sigma_Y^2, \kappa}$, $S_{\kappa, \beta}$; (c) ST_{β} in time and its components S_{β} , $S_{\sigma_Y^2, \beta}$, $S_{\kappa, \beta}$; (d) $ST_{\sigma_Y^2}$ in time and the mean of the QoI $\widehat{\sigma_Q}(b_0$ in (22)); (e) ST_{κ} in time and the mean of the QoI $\widehat{\sigma_Q}(b_0$ in (22)); (f) ST_{β} in time and the mean of the QoI $\widehat{\sigma_Q}(b_0$ in (22)).

beginning, while the principal sensitivity index associated with β becomes predominant at about $\mathcal{T} = 10$ (Figure 6c). The total sensitivity index of κ depicted in Figure 6b displays an opposite behavior: It is null at early times and after two increasing peaks it reaches a maximum at $\mathcal{T} \approx 30$. This is mainly due to the principal sensitivity index associated with κ for $\mathcal{T} \geq 10$, while before, the joint effect of σ_Y^2 and κ has major influence (Figure 6b).

In order to explore the physical meaning of these results we refer to Figures 6d–6f where we also report the zero-order coefficient a_0 , that is, the mean $\langle \widehat{Q} \rangle$. While the combined uncertainty in σ_Y^2 and β affects the ascending part of the mean $\langle \widehat{Q} \rangle$ profile (i.e., the front of the solute plume crossing the control plane), the tail (i.e., descending part of the curve after the peak of a_0) is influenced by the uncertainty in κ . This result is in agreement with the analysis carried out in Figure 3 where it is shown that the effects of κ are strongly manifested in the tailing edge of the solute plume. A slower mass release rate can potentially lead to prolonged tailing behavior in the temporal evolution of mean solute discharge at the control plane. The results reported in Figure 6 also show that the peak of the profile is influenced by the uncertainty in σ_Y^2 and κ . This is in line with the fact that the peak of the first moment of the solute discharge is strongly related to the spatial heterogeneity of the aquifer and to the mass release rate. Both physical heterogeneity and source zone mass release rate control the spatial distribution of the solute mass. The mass release rate controls the amount of mass being injected into the hydrogeological system. Therefore, the peak of the mean solute discharge is sensitive to both σ_Y^2 and κ . Our results also suggest that the recharge rate plays a role mainly in correspondence of the upward and downward branches of the profile around the peak of the mean $\langle \widehat{Q} \rangle$. These results are in agreement with (and extend) those presented in Figure 2 in which it is shown that the peak magnitude of the mean mass discharge is mainly affected by the variability in κ , while the time at which the peak is reached strongly depends on the value of β . As discussed in section 5.1, when β increases, the mean velocity of the aquifer is augmented, and we observe earlier solute breakthrough time at the control

plane. For such reasons, the sensitivity of $\langle Q \rangle$ toward β is larger when the leading edge of the plume crosses the control plane.

Next, we analyze the GSA results related to the second QoI, that is, $\widehat{\sigma}_Q$ (Equation 22). Figure 7a shows values the temporal variation of the total sensitivity index, principal sensitivity index, and second-order sensitivity indices associated with σ_Y^2 . Figures 7b and 7c reports for the sensitivity indices associated with κ and β respectively. Finally, Figures 7d–7f overlap the total sensitivity index of σ_Y^2 , κ , and β , respectively, with the mean (in the random parameter space) of $\widehat{\sigma}_Q$, provided by coefficient b_0 in (22).

Figure 7a shows that the total sensitivity index of σ_Y^2 rapidly decreases after early times, and increases again until it reaches unity at $\mathcal{T} = 25 - 30$. This behavior is an outcome of the combined effect of the uncertainty in σ_Y^2 and β at the very beginning, while the principal sensitivity index associated with σ_Y^2 becomes predominant for $\mathcal{T} > 25$, followed by the joint influence of σ_Y^2 and κ (Figure 7a). Figure 7c illustrates that the total sensitivity index of β also rapidly decreases and after a second peak at about $\mathcal{T} = 10$, it remains under a value of 0.1. The combined effect of the uncertainty in σ_Y^2 and β is maximum at the very beginning, where the total sensitivity index of β exhibits its major influence (Figure 7c). The total sensitivity index of κ increases rapidly and maintains values greater than of 0.3 with a minimum at about $\mathcal{T} = 30$. This is mainly related to the principal sensitivity index associated with κ , followed by the joint effect of σ_Y^2 and κ (Figure 7b).

Figures 7d–7f show that both the ascending (i.e., leading edge) and descending (i.e., trailing edge) parts of the mean $\widehat{\sigma}_Q$ profile (captured by the temporal dynamics of b_0) are influenced by the uncertainties in σ_Y^2 , κ , and β . However, we point out the parameter σ_Y^2 has a significant role in the leading and trailing edges of $\widehat{\sigma}_Q$. This is expected since the leading and trailing edges of the solute plume in a heterogeneous aquifer are subject to largest uncertainty (see Rubin, 1991). As a consequence, the degree of heterogeneity σ_Y^2 plays an important role in predicting the uncertainty in the solute plume's early and late time behavior. The uncertainty of β mainly affects the ascending part of b_0 , whereas the peak of b_0 is influenced by the uncertainty in κ . As noticed above, these results are in agreement with and extend the theoretical analysis reported in Figure 2. In agreement with the results shown in section 5.1, the recharge parameter β is important in controlling the early breakthrough times (see also Rubin & Bellin, 1994). Therefore β will have an impact on the uncertainty estimate of the solute discharge at early times. The uncertainty in κ is dominant in the peak of b_0 since κ controls the amount of mass injected into the subsurface.

According to our GSA results it emerges that κ and secondly σ_Y^2 are the parameters whose uncertainty has the greatest impact on the robustness of the mean and standard deviation of solute discharge prediction. This result is limited to the conditions explored in this work. Nevertheless, it shows how the approach can be used to better allocate resources toward uncertainty reduction in these parameters to increase the robustness of model predictions. This is critical to design mitigation measures (Ciriello, Guadagnini, et al., 2013; Ciriello et al., 2015) as well as to aid decision makers to prioritize limited financial resources to reduce the uncertainty in risk estimation (de Barros & Rubin, 2008).

6. Summary

There is a need to enhance our fundamental understanding of how contaminant source zone characteristics and aquifer recharge rates impact the risks associated with aquifer pollution. It is expected that climate change will impact aquifer recharge rates (O'Connell & Hou, 2015) which can potentially affect aquifer remediation campaigns as shown in Libera et al. (2019). At the same time, the contaminant release history has a profound impact on transport and is often unknown. Quantification of the impact of these factors in solute mass fluxes is challenging, especially in the presence of the physical heterogeneity of the subsurface formation.

Through the use of the Lagrangian framework (Cvetković et al., 1992; Dagan et al., 1992) and the findings reported in the literature (Bellin et al., 1994; Destouni et al., 2001), our study expands the work of de Barros (2018) to account for the effects of aquifer recharge on the low-order moments of the solute discharge. The proposed solution is application-oriented and can be used in preliminary screening analysis. The expressions provide fundamental physical insights of the role of natural (i.e., hydrological) and engineered (i.e., source zone) parameters in controlling the transport behavior of a solute plume and its uncertainty. In the first part of our work, we developed semianalytical expressions to compute the mean and variance of the solute mass discharge at an environmentally sensitive location, as functions of the parameters characterizing aquifer

heterogeneity, the (uniform) recharge applied over the aquifer, and the solute mass release rate at the source zone. The sensitive location is represented as a control plane normal to the mean flow direction and the flow field is assumed to be nonuniform and unidirectional. The derived semianalytical solutions are then employed to provide an insight on the role of aquifer recharge and solute mass release rates at the source zone on low-order statistics of the mass discharge at the control plane. This is achieved through a systematic analysis of $\langle Q \rangle$ and σ_Q for different (single point) values of the governing parameters.

Given the uncertainty present in multiple parameters characterizing the hydrogeological system, we perform a GSA in order to better understand the impact of uncertainty in the parameters σ_Y^2 , κ , and β on the robustness of predictions of the mean and standard deviation of the solute mass discharge. This is done to extend the theoretical analysis by enabling parameters to vary simultaneously over the full range of their plausible variation. Our results reveal the primary role of κ in affecting the peaks of the mean (in the random parameter space) profiles of $\langle Q \rangle$ and σ_Q , while β mainly affects the time at which the peaks are reached (i.e., the ascending part of both the profiles) and the dynamics of the leading edge of the solute plume. As for σ_Y^2 , our results illustrate that it affects the ascending part and the peak of the mean $\langle Q \rangle$ profile, while its impact is concentrated on the upward and downward branches around the peak (i.e., leading and trailing edge of the plume) when considering the mean σ_Q profile. Overall, it can be said that uncertainty in κ exhibits a major impact on model outputs, followed by uncertainty associated with σ_Y^2 . Therefore, to narrow the uncertainty in these parameters is important to increase the robustness of model predictions.

The results presented in this work are limited to the specific setting analyzed and the range of values explored. The semianalytical expressions for $\langle Q \rangle$ and σ_Q are limited to aquifers displaying low-to-mild levels of heterogeneity (i.e., $\sigma_Y^2 \lesssim 1$). The computational results illustrated are also limited to the mass release model adopted. Other source zone mass release models, such as the one provided in equation 8 of Selroos (1997a), can be incorporated in our study. Although we considered nonreactive advective dominated transport, the effects of local-scale dispersion as well as reactions can be incorporated in our semianalytical expressions for the solute mass discharge by adopting the particle-trajectory covariance and reaction functions reported in Fiori et al. (2002). The GSA results will depend on the statistical characterization of the uncertain parameters as well as the type of hydrogeological heterogeneity model adopted. Further research should be carried out in investigating how the uncertainty of these parameters are propagated to transport predictions in aquifers displaying high levels of heterogeneity (i.e., $\sigma_Y^2 > 1$). Under these conditions, the benefits of the PCE method employed in our work will become evident since the first-order approximation will not hold and a numerical Monte Carlo framework will have to be adopted. Nevertheless, the first-order solutions can be used for preliminary screening stages of the analysis and to obtain physical insights on the interaction of key parameters on groundwater contamination risk estimates.

Data Availability Statement

There are no data sharing issues given that this work is computational, and all numerical results are reported in the tables and figures.

Acknowledgments

V. C. acknowledges support from University of Bologna RFO (Ricerca Fondamentale Orientata) 2015 2016 and ALMAIDEA Grant 2017. FPJdB gratefully acknowledge the financial support by the National Science Foundation under grant 1654009. Both authors are thankful to the three reviewers, Editor, and Associate Editor for the time invested in the review process as well as the constructive comments that helped improve the overall quality of this work.

References

- Alapati, S., & Kabala, Z. J. (2000). Recovering the release history of a groundwater contaminant using a non-linear least-squares method. *Hydrological Processes*, *14*(6), 1003–1016.
- Andričević, R., & Cvetković, V. (1996). Evaluation of risk from contaminants migrating by groundwater. *Water Resources Research*, *32*(3), 611–621.
- Andričević, R., Daniels, J. I., & Jacobson, R. L. (1994). Radionuclide migration using a travel time transport approach and its application in risk analysis. *Journal of Hydrology*, *163*(1–2), 125–145.
- Bellin, A., Rubin, Y., & Rinaldo, A. (1994). Eulerian-Lagrangian approach for modeling of flow and transport in heterogeneous geological formations. *Water Resources Research*, *30*(11), 2913–2924.
- Bellin, A., & Tonina, D. (2007). Probability density function of non-reactive solute concentration in heterogeneous porous formations. *Journal of Contaminant Hydrology*, *94*(1), 109–125.
- Boso, F., Broyda, S., & Tartakovsky, D. (2014). Cumulative distribution function solutions of advection–reaction equations with uncertain parameters. *Proceedings of the Royal Society A: Mathematical Physical and Engineering Sciences*, *470*(2166), 20140189.
- Boso, F., & Tartakovsky, D. M. (2016). The method of distributions for dispersive transport in porous media with uncertain hydraulic properties. *Water Resources Research*, *52*, 4700–4712. <https://doi.org/10.1002/2016WR018745>
- Botter, G., Bertuzzo, E., Bellin, A., & Rinaldo, A. (2005). On the Lagrangian formulations of reactive solute transport in the hydrologic response. *Water Resources Research*, *41*, W04008. <https://doi.org/10.1029/2004WR003544>
- Butera, I., & Tanda, M. G. (1999). Solute transport analysis through heterogeneous media in nonuniform in the average flow by a stochastic approach. *Transport in Porous Media*, *36*(3), 255–291.

- Butera, I., & Tanda, M. G. (2003). A geostatistical approach to recover the release history of groundwater pollutants. *Water Resources Research*, 39(12), 1372. <https://doi.org/10.1029/2003WR002314>
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, 13(1), 206–222.
- Ciriello, V., Di Federico, V., Riva, M., Cadini, F., De Sanctis, J., Zio, E., & Guadagnini, A. (2013). Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer. *Stochastic Environmental Research and Risk Assessment*, 27, 945–954.
- Ciriello, V., Edery, Y., Guadagnini, A., & Berkowitz, B. (2015). Multimodel framework for characterization of transport in porous media. *Water Resources Research*, 51, 3384–3402. <https://doi.org/10.1002/2015WR017047>
- Ciriello, V., Guadagnini, A., Di Federico, V., Edery, Y., & Berkowitz, B. (2013). Comparative analysis of formulations for conservative transport in porous media through sensitivity-based parameter calibration. *Water Resources Research*, 49, 5206–5220. <https://doi.org/10.1002/wrcr.20395>
- Ciriello, V., Lauriola, I., Bonvicini, S., Cozzani, V., Federico, V. D., & Tartakovsky, D. M. (2017). Impact of hydrogeological uncertainty on estimation of environmental risks posed by hydrocarbon transportation networks. *Water Resources Research*, 53, 8686–8697. <https://doi.org/10.1002/2017WR021368>
- Ciriello, V., Lauriola, I., & Tartakovsky, D. M. (2019). Distribution-based global sensitivity analysis in hydrology. *Water Resources Research*, 55, 8708–8720. <https://doi.org/10.1029/2019WR025844>
- Crestaux, T., Maitre, L. O., & Martinez, J.-M. (2009). Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering and System Safety*, 94(7), 1161–1172.
- Cvetković, V., Shapiro, A. M., & Dagan, G. (1992). A solute flux approach to transport in heterogeneous formations: 2. Uncertainty analysis. *Water Resources Research*, 28(5), 1377–1388.
- Cvetkovic, V., Soltani, S., & Vigouroux, G. (2015). Global sensitivity analysis of groundwater transport. *Journal of Hydrology*, 531, 142–148.
- Dagan, G. (1984). Solute transport in heterogeneous porous formations. *Journal of Fluid Mechanics*, 145, 151–177.
- Dagan, G. (1989). *Flow and transport in porous formations*. Berlin, Heidelberg: Springer-Verlag.
- Dagan, G., Cvetkovic, V., & Shapiro, A. (1992). A solute flux approach to transport in heterogeneous formations: 1. The general framework. *Water Resources Research*, 28(5), 1369–1376.
- de Barros, F. P. J. (2018). Evaluating the combined effects of source zone mass release rates and aquifer heterogeneity on solute discharge uncertainty. *Advances in Water Resources*, 117, 140–150.
- de Barros, F. P. J., Ezzedine, S., & Rubin, Y. (2012). Impact of hydrogeological data on measures of uncertainty, site characterization and environmental performance metrics. *Advances in Water Resources*, 36, 51–63.
- de Barros, F. P. J., & Fiori, A. (2014). First-order based cumulative distribution function for solute concentration in heterogeneous aquifers: Theoretical analysis and implications for human health risk assessment. *Water Resources Research*, 50, 4018–4037. <https://doi.org/10.1002/2013WR0150>
- de Barros, F. P. J., & Rubin, Y. (2008). A risk-driven approach for subsurface site characterization. *Water Resources Research*, 44, W01414. <https://doi.org/10.1029/2007WR006081>
- Dentz, M., Le Borgne, T., Englert, A., & Bijeljic, B. (2011). Mixing, spreading and reaction in heterogeneous media: A brief review. *Journal of Contaminant Hydrology*, 120, 1–17.
- Dentz, M., & Tartakovsky, D. M. (2010). Probability density functions for passive scalars dispersed in random velocity fields. *Geophysical Research Letters*, 37, L24406. <https://doi.org/10.1029/2010GL045748>
- Destouni, G., Simic, E., & Graham, W. (2001). On the applicability of analytical methods for estimating solute travel time statistics in nonuniform groundwater flow. *Water Resources Research*, 37(9), 2303–2308.
- Ezzedine, S., & Rubin, Y. (1997). Analysis of the Cape Cod tracer data. *Water Resources Research*, 33(1), 1–11.
- Falta, R. W., Rao, P. S., & Basu, N. (2005). Assessing the impacts of partial mass depletion in DNAPL source zones: I. Analytical modeling of source strength functions and plume response. *Journal of Contaminant Hydrology*, 78(4), 259–280.
- Fiori, A., Berglund, S., Cvetković, V., & Dagan, G. (2002). A first-order analysis of solute flux statistics in aquifers: The combined effect of pore-scale dispersion, sampling, and linear sorption kinetics. *Water Resources Research*, 38(8), 1137. <https://doi.org/10.1029/2001WR000678>
- Fiori, A., & Dagan, G. (2000). Concentration fluctuations in aquifer transport: A rigorous first-order solution and applications. *Journal of Contaminant Hydrology*, 45(1), 139–163.
- Fiori, A., & Russo, D. (2008). Travel time distribution in a hillslope: Insight from numerical simulations. *Water Resources Research*, 44, W12426. <https://doi.org/10.1029/2008WR007135>
- Fiori, A., Zarlenga, A., Jankovic, I., & Dagan, G. (2017). Solute transport in aquifers: The comeback of the advection dispersion equation and the first order approximation. *Advances in Water Resources*, 110, 349–359.
- Foussereau, X., Graham, W., Akpoji, G., Destouni, G., & Rao, P. (2001). Solute transport through a heterogeneous coupled vadose-saturated zone system with temporally random rainfall. *Water Resources Research*, 37, 1577–1588.
- Ghanem, R. G., & Spanos, P. D. (1991). *Stochastic finite elements—a spectral approach*. Berlin: Springer. <https://doi.org/10.1007/978-1-4612-3094-6>
- Gotovac, H., Cvetkovic, V., & Andricevic, R. (2009). Flow and travel time statistics in highly heterogeneous porous media. *Water Resources Research*, 45, W07402. <https://doi.org/10.1029/2008WR007168>
- Henri, C. V., Fernández-García, D., & de Barros, F. P. J. (2015). Probabilistic human health risk assessment of degradation-related chemical mixtures in heterogeneous aquifers: Risk statistics, hot spots, and preferential channels. *Water Resources Research*, 51, 4086–4108. <https://doi.org/10.1002/2014WR016717>
- Im, J., Rizzo, C. B., & de Barros, F. P. J. (2020). Resilience of groundwater systems in the presence of bisphenol A under uncertainty. *Science of the Total Environment*, 727, 138363.
- Kapoor, V., & Kitanidis, P. K. (1998). Concentration fluctuations and dilution in aquifers. *Water Resources Research*, 34(5), 1181–1193.
- Kokkinaki, A., Werth, C. J., & Sleep, B. (2014). Comparison of upscaled models for multistage mass discharge from DNAPL source zones. *Water Resources Research*, 50, 3187–3205.
- Kreft, A., & Zuber, A. (1978). On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. *Chemical Engineering Science*, 33(11), 1471–1480.
- Li, L., & Graham, W. (1999). Stochastic analysis of solute transport in heterogeneous aquifers subject to spatiotemporal random recharge. *Water Resources Research*, 35(4), 953–971.
- Libera, A., de Barros, F. P. J., Faybishenko, B., Eddy-Dilek, C., Denham, K., Lipnikov, K., et al. (2019). Climate change impact on residual contaminants under sustainable remediation. *Journal of Contaminant Hydrology*, 226(103), 518.

- MacFarlane, D., Cherry, J., Gillham, R., & Sudicky, E. (1983). Migration of contaminants in groundwater at a landfill: A case study: 1. Groundwater flow and plume delineation. *Journal of Hydrology*, *63*(1–2), 1–29.
- Meyer, D. W., Tchelepi, H. A., & Jenny, P. (2013). A fast simulation method for uncertainty quantification of subsurface flow and transport. *Water Resources Research*, *49*, 2359–2379. <https://doi.org/10.1002/wrcr.20240>
- Michalak, A. M., & Kitaniadis, P. K. (2003). A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification. *Water Resources Research*, *39*(2), 1033. <https://doi.org/10.1029/2002WR001480>
- Moslehi, M., & de Barros, F. P. J. (2017). Uncertainty quantification of environmental performance metrics in heterogeneous aquifers with long-range correlations. *Journal of Contaminant Hydrology*, *196*, 21–29.
- Murakami, H., Chen, X., Hahn, M. S., Liu, Y., Rockhold, M. L., Vermeul, V. R., et al. (2010). Bayesian approach for three-dimensional aquifer characterization at the Hanford 300 area. *Hydrology and Earth System Sciences*, *14*(10), 1989–2001.
- Neupauer, R. M., Borchers, B., & Wilson, J. L. (2000). Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. *Water Resources Research*, *36*(9), 2469–2475.
- O'Connell, S., & Hou, D. (2015). Resilience: A new consideration for environmental remediation in an era of climate change. *Remediation Journal*, *26*(1), 57–67.
- Oladyshkin, S., de Barros, F. P. J., & Nowak, W. (2012). Global sensitivity analysis: A flexible and efficient framework with an example from stochastic hydrogeology. *Advances in Water Resources*, *37*, 10–22.
- Persson, K., & Destouni, G. (2009). Propagation of water pollution uncertainty and risk from the subsurface to the surface water system of a catchment. *Journal of Hydrology*, *377*(3–4), 434–444.
- Riva, M., Guadagnini, A., & Dell'Oca, A. (2015). Probabilistic assessment of seawater intrusion under multiple sources of uncertainty. *Advances in Water Resources*, *75*, 93–104.
- Rizzo, C. B., & de Barros, F. P. J. (2017). Minimum hydraulic resistance and least resistance path in heterogeneous porous media. *Water Resources Research*, *53*, 8596–8613. <https://doi.org/10.1002/2017WR020418>
- Rubin, Y. (1991). Transport in heterogeneous porous media: Prediction and uncertainty. *Water Resources Research*, *27*(7), 1723–1738.
- Rubin, Y. (2003). *Applied stochastic hydrology*. Oxford: Oxford University Press.
- Rubin, Y., & Bellin, A. (1994). The effects of recharge on flow nonuniformity and macrodispersion. *Water Resources Research*, *30*(4), 939–948.
- Rubin, Y., & Dagan, G. (1987). Stochastic identification of transmissivity and effective recharge in steady groundwater flow: 1. Theory. *Water Resources Research*, *23*(7), 1185–1192.
- Rubin, Y., & Dagan, G. (1992). Conditional estimation of solute travel time in heterogeneous formations: Impact of transmissivity measurements. *Water Resources Research*, *28*(4), 1033–1040.
- Sanchez-Vila, X., & Guadagnini, A. (2005). Travel time and trajectory moments of conservative solutes in three dimensional heterogeneous porous media under mean uniform flow. *Advances in Water Resources*, *28*(5), 429–439.
- Scanlon, B. R., Healy, R. W., & Cook, P. G. (2002). Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, *10*(1), 18–39.
- Schwede, R. L., & Cirpka, O. A. (2010). Stochastic evaluation of mass discharge from pointlike concentration measurements. *Journal of Contaminant Hydrology*, *111*(1), 36–47.
- Selroos, J.-O. (1997a). A stochastic-analytical framework for safety assessment of waste repositories: 1. Theory. *Groundwater*, *35*(3), 468–477.
- Selroos, J.-O. (1997b). A stochastic-analytical framework for safety assessment of waste repositories: 2. Application. *Groundwater*, *35*(5), 775–785.
- Severino, G., Tartakovsky, D., Srinivasan, G., & Viswanathan, H. (2012). Lagrangian models of reactive transport in heterogeneous porous media with uncertain properties, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, *468*(2140), 1154–1174.
- Shapiro, A. M., & Cvetkovic, V. D. (1988). Stochastic analysis of solute arrival time in heterogeneous porous media. *Water Resources Research*, *24*(10), 1711–1718.
- Sobol', I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling and Computing*, *1*, 407–414.
- Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, *55*, 271–280.
- Soga, K., Page, J., & Illangasekare, T. (2004). A review of NAPL source zone remediation efficiency and the mass flux approach. *Journal of Hazardous Materials*, *110*(1–3), 13–27.
- Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety*, *93*, 964–979.
- Tonina, D., & Bellin, A. (2008). Effects of pore-scale dispersion, degree of heterogeneity, sampling size, and source volume on the concentration moments of conservative solutes in heterogeneous formations. *Advances in Water Resources*, *31*(2), 339–354.
- Um, K., Hall, E. J., Katsoulakis, M. A., & Tartakovsky, D. M. (2019). Causality and Bayesian network PDEs for multiscale representations of porous media. *Journal of Computational Physics*, *394*, 658–678. <https://doi.org/10.1016/j.jcp.2019.06.007>
- Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q., & Birkholzer, J. T. (2014). Making sense of global sensitivity analyses. *Computers & Geosciences*, *65*, 84–94.
- Webster, M., Tatang, M. A., & McRae, G. J. (1996). Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model (*Tech. Rep.*) Cambridge, MA: MIT joint program on the science and policy of global change reports series No. 4. MIT.
- Wiener, N. (1938). The homogeneous chaos. *American Journal of Mathematics*, *60*(4), 897–936.
- Xie, Y., Cook, P. G., Simmons, C. T., Partington, D., Crosbie, R., & Batelaan, O. (2018). Uncertainty of groundwater recharge estimated from a water and energy balance model. *Journal of Hydrology*, *561*, 1081–1093.
- Xie, Y., Crosbie, R., Simmons, C. T., Cook, P. G., & Zhang, L. (2019). Uncertainty assessment of spatial-scale groundwater recharge estimated from unsaturated flow modelling. *Hydrogeology Journal*, *27*(1), 379–393.
- Xiu, D., & Karniadakis, G. E. (2002). The Wiener-Askey polynomial chaos for stochastic differential equations. *Journal of Scientific Computing*, *24*(2), 619–644.
- Xiu, D., & Tartakovsky, D. M. (2006). Numerical methods for differential equations in random domains. *Journal of Scientific Computing*, *28*(3), 1167–1185.