AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2015WR017047

Special Section:
The 50th Anniversary of Water Resources Research

Multimodel framework for characterization of transport in porous media

Valentina Ciriello[1], Yaniv Edery[2], Alberto Guadagnini[3,4], and Brian Berkowitz[2]

[1]Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali, Università di Bologna, Bologna, Italy, [2]Department of Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot, Israel, [3]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Milano, Italy, [4]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

Key Points:
- Multimodel interpretation of reactive chemical transport in porous media
- Quantification of parameter uncertainty in reactive transport settings
- Sensitivity-based calibration strategy relying on model proxies via gPC

Supporting Information:
- Supporting Information S1

Correspondence to:
V. Ciriello,
v.ciriello@unibo.it

Citation:
Ciriello, V., Y. Edery, A. Guadagnini, and B. Berkowitz (2015), Multimodel framework for characterization of transport in porous media, *Water Resour. Res.*, *51*, doi:10.1002/2015WR017047.

Received 3 FEB 2015
Accepted 12 APR 2015
Accepted article online 16 APR 2015

Abstract We consider modeling approaches to characterize solute transport in porous media, integrating them into a unique theoretical and experimental framework for model evaluation and data interpretation. To date, development of (conservative and reactive chemical) transport models and formulation of model calibration methods grounded on sensitivity-based collection of measurements have been pursued in parallel. Key questions that remain include: For a given set of measurements, which conceptual picture of the transport processes, as embodied in a mathematical model or models, is most appropriate? What are the most valuable space-time locations for solute concentration measurements, depending on the model selected? How is model parameter uncertainty propagated to model output, and how does this propagation affect model calibration? We address these questions by merging parallel streams of research—model formulation, reduction, calibration, sensitivity analysis, and discrimination—offering our view on an emerging framework that guides (i) selection of an appropriate number and location of time-dependent concentration measurements given a transport model and (ii) assessment (through discrimination criteria) of the relative benefit of applying any particular model from a set of several models. Our strategy is to employ metrics to quantify the relative contribution of each uncertain model parameter to the variability of the model output. We evaluate these metrics through construction of a surrogate (or "meta") transport model that has the additional benefit of enabling sensitivity analysis and model calibration at a highly reduced computational cost. We demonstrate the applicability of this framework, focusing on transport of reactive chemicals in laboratory-scale porous media.

1. Introduction

Mathematical models are typically developed to depict natural phenomena in a manner consistent with the behavior of observable state variables and, eventually, measurable model parameter values. The complexity of key physical, chemical, and biological processes taking place in natural systems hampers our ability to fully and uniquely characterize the system behavior. This is critically evident in settings involving flow and chemical transport processes in the subsurface, where model predictions are compromised by uncertainties arising from the lack of data and attempts to deterministically depict system heterogeneity.

With reference to predicting flow and transport behavior in porous and fractured media, application of models to quantify uncertainty has been studied broadly in the literature. In this context, reliability and accuracy of predictions hinge on both model selection and on model parameter estimation; the former should account, in particular, for components of structural uncertainty [*Carrera and Neuman*, 1986a, 1986b; *Samper and Neuman*, 1989a, 1989b; *National Research Council*, 2001; *Ye et al.*, 2004; *Tartakovsky*, 2007]. Various techniques have been proposed to address key elements and benefits associated with the joint use of multiple interpretive models, in a multimodel approach, as illustrated in the studies of, e.g., *Beven and Binley* [1992], *James and Oldenburg* [1997], *Samper and Molinero* [2000], *Beven and Freer* [2001], and *Gaganis and Smith* [2001]. The main hypothesis underlying these works is that it is possible to exploit model reproduction of past monitored system behaviors to assess the ability of a model to predict unobserved (future) system states. All of these studies require computationally intensive Monte Carlo (MC) simulations, which are typically performed by (randomly) sampling the model parameter space.

An alternative approach combines Bayesian model averaging (BMA) and maximum likelihood (ML) to assess the joint predictive uncertainty of multiple competing models [e.g., *Neuman*, 2002, 2003; *Ye et al*., 2004; *Neuman et al*., 2012]. This maximum likelihood version of the Bayesian Model Averaging (MLBMA) is fully consistent with model parameter estimation methods and can be compatible with both deterministic and stochastic models. As an additional feature, it allows updating of model posterior probabilities (which can be calculated according to Bayes' rule) and parameter estimates at an affordable computational cost, as data become available. Recent applications of MLBMA include quantifying the worth of additional data (before they are actually collected) for geostatistical characterization of aquifer heterogeneity, with (a) multiple variogram models (and eventually measured values) of system hydraulic parameters [*Neuman et al*., 2012; *Lu et al*., 2012] and (b) alternative conceptual models of flow in a randomly heterogeneous aquifer [*Xue et al*., 2014].

These ideas offer a basis to improve our ability to distinguish among different—and sometimes competing—mathematical formulations of flow and transport behavior in natural geological systems. Rigorous model selection criteria (MSC), or model information criteria, have been formulated to (a) rank models or (b) assign probabilistic posterior weights to each model in a multimodel interpretive framework. Among the available theoretical formulations for MSCs, we recall the AIC [*Akaike*, 1974] and AICc [*Hurvich and Tsai*, 1989], as well as the Bayesian criteria BIC [*Schwarz*, 1978] and KIC [*Kashyap*, 1982]. These criteria enable discrimination among alternative models on the basis of their (i) ability to reproduce system behavior and (ii) structural complexity in terms of number of parameters to estimate. ML can be used to estimate model parameters and the associated estimation uncertainty against available data on the state variable (and possibly, but not necessarily, of the parameters themselves). A detailed discussion on the theoretical and practical features of these MSCs is presented by, e.g., *Ye et al*. [2008] to which the reader is referred for additional details.

Application of multimodel approaches in the literature has been devoted mostly to predicting flow behavior in natural porous and/or fractured media. Consideration of transport processes in the above context is more difficult. Various studies have led to the critical conclusion that continuum-scale formulations used to interpret (conservative and reactive) transport experiments are typically dependent on the type and strength of physical and chemical processes occurring at the level of individual pores [see, e.g., *Rashidi et al*., 1996; *Cao and Kitanidis*, 1998; *Gramling et al*., 2002; *Tartakovsky et al*., 2008, 2009; *Edery et al*., 2010; *Sanchez-Vila et al*., 2010, and references therein]. An intrinsic weakness of continuum-scale interpretive formulations is an inadequate, averaged (or upscaled) description of pore-scale processes; under many conditions, such processes cannot simply be homogenized and transferred to a macroscale effective model. With particular reference to reactive transport settings, slow advection zones, where displacement is controlled mainly by diffusion, hinder instantaneous contacts among chemicals species. In this context, incomplete mixing has been observed in laboratory experiments [e.g., *Gramling et al*., 2002], where this effect is most relevant at the lowest concentration levels [e.g., *Rashidi et al*., 1996; *Cao and Kitanidis*, 1998; *Sanchez-Vila et al*., 2010]. Accounting for non-Fickian or anomalous transport observed at a macroscale [see, e.g., *Dagan and Neuman*, 1991; *Cushman and Ginn*, 1993; *Berkowitz and Scher*, 1997; *Haggerty et al*., 2000; *Cirpka and Kitanidis*, 2000; *Sánchez-Vila and Carrera*, 2004; *Berkowitz et al*., 2006; *Rubin et al*., 2012; *Edery et al*., 2014] has been accomplished by various mathematical conceptualizations, including particle-based (or particle tracking, PT) methods [e.g., *Gillespie*, 1977; *Lindenberg and Romero*, 2007; *Srinivasan et al*., 2007; *Palanichamy et al*., 2007; *Yuste et al*., 2008; *Edery et al*., 2010].

Ciriello et al. [2013a] first explored the applicability of MSC and global sensitivity analysis (GSA) to compare among different models of conservative transport at the laboratory scale. The use of GSA enabled these authors to identify, within an available data set, the subsets of observations carrying the highest information content to characterize the sensitivity of a model response to its uncertain input parameters. The modeling strategy proposed by these authors was then based on the use of data collected at such most sensitive (in space and time) locations to calibrate a given transport model. Identification of subsets of observations which, among the complete set of available data, are most informative for model parameter estimation is then critical, as evidenced by, e.g., *Tiedeman et al*. [2003, 2004] and *Barth and Hill* [2005a, 2005b]. *Ciriello et al*. [2013a] observed that, in general, the composition of such optimal data (sub-)sets depends on the selected model. Hence, the key point of the strategy introduced by *Ciriello et al*. [2013a] relies on combining information provided by (a) MSC and (b) sensitivity-based model calibration to

compare alternative models. This approach relies on global sensitivity analysis (GSA) [e.g., *Sobol*, 1993; *Saltelli et al*., 2000], which allows simultaneous consideration of all possible (and unknown) parameter values within their likely ranges of variability. This element is the main distinction between GSA and local sensitivity-based approaches and provides remarkable advantages in the context of model calibration [e.g., *Fajraoui et al*., 2011; *Ciriello et al*., 2013a, and references therein].

GSA and model calibration are typically computationally demanding, because they require a large number of runs with each model. Computational efforts can be reduced by replacing a given system model by a simplified (reduced-order) model that renders an acceptable approximation of the target system response. These reduced-order models are typically referred to as model proxies, and/or lower complexity, or surrogate models. In this context, the Polynomial Chaos Expansion (PCE) has recently gained increasing popularity (in the context of subsurface flow and transport, over many scales, see e.g., *Lin et al*. [2010], *Oladyshkin and Nowak* [2012], *Oladyshkin et al*. [2012, 2013], *Ciriello and Di Federico* [2013], *Ciriello et al*. [2013b], *Laloy et al*. [2013], *Ashraf et al*. [2013], *Sun et al*. [2013], *Formaggia et al*. [2013], *Liao and Zhang* [2013, 2014], *Porta et al*. [2014], *Rajabi et al*. [2015], *Esfandiar et al*. [2015], and *Riva et al*. [2015, and references therein]). In this approach, the (finite-variance) response surface of a given model is projected onto a probabilistic space, generated by a basis of polynomials orthonormal with respect to the joint probability density function (pdf) of the uncertain model parameters [*Wiener*, 1938; *Ghanem and Spanos*, 1991]. Once the PCE approximation of the model response has been constructed, the first two moments (i.e., ensemble mean and variance) and global sensitivity measures associated with the target state variable may be defined analytically as functions of the coefficients of the expansion. Approaches based on PCE can be considered as well established to perform GSA [*Sudret*, 2008, and references therein] and have been tested under several flow and transport scenarios in porous media [e.g., *Fajraoui et al*., 2011; *Formaggia et al*., 2013; *Ciriello and Di Federico*, 2013; *Ciriello et al*., 2013b; *Porta et al*., 2014].

In this work, we extend the methodology proposed by *Ciriello et al*. [2013a] to consideration of transport of reactive chemicals in laboratory-scale porous media. Reactive transport is an emerging field of research giving rise to various models; comparison among models and parameter relationships within each model provides insight into mechanisms of transport and reactions. We consider two representative mathematical models. The first one employs a continuum approach in which the macroscopic system behavior is described by means of a typical Advection-Dispersion-Reaction equation (ADRE), which assumes Fickian transport, with continuum-scale chemical reactions employing a kinetic reaction rate. This rate is proportional to the local reactant concentration, with a time-dependent coefficient being employed to account for pore-scale mixing [*Sanchez-Vila et al*., 2010]. The second model is based on a particle-tracking approach within the continuous time random walk (CTRW) framework [*Edery et al*., 2010], which accounts for non-Fickian transport, with chemical reactions based on a probability of local particle-particle interactions (where particles are representative of amounts of dissolved chemical species). In this scheme, chemical transport and the degree of mixing are modeled using an appropriate probability density function of particle transition rates [*Berkowitz et al*., 2006]. Given the computational cost associated with these models, we merge the methodology of *Ciriello et al*. [2013a] with that of *Porta et al*. [2014], *Esfandiar et al*. [2015], and *Riva et al*. [2015], where the construction of a PCE-based surrogate model serves as a basis for GSA and model calibration.

We discuss each step involved in our strategy and explore the potential of the approach by (re-) analysis of the reactive transport experiments of *Gramling et al*. [2002], where a simple bimolecular reaction is considered. We show how parallel streams of research (modeling, calibration, GSA, and MSC concepts) can be blended seamlessly into a unique framework that guides the selection of sensitivity-based sets of concentration measurements to calibrate subsurface transport models. Moreover, the framework enables identification (through MSC) of the relative advantages of a given model within a collection of available models.

2. Transport Modeling in Porous Media

While experimental analyses of conservative solute transport in laboratory-scale porous media are abundant in the literature [e.g., *Silliman and Simpson*, 1987; *Silliman et al*., 1987; *Rashidi et al*., 1996; *Berkowitz et al*., 2000; *Raje and Kapoor*, 2000; *Gramling et al*., 2002; *Jose and Cirpka*, 2004; *Zinn et al*., 2004; *Rahman et al*., 2005], these are chiefly focused on the use of observed data to calibrate a single mathematical model.

Ciriello et al. [2013a] analyzed conservative transport experiments documented by *Gramling et al*. [2002] in a multimodel framework. For this purpose, three widely employed transport models were selected: (i) the advection-dispersion equation (ADE), (ii) a dual-porosity model, and (iii) CTRW. The main idea introduced in *Ciriello et al*. [2013a] is the application of formal model selection criteria, not necessarily relying on the entire available data set, but rather on subsets of observations that are identified through GSA as being the most sensitive for a given model. As such, model parameters are estimated following a variance-based sensitivity analysis, so that one can assess the ability of each model to interpret subsets of observations that carry high levels of information on system variability. A final analysis step, where the resulting calibrated model predictions are compared against an augmented data set, provides information on the ability of each model to interpret the particular dynamic processes. *Ciriello et al*. [2013a] concluded that the ADE ranked as the best model when approximating a subset of most sensitive observations, the latter being determined through GSA. However, comparison of predictions from calibrated models against an augmented data set showed the CTRW to offer the best overall performance. These results suggest that the sensitive observations related to CTRW contribute to the model predictive capability more than the most sensitive counterparts associated with the other tested models.

Conceptualizations of solute transport at a continuum scale are typically grounded on the assumption that spreading can be described through a Fickian analogy. In this case, a (spatially and temporally constant) set of dispersivity coefficients is estimated through observed solute concentrations. On the other hand, several laboratory experiments document the occurrence of anomalous transport, the latter terminology typically being used to identify system behavior departing from the Fickian analog [e.g., *Silliman and Simpson*, 1987; *Berkowitz and Scher*, 1995; *Levy and Berkowitz*, 2003; *Bromly and Hinz*, 2004; *Kuntz et al*., 2011]. The occurrence of spatially distributed slow advection zones, which are visited by solute particles and within which diffusive processes strongly influence transport, tends to markedly affect chemical migration. These features are critical both for conservative transport and in the presence of reacting chemicals. With reference to reactive transport, these zones constitute a limiting factor that controls the way reactants come into contact in the system [e.g., *Raje and Kapoor*, 2000; *Gramling et al*., 2002]. As a consequence, it is critical to properly account for the effects of incomplete mixing of reactants. Use of continuum-scale models and failure to properly account for these dynamics might lead to severe over- (or under-) estimation of solute concentrations, reaction rates, and reaction products. Typically, these effects are most significant for the lowest concentration levels [e.g., *Rashidi et al*., 1996; *Cao and Kitanidis*, 1998; *Sanchez-Vila et al*., 2010].

In this context, it is clear that one can consider several alternative modeling formulations to interpret observed transport scenarios. Once the transport scenario and a set of models are selected, the key parameters linked to each of the models must be identified. Complexity of model calibration generally increases with the number of parameters. As stated in section 1, GSA methodologies provide valuable approaches to quantify the relative importance of each model parameter. This information, in turn, enables one to assess the degree of reliability that can be associated with each estimated parameter. In this sense, the critical importance of GSA is also clear in the design phase of an experiment [*Ciriello et al*., 2013a].

Following calibration, discrimination/ranking of the selected models can be performed by employing criteria such as those referenced in section 1. These criteria allow evaluation of the model performance in a relative sense, on the basis of the agreement between model predictions and hydrologic observations through maximum likelihood (ML) estimate of model parameters.

We consider here the data set of *Gramling et al*. [2002], which documents results of laboratory-scale, conservative, and reactive transport experiments. As mentioned above, the conservative experiment was examined in a multimodel framework by *Ciriello et al*. [2013a]. The reactive transport experiment of *Gramling et al*. [2002] was performed in a rectangular glass chamber (36 cm long) filled with crushed cryolite grains (average diameter 0.13 cm). The porous material (with an estimated porosity equal to 0.36) was first saturated with $EDTA^{4-}$ and the reactant $CuSO_4$ was then injected into the system. A bimolecular irreversible reaction among the chemical species Cu^{2+} and $EDTA^{4-}$ occurred in the pore space, yielding $CuEDTA^{2-}$ as a reaction product. The concentration of the reaction product was reported at 4 times ($\tau_1 = 619$ s; $\tau_2 = 916$ s; $\tau_3 = 1114$ s; $\tau_4 = 1510$ s), for a flow rate of 2.67 mL/min. This is the data set referred to in the remainder of the paper. We focus on the way two existing models, which account for incomplete mixing, can be employed in a multiple-model analysis framework. One model (hereinafter ADRE-KM) is the empirical ADRE formulation proposed by *Sanchez-Vila et al*. [2010], according to which, as stated in section 1, a kinetic

model (KM) is employed; this involves a reaction rate taken to be linearly proportional to the local reactant concentration [*Hering and Morel*, 1988]. The second model is based on a particle tracking form of CTRW, as presented by *Edery et al*. [2010]; we hereafter refer to this model as CTRW-PT. A summary of the key points associated with the mathematical framework underlying these two models is provided as supporting information.

The following section is devoted to a complete illustration of the methodology we propose for the consistent analysis (and design) of model-driven chemical transport experiments. Key questions that can be addressed include: (a) which conceptual picture of the (reactive) transport processes, as embodied in a mathematical model (or models), is most effective to interpret a given set of observations? (b) What space and time locations for chemical concentration measurements provide the most valuable information, depending on the choice of model? and (c) How is model parameter uncertainty propagated to model output, and how does this propagation affect model calibration? We employ an approach based on the seamless fusion of model formulation, calibration, and discrimination, as well as rigorous sensitivity analysis, to define a coherent framework for the characterization of chemical transport in porous media.

3. Methodological Framework

Our methodological framework relies on Polynomial Chaos Expansion (PCE) theory, which is employed here as a basis to construct proxies of given transport models whose parameters are affected by uncertainty. The theory underlying PCE was introduced by *Wiener* [1938] to characterize Gaussian processes. Nevertheless, adoption of this approach to engineering applications is quite recent, being due to *Ghanem and Spanos* [1991] in the context of the stochastic finite element method (SFEM) [see also *Matthies*, 2008, and references therein]. The main idea of the approach relies on projection of the selected model response (i.e., the state variable of interest) onto a probabilistic space (which is usually termed *Polynomial Chaos*) to derive a polynomial approximation of the response surface associated with the model. *Xiu and Karniadakis* [2002] extended the approach to non-Gaussian distributions by introducing the Askey family of hypergeometric polynomials, on which the so-called *generalized Polynomial Chaos* (gPC) is based.

This methodology has received increasing attention in the recent years due to its versatility of use in characterizing diverse physical problems. Recent examples of application of the gPC in the hydrogeological context include, e.g., GSA, uncertainty quantification, and Bayesian inference [*Ciriello et al*., 2013b; *Laloy et al*., 2013; *Ashraf et al*., 2013; *Sun et al*., 2013; *Liao and Zhang*, 2013, 2014; *Porta et al*., 2014; *Rajabi et al*., 2015; *Riva et al*., 2015]. A critical limitation of the approach relates to its ability to efficiently treat a large number of uncertain model parameters, i.e., if the number of model parameters subject to uncertainty increases together with the variance and the complexity of the model response surface (representing the given model output), then the computational cost required to construct a gPC approximation may be significant [e.g., *Crestaux et al*., 2009]. Nevertheless, within the range of theoretical applicability of the gPC, the method has generally provided accurate results even for complex modeling scenarios [e.g., *Ciriello et al*., 2013b; *Porta et al*., 2014; *Esfandiar et al*., 2015; *Riva et al*., 2015].

As mentioned in sections 1 and 2, *Ciriello et al*. [2013a] suggested the adoption of the gPC approach as the basis of a strategy to combine sensitivity-based model calibration and model selection criteria, to compare alternative conservative transport models. Here we extend this strategy to more complex transport scenarios in porous domains, by considering reactive behavior of solutes. Given the computational cost associated with reactive transport models, we propose to employ the gPC technique not only for GSA, as in *Ciriello et al*. [2013a], but also in the context of model calibration, along the lines of *Esfandiar et al*. [2015]. Figure 1 depicts a sketch of the workflow associated with the methodology we propose. In the following, we describe the details of the theoretical framework of our methodology. We do so upon recalling some of the basic concepts illustrated by *Ciriello et al*. [2013a].

3.1. Generalized Polynomial Chaos Representation of a System Model

The response surface of a selected model, $Y=f(x,t,\mathbf{X})$, can be approximated by resorting to the gPC technique if Y has finite variance. Here vector $\mathbf{X}$ collects the model parameters whose values are considered as uncertain. The gPC approximation, $\tilde{Y}$, of Y is given by a finite series of polynomials, orthogonal with respect

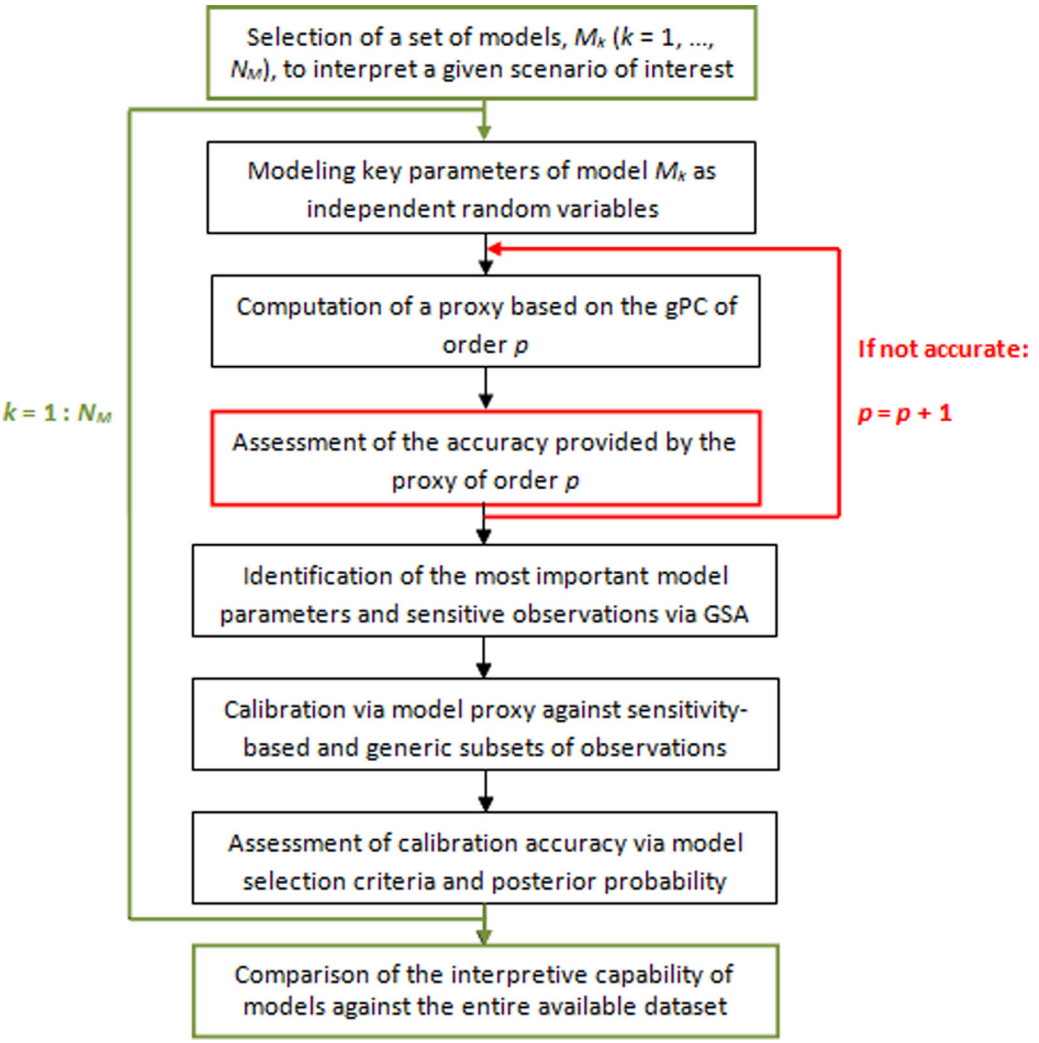

Figure 1. Sketch of the workflow associated with the proposed methodology.

to the joint probability density function of the parameters collected in **X** [*Wiener*, 1938; *Ghanem and Spanos*, 1991], i.e.,

$$Y(x,t,\mathbf{X}) \cong \tilde{Y}(x,t,\mathbf{\Omega}) = \sum_{j=0}^{P-1} a_j(x,t)\Psi_j(\mathbf{\Omega}). \tag{1}$$

In (1), $P=(M+p)!/(M!\ p!)$ represents the number of terms included in the expansion, which depends on the largest degree retained in the expansion, p, and on the number of model parameters, M; Ψ_j denotes a multivariate orthogonal polynomial and a_j is a deterministic coefficient of the expansion. Vector $\mathbf{\Omega}$ contains a set of independent random variables whose distribution is linked to the selected polynomial basis [*Xiu and Karniadakis*, 2002]. Variables collected in $\mathbf{\Omega}$ are related to the corresponding variables in **X** through an isoprobabilistic transformation [*Sudret*, 2008]. Computation of the deterministic set of coefficients a_j in (1) can be performed by means of a nonintrusive regression-based approach. The optimum set of regression points in the random parameter space is determined on the basis of the probabilistic collocation method [*Webster et al.*, 1996; *Huang et al.*, 2007]. This strategy employs the roots of the polynomial of 1 order higher than p, to ensure appropriate sampling of the region associated with largest probability in the parameter space, according to the adopted distributions of the parameters. Combinations of the roots of the polynomial of order $(p + 1)$ are taken as the optimum set of regression (or collocation) points for the polynomial of

order p [*Webster et al.*, 1996; *Sudret*, 2008; *Ciriello et al.*, 2013a, b]. Minimization, with respect to the unknown coefficients in (1), of the variance of the difference between the solution given by the original model and the surrogate model response calculated at the collocation points in the parameter space renders the values of a_j ($j = 0, 1, \ldots, P-1$).

Once the gPC representation is defined, at each space-time location (x,t) of interest, global sensitivity measures such as the Sobol indices [*Sobol*, 1993]

$$S_{\boldsymbol{\delta}} = \sum_{\boldsymbol{\alpha} \in \phi_{\boldsymbol{\delta}}} a_{\boldsymbol{\alpha}}^2 E\left[\Psi_{\boldsymbol{\alpha}}^2\right] / V\left(\tilde{Y}\right) \tag{2}$$

can be analytically determined. In (2), $\phi_{\boldsymbol{\delta}} = \left\{\boldsymbol{\alpha} \in N^M : k \in \boldsymbol{\delta} \Longleftrightarrow \alpha_k \neq 0, k = 1, \ldots, M\right\}$ denotes a set of multi-indices depending exactly on the subset of random model parameters identified by $\boldsymbol{\delta} = \{i_1, \ldots, i_s\}$, with $s = 1, \ldots, M$. The computation of $E\left[\Psi_{\alpha}^2\right]$ can be performed following, e.g., *Abramowitz and Stegun* [1970]. Terms a_0 and $V\left(\tilde{Y}\right) = Var\left[\sum_{j=0}^{P-1} a_j \Psi_j(\boldsymbol{\Omega})\right] = \left[\sum_{j=1}^{P-1} a_j^2 E\left[\Psi_j^2(\boldsymbol{\Omega})\right]\right]$ in (2), respectively, represent the mean and the total variance of $\tilde{Y}$ [*Sudret*, 2008]. Sobol indices defined in (2), considering all the possible sets $\boldsymbol{\delta}$ of model parameters, sum up to unity. For each model parameter, one can use (2) to compute (i) the principal sensitivity index (which is associated with $s = 1$), representing the reduction in the variance of the response if that parameter is not uncertain, and (ii) the total sensitivity index, related to the expected residual variance of the response if only the selected parameter is uncertain [*Sobol*, 1993; *Saltelli et al.*, 2000]. Note that if the principal and total sensitivity indices show similar values, then higher-order effects accounting for interactions among parameters are negligible [*Sobol*, 1993].

3.2. Sensitivity-Based Model Calibration

A sensitivity-based model calibration requires computation of the sensitivity indices (2) at each space-time location at which an observation of the system response is available, i.e., $(x,t)_i^{obs}$, $1 \leq i \leq N$, N being the number of available observations. This allows identification of the observations corresponding to the highest information content for a given selected model [*Ciriello et al.*, 2013a]. Let $\mathbf{Y}^{obs} = \left[Y_1^{obs}, \ldots Y_N^{obs}\right]$ be the vector including the complete set of available observations and $\mathbf{B}_Y$ the covariance matrix of measurement errors, here considered to be diagonal with nonzero terms equal to the observation error variance σ_i^2 [*Carrera and Neuman*, 1986a, 1986b]. Performing model calibration in a maximum likelihood (ML) framework requires calculation of model outputs at each $(x,t)_i^{abs}$ for a number of times until one reaches a minimum of the target objective function. As this can be computationally demanding, especially in the case of complex models, one can drastically reduce the computational cost by replacing the full system model $Y = f(x,t,\mathbf{X})$ with its gPC approximation (or surrogate model) evaluated at each $(x,t)_i^{obs}$. The ML estimate $\hat{\mathbf{X}}$ of the vector of the M model parameters can be obtained by minimizing, with respect to $\mathbf{X}$, the negative log likelihood criterion [*Carrera and Neuman*, 1986a, 1986b; *Bianchi Janetti et al.*, 2012, and references therein]

$$NLL = \sum_{i=1}^{N} \frac{J_i}{\sigma_i^2} + \ln|\mathbf{B}_Y| + N \ln(2\pi), \tag{3}$$

where $J_i = \left(Y_i^{obs} - \tilde{Y}_i\right)^2$, $\tilde{Y}_i$ being the output of the surrogate model at $(x,t)_i^{obs}$, i.e.,

$$\tilde{Y}_i = \sum_{j=0}^{P-1} a_j (x,t)_i^{obs} \Psi_j(\boldsymbol{\Omega}). \tag{4}$$

Following section 3.1, the vector $\tilde{\mathbf{Y}} = \left[\tilde{Y}_1, \ldots \tilde{Y}_N\right]$ of the surrogate model outputs at locations where measurements are available may be obtained as

$$\tilde{\mathbf{Y}} = \begin{bmatrix} \tilde{Y}_1 \\ \tilde{Y}_2 \\ \ldots \\ \tilde{Y}_N \end{bmatrix} = \mathbf{A}^{obs} \cdot \boldsymbol{\Psi} = \begin{bmatrix} a_0^1 & a_1^1 & \ldots & a_{P-2}^1 & a_{P-1}^1 \\ a_0^2 & a_1^2 & \ldots & a_{P-2}^2 & a_{P-1}^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_0^N & a_1^N & \ldots & a_{P-2}^N & a_{P-1}^N \end{bmatrix} \cdot \begin{bmatrix} \Psi_0(\boldsymbol{\Omega}) \\ \Psi_1(\boldsymbol{\Omega}) \\ \ldots \\ \Psi_{P-2}(\boldsymbol{\Omega}) \\ \Psi_{P-1}(\boldsymbol{\Omega}) \end{bmatrix}, \tag{5}$$

where we denote $a_j^i = a_j(x,t)_i^{obs}$ for brevity. Note that, once the surrogate model has been constructed, calculation of (5) does not require additional computational cost, because the terms composing the matrix of coefficients, $\mathbf{A}^{obs}$, have been already computed. We further note that the form of the approximation is the same at each $(x,t)_i^{obs}$, i.e., the vector $\mathbf{\Psi}$ does not depend on the specific space-time location considered, thus rendering the approach highly convenient.

Minimization of (3) using, e.g., the iterative Levenberg-Marquardt algorithm as embedded in the code PEST [*Doherty*, 2002], may then be easily performed at negligible computational cost, by taking advantage of the observation that the constructed surrogate model is expressed in a simple polynomial form. The computational effort involved with model calibration may be further decreased by considering only reduced sets of observations that correspond to the largest sensitivity index values associated with model parameters. Given a sensitivity-based subset of observations defined in this way, with $B < N$ denoting dimension, the system (5) reduces to

$$\tilde{\mathbf{Y}}_{N\times 1} = \mathbf{A}^{obs}_{N\times P} \cdot \mathbf{\Psi}_{P\times 1} \rightarrow \tilde{\mathbf{Y}}_{B\times 1} = \mathbf{A}^{SBobs}_{B\times P} \cdot \mathbf{\Psi}_{P\times 1}. \tag{6}$$

We indicate space-time locations corresponding to the observations included in the sensitivity-based subset as $(x,t)_i^{SB_{obs}}$; the matrix $\mathbf{A}^{SBobs}$ includes the gPC coefficients computed at each $(x,t)_i^{SB_{obs}}$.

3.3. Comparison of Alternative Interpretive Models

When multiple models are considered, one may repeat the steps described in sections 3.1–3.3 for each model formulation. In this case, once the parameters associated with each model are calibrated, either on the basis of the entire set of available observations or by considering only sensitivity-based subsets, the alternative formulations can be ranked by various criteria [e.g., *Neuman*, 2003; *Ye et al*., 2004, 2008; *Bianchi Janetti et al*., 2012, and references therein], including

$$AIC = NLL + 2M, \tag{7}$$

$$AIC_c = NLL + 2M + \frac{2M(M+1)}{N-M-1}, \tag{8}$$

$$KIC = NLL - M\ \ln(2\pi) - \ln|\mathbf{Q}|, \tag{9}$$

where the *Akaike* information criterion, *AIC*, is due to *Akaike* [1974], AIC_c to *Hurvich and Tsai* [1989], and *KIC* to *Kashyap* [1982]. Note that the lowest value of a given model identification criterion indicates the most favored model (according to the criterion itself) at the expenses of the remaining models. In (9), $\mathbf{Q}$ represents the Cramer-Rao lower-bound approximation for the covariance matrix of parameter estimates, i.e., the inverse expected Fisher information matrix (see *Ye et al*. [2008] for details). In a multimodel analysis, model discrimination criteria (7)–(9) can also be employed to assign posterior probability weights to the diverse model considered. The posterior probability related to model M_k ($k = 1, \ldots, N_M$, with N_M the number of available process models) is calculated as [*Ye et al*., 2008]

$$p(M_k|\mathbf{Y}^{obs}) = \frac{\exp\left(-\frac{1}{2}\Delta IC_k\right)p(M_k)}{\sum_{i=1}^{N_M}\exp\left(-\frac{1}{2}\Delta IC_i\right)p(M_i)}. \tag{10}$$

Here $\Delta IC_k = IC_k - IC_{min}$, with IC_k being either *AIC* (7), AIC_c (8), or *KIC* (9) and $IC_{min} = \min\{IC_k\}$ its minimum value across the models considered; $p(M_k)$ is the prior probability associated with each model. If no prior information is available, then $p(M_k) = 1/N_M$, so that all models are associated with equal prior probability. Equation (10) can also be employed in the context of Bayesian model averaging to weight each model prediction.

Note that model selection criteria and the posterior probability provide a rigorous way to compare alternative formulations based on the sets of observations selected for model calibration. These sets generally contain different observations for each model when they are constructed by the sensitivity-based approach described in section 3.2. Note that from a collection of tested models, a model identified as "best"—based on its ability to approximate the corresponding set of most sensitive observations—might not prove to be "best" when its predictive ability is then tested against an augmented data set [*Ciriello et al*., 2013a].

Table 1. Model Parameters and Adopted Sampling Distributions

Parameter	Model	Distribution	Mean	Standard Deviation
Effective velocity (v)	ADRE-KM	Normal	1.21×10^{-4} m/s	1.00×10^{-6} m/s
Longitudinal dispersivity (α_L)	ADRE-KM	Lognormal	1.45×10^{-3} m	4.50×10^{-4} m
Coefficient of the reaction rate (β_0)	ADRE-KM	Normal	240 L/(mol s^{1-m})	48 L/(mol s^{1-m})
Exponent of the reaction rate (m)	ADRE-KM	Normal	9.3×10^{-1}	1.86×10^{-2}
Transport velocity (v_ψ)	CTRW-PT	Normal	1.21×10^{-4} m/s	1.00×10^{-6} m/s
Generalized dispersion coefficient (D_ψ)	CTRW-PT	Normal	1.75×10^{-7} m^2/s	5.44×10^{-8} m^2/s
Exponent of TPL distribution (β)	CTRW-PT	Normal	1.96	9.80×10^{-2}
Cutoff time (t_2)	CTRW-PT	Lognormal	10^5 s	5.0×10^4 s
Reaction radius (R)	CTRW-PT	Normal	3.0×10^{-3} m	5.0×10^{-4} m

4. Application and Discussion

Here we apply the methodological framework illustrated in section 3 to the set of data from the reactive transport experiment (mentioned in section 2) [*Gramling et al*., 2002]. The data set comprises depth-averaged concentration measurements of reaction product in a laboratory flow cell, measured at 4 times ($\tau_1 = 619$ s; $\tau_2 = 916$ s; $\tau_3 = 1114$ s; $\tau_4 = 1510$ s) from the beginning of the experiment.

The first step of our methodology, which constitutes the basis for a sensitivity-based model calibration, relies on computation of Sobol indices (see section 3.1). These indices are associated with each model parameter at the space-time locations at which observations are available. The parameters of each model we analyze are considered as the only sources of uncertainty. As no information or theoretical evidence indicates parameter correlation, the parameters are treated as statistically independent random variables, described by the probability distributions listed in Table 1. In principle, parameter correlations can be considered [see, e.g., *Li et al*., 2009] if marginal probability density functions of the uncertain parameters and the associated correlation matrix are known. Four flow and transport parameters need to be estimated for the ADRE-KM (see supporting information): the effective velocity, v; the longitudinal dispersivity, α_L; and the coefficient and exponent of the reaction rate, denoted as β_0 and m, respectively. Five parameters are considered for the CTRW-PT model (see supporting information): the transport velocity, v_ψ; the generalized dispersion coefficient, D_ψ; the exponent of the TPL distribution, β; the cutoff time, t_2; and the reaction radius, R. Each model parameter is associated with (i) a mean value, assessed on the basis of calibration results obtained by *Gramling et al*. [2002] and preliminary calibration against the complete available data set, and (ii) a value of standard deviation adequate to ensure that relatively wide intervals in the parameter space are explored (see Table 1). We assume that most of the parameters are described by normal distributions. A lognormal density is preferred for parameters whose selected range of variability could entail unphysical negative values. Average values associated with such distributions are selected on the basis of preliminary manual model calibration. The variance associated with each parameter distribution is selected to allow exploration of a wide range of parameter combinations during model calibration. In this context, the GSA results are only minimally affected by the selected functional format of the parameter distributions [see also *Ciriello et al*., 2013a].

Considering the probability distributions adopted for model parameters, Hermite polynomials are selected as a basis to generate a second-order approximation of the model response surface [*Xiu and Karniadakis*, 2002]. As shown in section 3, adoption of such polynomials enables us to derive analytically the Sobol indices employed in the GSA [e.g., *Sudret*, 2008]. We remark that selection of the order of the gPC is a critical step in applying the approach. Note that (a) increasing the order of the gPC might improve the accuracy of the surrogate model but yields a considerably increased number of collocation points, with a decrease in computational efficiency; and (b) assessing the accuracy associated with a given gPC order requires implementing the solution at higher orders. Computation of the coefficients of the gPC requires, at the order we consider, a number of full model runs equal to (i) 15 for the ADRE-KM, and (ii) 21 for CTRW-PT. This allows considerable efficiency in terms of computational cost. Construction of a gPC of order 3 would require a number of full model runs equal to (i) 71 for the ADRE-KM and (ii) 116 for CTRW-PT. We verified that a second-order approximation provides a good representation of the response surface associated with the full model solutions. We do so by comparing the results obtained through the proxy and original (full) system models with the same input parameter sets. The latter were randomly selected in the parameter space

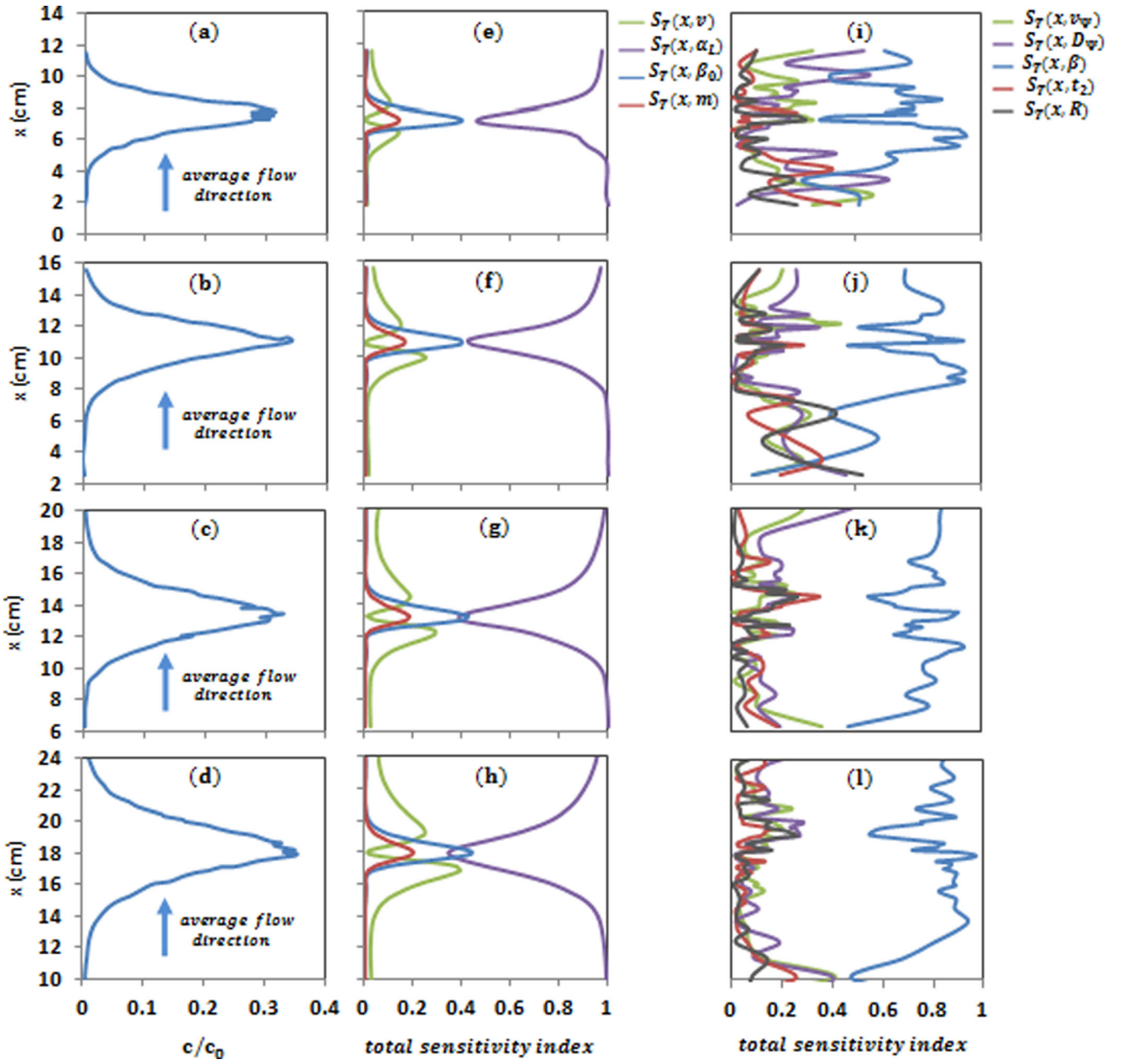

Figure 2. (a–d) Spatial distributions of normalized concentrations, c/c_0 (with c_0 the inflow concentration of reactant $CuSO_4$ at the upstream boundary at $x = 0$ and the initial concentration of reactant $EDTA^{4-}$ inside the domain) of the reaction product $CuEDTA^{2-}$ documented by *Gramling et al.* [2002]; total sensitivity indices resulting from the (GSA of the) selected models: (e–h) ADRE-KM; and (i–l) CTRW-PT. Note that the vertical axes denote the main direction of transport in the flow cell, as denoted by the arrow in Figures 2a–2d, while the horizontal axes denote relative concentration or sensitivity index. Each row $i = 1, \ldots, 4$ depicts results associated with a given time, corresponding to observation times τ_i reported by *Gramling et al.* [2002], i.e., $\tau_1 = 619$ s; $\tau_2 = 916$ s; $\tau_3 = 1114$ s; $\tau_4 = 1510$ s.

at locations not coinciding with those of the collocation points (not shown). Below, we provide additional comments on this point.

Figure 2 depicts the available spatial distributions of normalized concentrations, c/c_0 (c_0 being the inflow concentration of reactant $CuSO_4$ at the upstream boundary at $x = 0$ and the initial concentration of reactant $EDTA^{4-}$ inside the domain) of the reaction product $CuEDTA^{2-}$ together with the values calculated for the total sensitivity indices associated with each model parameters at all space-time locations at which observations are available. Table 2 lists the average values of the total sensitivity indices associated with the

Table 2. Average Values of the Total Sensitivity Indices Associated With the Two Models and Calculated on the Complete Set of Available Concentration Data

Parameter	Model	Mean Values of S_T
v	ADRE-KM	0.110
α_L	ADRE-KM	0.751
β_0	ADRE-KM	0.104
m	ADRE-KM	0.043
v_ψ	CTRW-PT	0.129
D_ψ	CTRW-PT	0.161
β	CTRW-PT	0.741
t_2	CTRW-PT	0.091
R	CTRW-PT	0.092

parameters of the two selected models. As a preliminary conclusion, these results highlight that α_L and β display the highest sensitivity along the flow cell, respectively, for ADRE-KM and CTRW-PT. This is in accord with the observation that α_L and β, respectively, drive the spread (in the ADRE-KM formulation) and the overall dispersive nature of transport, particularly the tails (in the CTRW-PT formulation) of the reactants and reaction product, thus controlling the way reactants come into contact and react.

With reference to the ADRE-KM model parameters, one can conclude that:

1. The two parameters linked to the reaction rate, i.e., β_0 and m, exhibit strongest influence at spatial locations corresponding to the peak of concentration at the various observation times; their rate of decrease in space is qualitatively similar to that of the relative concentration profiles.
2. Longitudinal dispersivity, α_L, is important along the entire flow cell, albeit with significantly diminished influence at locations where the highest concentrations are recorded. This result is consistent with findings of *Ciriello et al*. [2013a], in the context of conservative transport analysis, and reflects the fact that dispersivity drives spreading of concentration around the solute center of mass.
3. The total Sobol index associated with fluid effective velocity displays two peaks at spatial locations approximately corresponding to change of concavity of concentration profiles. This result corresponds to the observation that displacement of solute center of mass is typically governed by advective processes; these spatial locations are relatively close to the center of mass of the two reactants which are located upstream and downstream of the peak displayed by the reaction product.
4. The sum of the total sensitivity indices associated with the four parameters of the model is about unity at each space-time location, indicating negligible second-order effects due to joint effects of parameters.

With reference to the CTRW-PT model parameters, Figure 2 reveals that:

1. The exponent of the TPL distribution, β, is the most influential parameter along almost the entire length of the column, its relative importance increasing with time. A qualitatively similar result was obtained by *Ciriello et al*. [2013a] in their analysis of conservative transport experiments. This implies that obtaining accurate estimates of β is key to the proper calibration of CTRW-PT, while the influence of the actual values of the remaining model parameters on the variability of the model results tends to decrease significantly with time. This result is consistent with the physical meaning of β, which governs, in particular, the extreme values of the distribution of concentration fluctuations and, ultimately, the way reactants come into contact and react.
2. The behavior of the total sensitivity indices associated with the other four model parameters indicates that the magnitude of their contribution to the variance of the model response (i.e., the reaction product concentration) ranges in space between approximately 0–40%.
3. The sum of the total sensitivity indices is larger than unity at a non-negligible number of space-time locations. This suggests that the GSA of this model requires an accounting of the combined effect of parameter pairs on the variance of the model response. The relevance of this effect can be quantified through the second-order Sobol indices. For this specific case study, the contribution to the variance of the response of each second-order Sobol index, averaged over all the space-time locations of the data set, was found to vary between 1.5 and 3.5% (not shown).

Our methodology relies on GSA to identify subsets of observations containing sufficient information for model calibration. In this context, it is the local magnitude of the Sobol indices which provides a clear indication of the space-time locations at which a model is most sensitive to its parameter values. Therefore, we classify as sensitive the observations corresponding to the largest values of the total sensitivity indices.

Table 3. Calibration Sets for the Three Selected Models

Subset	Description	Number of Observations
1	Observations from concentration profile at τ_1 and τ_3	86
2	Most sensitive observations from concentration profiles at τ_1 and τ_3	20

We investigate the potential usefulness and robustness of a sensitivity-based model parameter estimation approach by calibrating the two models with two subsets of observations (Table 3). The first subset (termed hereinafter as subset 1) is composed of 86 observations uniformly distributed in space and selected from the relative concentration curves measured at times τ_1 and τ_3. The second subset (hereinafter termed as subset 2) is composed of the 20 most sensitive observations identified among those belonging to subset 1. This subset comprises about 25% of the elements of subset 1. Note that subset 1 is not based on GSA and is the same for both models. The entries of subset 2 in general differ for the two models, because of the diverse space-time distributions of the total Sobol indices associated with the models.

We then estimate the parameters of the selected models through minimization of the negative log likelihood (NLL) criterion for each data subset. We do so by employing the iterative Levenberg-Marquardt algorithm as embedded in the code PEST [*Doherty*, 2002]. As illustrated in section 3, this step relies on calibration of the model proxies resulting from the application of the gPC technique, already employed for GSA assessment in the previous step. This enables us to markedly reduce computational costs associated with model calibration, which may be significant for reactive transport problems.

Table 4 lists the results of model calibrations against subsets 1 and 2. Results are listed in terms of the estimated value of each parameter, denoted as C in Table 4, Gaussian 95% confidence intervals (in parentheses, computed on the basis of the same linearity assumption adopted in PEST for parameter estimation [*Doherty*, 2002]), and the ratio, $\Re$, of the difference between the upper (U) and lower (L) limits identifying the 95% uncertainty bounds and the estimated value, C, i.e., $\Re = (U - L)/C$. We observe that parameter estimates are generally associated with small values of $\Re$, exceptions being given by parameters β_0 and m of the ADRE-KM. The quality of these parameter estimates appears to be relatively poor. This is consistent with the observation that the model is sensitive to these parameters only in a relatively small range of spatial locations, at which dispersivity, α_L, still has a primary role. Hence, it can be concluded that these parameters do not affect significantly the model performance against both calibration subsets [*Ciriello et al*., 2013a].

It is also interesting to note that the estimated values of the parameters which GSA identifies as the most sensitive of each model, i.e., α_L for the ADRE-KM and β for the CTRW-PT, do not change significantly when

Table 4. Calibrated Values, C, of Model Parameters, Gaussian 95% Confidence Intervals, in Parentheses, and Ratio, $\Re$, of the Difference Between the Upper and Lower Limits Identifying the 95% Estimate Confidence Limits and C

		Subset 1		Subset 2	
Parameter	Model	C	$\Re$	C	$\Re$
v	ADRE-KM	1.24E-04 (1.23E-04, 1.25E-04)	0.02	1.26E-04 (1.22E-04, 1.30E-04)	0.06
α_L	ADRE-KM	1.07E-03 (0.92E-03, 1.22E-03)	0.28	1.07E-03 (0.59E-03, 1.55E-03)	0.90
β_0	ADRE-KM	190 (−446.33, 826.34)	6.70	190 (−587.35, 967.35)	8.18
m	ADRE-KM	9.53E-01 (5.10E-01, 13.99E-01)	0.94	9.44E-01 (4.01E-01, 14.86E-01)	1.15
v_ψ	CTRW-PT	1.21E-04 (1.20E-04, 1.21E-04)	0.01	1.20E-04 (1.19E-04, 1.20E-04)	0.01
D_ψ	CTRW-PT	1.64E-07 (1.48E-07, 1.80E-07)	0.20	1.45E-07 (1.17E-07, 1.73E-07)	0.39
β	CTRW-PT	1.95 (1.94, 1.97)	0.01	1.97 (1.96, 1.99)	0.02
t_2	CTRW-PT	116804.50 (81614, 151995)	0.60	104347.04 (73032, 135662)	0.60
R	CTRW-PT	3.51E-03 (3.29E-03, 3.75E-03)	0.13	2.89E-03 (2.62E-03, 3.17E-03)	0.19

Table 5. Results From Model Calibration and Identification Criteria[a]

	ADRE-KM				
Subset	*NLL*	*AIC*	AIC_c	*KIC*	Posterior Probability
1	−384.2	−376.2	−375.7	−348.3	0.00
2	**−115.2**	**−107.2**	**−104.6**	**−84.8**	**0.91**
	CTRW-PT				
Subset	*NLL*	*AIC*	AIC_c	*KIC*	Posterior Probability
1	**−435.0**	**−425.0**	**−424.3**	**−383.6**	**1.00**
2	−112.5	−102.5	−98.2	−63.8	0.09

[a]The smallest values for each subset are emphasized in bold.

calibrating on subset 1 instead of subset 2. This result supports the basic idea that a minimum set of selected sensitive observations allows capture of the parameter behavior. Moreover, we observe that parameters which GSA has revealed as less influential display remarkably different estimated values depending on the calibration data set considered. We remark that these results are consistent with the way sensitivity-based subsets of observations are constructed, i.e., favoring parameters associated with high values of total sensitivity indices at a non-negligible number of space-time locations.

Once parameter estimates are obtained, we investigate the accuracy with which the two models represent the elements of the calibration subsets. To this end, Table 5 lists values of (i) *NLL* (3), (ii) model selection criteria (7)–(9), and (iii) posterior probability (10), computed for each model in both calibration scenarios. We observe that all of these quantities indicate the ADRE-KM as the favored model when the sensitivity-based subset (subset 2) is used for calibration. We further note that the goodness of the performance of the two models is quite similar in this case; only the value of posterior probability clearly favors the ADRE-KM. CTRW-PT is considered as preferable in the case of subset 1, where model quality metrics are different from those related to ADRE-KM and the performance of the CTRW-PT model is clearly preferred.

As a final step, we assess the interpretive power of each model though a validation step, which entails comparing model predictions (based on the parameters estimated according to the procedure detailed above) against the entire data set of *Gramling et al.* [2002]. Figures 3a and 4a depict solutions provided by the two full system models (ADRE-KM and CTRW-PT, respectively), with parameter estimates based on gPC for the case of subset 1 (observations belonging to subset 1 are highlighted in red in the figures). The solutions are seen to match the measurements. Figures 3b and 4b depict the corresponding model solutions with parameter estimates based on gPC and relying on subset 2. The sensitive observations (which differ for the two models) constituting the elements of subset 2 are highlighted in red in the figures. Model predictions are also seen to be accurate, which is remarkable considering (i) the limited number of data employed for calibration and (ii) the adoption of proxies for parameter estimation. Accuracy of the approximations represented in Figures 3 and 4 is quantified in Table 5, which lists the mean square error (MSE) calculated between model predictions and available data. The CTRW-PT model always renders the lowest values of MSE at all the observation times for calibration subset 2. This suggests that, even as the ADRE-KM provides a high-quality fit of its sensitive observations (see Table 5), its predictive power is less than that of CTRW-PT when considering the entire experimental data set. Sensitive observations associated with CTRW-PT tend to provide a stronger basis for the model predictive capability than do their counterparts associated with the ADRE-KM. This result is consistent with the varying spatial distributions of the most sensitive observations associated with the two models (Figures 3b and 4b). These spatial distributions are linked to the different mathematical formulations of the two models and to the way model parameters govern the phenomena we investigate. The parameters of CTRW-PT tend to describe the overall behavior of the concentration profile, in terms of peak value, spreading and tailing. As a consequence, the most sensitive observations associated with CTRW-PT are distributed in a relatively uniform fashion along several parts of the concentration profile. In contrast, the ADRE-KM tends to describe the concentration profile in terms of two key moments, i.e., its mean and variance (the latter related to the spread). As a consequence, the most sensitive observation for this model are mainly concentrated around the center of mass and close to the tails of the concentration profile.

As an additional element, Table 6 lists the difference ΔMSE between the MSE values associated with subset 2 and 1. We observe that for the CTRW-PT model the interpretation of the complete data set generally

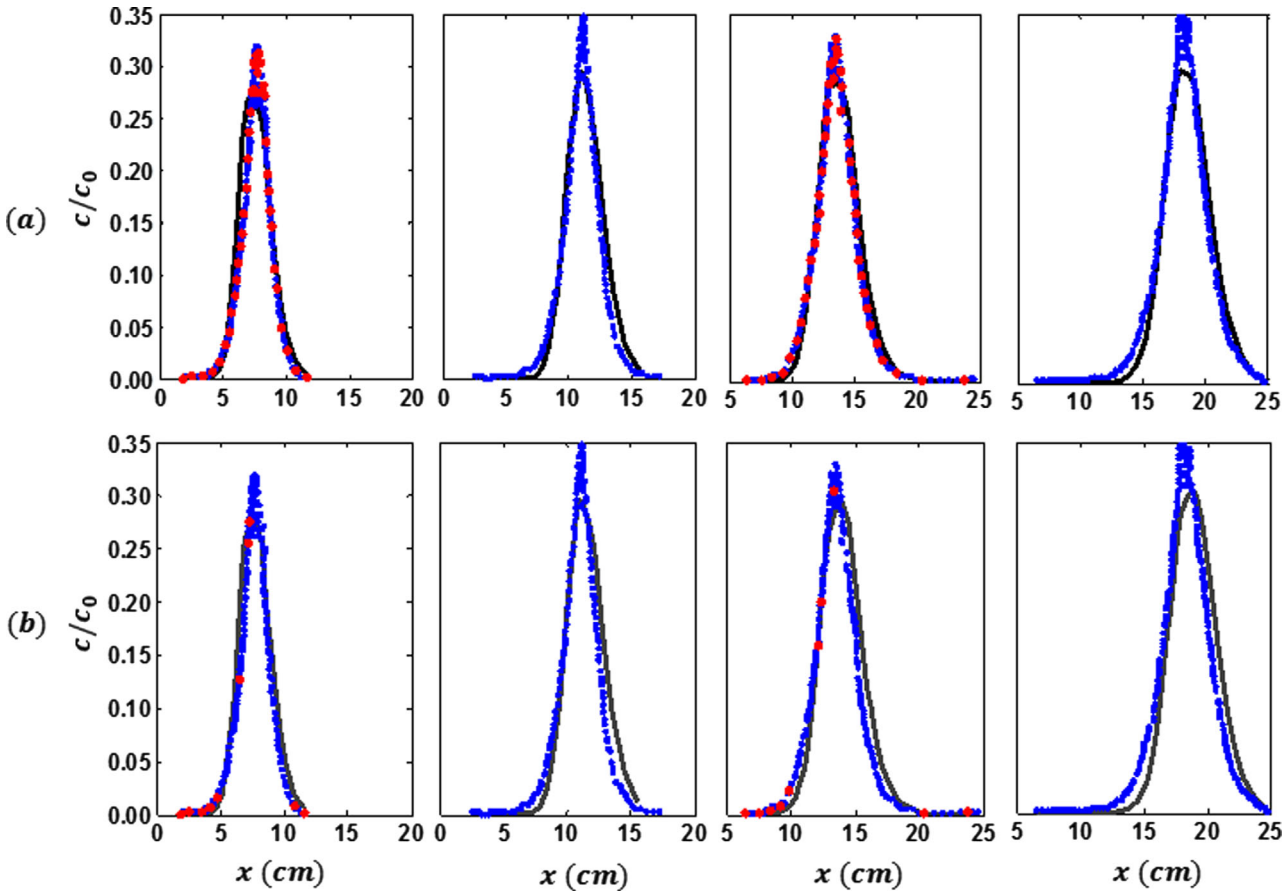

Figure 3. ADRE-KM model approximation (black lines) of the entire data set of *Gramling et al.* [2002] (blue dots), with model parameters estimated via gPC against (a) subset 1 and (b) subset 2 (see Table 4). Observations belonging to the subsets are represented by red dots.

improves when sensitivity-based model calibration is performed, as seen from the negative values of ΔMSE. The opposite is observed to occur for the ADRE-KM, for which the MSE associated with predictions based on calibration subset 2 is higher at most times (with the exception of the earliest sampling time). This suggests that when a model appears to be suitable to capture the behavior of a selected system, calibration against extended subsets (including also observations to which model parameters are not sensitive) may compromise the accuracy of predictions. This observation provides additional support for the effectiveness of our proposed methodology.

We remark that the use of proxies (in our case, the gPC technique) to perform GSA and model calibration should be accompanied by assessment of the accuracy of these approximations against the results of the full system model. Figure 5 depicts scatter plots between results obtained by the full model and the corresponding proxy at all space-time location where observation are available, relying on parameter estimates obtained by model calibration against subsets 1 (Figures 5a and 5b) and 2 (Figures 5c and 5d). We observe that results for both models are clustered around a 45° regression line. Spreading of points around the regression line is virtually negligible for the ADRE-KM. The coefficient of determination for the CTRW-PT is also very high, denoting very good performance of the model proxy. The latter observation, together with the facts that (i) the CTRW-PT results are by their nature influenced by statistical fluctuations and (ii) computational time can increase dramatically with the number of particles, does not justify use of higher order approximations in the construction of the gPC, at least for the purpose of the current application.

We end the analysis by emphasizing that the gPC renders approximations of the model response surface (in the random parameter space) at each space-time location of interest and, as shown in section 3, the coefficients of the polynomial expansion depend on space and time. Hence, each point depicted in Figure 5

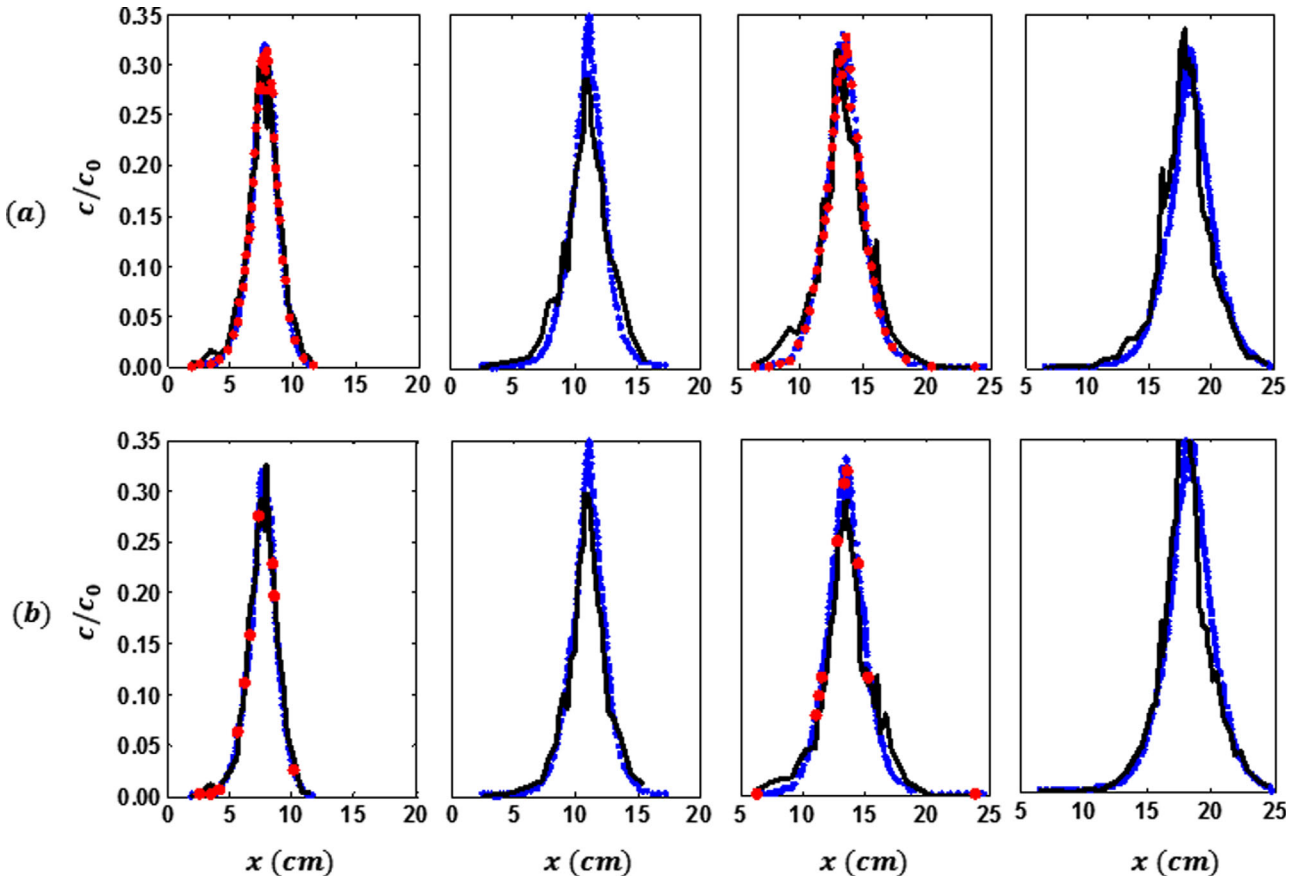

Figure 4. CTRW-PT model approximation (black lines) of the entire data set of *Gramling et al.* [2002] (blue dots) with model parameters estimated via gPC against (a) subset 1 and (b) subset 2 (see Table 4). Observations belonging to the subsets are represented by red dots.

is associated to a unique set of coefficients multiplying the orthogonal polynomials of the second-order gPC that we employed in this case study. If poor accuracy of the proxy is evidenced at a given order of expansion, one can still employ the methodological framework we propose by noting that an increase in the polynomial order is typically accompanied by an increase of the overall computational cost [e.g., *Ciriello et al.*, 2013a, b].

Note that the methodology described here is relevant for integration of experimental design and model application, at both laboratory and field scales. For any given model, the methodology can identify (i) the most influential parameters, and, significantly, (ii) the space and time locations where measurements

Table 6. Results of Model Validation in Terms of Mean Square Error (MSE) for Each of the Four Observation Times τ_i ($i = 1, 2, 3, 4$) and Observation Subsets 1 and 2[a]

		MSE		
Observation Time	Model	Subset 1	Subset 2	ΔMSE
τ_1	ADRE-KM	1.19E-03	8.63E-04	−3.27E-04
	CTRW-PT	**4.75E-04**	**3.64E-04**	−1.11E-04
τ_2	ADRE-KM	**6.71E-04**	1.24E-03	5.69E-04
	CTRW-PT	9.01E-04	**7.03E-04**	−1.98E-04
τ_3	ADRE-KM	**5.05E-04**	1.09E-03	5.85E-04
	CTRW-PT	8.79E-04	**9.22E-04**	4.30E-05
τ_4	ADRE-KM	**6.90E-04**	1.74E-03	1.05E-03
	CTRW-PT	1.62E-03	**1.41E-03**	−2.10E-04
All times jointly	ADRE-KM	**7.59E-04**	1.26E-03	5.01E-04
	CTRW-PT	9.06E-04	**8.90E-04**	−1.60E-05

[a]The smallest values for each model are in bold. ΔMSE is the difference between the MSE values associated with subset 2 and 1.

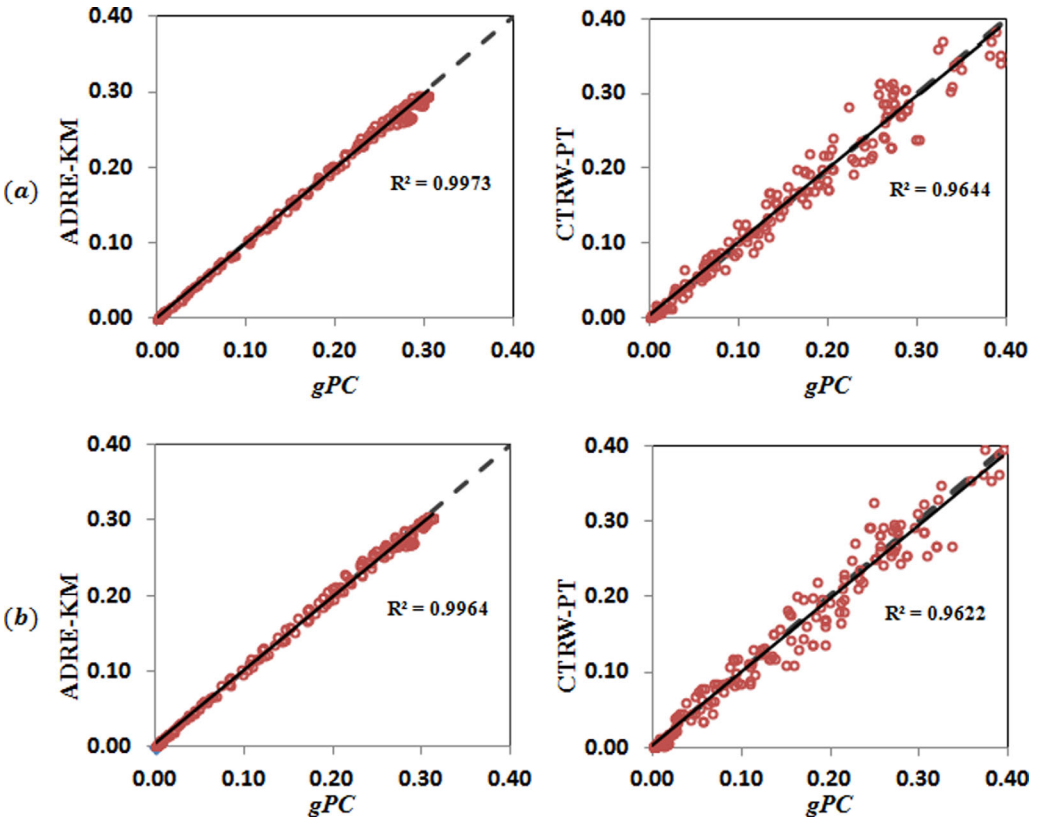

Figure 5. Comparison of concentrations computed by the full models and the corresponding gPC proxies, at the space-time locations in which observations are available, and for values of parameters obtained by calibration against (a) subset 1 and (b) subset 2 (see Table 4). The coefficient of determination, R^2, associated with the regression (continuous) line is also shown together with the 45 degrees (broken) line.

provide the most valuable information (in terms of model calibration). This information defines, at least ideally, where measuring efforts should be concentrated and the amount of data that should be collected. Moreover, by analysis of subsets of measurements, applicability of various models can be compared.

5. Conclusions

Quantification of transport phenomena in porous and fractured media involves three key questions: (i) For a given set of measurements, which conceptual picture of the (reactive) transport processes, as embodied in a mathematical model or models, is most appropriate? (ii) For a given model, what are the most valuable space and time locations for solute concentration measurements? (iii) How is model parameter uncertainty propagated to model output, and how does this propagation affect model calibration? To address these questions, we merged aspects of model formulation, calibration, global sensitivity analysis and discrimination, illustrating application of a methodology for sensitivity-based parameter estimation. The proposed methodology is general and may be employed in different contexts for model assessment and parameter estimation.

We have shown how this GSA-based approach addresses the questions defined above. Here measurements from a reactive transport experiment were employed to demonstrate the methodology and examine the interpretive capability of two reactive transport models (ADRE-KM and CTRW-PT). In the specific case study, we find that while the ADRE-KM fits the most sensitive observation points closely, the CTRW-PT offers improved predictive capabilities when results from both models are compared against the entire experimental data set (in the context of model validation). Moreover, we show in this application that the most sensitive observations associated with the CTRW-PT provide a stronger basis for model prediction than the corresponding most sensitive observations for the ADRE-KM.

Acknowledgments
V.C. acknowledges partial support from Marco Polo Program 2011 of the University of Bologna. B.B. and Y.E. gratefully acknowledge the support of a research grant from the Crystal Family Foundation. B.B. holds the Sam Zuckerberg Professorial Chair in Hydrology. A.G. acknowledges funding from MIUR (Italian Ministry of Education, Universities and Research—PRIN2010-11; project: "Innovative methods for water resources under hydro-climatic uncertainty scenarios." The numerical results used to produce the results of this manuscript are freely available upon request to the corresponding author.

References

Abramowitz, M., and I. Stegun (1970), Handbook of Mathematical Functions, Dover, N. Y.
Akaike, H. (1974), A new look at statistical model identification, *IEEE Trans. Autom. Control*, *19*, 716–723.
Ashraf, M., S. Oladyshkin, and W. Nowak (2013), Geological storage of CO_2: Application, feasibility and efficiency of global sensitivity analysis and risk assessment using the arbitrary polynomial chaos, *Int. J. Greenhouse Gas Control*, *19*, 704–719.
Barth, G. R., and M. C. Hill (2005a), Numerical methods for improving sensitivity analysis and parameter estimation of virus transport simulated using sorptive-reactive processes, *J. Contam. Hydrol.*, *76*, 251–277.
Barth, G. R., and M. C. Hill (2005b), Parameter and observation importance in modeling virus transport in saturated systems—Investigations in a homogeneous system, *J. Contam. Hydrol.*, *80*, 107–129.
Berkowitz, B., and H. Scher (1995), On characterization of anomalous dispersion in porous and fractured media, *Water Resour. Res.*, *31*(6), 1461–1466.
Berkowitz, B., and H. Scher (1997), Anomalous transport in random fracture networks, *Phys. Rev. Lett.*, *79*, 4038–4041.
Berkowitz, B., H. Scher, and S. E. Silliman (2000), Anomalous transport in laboratory-scale, heterogeneous porous media, *Water Resour. Res.*, *36*(1), 149–158.
Berkowitz, B., A. Cortis, M. Dentz, and H. Scher (2006), Modeling non-Fickian transport in geological formations as a continuous time random walk, *Rev. Geophys.*, *44*, RG2003, doi:10.1029/2005RG000178.
Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.
Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, *249*, 11–29.
Bianchi Janetti, E., I. Dror, M. Riva, A. Guadagnini, and B. Berkowitz (2012), Estimation of single-metal and competitive sorption isotherms through maximum likelihood and model quality criteria, *Soil Sci. Soc. Am. J.*, *76*, 1229–1245.
Bromly, M., and C. Hinz (2004), Non-Fickian transport in homogeneous unsaturated repacked sand, *Water Resour. Res.*, *40*, W07402, doi:10.1029/2003WR002579.
Cao, J., and P. K. Kitanidis (1998), Pore-scale dilution of conservative solutes: An example, *Water Resour. Res.*, *34*(8), 1941–1949.
Carrera, J., and S. P. Neuman (1986a), Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information, *Water. Resour. Res.*, *22*(2), 199–210.
Carrera, J., and S. P. Neuman (1986b), Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms, *Water Resour. Res.*, *22*(2), 211–227.
Ciriello, V., and V. Di Federico (2013), Analysis of a benchmark solution for non-Newtonian radial displacement in porous media, *Int. J. Non-Linear Mech.*, *52*, 46–57.
Ciriello, V., A. Guadagnini, V. Di Federico, Y. Edery, and B. Berkowitz (2013a), Comparative analysis of formulations for conservative transport in porous media through sensitivity-based parameter calibration, *Water Resour. Res.*, *49*, 5206–5220, doi:10.1002/wrcr.20395.
Ciriello, V., V. Di Federico, M. Riva, F. Cadini, J. De Sanctis, E. Zio, and A. Guadagnini (2013b), Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer, *Stochastic Environ. Res. Risk Assess.*, *27*, 945–954.
Cirpka, O. A., and P. K. Kitanidis (2000), An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport, *Water Resour. Res.*, *36*(5), 1209–1220.
Crestaux, T., O. Le Maître, and J. M. Martinez (2009), Polynomial chaos expansion for sensitivity analysis, *Reliab. Eng. Syst. Safety*, *94*(7), 1161–1172.
Cushman, J. H., and T. R. Ginn (1993), On dispersion in fractal porous media, *Water Resour. Res.*, *29*(10), 3513–3515.
Dagan, G., and S. P. Neuman (1991), Nonasymptotic behavior of a common Eulerian approximation for transport in random velocity fields, *Water Resour. Res.*, *27*(12), 3249–3256.
Doherty, J. (2002), *PEST: Model Independent Parameter Estimation, User Manual*, 4th ed., Watermark Numer. Comput., Corinda, Queensland, Australia.
Edery, Y., H. Scher, and B. Berkowitz (2010), Particle tracking model of bimolecular reactive transport in porous media, *Water Resour. Res.*, *46*, W07524, doi:10.1029/2009WR009017.
Edery, Y., A. Guadagnini, H. Scher, and B. Berkowitz (2014), Origins of anomalous transport in heterogeneous media: Structural and dynamic controls, *Water Resour. Res.*, *50*, 1490–1505, doi:10.1002/2013WR015111.
Esfandiar, B., G. Porta, S. Perotto, and A. Guadagnini (2015), Impact of space-time mesh adaptation on solute transport modeling in porous media, *Water Resour. Res.*, *51*, 1315–1332, doi:10.1002/2014WR016569.
Fajraoui, N., F. Ramasomanana, A. Younes, T. A. Mara, P. Ackerer, and A. Guadagnini (2011), Use of global sensitivity analysis and polynomial chaos expansion for interpretation of nonreactive transport experiments in laboratory-scale porous media, *Water Resour. Res.*, *47*, W02521, doi:10.1029/2010WR009639.
Formaggia, L., A. Guadagnini, I. Imperiali, V. Lever, G. Porta, M. Riva, A. Scotti, and L. Tamellini (2013), Global sensitivity analysis through polynomial chaos expansion of a basin-scale geochemical compaction model, *Comput. Geosci.*, *17*(1), 25–42.
Gaganis, P., and L. Smith (2001), A Bayesian approach to the quantification of the effect of model error on the predictions of groundwater models, *Water Resour. Res.*, *37*(9), 2309–2322.
Ghanem, R. G., and P. D. Spanos (1991), *Stochastic Finite Elements—A Spectral Approach*, Springer, Berlin.
Gillespie, D. (1977), Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, *25*, 2340–2361.
Gramling, C. M., C. F. Harvey, and L. C. Meigs (2002), Reactive transport in porous media: A comparison of model prediction with laboratory visualization, *Environ. Sci. Technol.*, *36*, 2508–2514.
Haggerty, R., S. A. McKenna, and L. C. Meigs (2000), On the late-time behavior of tracer test breakthrough curves, *Water Resour. Res.*, *36*(12), 3467–3479.
Hering, J. G., and F. M. M. Morel (1988), Kinetics of trace-metal complexation: Role of alkaline-earth metals, *Environ. Sci. Technol.*, *22*(12), 1469–1478.
Huang, S., M. Sankaran, and R. Ramesh (2007), Collocation-based stochastic finite element analysis for random field problems, *Probab. Eng. Mech.*, *22*, 194–205.
Hurvich, C. M., and C. L. Tsai (1989), Regression and time series model selection in small sample, *Biometrika*, *76*(2), 297–307.
James, A. L., and C. M. Oldenburg (1997), Linear and Monte Carlo uncertainty analysis for subsurface contaminant transport simulation, *Water Resour. Res.*, *33*(11), 2495–2508.

Jose, S. C., and O. A. Cirpka (2004), Measurement of mixing-controlled reactive transport in homogeneous porous media and its prediction from conservative tracer test data, *Environ. Sci. Technol.*, *38*, 2089–2096.
Kashyap, R. L. (1982), Optimal choice of AR and MA parts in autoregressive moving average models, *IEEE Trans. Pattern Anal.*, *4*(2), 99–104.
Kuntz, B. W., S. Rubin, B. Berkowitz, and K. Singha (2011), Quantifying solute transport at the Shale Hills Critical Zone Observatory, *Vadose Zone J.*, *10*, 843–857.
Laloy, E., B. Rogiers, J. A. Vrugt, D. Mallants, and D. Jacques (2013), Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov Chain Monte Carlo simulation and polynomial chaos expansion, *Water Resour. Res.*, *49*, 2664–2682, doi:10.1002/wrcr.20226.
Levy, M., and B. Berkowitz (2003), Measurement and analysis of non-Fickian dispersion in heterogeneous porous media, *J. Contam. Hydrol.*, *64*(3–4), 203–226.
Li, W., Z. Lu, and D. Zhang (2009), Stochastic analysis of unsaturated flow with probabilistic collocation method, *Water Resour. Res.*, *45*, W08425, doi:10.1029/2008WR007530.
Liao, Q., and D. Zhang (2013), Probabilistic collocation method for strongly nonlinear problems: 1. Transform by location, *Water Resour. Res.*, *49*, 7911–7928, doi:10.1002/2013WR014055.
Liao, Q., and D. Zhang (2014), Probabilistic collocation method for strongly nonlinear problems: 2. Transform by displacement, *Water Resour. Res.*, *50*, 8736–8759, doi:10.1002/2014WR016238.
Lin, G., A. M. Tartakovsky, and D. M. Tartakovsky (2010), Uncertainty quantification via random domain decomposition and probabilistic collocation on sparse grids, *J. Comput. Phys.*, *229*, 6995–7012.
Lindenberg, K., and A. H. Romero (2007), Numerical study of $A + A \rightarrow 0$ and $A + B \rightarrow 0$ reactions with inertia, *J. Chem. Phys.*, *127*, paper no. 174506, doi:10.1063/1.2779327.
Lu, D., M. Ye, S. P. Neuman, and L. Xue (2012), Multimodel Bayesian analysis of data-worth applied to unsaturated fractured tuffs, *Adv. Water Resour.*, *35*, 69–82.
Matthies, H. G. (2008), Stochastic finite elements: Computational approaches to stochastic partial differential equations, *ZAMM*, *88*(11), 849–873.
National Research Council (2001), *Conceptual Models of Flow and Transport in the Fractured Vadose Zone*, Natl. Acad. Press, Washington, D. C.
Neuman, S. P. (2002), Accounting for conceptual model uncertainty via maximum likelihood Bayesian model averaging, *Acta Univ. Carolinae Geol.*, *46*(2–3), 529–534.
Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models, *Stochastic Environ. Res. Risk Assess.*, *17*(5), 291–305, doi:10.1007/s00477-003-0151-7.
Neuman, S. P., L. Xue, M. Ye, and D. Lu (2012), Bayesian analysis of data-worth considering model and parameter uncertainties, *Adv. Water Resour.*, *36*, 75–85.
Oladyshkin, S., and W. Nowak (2012), Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion, *Reliab. Eng. Syst. Safety*, *106*, 179–190.
Oladyshkin, S., F. P. J. de Barros, and W. Nowak (2012), Global sensitivity analysis: A flexible and efficient framework with an example from stochastic hydrogeology, *Adv. Water Resour.*, *37*, 10–22.
Oladyshkin, S., H. Class, and W. Nowak (2013), Bayesian updating via bootstrap filtering combined with data-driven polynomial chaos expansions: Methodology and application to history matching for carbon dioxide storage in geological formations, *Comput. Geosci.*, *17*(4), 671–687.
Palanichamy, J., T. Becker, M. Spiller, J. Kangeter, and S. Mohan (2007), Multicomponent reaction modelling using a stochastic algorithm, *Comput. Visualization Sci.*, *12*, 51–61.
Porta, G., L. Tamellini, V. Lever, and M. Riva (2014), Inverse modeling of geochemical and mechanical compaction in sedimentary basins through Polynomial Chaos Expansion, *Water Resour. Res.*, *50*, 9414–9431, doi:10.1002/2014WR015838.
Rahman, Md. A., C. J. Surabhin, W. Nowak, and O. A. Cirpka (2005), Experiments on vertical transverse mixing in a large-scale heterogeneous model aquifer, *J. Contam. Hydrol.*, *80*(3–4), 130–148.
Rajabi, M. M., B. Ataie-Ashtiani, and C. T. Simmons (2015), Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations, *J. Hydrol.*, *520*, 101–122.
Raje, D. S., and V. Kapoor (2000), Experimental study of bimolecular reaction kinetics in porous media, *Environ. Sci. Technol.*, *34*, 1234–1239.
Rashidi, M., L. Peurrung, A. F. B. Tompson, and T. J. Kulp (1996), Experimental analysis of pore-scale flow and transport in porous media, *Adv. Water Resour.*, *3*, 163–180.
Riva, M., A. Guadagnini, and A. Dell'Oca (2015), Probabilistic assessment of seawater intrusion under multiple sources of uncertainty, *Adv. Water Resour.*, *75*, 93–104.
Rubin, S., I. Dror, and B. Berkowitz (2012), Experimental and modeling analysis of coupled non-Fickian transport and sorption in natural soils, *J. Contam. Hydrol.*, *132*, 28–36.
Saltelli, A., S. Tarantola, and F. Campolongo (2000), Sensitivity analysis as an ingredient of modeling, *Stat. Sci.*, *15*(4), 377–395.
Samper, F. J., and S. P. Neuman (1989a), Estimation of spatial covariance structures by adjoint state maximum likelihood cross validation: 1. Theory, *Water Resour. Res.*, *25*(3), 351–362.
Samper, F. J., and S. P. Neuman (1989b), Estimation of spatial covariance structures by adjoint state maximum likelihood cross validation: 2. Synthetic experiments, *Water Resour. Res.*, *25*(3), 363–371.
Samper, J., and J. Molinero (2000), Predictive uncertainty of numerical models of groundwater flow and solute transport, *Eos Trans. AGU*, *81*(48), Fall Meet. Suppl., Abstract H11G-04.
Sánchez-Vila, X., and J. Carrera (2004), On the striking similarity between the moments of breakthrough curves for a heterogeneous medium and a homogeneous medium with a matrix diffusion term, *J. Hydrol.*, *294*(1–3), 164–175.
Sanchez-Vila, X., D. Fernàndez-Garcia, and A. Guadagnini (2010), Interpretation of column experiments of transport of solutes undergoing an irreversible bimolecular reaction using a continuum approximation, *Water Resour. Res.*, *46*, W12510, doi:10.1029/2010WR009539.
Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.*, *6*(2), 461–464.
Silliman, S. E., and E. S. Simpson (1987), Laboratory evidence of the scale effect in dispersion of solutes in porous media, *Water Resour. Res.*, *23*(8), 1667–1673.
Silliman, S. E., L. F. Konikow, and C. I. Voss (1987), Laboratory investigation of longitudinal dispersion in anisotropic porous media, *Water Resour. Res.*, *23*(11), 2145–2151.
Sobol, I. M. (1993), Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. 1*, 407–414.
Srinivasan, G., D. M. Tartakovsky, B. A. Robinson, and A. B. Aceves (2007), Quantification of uncertainty in geochemical reactions, *Water Resour. Res.*, *43*, W12415, doi:10.1029/2007WR006003.

Sudret, B. (2008), Global sensitivity analysis using polynomial chaos expansions, *Reliab. Eng. Syst. Safety*, *93*, 964–979.
Sun, A. Y., M. Zeidouni, J. P. Nicot, Z. Lu, and D. Zhang (2013), Assessing leakage detectability at geologic CO_2 sequestration sites using the probabilistic collocation method, *Adv. Water Resour.*, *56*, 49–60.
Tartakovsky, A. M., G. Redden, P. C. Lichtner, T. D. Scheibe, and P. Meakin (2008), Mixing-induced precipitation: Experimental study and multiscale numerical analysis, *Water Resour. Res.*, *44*, W06S04, doi:10.1029/2006WR005725.
Tartakovsky, A. M., G. D. Tartakovsky, and T. D. Scheibe (2009), Effects of incomplete mixing on multicomponent reactive transport, *Adv. Water Resour.*, *32*(11), 1674–1679.
Tartakovsky, D. M. (2007), Probabilistic risk analysis in subsurface hydrology, *Geophys. Res. Lett.*, *34*, L05404, doi:10.1029/2007GL029245.
Tiedeman, C. R., M. C. Hill, F. A. D'Agnese, and C. C. Faunt (2003), Methods for using groundwater model predictions to guide hydrogeologic data collection, with application to the Death Valley regional ground-water flow system, *Water Resour. Res.*, *39*(1), 1010, doi:10.1029/2001WR001255.
Tiedeman, C. R., D. M. Ely, M. C. Hill, and G. M. O'Brien (2004), A method for evaluating the importance of system state observations to model predictions, with application to the Death Valley regional groundwater flow system, *Water Resour. Res.*, *40*, W12411, doi:10.1029/2004WR003313.
Webster, M., M. A. Tatang, and G. J. McRae (1996), Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model, *Tech. Rep. 4*, MIT Joint Program on the Sci. and Policy of Global Change. [Available at http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt4.pdf.]
Wiener, N. (1938), The homogeneous chaos, *Am. J. Math.*, *60*, 897–936.
Xiu, D., and G. E. Karniadakis (2002), The Wiener-Askey polynomial chaos for stochastic differential equations, *J. Sci. Comput.*, *24*(2), 619–644.
Xue, L., D. Zhang, A. Guadagnini, and S. P. Neuman (2014), Multimodel Bayesian analysis of groundwater data worth, *Water Resour. Res.*, *50*, 8481–8496, doi:10.1002/2014WR015503.
Ye, M., S. P. Neuman, and P. D. Meyer (2004), Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, *40*, W05113, doi:10.1029/2003WR002557.
Ye, M., P. D. Meyer, and S. P. Neuman (2008), On model selection criteria in multimodel analysis, *Water Resour. Res.*, *44*, W03428, doi:10.1029/2008WR006803.
Yuste, S. B., J. J. Ruiz-Lorenzo, and K. Lindenberg (2008), Coagulation reactions in low dimensions: Revisiting subdiffusive A+A reactions in one dimension, *Phys. Rev. E*, *80*, 051114.
Zinn, B., L. C. Meigs, C. F. Harvey, R. Haggerty, W. J. Peplinski, and C. F. Von Schwerin (2004), Experimental visualization of solute transport and mass transfer processes in two-dimensional conductivity fields with connected regions of high conductivity, *Environ. Sci. Technol.*, *38*(14), 3916–3926.