

Archived at the Flinders Academic Commons: http://dspace.flinders.edu.au/dspace/

'This is the peer reviewed version of the following article: Rajabi, M. M., Ataie-Ashtiani, B., & Simmons, C. T. (2018). Model-data interaction in groundwater studies: Review of methods, applications and future directions. Journal of Hydrology, 567, 457–477. https:// doi.org/10.1016/j.jhydrol.2018.09.053

which has been published in final form at https://doi.org/10.1016/j.jhydrol.2018.09.053

© 2018 Elsevier BV. This manuscript version is made available under the CC-BY-NC-ND 4.0 license: http://creativecommons.org/licenses/by-nc-nd/4.0/

Accepted Manuscript

Research papers

Model-Data Interaction in Groundwater Studies: Review of Methods, Applications and Future Directions

Mohammad Mahdi Rajabi, Behzad Ataie-Ashtiani, Craig T. Simmons

PII:S0022-1694(18)30747-9DOI:https://doi.org/10.1016/j.jhydrol.2018.09.053Reference:HYDROL 23150

To appear in: *Journal of Hydrology*



Please cite this article as: Mahdi Rajabi, M., Ataie-Ashtiani, B., Simmons, C.T., Model-Data Interaction in Groundwater Studies: Review of Methods, Applications and Future Directions, *Journal of Hydrology* (2018), doi: https://doi.org/10.1016/j.jhydrol.2018.09.053

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

REVIEW ARTICLE

Model-Data Interaction in Groundwater Studies: Review of

Methods, Applications and Future Directions

Mohammad Mahdi Rajabi ^{a,*} Behzad Ataie-Ashtiani ^{b, c}, Craig T. Simmons ^c

* Corresponding author

^a Civil and Environmental Engineering Faculty, Tarbiat Modares University, PO Box 14115-397, Tehran, Iran.

^b Department of Civil Engineering, Sharif University of Technology, PO Box 11155-9313, Tehran, Iran

° National Centre for Groundwater Research & Training, College of Science & Engineering, Flinders University, GPO Box 2100, Adelaide, South Australia 5001, Australia

E-mail addresses: <u>mmrajabi@modares.ac.ir</u> (M.M. Rajabi), <u>ataie@sharif.edu</u> (B. Ataie-Ashtiani), <u>craig.simmons@flinders.edu.au</u> (C.T. Simmons). First Submitted: July 2018

Revision submitted: September 2018

Abstract

We define model-data interaction (MDI) as a two way process between models and data, in which on one hand data can serve the modeling purpose by supporting model discrimination, parameter refinement, uncertainty analysis, etc., and on the other hand models provide a tool for data fusion, interpretation, interpolation, etc. MDI has many applications in the realm of groundwater and has been the topic of extensive research in the groundwater community for the past several decades. This has led to the development of a multitude of increasingly sophisticated methods. The progress of data acquisition technologies and the evolution of models are continuously changing the landscape of groundwater MDI, creating new challenges and opportunities that must be properly understood and addressed. This paper aims to review, analyze and classify research on MDI in groundwater applications, and discusses several related

aspects including: (1) basic theoretical concepts and classification of methods, (2) sources of uncertainty and how they are commonly addressed, (3) specific characteristics of groundwater models and data that affect the choice of methods, (4) how models and data can interact to provide added value in groundwater applications, (5) software and codes for MDI, and (6) key issues that will likely form future research directions. The review shows that there are many tools and techniques for groundwater MDI, and this diversity is needed to support different MDI objectives, assumptions, model and data types and computational constraints. The study identifies eight categories of applications for MDI in the groundwater literature, and highlights the growing gap between MDI practices in the research community and those in consulting, industry and government.

Keywords: Model-data interaction; Groundwater modeling; Uncertainty analysis; Data assimilation; Data fusion.

1. Introduction

The complexities of physics-based numerical models of groundwater flow and contaminant transport have grown substantially since the 1990s, moving from early two-dimensional steady state homogeneous or layer cake models to dynamic three-dimensional models capable of simulating highly heterogeneous formations and complex phenomena such as coupled and multiscale processes. This growing sophistication has resulted in models with significantly more data requirements compared to the past. In the last three decades, we have also seen a significant increase in the role of groundwater models in decision-making processes (Refsgaard et al., 2010). Models are now used throughout the world to assist decision making on issues such as optimal groundwater extraction (e.g. Ketabchi and Ataie-Ashtiani, 2015a, b; Triki et al., 2017), groundwater contamination response (e.g. Ritzel et al., 1994), aquifer recharge and recovery

(Drumheller et al., 2017), groundwater remediation and cleanup (e.g. Bayer and Finkel, 2004 and 2007; Singh and Minsker, 2008) and determination of wellhead protection areas (e.g. Wheater et al., 2000; Feyen et al., 2001). The increasing role of models in decision making means that models are required to provide more reliable predictions with proper estimation of prediction uncertainties. Regardless of a model's level of complexity, this can only be achieved if the appropriate data are available and effectively incorporated into the modeling process.

In parallel, hydrogeology has also become increasingly data rich due to the significant progress made in technologies that enable the collection, transfer and storage of data (Fasbender, 2008; Barnhart et al., 2010). Advances in geophysical, remote sensing (e.g. satellite imaging), smart meter and field sensor technologies have led to a massive upsurge in data quantity and diversification of data types. Moreover, thanks to the impressive developments in communication technologies such as wireless sensor networks (WSNs), data can now be delivered at extremely higher rates. The past two decades have also seen the creation and expansion of many national and regional databases for geologic and groundwater related data (e.g. United States Geological Survey groundwater database:

https://water.usgs.gov/ogw/data.html, South Africa national groundwater archive: http://www.dwa.gov.za/groundwater/nga.aspx, etc.), in which a massive amount of data is being continuously collected (Szalkai et al., 2007). These data-related developments have created new challenges, such as online data, heterogeneous data and massive data issues, which are wellknown challenges in some other fields of science, but were not previously encountered by the groundwater modeling community (Hayley, 2017).

These trends have changed our conception of the relationship between models and data. Traditionally, this relationship was mainly concerned with employing data for model calibration

and validation. But in a broadened and increasingly popular perspective that we refer to as model-data interaction (MDI), this relationship is viewed as a two way process in which on one hand, data can serve the modeling purpose by supporting model refinement, discrimination, uncertainty analysis, etc., and on the other hand, models can provide a framework for guidance of data collection and data analysis by assisting data fusion, interpretation, interpolation, etc. To the author's knowledge, the term MDI has not been previously used in the groundwater literature to describe this two way process, but it has been used in some other fields of science and engineering (e.g. Norby et al., 2016).

From this perspective, MDI offers great potential, a number of which we review below, such as:

- Sources of data in groundwater studies are numerous, but no single data source can provide a complete picture of a groundwater system (Linde et al., 2017), and therefore multiple data sources must be integrated. In many instances, the nature and spatiotemporal scale of hydrogeologic data is so diverse that they are not readily related to one another, making data integration very difficult (Porter et al., 2000). In these increasingly familiar situations, MDI provides a unique framework for integrating multiple types of data, and for solving related problems such as data conflict detection and resolution, and outlier detection (Brunner et al., 2017).
- With the progress of measurement technologies, we are approaching a state where we are no longer hindered by the power to collect more data, but by our ability to extract valuable information from available data. In these circumstances, MDI opens new opportunities for information extraction from different data types.
- MDI can improve the performance of groundwater models and open the model structure selection and parameter estimation processes to new and previously untapped sources of

information (Rajabi and Ataie-Ashtiani, 2016). This can alleviate the data scarcity problem frequently encountered in model construction efforts. Moreover, there is a growing recognition of the importance of characterizing the uncertainties in our models and data (Beven, 2010). In this sense, MDI provides a consistent tool for examining all sources of uncertainty in a common framework.

Data collection in groundwater studies is currently guided mostly by intuition and expert knowledge or by practical considerations rather than a quantitative understanding of what data will most reduce uncertainties and how much data are required to do so (Kikuchi, 2015). In this context, MDI techniques can be valuable tools in the guidance of groundwater data collection, setting the stage for a shift of paradigm in this area.

MDI has been the topic of extensive research in the groundwater community for the past several decades, leading to the development of a large number of increasingly sophisticated algorithms. These studies have transformed MDI from the simple direct insertion of data into models, to advanced sequential data assimilation techniques (Moradkhani et al., 2005; Liu et al., 2012). A significant portion of the progress in groundwater MDI has originated from developments in other fields of science and technology such as meteorology, atmospheric sciences, oceanography, robotics, defense, aerospace, etc. (examples of reviews include Ghil and Malanotte-Rizzoli, 1991; Wang et al., 2009; Peng et al., 2011).

Several issues have motivated us to write a review paper on groundwater MDI. These include: (1) recent trends that have changed the landscape of MDI in the groundwater literature, creating new challenges and opportunities that must be properly understood and addressed, (2) the multitude of mathematical algorithms presented in the literature for MDI that have made the topic confusing and hard to penetrate even for the experienced modeler, and (3) lack of a

comprehensive review on the subject. Although a number of notable reviews on related terms such as inverse modeling (e.g. Carrera et al., 2005, 2010; Zhou et al., 2014) and uncertainty analysis (e.g. Refsgaard et al., 2012; Wu and Zeng, 2013; Linde et al., 2017) are available, the more general topic of MDI in the groundwater literature has not been yet reviewed. The key objectives of this paper can be summarized as follows: (1) to review the basic theoretical concepts and present a reasonably simple classification of different frameworks and algorithms for anyone interested in obtaining an overview of MDI, (2) to analyze the pros, cons, and underlying assumptions of key algorithms with the intention of helping researchers and practitioners choose appropriate algorithms for various situations, (3) to highlight different ways that models and data can interact to provide added value in groundwater applications, (4) to identify codes and software previously used for MDI in the groundwater literature and explain key trends in this regard, and (5) to identify key issues and gaps that will likely form future directions. The outline of the paper is demonstrated in the **Fig. 1**.

It should be noted that almost every groundwater modeling study in the literature involves some component of MDI, because as in any other field of engineering, "models without data are fantasy" (Nisbet et al., 2014). So it is rather impossible to be exhaustive in this review. But we focus our review on a number of significant papers published largely over the last two decades, which deal with novel, formal mathematical algorithms for MDI. The number of citations and our (subjective) judgment of the quality and significance of these papers has played a key role in their selection.

2. Review of Theoretical Concepts and Techniques

Groundwater models are generic computer codes that use numerical algorithms to solve the partial differential equations that govern groundwater flow and (also often) contaminant

transport. These generic codes become site-specific models when data obtained from a particular geographical area is used to:

- Characterize the structure and parameters of the model. These tasks are referred to as model structure identification and parameter estimation (or model calibration) respectively.
- Test or validate the model in a procedure often referred to as model verification or validation. The term model selection is also used in this regard if alternative plausible models exist (Konikow and Bredehoeft, 1992; Oreskes et al., 1994; Beven, 2018).

The resulting models are used by research, government and consulting communities for a variety of objectives including understanding of hydrogeological systems and their interacting components across different scales, replication of past system behavior, forecast of system responses to potential future stresses and management practices, or to guide future data collection endeavors (Beven and Young, 2013; Doherty and Simmons, 2013). As it is not possible for models to characterize groundwater systems perfectly, model outputs should be associated with uncertainty intervals, making uncertainty analysis an important part of the entire modelling process (Freeze et al., 1990 and 1992). Data can be used in a systematic way to refine both model predictions and their uncertainty estimates within the framework of different modeling objectives. This procedure is often called state estimation. Model verification or selection, parameter estimation, guidance of future data collection and state estimation encompass different aspects of MDI.

The MDI process results in transformation of the initial datasets into new sets of data. For instance, assume that we are using sparse hydraulic conductivity (K) data obtained from pumping tests in conjunction with hydraulic head (h) data measured in a number of observation

wells, to estimate K values for a groundwater model through model calibration. The original K values (obtained from the pumping tests) are rarely left unchanged by the time the model is calibrated. The calibration hides some of the unknowns (e.g. heterogeneity) but also affects the representative scales, as K values for the pumping tests may, for instance, represent a smaller region than the one being used in the calibration process. Moreover, the calibration process uses sparse measurements of K to generate estimates that cover the entire modeling domain. It also combines the information contained in K and h data to estimate the hydraulic conductivity. These inevitable outcomes of MDI can be leveraged for a variety of purposes including customizing data resolution, improving data coverage and fusing different forms of data.

2.1. Framework of Model-Data Interaction

From the above discussions we can infer that MDI is not a one-time occurrence but a continuous process that has many aspects and applications. Regardless of the intended application, MDI rests on three foundations: forward model, data and synthesis method (Raupach et al., 2005). A forward model of the system (denoted by f) propagates the state (e.g. piezometric heads, chemical concentrations, temperatures, etc.) at time Y_t to time Y_{t+1} based on: (1) a set of (often assumed to be) time-invariant input parameters θ which include physical descriptions of subsurface characteristics (e.g. porosity, hydraulic conductivity, etc.) and contaminant transport and transformation characteristics (e.g. dispersivity, sorption factor, chemical reaction rates, etc.), and (2) time-dependent external forcing terms u_t which include surface recharge, lateral inflow, river-aquifer interactions, etc. This notion can be formulated as (Leisenring and Moradkhani, 2011):

$$Y_{t+1} = f(Y_t, u_t, \theta) + \omega_t$$

where the generally state-dependent and stochastic model error vector is represented by ω_t and the conceptual and mathematical model structure is embedded in f. This notion is schematically illustrated in **Fig. 2**.

Data (denoted by D) can be related to direct or indirect measurements (hard data) or qualitative assessments (soft data) of state variables (D_Y), external forcing terms (D_u), parameters (D_θ) or model structure (D_f) (Moradkhani et al., 2005; Liu and Gupta, 2007). **Table 1** summarizes common types of data in groundwater MDI. The measurement model or observation equation h_t is the mapping from the parameters, states and external forcing terms to the various types of data (Liu et al., 2012), formulated as:

$$D = h_t(Y_t, u_t, \theta) + \eta_t$$

(2)

where the stochastic term η_t represents observation error which is an inherent part of the measurement model. The mean values of ω_t and η_t denote systematic bias and their covariance signifies the uncertainty in the model predictions and observations (Liu and Gupta, 2007). The mean and covariance of ω_t and η_t are generally not directly observable, and are also difficult to estimate by indirect methods. Hence, they are often assigned presumptive values. A common practice is to assign uncorrelated Gaussian distribution with zero mean (i.e. white noise) to ω_t and η_t (e.g. Fu and Gómez-Hernández, 2009; Laloy et al., 2013).

The synthesis method is a formal mathematical algorithm that combines the model (f) and data (*D*) by varying some properties of the model to give their optimal combination, accounting for the associated uncertainties (Wang et al., 2009; Keenan et al., 2011). In general, a host of methods exist for synthesis with widely varying degrees of sophistication. It is important to realize that due to the diversity of model and data types, and the various objectives and simplifying assumptions that are subsumed under MDI, no universally best synthesis method exists (Porter et

al., 2000; Peng et al., 2011; Zhou et al, 2014). However, groundwater MDI has several specific characteristics that affect the choice of a synthesis method: (1) groundwater models are moderately to highly nonlinear (Samuel and Jha, 2003; Schöniger et al., 2012; Wallis et al., 2014; Siade et al., 2017), (2) they result in state variables (e.g. heads and concentrations) that are time-dependent (Eigbe et al., 1998), (3) they are commonly CPU-intensive and have high computational demands (most notably in real-world applications), as they rely on large systems of equations that need to be solved for each model run (Carrera et al., 2010; Rajabi et al., 2015b), (4) they are inherently high-dimensional with respect to parameters, external forcing terms and states, unless strict parameterization is employed (Siade et al., 2017), (5) their parameter and state variables exhibit non-Gaussianity in certain cases such as strong spatial heterogeneity of hydraulic conductivity, contaminant transport with sharp fronts as in advection dominated processes, and curvilinear crispy geometries (as in river beds) (Hendricks Franssen et al., 2009; Hendricks Franssen and Kinzelbach, 2009; Schöniger et al., 2012; Crestani et al., 2013), (6) they are controlled by physical properties (e.g. permeability) that are characterized by a high degree of heterogeneity, possess scales of variation spanning several orders of magnitude, and are scaledependent (Hill and Tiedeman, 2006; Kerrou et al., 2008), (7) complex interactions exist between different inputs of a groundwater model, e.g. between hydraulic conductivity and dispersivity (Gelhar et al., 1992), (8) groundwater related observations are often sparse and incomplete as they include an inadequate set of points in space and time to fully characterize the groundwater system (Rajabi and Ataie-Ashtiani, 2016), and (9) groundwater related observations are commonly indirect and have scaling differences with model parameters, states and external forcing terms (Carrera et al., 2005; Liu and Gupta, 2007).

2.2. Uncertainty in Model-Data Interaction

MDI has proven to be promising in improving understanding, quantification, communication and reduction of uncertainty in both models and data (Freeze et al., 1992; Porter et al., 2000; Liu et al., 2012). In many cases, uncertainty analysis is embedded in the MDI process. Sources of uncertainty in groundwater modeling can be classified into the following four interconnected categories:

- Structural uncertainty which arises from the fact that models are inherently simplified and imperfect approximations of complex real-world processes (Neuman, 2003). Structural uncertainty includes conceptual uncertainty (e.g. characterization of heterogeneity patterns, type of boundary conditions, time regime, etc.), and mathematical uncertainty (arising from alternatives in the mathematical implementation of conceptual models) (Bredehoeft, 2005; Ma and La Pointe, 2011).
- Parameter uncertainty which results from the fact that groundwater model parameters are often aggregate quantities that should be estimated from sparse and sometimes indirect measurements or qualitative assessments across the heterogeneous geological domain (Yeh, 1986; Rajabi and Ataie-Ashtiani, 2014).
- Data uncertainty which arises from instrument uncertainty due to imperfect measurement devices, and representativeness uncertainty due to inconsistency in spatial and/or temporal scales between the measured variables and the associated model variable (Liu and Gupta, 2007).
- 4. Extrapolation uncertainty which stems from the temporal prediction errors in the estimation of future external forcing terms from past data (Dessai and Hulme, 2007).

When the past state of the groundwater system is being simulated, structural, parameter and data uncertainties collectively propagate through the model and result in model uncertainty.

Alternatively, when the future state is being predicted, all four sources of uncertainty affect predictive uncertainty (Liu and Gupta, 2007). Throughout this paper, we will refer to both as model output uncertainty. The above concepts are illustrated in **Fig. 3**.

There is some dispute regarding which source dominates model uncertainty in groundwater simulations. Some references suggest parameter uncertainty (e.g. Smith and Schwartz, 1981; Hendricks Franssen and Kinzelbach, 2009; Schöniger et al., 2012), and others suggest structural uncertainty (e.g. Bredehoeft, 2005; Højberg and Refsgaard, 200**5**; Refsgaard et al., 2012). This dispute is partly due to the different approaches employed in handling uncertainty. Nevertheless, there is a wide consensus that: (1) structural uncertainty is the most difficult to quantify (Refsgaard et al., 2006), (2) ideally all sources of uncertainty should be simultaneously addressed to avoid misleading uncertainty predictions (Pappenberger and Beven, 2006; Linde et al., 2017), (3) different uncertainty sources have distinct error characteristics and hence require different approaches to deal with (Rajabi and Ataie-Ashtiani, 2016), and (4) a challenging and important aspect of quantifying model output uncertainty is to account for the effect of interaction among different uncertainty sources (Liu et al., 2012).

A review of the literature shows that there are basically two approaches to handling different sources of uncertainty in groundwater MDI. In the first approach, different sources of uncertainty are explicitly lumped together and mapped into model parameters. Some have noted that this method can create bias in uncertainty estimation (Kerrou et al., 2008). The second approach is to address different sources of uncertainty separately, according to their specific characteristics. The most common way to do this is to characterize parameter, data and extrapolation uncertainties as continuous probability distributions (often normal, log-normal or uniform) (Rajabi et al., 2015b). Structural uncertainty is commonly described by discrete scenarios for the

model structure, with probabilities assigned to the scenario's existence (Refsgaard et al., 2012). This results in multi-model ensembles which are formed, for example, by defining different parameterizations of the aquifer system (e.g. Feyen et al., 2001). Since the selection of individual models in the ensemble is mostly based on expert insight rather than formal methods, this approach is quite subjective and the scenarios may not represent a complete sampling of the model space (Beven, 1993; Liu et al., 2012). Hence obtaining reliable uncertainty estimates becomes a matter of both chance and experience. Nonetheless, since no other commonly accepted approach exists, this remains the dominant approach.

2.3. Review of Synthesis Methods

In this sub-section we review and classify a number of key synthesis methods previously used in the groundwater literature for MDI. A map of these methods is provided in **Fig. 1** to guide the interested reader through this sub-section. The review is meant to provide an overview of methods, and so the mathematical descriptions are kept to a minimum and references are provided for further reading.

2.3.1. Manual Insertion, Direct and Indirect Methods

In the simplest observation scenario, the model structure and values of all model parameters (θ), external forcing terms (u_t) and initial states (Y_t) are measured or inferred from data. In other words, adequate data is available to fully characterize the groundwater system, perform model simulations by manual insertion, and estimate the state variables at a later time (Y_{t+1}) at all points in the discretized domain of the problem (Kerrou et al., 2008). The resulting simulations can be augmented with forward uncertainty propagation analysis (UPA) to characterize the effect

of measurement noise on the uncertainty in state estimations (e.g. Rajabi and Ataie-Ashtiani, 2014; Rajabi et al., 2015a).

An alternative observation scenario is that data on some model parameters and external forcing terms is sparse, but the state (*Y*) is known exhaustively at all nodes of the discretized domain at times *t* and *t* + 1. In this scenario the unknown variables of the model can be estimated by the direct approach of solving the inverse problem, formulated simply as $\theta = f^{-1}(Y_t, u_t, Y_{t+1})$, often neglecting the stochastic ω_t term (Neuman, 1973). Only if the relationship between model inputs and state variables are assumed to be linear, can the solution be calculated by closed-form matrix expressions (Wang et al., 2009). Even so, a key challenge of the direct method is the ill-posedness and the singularity of the matrices involved in the numerical formulation (Zhou et al., 2014). Some researchers have tried to alleviate this problem by building overdetermined systems of equations (i.e. systems that have more equations than unknowns) (e.g. Ponzini and Lozej, 1982).

These two scenarios (namely having adequate data to fully characterize either all model inputs or the state variables) are mostly relevant to hypothetical and laboratory scale problems and are rare, if not nonexistent, in real-world cases where data on various model inputs and state variables are typically sparse due to physical and financial limitations (Rajabi and Ataie-Ashtiani, 2016). In the case of data scarcity, the indirect approach should be applied. There are two broad classes of indirect approaches in MDI: frequentist (also known as classical or non-Bayesian) and Bayesian (or probabilistic). We will use this classification in the subsequent subsections, although other classifications for indirect approaches also exist, which include batch vs. sequential methods (e.g. Wang et al, 2009) and optimization vs. sampling methods (e.g. Zhou et al., 2014).

2.3.2. Frequentist Approach

The frequentist approach assumes that model variables (e.g. parameters θ) are possibly unknown but have deterministic values, and the only source of randomness is data uncertainty. Based on this assumption, the frequentist approach tries to construct a point estimate for each unknown variable, often along with its confidence interval that quantifies the accuracy of the estimation process (Bernardo and Smith, 2001). These estimated values are then used for deterministic state reconstruction or prediction. The frequentist approach assumes no prior distribution for the unknown variables, but many algorithms used in the frequentist approach allow for use of prior information to provide initial values and lower or upper bounds on the variations of the unknown parameters, or to penalize departures from prior estimates (Alcolea et al., 2006).

The most popular method in the framework of the frequentist approach, and historically the first widely successful inverse method in groundwater applications is maximum likelihood estimation (MLE) (Carrera and Neuman, 1986a). MLE it is not based on the linearity assumption, and does not depend on any approximation for the relationship between the state variables and the model inputs, and it can also incorporate many types of data. MLE estimates parameters θ (or external forcing u_t) in such a way that the model response (i.e. state variable) fits the data D_Y in some optimal sense. In other words, the most likely parameter values are those that maximize the likelihood of observing the data:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} L(\theta; D_Y) \tag{3}$$

Where $argmax_{\theta}$ refers to the input parameters θ (or similarly u_t), called arguments, at which the function outputs are as large as possible, and *L* denotes likelihood. A common way to formulate the likelihood function can be expressed as follows (Zhou et al., 2014):

$$L(\theta; D_Y) \propto exp\left\{-\frac{1}{2}\sum_{i=1}^{N_{obs}} (f_i(Y_{t-1}, u_t, \theta) - D_{Y_i})^T C_i^{-1} (f_i(Y_{t-1}, u_t, \theta) - D_{Y_i})\right\}$$
(4)

where C_i^{-1} is the corresponding covariance of the observation errors and N_{obs} is the number of measurements. In order to facilitate statistical analysis, the measurement error is often assumed to be Gaussian in MLE (McLaughlin and Townley, 1996).

Least squares estimation (LSE) is a special case of MLE based on the assumption that errors in data are independent and normally distributed with constant (unknown) variance. Nonlinear LSE algorithms were first applied to distributed parameter groundwater problems when numerical models became widely available in the 1960s and 1970s. In LSE the sum of squared errors (L_2 norm) between the observed values and values predicted by the model are minimized. The ordinary least squares estimator is defined as:

$$\hat{\theta}_{LES} = argmin_{\theta} \sum_{i=1}^{n} [D_{Y_i} - f_i(Y_{t-1}, u_t, \theta)]^2$$
(5)

Partly due to its simplicity and the ease of implementation, LSE is widely used for parameter estimation in groundwater modeling applications and has been employed in several highly popular inverse codes (see section 4). In weighted LSE, weighting factors are used to express the relative magnitude of values and confidence in data. A regularization or plausibility term can be included in the objective function above, in order to ensure stability of the optimization problem (Medina and Carrera, 1996). LSE is highly sensitive to outliers and hence some studies employ other frequentist objective functions such as absolute value of the difference between measured and computed values (L_1 norm) (Woodbury et al., 1987), Nash-Sutcliffe index (e.g. Mugunthan et al., 2005) and Minkowski distance function (e.g. Zhou et al., 2012).

MLE, LSE and other similar frequentist optimization problems fall in the field of calculus of variations and so they are sometimes referred to as variational methods (Rayner et al., 2016). Optimization algorithms used in the context of the frequentist approach include:

- 1. Local optimization methods which often quickly converge to the optimum if the search is started from a point in sufficient proximity of the optimum, but only guarantee local convergence as they do not have a mechanism to escape from local optima (Mugunthan et al., 2005). Local methods used in the context of groundwater MDI include: (1.1) Derivative-based (or gradient-based) methods such as the widely used Levenberg-Marquardt (LM) algorithm and its modifications (e.g. Nowak and Cirpka, 2004) and the conjugate gradient method (e.g. Carrera and Neuman, 1986b), which are both highly efficient especially for low dimensional problems and have been widely incorporated in popular inverse codes. These methods can fail if the objective function is discontinuous (or their derivatives are discontinuous), non-smooth, multi-modal or ill-conditioned. (1.2) Derivative-free methods which are used when derivative information is unavailable, unreliable or impractical to obtain due to the above mentioned problems (Rios and Sahinidis, 2013). Examples of these methods are pattern search (Hooke and Jeeves, 1961) including generalized pattern search and mesh adaptive direct search (e.g. Zhou et al., 2012; Haddad et al. 2013) and Nelder-Mead simplex algorithm (Nelder and Mead, 1965) (e.g. Lambot et al., 2002).
- 2. Global optimization methods which are more likely to find the global optima for the objective functions in comparison to local methods (Wang et al., 2009). Global methods can be further classified into: (2.1) deterministic (or exact) methods, such as multilevel coordinate search (Huyer and Neumaier, 1999) (e.g. Lambot et al., 2002, 2004), and

branch-and-bound method (Huyer and Neumaier, 2008) (e.g. Sun et al., 2006), and (2.2) stochastic search algorithms which rely on probabilistic search rules to find good solutions and can locate the neighborhood of the global optima relatively fast, but their efficiency comes at the cost of computational effort and inability to guarantee global optimality. Stochastic optimization methods used in the context of groundwater MDI are mostly Meta-heuristic methods. Meta-heuristics are a group of both local and global optimization algorithms that are inspired by natural processes (Ketabchi and Ataie-Ashtiani, 2015c). These include evolution as in genetic algorithm (Holland, 1975) (e.g. Samuel and Jha, 2003) and derandomized evolution strategies (Hansen and Ostermeier, 2001) (e.g. Bayer and Finkel, 2004), social behavior of biological organisms as in ant colony optimization (Dorigo and Stützle, 2004) (e.g. Abbaspour et al., 2001), particle swarm optimization (Kennedy and Eberhart, 1995) (e.g. Haddad et al. 2013) and controlled cooling associated with a physical process as in simulated annealing (e.g. Tsai et al., 2003). Meta-heuristic algorithms can find acceptable solutions in a reasonable time in both complex and large search domains (e.g. Ketabchi and Ataie-Ashtiani, 2015c).

Local and global methods can be hybridized into efficient optimization methods. For example LM can be combined with a stochastic quasi-Monte Carlo algorithm to search for global optimal values (Peng et al., 2011).

2.3.3. Bayesian Approach

The Bayesian approach (also referred to as Bayesian Inference) has two key distinctions with the frequentist approach: first it formally considers model inputs and outputs as random variables and hence formulates the entire problem in a probabilistic framework (Fasbender et al., 2008), and second, it allows for formal consideration of prior information in the inference process

(Wang et al., 2009). As discussed throughout the rest of the paper, these two key distinctions provide a convenient framework for uncertainty analysis, data fusion, regularization, data worth analysis and incorporation of soft data in groundwater applications. We shall first present the formulation of Bayesian inference as a parameter estimation problem, and then generalize it to other model components. We denote the prior distribution of the parameter set θ with $P(\theta)$ which is prior beliefs on the parameter values before employing a specific dataset. We also represent the likelihood function that characterizes the likelihood of the observations set D_Y given a certain parameter set θ by $P(D_Y | \theta)$. The distribution of observation D_Y given θ is tied to the measurement model. In Bayesian inference, the posterior distribution $P(\theta | D_Y)$ is obtained through the application of Bayes theorem (Gamerman and Lopes, 2006):

$$P(\theta \mid D_Y) = \frac{P(D_Y \mid \theta).P(\theta)}{P(D_Y)}$$
(6)

where $P(D_Y)$ is the proportionality (or normalization) constant which characterizes the evidence or the marginal probability of the data, and can be computed from (Gamerman and Lopes, 2006):

$$P(D_Y) = \int P(D_Y | \theta) . P(\theta) d\theta$$
(7)

The lack of prior information can be expressed by using a non-informative prior (Gelman et al., 2013). In the case where there are multiple measurements (e.g. measurements at different times (1, 2,...,*T*) denoted by $D_{Y_{1:T}}$), the joint likelihood of all measurements is the product of distributions of individual measurements if the measurements are assumed to be conditionally independent. The resulting posterior distribution is (Box and Tiao, 2011):

$$P(\theta \mid D_{Y_{1:T}}) = \frac{1}{Z} P(\theta) \prod_{k=1}^{T} P(D_{Y_k} \mid \theta)$$
(8)

where *Z* is the normalization constant given by:

$$Z = \int P(\theta) \prod_{k=1}^{l} P(D_{Y_k} | \theta) d\theta$$
(9)

The predictive posterior distribution is the distribution of not yet observed state variables when all the information in the observed measurements and the model is used. The predictive posterior distribution can be estimated as follows (Box and Tiao, 2011):

$$P(Y_{T+n} \mid D_{Y_{1:T}}) = \int P(Y_{T+n} \mid \theta) . P(\theta \mid D_{Y_{1:T}}) d\theta$$
(10)

The Bayesian posterior distribution can be reduced to point estimates through a host of methods which include choosing the mean or maximum of the posterior distribution. The latter approach is often called the maximum a posteriori (MAP) estimate (e.g. Kowalsky et al., 2004). MLE can be seen as a MAP estimate with uniform prior $P(\theta) \propto 1$ on the parameter θ . In other words, the Bayesian approach provides a formal way of including prior information and regularization terms into MLE (Särkkä, 2013).

The same notion can be applied to the estimation of all model components, based on the sequence of conditional dependence described as $\{f \rightarrow (\theta \text{ and } u_t) \rightarrow Y_t\}$ (dependent variables appear at the end of the arrows) (Liu and Gupta, 2007):

$$P(f \mid D_Y) = \frac{P(D_Y \mid f).P(f)}{P(D_Y)}$$
(11-1)

$$P(\theta \mid D_{Y}, f) = \frac{P(D_{Y} \mid f, \theta) . P(\theta \mid f) P(f)}{P(D_{Y}, f)}$$
(11-2)

$$P(u_{t} \mid D_{Y}, f, \theta) = \frac{P(D_{Y} \mid u_{t}, f, \theta) . P(u_{t} \mid f, \theta) . P(\theta \mid f) . P(f)}{P(D_{Y}, f, \theta)}$$
(11-3)

$$P(Y_{t} \mid D_{Y}, f, \theta, u_{t}) = \frac{P(D_{Y} \mid Y_{t}, f, \theta, u_{t}) . P(Y_{t} \mid f, \theta, u_{t}) . P(u_{t} \mid f, \theta) . P(\theta \mid f) . P(f)}{P(D_{Y}, f, \theta, u_{t})}$$
(11-4)

2.3.3.1. Batch vs. Recursive Bayesian Inference

When various types of data obtained at different time steps are simultaneously taken into account as a single whole dataset, the resulting posterior distribution is denoted by batch Bayesian estimation. This type of solution to Bayesian inference is very common in groundwater applications (e.g. Hassan et al., 2009; Laloy et al., 2013; Rajabi and Ataie-Ashtiani, 2016). However, this full posterior formulation has the disadvantage that each time we obtain a new measurement, the full posterior distribution must be recomputed. This is particularly a problem in dynamic estimation where measurements are typically obtained one at a time and we would want to compute the best possible estimate after each measurement (Särkkä, 2013). So alternatively, if we treat the posterior distribution obtained from data for the previous time step as the prior for the current time step, the result is called recursive or sequential Bayesian estimation or Bayesian filtering. The recursive formulation of Bayesian estimation has several useful properties: (1) parameter estimates can be updated gradually as soon as new observations arrive, paving the way for online learning. This is particularly useful when the problem is sequential by nature (Wang et al., 2009) as in a groundwater plume with time-variable source or when data is collected gradually and it needs to be incorporated into model estimations without having to solve the problem from the start (El Gharamti et al., 2013; Zhou et al., 2014). (2) It

allows for reducing the computational size of the problem when the problem is computationally expensive (Bruhwiler et al., 2005). (3) It allows for considering the temporal evolution of parameter values, i.e. parameters are assumed to be time-dependent stochastic processes (hence θ (*t*) instead of θ) (Särkkä, 2013). (4) It does not require storage of all past information about the states and parameters (Moradkhani et al., 2005; El Gharamti et al., 2013).

2.3.3.2. Numerical Methods for Batch Bayesian Inference

In practical problems involving the estimation of parameters, external forcing terms or model states, exact analytical solutions for the continuous posterior distribution in batch Bayesian inference are available for very limited combinations of model forms and probability functions, such as the normal linear model (Qian et al., 2003; Khaleghi et al., 2013). This necessitates the use of numerical approximation techniques. Markov chain Monte Carlo (MCMC) methods (or samplers) are a general class of strategies that provide a powerful tool for numerical approximation of the posterior distribution in batch Bayesian inference. MCMC came into widespread use as a tool for Bayesian inference in science and engineering with the work of Tanner and Wong (1987) and Gelfand and Smith (1990). They were introduced into the groundwater literature by Oliver et al. (1997), and have been the dominant method for numerical approximation of batch Bayesian inference in groundwater modeling applications ever since. As the name suggests, MCMC methods employ Monte Carlo concepts in the framework of Markov chains. They attempt to generate Monte Carlo samples from the posterior distributions conditioned on the observations, by a special sequential process in which each new sample depends on the properties of the sample drawn immediately before it (and not the more early ones), thus creating a Markovian chain of samples. In MCMC, these samples can be drawn from a distribution even if all that is known about the distribution is how to estimate the probability

density for the samples (Gamerman and Lopes, 2006; van Ravenzwaaij et al., 2016), which is a very beneficial aspect of MCMC as the posterior distributions are unknown at the onset of calculations. If the Markovian chain is sufficiently large (often on the order of 10³ to 10⁵ samples in previous groundwater applications e.g. Hassan et al., 2009; Laloy et al., 2013; Rajabi and Ataie-Ashtiani, 2016), it will eventually converge to the stationary, posterior distribution of the parameters (Andrieu et al., 2003). Estimation of the probability density for each sample constitutes a forward model simulation and hence when dealing with CPU-intensive groundwater numerical models, MCMC methods become extremely demanding in terms of the required computations (Hendricks Franssen and Kinzelbach, 2008).

The main difference amongst various MCMC methods is how sampling is implemented based on these general concepts. There are two central notions used for building samplers (Tierney and Mira, 1999): (1) dimension reduction by conditioning, which is the foundation of the Gibbs sampler (e.g. Michalak, 2008), and (2) proposal and rejection, which is the basis of the Metropolis-Hastings (e.g. Laloy et al., 2013) and similar algorithms such as adaptive Metropolis (AM) (e.g. Hassan et al., 2009; Wu et al., 2011) and delayed rejection adaptive Metropolis (DRAM) (e.g. Rajabi and Ataie-Ashtiani, 2016). These two notions have also been combined into what is called Metropolis-within-Gibbs samplers (e.g. Cui et al., 2013). Generally speaking, there is no universally optimal MCMC algorithm and the choice of an appropriate MCMC algorithm depends on the specific nature of the problem. Nevertheless, some researchers have proposed MCMC algorithms that seem to offer great potential for increasing the computational efficiency in comparison with more traditional MCMC algorithms. These novel MCMC algorithms often rely on increasing the acceptance rates of proposals (e.g. Vrugt et al., 2009),

dimensionality reduction of the forward simulations (e.g. Laloy et al., 2013) or parallelization of MCMC computations (e.g. Laloy and Vrugt, 2012; Joseph and Guillaume, 2013).

An alternative to MCMC is importance sampling (IS) (e.g. Ng et al., 2009). IS is an algorithm for generating weighted samples (i.e. particles) from the posterior distribution. The main difference between IS and MCMC is that each of the particles in IS has an associated weight, which reflects its ability to match observations (Lu and Zhang, 2003; Särkkä, 2013). Similar to MCMC, IS does not rely on the implicit Gaussian and linearity assumptions. But IS can be inefficient if the unconditional probability density of the states is a weak approximation of the conditional density, and can especially become computationally infeasible when the state vector is high-dimensional and particles are computationally demanding (Daum and Huang, 2003).

2.3.3.3. Bayesian Filtering

The objective of Bayesian filtering is to estimate the time-varying state of the system (or parameters and external forcing terms as described in section 3) which is observed through sparse and noisy measurements (Chen, 2003). In Bayesian filtering, the marginal posterior distributions (or filtering distributions) of the state variables at time step k (i.e. Y_k) are computed using the history of measurements up to and including the time step k. Bayesian filtering generally includes two stages: a prediction stage and an update stage. In the prediction stage, the distribution of the state Y_k is computed based on measurements in previous time steps by (Doucet et al., 2001; Chen, 2003):

$$P(Y_k \mid D_{Y_{1:k-1}}) = \int P(Y_k \mid Y_{k-1}) P(Y_{k-1} \mid D_{Y_{1:k-1}}) dY_{k-1}$$
(12)

In the update stage, the new measurements D_{Y_k} is used to update the distribution obtained in the prediction stage by applying Bayes rule:

$$P(Y_k \mid D_{Y_{1:k}}) = \frac{1}{Z_k} P(D_{Y_k} \mid Y_k) . P(Y_k \mid D_{Y_{1:k-1}})$$
(13)

where Z_k is the normalization constant, given by:

$$Z_{k} = \int P(D_{Y_{k}} | Y_{k}) P(Y_{k} | D_{Y_{1:k-1}}) dY_{k}$$
(14)

2.3.3.4. Closed-form Solutions for the Bayesian Filter

A few classes of filtering problems have closed-form solutions, with the Kalman filter (KF) (Kalman, 1960) being the most popular one. KF is based on a linear and Gaussian assumption for the state space models, described as follows (Eigbe et al., 1998):

$$Y_k = \Phi_{k-1} Y_{k-1} + q_{k-1}$$
 (15-a)
 $D_{Y_k} = H_k Y_k + r_k$ (15-b)

Where $q_{k-1} \sim N(0, Q_{k-1})$ is the process noise, $r_k \sim N(0, R_k)$ is the measurement noise, and the initial distribution is assumed to be $Y_0 \sim N(m_0, \psi_0)$. Φ_{k-1} is the transition matrix describing the forward model and H_k is the measurement model matrix. Starting from the initial distribution of the state, the prediction stage of the KF tries to estimate the mean and covariance of the state for the next time step by (Eigbe et al., 1998):

$$\tilde{m}_{k} = \Phi_{k-1} m_{k-1}$$
 (16-a)
 $\tilde{\psi}_{k} = \Phi_{k-1} \psi_{k-1} \Phi_{k-1}^{T} + Q_{k-1}$ (16-b)

The update stage subsequently revises these estimations by employing data for time step k (D_{Y_k}) as follows (Eigbe et al., 1998):

$$m_{k} = \tilde{m}_{k} + L_{k}(D_{Y_{k}} - H_{k}\tilde{m}_{k})$$

$$\psi_{k} = \tilde{\psi}_{k} - L_{k}(H_{k}\tilde{\psi}_{k}H_{k}^{T} + R_{k})L_{k}^{T}$$
(17-b)

where L_k is the Kalman gain matrix (Wang et al., 2009). KF can be further simplified by empirical forcing of the model fields toward the observed values which is called nudging (Lahoz et al., 2007).

The use of KF in groundwater MDI problems dates back to the 1970s (e.g. McLaughlin 1976), and it has been applied to a range of applications ever since. However, the popularity of the classic KF in groundwater literature is considerably less than the more general domain of hydrology, where KF is commonly applied to rainfall-runoff modeling, flood forecasting, rainfall prediction, etc. (see Liu et al., 2012). The reason is that KF, in its most basic form, has several key limitations which are particularly problematic in groundwater applications: (1) it is based on the linearity assumption whereas groundwater flow and solute transport models are nonlinear (Zhou et al., 2014), (2) it cannot break down the computations over space the way it does over time, and hence cannot account for the spatial physical and statistical relationships of groundwater systems (Porter et al., 2000), (3) there is often not enough information about error structures to fill the large matrixes of Eq. (17-b) with meaningful numbers (Reichle, 2008), and (4) it is computationally expensive for high-dimensional problems and is hence strongly restricted to small size problems (Hendricks Franssen and Kinzelbach, 2008).

Variational Bayesian methods (VBMs), notably the sequential 4DVAR method and its batch counterparts 1DVAR and 3DVAR (where 1D and 3D refer to one or three spatial dimensions, the 4th dimension in 4DVAR is time (Reichle, 2008), and VAR stands for variational data assimilation), can be viewed as simplifications of the KF, as they do not propagate the state covariance matrix explicitly (Liu et al., 2012). VBMs have relatively low computational demand and are preferred for use with computationally expensive models and large-dimensional problems with poorly related nonlinear observations (Rawlins et al., 2007). However, VBMs do not provide an estimate of predictive uncertainty by themselves (Peng et al., 2011). VBMs have been used extensively in the numerical weather prediction community. One of the rather few applications in groundwater MDI is Kabir et al. (2017) which employs 3DVAR and 4DVAR for state estimation in a hypothetical test case. More applications of VBMs are expected in future groundwater studies.

2.3.3.5. Numerical Approximations for the Bayesian Filter

Due to the limiting nature of closed-form solutions and the intractability of the Bayesian filtering equations (Särkkä, 2013), several numerical approximation techniques have been proposed. The most popular of such techniques in the groundwater literature are briefly reviewed in the following paragraphs.

Extended Kalman filter (ExKF) (Gelb, 1974) is a numerical approximation method which forms a Taylor series expansion at the nominal (or MAP) solution to provide a linear approximation of the nonlinear and non-Gaussian state space models. ExKF has been applied to groundwater MDI problems by e.g. Leng and Yeh (2003) and Yeh and Huang (2005). ExKF has several major setbacks that affect its application to groundwater MDI: (1) the covariance approximation deteriorates with time, particularly when the Taylor series approximation is poor as in highly

heterogeneous fields (Zhou et al., 2014), (2) it can be computationally very demanding due to the error covariance propagation (Evensen, 2003) especially for finely discretized groundwater models, and (3) it can lead to unstable results or even divergence when the nonlinearity in the system is strong (Miller et al., 1994; Moradkhani et al., 2005). Due to these setbacks, the ExKF has lost popularity in the groundwater literature.

Ensemble Kalman filter (EnKF) is another method which was first introduced by Evensen (1994) and later modified by Burgers et al. (1998), and is a Monte Carlo variant of the KF. In EnKF, the probability distribution of state (Y_k) is represented by ensemble of realizations [$Y_k^1, Y_{k,...,Y_k^N}^2$]. This ensemble is built by sampling from the known distribution of Y_0 in the first step of computations. Each of these realizations is then separately propagated through time in the subsequent steps by a two-stage prediction/update procedure. In each time step, the mean and covariance can be explicitly computed based on the realizations, noting that the prediction stage involves a model simulation for each realization in the ensemble. Observations are treated as random variables by generating an observation ensemble with a mean equal to the actual observation at each time and using a predefined covariance. The EnKF also represents the model errors by generating perturbations at each time step (Evensen, 2003, 2009).

The EnKF has a number of advantages which are particularly desirable in groundwater MDI problems: (1) it can handle modestly nonlinear state space models (Reichle et al, 2008), (2) it can be used to deal with high-dimensional problems (Peng et al., 2011), (3) it is flexible in its treatment of errors in model dynamics and parameters (Reichle, 2008), and (4) the necessary covariance between states at any given time step are estimated efficiently from a limited ensemble of stochastic realizations without requiring sensitivity analyses (Schöniger et al.,

2012). In general, the performance of EnKF is affected by the choice of the ensemble size and generation method, model characteristics and analysis scheme (Moradkhani et al., 2005). Due to these advantages and its ease of implementation, model-independence, and robustness in solving different types of problems encountered in groundwater applications, the EnKF has become more popular than any other method in addressing problems of a sequential nature in the groundwater literature. Examples for the use of the EnKF in groundwater applications include Chen and Zhang (2006), Drécourt et al. (2006), Liu et al. (2008), Hendricks Franssen et al. (2011), El Gharamti et al. (2013), Erdal and Cirpka (2017) and Xu and Gómez-Hernández, (2018). However, the basic form of EnKF has the key disadvantage of being based on the Gaussian assumption and non-normally distributed state variables. Highly nonlinear dynamics can result in increasing underestimation of variance over time (called filter inbreeding) and filter divergence in the EnKF (Hendricks Franssen and Kinzelbach, 2009). A number of studies in the groundwater literature have tried to make EnKF applicable to non-Gaussian models. For example, Schöniger et al. (2012) proposed the use of nonlinear, monotonic transformations to the observed states, rendering them Gaussian. Xu and Gómez-Hernández (2016b) employed the normal-score EnKF to jointly estimate non-Gaussian aquifer parameters by assimilating three kinds of state variables. Another important disadvantage of EnKF is that the size of the ensemble can be computationally prohibitive for CPU-intensive groundwater models, because hundreds of ensemble members are usually needed for reliable updating without filter inbreeding (Liu et al., 2012).

The EnKF has a number of variants such as the ensemble square root Kalman filter (EnSRF) (Whitaker and Hamill, 2002), which uses the traditional Kalman gain for updating the ensemble mean, but uses a reduced Kalman gain to update deviations (Leisenring and Moradkhani, 2011).

The use of EnSRF and other variants of the EnKF such as the adaptive ensemble Kalman filter and the hybrid adaptive ensemble Kalman filter in the groundwater literature is rare (as one of the few examples see Rajib et al., 2017).

Sequential Monte Carlo (SMC) methods (Gordon et al., 1993), also known as particle filters and bootstrap filters, are another group of numerical approximation techniques that represent the posterior distribution in recursive Bayesian inference as a set of Monte Carlo samples with associated weights, which are termed particles. The particles sample the state (and/or parameter and external forcing) space according to a given prior distribution. The particles are then propagated forward in time, with SMC performing updates on particle weights instead of state variables (Peng et al., 2011). This has the advantage of reducing numerical instability (Ristic et al., 2003). SMC is applicable to any state space model with any format and statistical distribution, whether linear or nonlinear, and Gaussian or non-Gaussian (Han and Li, 2008). This is a highly advantageous characteristic in certain groundwater problems, such as the inverse estimation of hydraulic conductivities for non-multi-Gaussian media. However, the basic form of SMC has two key undesirable characteristics: (1) SMC tends to assign very large weights to the few particles with strong data support, leading to severe reduction of the effective sample size and hence deterioration of the statistics, and (2) applying SMC to a high-dimensional state and parameter space requires a very large number of particles and hence a very large amount of CPU time (Snyder et al., 2008; Hendricks Franssen and Kinzelbach, 2009; Schöniger et al., 2012). Resampling techniques like sequential importance resampling can reduce these problems to some extent (Van Leeuwen, 2009; Leisenring and Moradkhani, 2011). SMC may outperform EnKF when the number of particles is sufficiently large (> 100). But the often larger number of simulations required by SCM compared to EnKF (Liu et al., 2012) has limited its application to

CPU-intensive real-world groundwater models. Example of SMC application in groundwater MDI problems includes Chang et al (2012) and Abbaszadeh et al. (2018). SMC is popular in some engineering fields such as tracking and signal processing (Djuric et al., 2003), and it has also been applied to many hydrologic problems (Zhou et al., 2006; Smith et al., 2008). SMC can be combined with MCMC (see Moradkhani et al., 2012) so that a desirable performance can be achieved with a small, manageable ensemble size.

Apart from the three families of methods described above, other techniques for numerical approximations of the Bayesian filter exist but are rarely applied to groundwater MDI. An example is the unscented Kalman filter (UnKF) (e.g. Chang and Sayemuzzaman, 2014), which presumes that the state space is unimodal, symmetric and unbound. Despite the more relaxed assumptions of the UnKF, it has received little attention in the groundwater literature to date, and so we expect more applications in the future.

2.3.3.6. Bayesian Smoothing

While Bayesian filters in their basic form only compute estimates of the current state of the system given the history of measurements, Bayesian smoothers can be used to reconstruct states that happened before the current time. Smoothing distributions computed by the Bayesian smoothers are the marginal distributions of the state Y_k given a certain interval of measurements $D_{Y_{1:k}}$ (k = 1, 2, ..., T) in which T > k (Kitagawa, 1987):

$$P(Y_k \mid D_{Y_{1;T}}) = P(Y_k \mid D_{Y_{1;k}}) \int \left[\frac{P(Y_{k+1} \mid Y_k) . P(Y_{k+1} \mid D_{Y_{1;T}})}{P(Y_{k+1} \mid D_{Y_{1;k}})} \right] dY_{k+1}$$
(18)

where $P(Y_k | D_{Y_{1;k}})$ is the filtering distribution of the time step *k*. The integration is replaced with summation if some of the state components are discrete (Särkkä, 2013). Similar to Bayesian

filtering, there are both closed-form and numerical solutions for the Bayesian smoothing equation. Closed-form solutions include the Kalman smoother (KS) or the Rauch–Tung–Striebel smoother (RTSS) (Rauch et al., 1963) for linear Gaussian state space models. Numerical solutions include the extended Rauch–Tung–Striebel smoother (ExRTSS) (Cox, 1964; Sage and Melsa, 1971), which assumes Gaussian approximation to the smoothing distribution and the ensemble smoother (ES) (Van Leeuwen and Evensen, 1996). Among these, the most commonly used method in the groundwater literature is ES. Similar to EnKF, the ES employs an ensemble of realizations obtained through MCS, but it is based on one-in-all conditioning and can assimilate all the measurements at once, using a single analysis step rather than the stepwise conditioning of the EnKF. So the ES (and other Bayesian smoothers) are also classified as batch methods in some references. However note that employing the ES instead of the EnKF increases the nonlinearity of the parameter update step during data assimilation, and also the need for iterations of the algorithm (Chen and Oliver, 2013). Examples for the use of ES in the groundwater literature includes Bailey and Baù (2010, 2012), Chang et al. (2017) and White (2018).

3. Applications

Applications of MDI in groundwater literature can be classified into eight categories: parameter estimation and uncertainty quantification, state estimation and uncertainty quantification, stateparameter estimation and uncertainty quantification, model selection and structural uncertainty analysis, guidance of data collection and data worth analysis, improving data coverage, customizing data resolution, and fusion of heterogeneous data. These applications are not mutually exclusive, and two or more applications may be intended simultaneously. In this

section, we review these applications considering their objectives, basic theoretical concepts, types of algorithms used in each application, and significant trends in the groundwater literature.

3.1. Parameter Estimation and Uncertainty Quantification

Parameter estimation is referred to as parameter fitting (optimization, inference or tuning), model calibration or inverse modeling in the groundwater literature (Carrera et al., 2005 and 2010; Ataie-Ashtiani et al., 2013). Its objective is to employ available data (either directly related to the parameter of interest, or associated with state variables) to provide reasonable estimates of model parameters and external forcing terms over a past period, conditional on the specification of model structure, so that the model makes sufficiently accurate predictions of the true state of the system (Liu and Gupta, 2007). Parameter estimation often also includes approximation of associated uncertainties.

In both research and practical groundwater modeling studies, parameter estimation has long been the dominant method for model improvement by conditioning to data and is often considered a necessary step in groundwater modeling (Hill and Tiedeman, 2006). It is by far the most explicitly mentioned application of MDI in the groundwater literature. There are several reasons for this profound attention toward parameter estimation in the groundwater literature: first, no reliable prediction can be made without proper estimation of model parameters and external forcing terms (Zhou et al., 2014), and second, it is often much easier to condition parameter and external forcing values to data, as compared to model structure (Refsgaard et al., 2006 and 2012).

Groundwater parameter estimation often involves the estimation of hydraulic conductivity (or related parameters such as permeability or transmissivity) and its spatial distribution (e.g. Hendricks Franssen and Kinzelbach, 2009; Hendricks Franssen, et al., 2009; Ataie-Ashtiani et al.

2013) as it highly affects model outputs for the prediction of both groundwater flow and transport of contaminants, and also because it may vary over many orders of magnitude in a relatively small volume of the media (McLaughlin and Townley, 1996). However, parameter estimation may also focus on other parameters and external forcing terms such as the recharge rate (Ng et al., 2009; Hendricks Franssen et al., 2004, 2008; Hendricks Franssen and Kinzelbach, 2008), discharges, fluxes, leakage coefficients and piezometric heads on designated boundaries (e.g. Liu et al., 2009; Hendricks Franssen et al., 2011; Irsa and Zhang, 2012), pollutant source location and release histories (e.g. Sun et al. 2006; Hendricks Franssen et al., 2011), and dispersivity (e.g. Ataie-Ashtiani et al., 2013), or a combination of different parameters (e.g. Xu and Gómez-Hernández, 2016a, b, 2018). Amongst these, the estimation of recharge rate is of key interest in arid and semi-arid regions due to its more unpredictable nature in these areas (Hendricks Franssen et al., 2008).

Parameter estimation has been a topic of intensive research for the past decades and several notable review papers are available on the subject, including McLaughlin and Townley (1996), Carrera et al. (2005 and 2010), Vrugt et al. (2008) and Zhou et al (2014). As described in all these review papers, historically, the common approach for the estimation of groundwater model parameters until the end of the 1990s was trial-and-error, in which parameter values are manually changed until a reasonable match between model predictions and data is achieved. This approach is highly dependent on the subjective judgment of the modeler and by no means guarantees the optimal choice of parameter values (Ataie-Ashtiani et al., 2013). These key setbacks have resulted in increasing use of automated parameter estimation based on frequentist, nonlinear, batch methods. Despite the many advantages of using these methods instead of trial-and-error parameter estimation, they have several limitations: (1) these methods employ
objective functions that are based on collective measures of uncertainty and ignore the special characteristics of the individual components of uncertainty such as structural and data uncertainties (Liu et al., 2012; Zhou et al 2014), (2) pursuing a single optimal parameter set can create bias in model predictions (Hendricks Franssen et al., 2009), and (3) they do not have the ability to gradually reduce parameter (and hence prediction) uncertainty as new data become available (Liu and Gupta, 2007). In recognition of these limitations, there has been growing interest in the use of Monte Carlo-based Bayesian techniques for parameter estimation. Algorithms that have been of key interest include EnKF (e.g. Hendricks Franssen et al., 2011) and ES (e.g. Bailey and Baù, 2010, 2012). In Monte Carlo-based methods, likely realizations of the input parameters are conditioned to the available measurements with geostatistical techniques such as sequential Gaussian simulation (Gómez-Hernández and Journel, 1993) (if local measurements of the desired parameters are available) or co-located co-simulation (Almeida and Frykman, 1994; Hendricks Franssen et al., 2008).

The estimation of heterogeneous fields of hydraulic conductivity and other similar parameters (e.g. transmissivity, storage coefficient and recharge rate) from sparse data is commonly a highly underdetermined problem and is prone to ill-posedness, non-uniqueness and instability. To alleviate these problems, the estimation of such parameter fields is often done by employing a parameterization technique that limits the number of unknown variables and provides the modeler with only enough heterogeneity required to simulate observations of past system behavior (Sepúlveda and Doherty, 2015). These include non-geostatistical methods such as the classical zonation method (Carrera and Neuman, 1986a, b), and geostatistics-based methods such as the pilot points method (de Marsily, 1978), regularized pilot points method (Alcolea et al., 2006), sequential self-calibration (Gómez-Hernánez et al., 1997; Hendricks Franssen et al.,

1999), and ridge function (Mantoglou, 2003). Geostatistical methods may turn the parameter estimation problem into the estimation of the geostatistical variables of the unknown parameters (e.g. range, nugget, sill, etc.). Using these techniques always results in loss of detail and may produce overly smoothed parameter fields (Moore and Doherty, 2006). Hence much work is still being done by the groundwater research community to develop new methods.

3.2. State Estimation and Uncertainty Quantification

State estimation involves characterization of the past (retrospective), present or future (forecast) state of the groundwater system and their uncertainties, by combining state information from both the model and available data (Liu and Gupta, 2007; Liu et al., 2012). This may include, for example, reconstructing spatial flow and contaminant plume fields. State estimation which is also commonly referred to as data assimilation, is typically based on specification of the model structure and parameters in advance, and the estimations are solely applied to the state variables. Hence state estimations are conditional on the specific model structure(s) and parameter values (Moradkhani et al., 2005; Schöniger et al., 2012).

In the groundwater literature, state estimation is often formulated as a filtering or smoothing problem and is commonly interrelated with three other application of MDI: improving data coverage, customizing data resolution and data fusion. The commonly applied methodologies for solving the problem via filtering in recent groundwater literature is the EnKF and SMC methods (e.g. Bailey and Baù, 2012). Dual state-state estimation is a term used for the concurrent estimation of groundwater flow and contaminant states (El Gharamti et al., 2013).

3.3. Simultaneous State-Parameter Estimation and Uncertainty Quantification

There are basically two formulations for the simultaneous estimation of states and parameters in the groundwater literature: (1) the joint (or state augmentation) approach, and (2) the dual

estimation approach. The standard joint approach simultaneously estimates state and parameters as a single augmented vector (e.g. Chen and Zhang 2006; Hendricks Franssen and Kinzelbach 2008, 2009; Liu et al. 2008; Hendricks Franssen et al., 2011). The joint approach is very susceptible to instability and intractability as a result of increase in the number of unknown variables, especially in highly nonlinear systems (Moradkhani et al., 2005). The alternative dual formulation is based on two interactive parallel filters: a filter for the parameters and another for the states, with the parameters undergoing an artificial evolution (i.e. random walk) while waiting to be updated indirectly by the state variables data (e.g. El Gharamti et al., 2013). An example of dual estimation is the dual extended Kalman filter (dual-ExKF) (Thiemann et al., 2001).

3.4. Model Selection and Structural Uncertainty Analysis

Model selection, also known as model discrimination or identification, is a key part of multimodel approaches for the consideration of structural uncertainty (Höge et al., 2018). It involves using data to choose amongst various independent plausible alternative model structures (including governing equations, heterogeneity patterns, type of boundary conditions etc.) that best describe the relationship between model inputs and outputs (Refsgaard et al., 2006; Gupta et al., 2012). Model selection may also include assigning probabilities to the chosen model structures based on their ability to reproduce the available data and then combining predictions made by the chosen model structures to form a reliable description of the total prediction uncertainty. MDI for model selection and structural uncertainty analysis in the groundwater literature is mostly performed using one of the following Bayesian methods: generalized likelihood uncertainty estimation (GLUE) (Beven and Binley, 1992), Bayesian model averaging (BMA) (Draper et al., 1995) and maximum likelihood Bayesian model averaging (MLBMA)

(Neuman, 2003) or a hybridization of at least two of them (e.g. Rojas et al., 2010a). These methods and some of their applications are briefly reviewed in the following sub-sections.

3.4.1. Generalized Likelihood Uncertainty Estimation

GLUE is a conditional MCS technique that involves sequential implementation of the following steps: (1) defining alternative candidate model structures, (2) assigning appropriate prior parameter uncertainty distributions for each model structure, (3) performing MCS based on samples drawn from the parameter distributions of each candidate, (4) using a likelihood measure such as the Gaussian (Romanowicz et al., 1994), efficiency (Freer et al., 1996) and fuzzy type (Jensen, 2003) measures, to assess the resemblance of each simulation output with data on systems states, (5) selecting candidates with likelihoods above a specified threshold as "behavioral" models and setting the other candidates aside, (6) calculating weights for each behavioral model by normalizing the corresponding likelihood values in a way that all the weights sum up to one, and (7) estimating the likelihood-weighted probability distribution of model outputs (Beven and Freer, 2001; Liu and Gupta, 2007). In GLUE the likelihoods can be updated sequentially as new data becomes available. GLUE allows for the explicit assessment of model structure and parameter uncertainties, but does not account for data uncertainty. Applications of GLUE in groundwater MDI are numerous, with the majority of applications focusing on the uncertainty in geological structures and hydraulic conductivity patterns and to a lesser extent on recharge patterns, assuming that other components of the model structure (e.g. governing equations) are without uncertainty (e.g. Feyen et al., 2001; Morse et al., 2003; Hassan et al., 2008). This is also the case for the other two methods discussed in this sub-section (BMA and MLBMA). The GLUE methodology has been criticized for not having a likelihood function that is consistent with probability theory, and instead relying on less formal likelihoods that are

defined by the user without satisfying Bayes theorem, resulting in a loss of consistency in learning (Mantovan and Todini, 2006; Montanari et al., 2009).

3.4.2. Bayesian Model Averaging

BMA is a more formal Bayesian approach for combining predictions made by multiple model structures. The BMA predictive distribution of an output of interest *Y* given data *D*, is estimated by (Hoeting et al., 1999):

$$P(Y \mid D) = \sum_{i=1}^{n_m} P(Y \mid D, f) P(f \mid D)$$
(19)

where n_m is the number of alternative model structures. In BMA the predictive distribution of *Y* is an average of its prediction distributions for each alternative model structure $(P(Y \mid D, f))$ weighted by its posterior model probability $(P(f \mid D))$. BMA can be used to differentiate between prediction uncertainties arising from individual models; and is able to identify unfavorable models for model selection. Examples for the use of BMA for MDI in the groundwater literature include Tsai and Li (2008), Li and Tsai (2009), Ye et al. (2010) and Tsai (2010). BMA is computationally very demanding especially when applied to CPU-intensive groundwater models.

3.4.3. Maximum Likelihood Bayesian Model Averaging

MLBMA was developed in an attempt to make BMA computationally feasible (Neuman, 2003). It is in fact an approximation to BMA that relies on producing maximum likelihood parameter estimations and then expanding around these values by MCS. MLBMA subsequently approximates the posterior model probabilities using, e.g. the Kashyap information criterion (Kashyap, 1982) or the Bayesian information criterion (BIC) (Schwarz, 1978). MLBMA is

capable of dealing with lack of prior information on model parameters (Ye et al., 2005). Examples of groundwater studies using MLBMA for MDI include Neuman (2003), Ye et al. (2004 and 2005), and Lu et al. (2015). Despite the computational convenience, several shortcomings have been mentioned for MLBMA in the literature. These include: (1) it relies on the calibration of parameters for each alternative model structure hence creating the risk of biased parameter estimates that tend to compensate for errors in the model structure (Rojas et al., 2010b), and (2) prediction of state variables not included in the data used for calibration may become biased and may underestimate the effects of model structural uncertainty (Refsgaard et al., 2006; Troldborg et al., 2007).

3.5. Guidance of Data Collection and Data worth Analysis

Groundwater data collection campaigns are costly and almost always prone to logistical and financial constraints. This implies the need to develop vigorous methodologies for the optimal collection of data. Besides (or complementary) to the informal or subjective methods guided by professional experience and judgment, and the formal or objective methods based on pure geostatistics (e.g. kriging frameworks); simulation-based methods can play a key role in the guidance of data collection in groundwater studies. In this regard, models help in the identification of the most informative data to collect with respect to a specific objective. This process is referred to as experimental design, value-of-information or data worth analyses (DWA). Simulation-based DWA has been used in the literature for the identification of the optimal number (e.g. Norberg and Rosén, 2006) and location of observation wells (e.g. Siade et al., 2017), frequency of sampling (e.g. Kollat et al., 2011), tracer test design including choice of the injection rate, duration of test, and type of tracer (e.g. Wallis et al., 2014), choosing between different types of data (e.g. choosing between conductivity, piezometric heads and travel times

data as in Fu and Gómez-Hernández, 2009; or between concentration and temperature data as in Dausman et al., 2010). In the latter case, the worth of different types of data with different measurement accuracies can be analyzed by using simulation-based Pareto methodology (see Brunner et al., 2012).

Simulation-based DWA can be carried out in at least two contexts: (1) it can be used to compare alternative future data collection schemes at a given stage in a phased survey, and (2) it can be used to decide when to stop or limit a staged data collection program to reduce data redundancy (i.e. observations having similar information content) (Freeze et al., 1992; Khader and McKee, 2014). In both of these contexts, a precondition for the use of model-based DWA is that some minimal groundwater exploration that allows for the development of a justifiable model has taken place (Kikuchi, 2017).

Simulation-based DWA is essentially built on a strategy to quantify the worth of measurements. The most common strategy is to compute a measurement's ability to reduce the uncertainty or error variance of key model predictions (e.g. heads, contaminant concentrations or travel times, magnitude of plume spreading, etc.) which affect management decisions (e.g. Freeze et al., 1992; Cirpka et al., 2004; Dauşman et al., 2010; Wallis et al., 2014; Kikuchi et al., 2015; Wöhling et al., 2016; Siade et al., 2017). In context (1) above, this is often done in the following two steps which are built upon Bayesian statistics and are referred to as Bayesian DWA or Bayesian experimental design. In the first step, a prior analysis is performed where the uncertainty in key model predictions is quantified on the basis of currently available data. In the second step which is called pre-posterior analysis, the uncertainty in key model predictions is re-calculated by assuming that new measurements are carried out in a specific timeframe on a set of locations in the proposed data collection program. Obviously, the values that are actually going to be

measured at the proposed locations are unknown. But what is known is that measurement will reduce the uncertainty at the proposed points to zero (or to some measurement uncertainty) for the intended parameters or state variables. This is sufficient to allow for the calculation of the uncertainty reduction resulting from the proposed measurements with respect to the uncertainties calculated in the prior analysis. Hence in the pre-posterior analysis, the magnitude of predictive uncertainties are considered in a relative rather than absolute sense. Some references refer to this procedure as a form of sensitivity analysis (Finsterle, 2015). The pre-posterior probabilities are often calculated through linear uncertainty analysis (e.g. Dausman et al., 2010), numerous variants of MCMC (e.g. Fu and Gómez-Hernández, 2009), null-space Monte Carlo (NSMC) (e.g. Siade et al., 2017) or EnKF (e.g. Kollat et al., 2011). Repeating pre-posterior analysis for various data collection alternatives in the framework of optimization (e.g. by GA as in Wöhling et al., 2016) or scenario analysis (e.g. Fu and Gómez-Hernández, 2009) allows for the selection of the optimal design.

In the context (2) above, a prior analysis is performed based on data collected in some initial stages of the data collection program (initial dataset). Then pre-posterior analysis is replaced by posterior analysis in which uncertainty in key model predictions is calculated based upon augmenting different subsets of data collected in the subsequent stages (with known measurement values) to the initial dataset. By comparing the outcome of posterior analysis for different subsets of data, one can choose to stop parts of the measurement program that result in the least impact on reducing model prediction uncertainties. Note that prior, pre-posterior and posterior analysis all require propagation of uncertainty from model parameters and observation data to model forecasts (Leaf, 2017).

Some studies go a step further and characterize the worth of data by quantifying the benefits of reducing model prediction uncertainty in terms of risk reduction or monetary costs of economic regret resulting from making wrong decisions in the context of remedial or aquifer exploitation decision making (e.g. Feyen and Gorelick, 2005; Norberg and Rosén, 2006; Neuman et al., 2012). A common formulation for this notion is to maximize some form of the following objective function (ς):

$$\varsigma = B - C_i - \gamma P_f C_f$$

(20)

where *B* is the benefit (e.g. profit from water sales, etc.), C_i is the investment cost (including the cost of measurements), γ is the risk aversion factor, P_f is the probability of failure (e.g. failing to abide by a regulation/policy, etc.), and C_f is the cost of failure (e.g. fines, waste of groundwater resources, etc.). This strategy has two key interconnected advantages: first, it is based on a more direct assessment of cost-effectiveness, and second, communicating the results with stakeholders often becomes easier. The downside to such an approach is that it requires a quantitative definition of the cost of being wrong due to lack of data, which may not always be known or easy to quantify.

In simulation-based methods, DWA may also be quantified based on a measurements ability to improve the estimation of unmeasurable (or hard to measure) parameters. This can be done by assessing the sensitivity of potential measurements (i.e. states) to model parameters through model sensitivity analysis (e.g. Ataie-Ashtiani et al., 2013). An alternative approach is to analyze the measurements ability to reduce parameter estimation uncertainty in a Bayesian framework. In the latter case, the objective function is usually derived from the covariance matrix of the parameters based on A-optimality (minimizing the trace of the covariance matrix, e.g. Hsu and

Yeh, 1989), D-optimality (minimizing the determinant of the covariance matrix, e.g. Catania and Paladino, 2009; Siade et al., 2017), E-optimality (minimizing the eigenvalue of the covariance matrix, e.g. Nordqvist, 2000), or expected Shannon information gain (i.e. relative entropy, e.g. Zhang et al., 2016). See Nowak (2010) for a review of these criteria.

Traditionally, DWA studies have relied on a single conceptual/mathematical model of the groundwater system, making the predictions prone to statistical bias and underestimation of uncertainty (Xue et al., 2014). A more recent approach in the literature is to perform predictions by means of multiple models and then characterize data worth as the contribution of a set of measurements to: (1) the resulting multi-model prediction uncertainty, or (2) model selection/discrimination among a set of viable alternative models. This is often done within a BMA (e.g. Pham and Tsai, 2016) or MLBMA (e.g. Neuman et al., 2012; Xue et al., 2014) framework. In case (1) data worth can be defined, for example, as the difference between the trace of the posterior covariance with and without a set of real measurements (in posterior analysis) or randomly chosen estimates of potential measurements (in pre-posterior analysis) (Neuman et al, 2012; Xue et al., 2014). In (2) the worth of data can be assessed, for example, based on the number of conceptualizations retained in the ensemble after a specific subset of data is considered.

3.6. Improving Data Coverage

In groundwater studies models can be used to interpolate and extrapolate data to provide spatial and temporal coverage of the desired domain. A classic example is the use of groundwater models for the creation of water table or concentration contour maps from sparse and irregularly distributed field measurements. Although the use of classic geostatistical interpolation methods such as inverse distance weighting and kriging variants are common for this purpose, these

methods are known to be vulnerable to outliers, may contradict obvious characteristics of the groundwater system and often fail to represent complex variations between relatively distant measurement points (Fasbender et al., 2008; Buchanan and Triantafilis, 2009). These shortcomings can be alleviated through the use of models, because the physical constraints imposed by models offer additional valuable information in the development of contour maps. The methodology is straightforward: the model is calibrated using the sparse field measurements of water table and/or concentration, then the current state of the groundwater system is reconstructed by the model and the results are used to generate the contour maps. Data generalization can also be performed by solving Bayesian filtering or smoothing state estimation problems. Use of models in such Bayesian frameworks allows for the incorporation of auxiliary data and geostatistics in the data generalization process (for example, see Peeters et al., 2010). Despite these advantages, models are rarely developed for the single purpose of creating a groundwater contour map for two main reasons: first, the time and effort needed to create and calibrate the model compared to the use of pure geostatistics are considerably higher, and second, commonly a mismatch remains between observed and simulated values in the measurement points.

3.7. Customizing Data Resolution

Models can be used to upscale (or similarly downscale) data. Upscaling refers to the transformation of data collected at a fine scale onto a coarser scale. A typical example in groundwater studies is the upscaling of hydraulic conductivity data, also referred to as the estimation of effective, equivalent, interpreted, homogenized or block hydraulic conductivity (Sanchez-Vila et al., 2006). Another parameter commonly upscaled in groundwater studies is dispersivity (de Barros and Dentz, 2016). The basic notion behind model-based data upscaling in

groundwater studies is that quantities such as flows and hydraulic head gradients computed by a model at a coarse scale block should match the corresponding average values of these quantities modeled at the fine scale blocks that form the course block and pertain to the scale of data acquisition. In other words, the upscaled quantity is chosen so that the upscaled blocks can reproduce the behavior of the heterogeneous medium through modeling. The advantage of this model-based method in comparison to pure geostatistics is that it is not limited to a specific spatial distribution pattern, degree of variability or aquifer geometry. A pioneering study in this regard is Gómez-Hernández (1991), followed by studies such as Zhou et al. (2010), who extended the methodology of Gómez-Hernández (1991) to three dimensions, Li et al. (2012) which coupled this upscaling method with inverse modeling through the EnKF, Fernàndez-Garcia et al. (2009) and Li et al. (2011) who applied transport models to upscaling, and Godoy et al. (2018) which employed the Laplacian-with-skin method for upscaling of hydraulic conductivity.

3.8. Fusion of Heterogeneous Data

Hydrogeological site investigations involve the collection of an array of different types of data which should be combined to form a unified picture of the aquifer. For several reasons, hydrogeological data fusion is often a difficult task: (1) various types of data may relate to different aspects of the system and hence do not share the same nature. These inherently different forms of data cannot be readily related to each other. A typical example of this notion is pollutant concentration and hydraulic conductivity data. (2) Even data of the same nature generally do not share the same quality and may have dissimilar spatial and temporal scales and degrees of uncertainty and imperfection. For example in many site investigations, data from traditional characterization and monitoring methods (such as core analyses and hydraulic tests) are

supplemented with coverage of greater density from indirect geophysical surveys, and soft data based on expert knowledge and field questionnaires. The degree of uncertainty associated with each of these types of data is very different from the other. This is known as data heterogeneity (Khaleghi et al., 2013). (3) There is the typical problem of handling inconsistency and conflict in temporally or spatially overlapping data.

For decades, hydrogeologists have relied on a combination of expert knowledge and geostatistical methods for formal data fusion. But most existing geostatistical methods (such as classical kriging and cokriging) are limited in their ability to simultaneously account for large numbers of information sources (Porter et al., 2000; Fasbender, 2008) and hence research on the improvement of geostatistical data fusion methods is ongoing (e.g. Hosseini and Kerachian, 2017). Model-based data fusion (also called data fusion modeling in the groundwater literature, e.g. Porter et al., 2000) is a very powerful tool for solving the problems associated with data fusion in hydrogeology. Models provide prior knowledge of the physical relationships between dissimilar datasets, and data fusion algorithms use these relationships to extract integrated information from the measured data. We classify previous work on model-based data fusion in the groundwater literature into three categories, which are described in the following subsections.

3.8.1 Fusion of Different Types of Direct Measurements

This is the most common form of data fusion in groundwater studies, partly because traditional frequentist parameter estimation methods implemented in codes such as PEST, MODFLOWP, UCODE, iTHOUGH2, etc., embody this form of data fusion. These frequentist methods use direct measurements of model parameters (e.g. hydraulic conductivity, porosity, dispersivity, etc.) to provide initial estimates and bounds for each of these parameters, and then

simultaneously employ different forms of data related to direct measurements of model state variables (e.g. head, concentration, travel time, etc.) to update these initial estimates. Hence, both the resulting parameter estimates and model predictions are based on the integration of different forms of data. But as previously discussed, these methods do not provide a proper estimation of uncertainties. The use of Bayesian fusion techniques alleviates this problem. In Bayesian fusion, data on model parameters is employed to build prior estimates for these parameters in the form of probability distributions, and the data on state variables are used to update these priors and obtain the posterior probability distributions of model parameters. Bayesian fusion has been implemented through MCMC algorithms for the fusion of non-sequential data (e.g. Hassan et al., 2009; Laloy et al., 2013) and variants of the KF have been employed for the fusion of sequential data (e.g. Porter et al., 2000; Bailey and Baù, 2012) in the groundwater literature.

3.8.2. Fusion of Direct and Indirect Measurements

Direct point measurements of hydrogeologic data are commonly limited because their acquisition is expensive, time-consuming and invasive (Rajabi and Ataie-Ashtiani, 2016). Several types of indirect measurement data can be used to alleviate this problem. Two of the most commonly used are geophysical data and remote sensing (RS) data. It is well known that the use of geophysical data (such as ground-penetrating radar data, electrical resistance tomography data, etc.), in conjunction with direct measurement of hydrogeologic variables, can substantially improve characterization of subsurface variability (Kowalsky et al., 2005 and 2006; Yeh et al., 2007). However, geophysical methods provide data on geophysical properties in the subsurface that are nonlinearly related to the variables of interest, and the standard relationships commonly used to infer these variables from geophysical data may induce artifacts that cannot be interpreted from a hydraulic viewpoint (Camporese et al., 2011). Groundwater models can be

used to solve this problem through for example, coupled hydrogeophysical inversion where the hydraulic and geophysical equations are considered as a coupled system (Pollock and Cirpka, 2012). A common approach for fusing geophysical data with direct measurements of hydrogeologic variables in this context is to employ geophysical data as part of the measurement model as follows (Kowalsky et al., 2004):

$$D_{aug} = \begin{bmatrix} D_H \\ D_{GP} \end{bmatrix} = \begin{bmatrix} h_{tH}(Y_t, u_t, \theta) \\ h_{tGP}(Y_t, u_t, \theta) \end{bmatrix} + \begin{bmatrix} \eta_{tH} \\ \eta_{tGP} \end{bmatrix}$$
(21)

Where D_{aug} is an augmented vector of all observational data, h_{tH} is the measurement model that maps (Y_{t}, u_{t}, θ) to the hydrological measurements (D_{H}) , h_{tGP} is the measurement model that maps (Y_{t}, u_{t}, θ) to the geophysical data (D_{GP}) , and η_{tH} and η_{tGP} are the observation stochastic error vectors for the hydrological and geophysical data respectively.

RS data, usually combined with geographic information systems, can also be a potentially useful source of information for groundwater modeling. This is especially true in regional-scale modeling in areas where other forms of data are scarce. Current air and satellite-based RS technologies can penetrate the ground for only a few centimeters, but this is enough to provide data for the inference of surface forcing and some geologic properties, making RS a potentially valuable source of information in the study of shallow groundwater (Becker, 2006). RS data is often supplemented with ground control measurements in order to scale RS data to the variable of interest or to estimate the error statistics (i.e. uncertainty) of the RS data (Hendricks Franssen et al., 2008). Examples for the use of RS data in groundwater MDI include:

- Estimates of values and spatial patterns of groundwater model recharge rates from satellite images pertaining to precipitation, actual evapotranspiration, etc. (e.g. Brunner et al., 2004; Hendricks Franssen et al., 2008),
- Use of RS based digital elevation models (DEMs) to constrain groundwater flow model outputs and avoid erroneous artesian piezometric head values (Hendricks Franssen et al., 2008),
- Employing geologic maps derived from RS data for groundwater prospecting, e.g. providing information about hydraulic conductivities, water reserves of water bearing formations and identification of faults and fracturing for groundwater modeling (Waters et al. 1990),
- Using RS imagery to identify and characterize boundary conditions for groundwater models. This may include identification of streams, lakes, wetlands, seepage areas, recharge and evapotranspiration zones, or dynamic monitoring of stream headwater (Becker, 2006).

3.8.3. Fusion of Field Measurements and Expert Knowledge

Expert knowledge has long been identified as a key source of information for groundwater modeling because experts have the ability to interpret complex and ambiguous evidence based on their broader experiences (O'Hagan, 2012). This has made the fusion of soft data pertaining to expert knowledge and hard data obtained from field measurements, an important topic in groundwater studies. The informal use of expert knowledge in groundwater model conceptualization and parameter estimation is highly common. Formal mechanisms in this context are mostly based on the Bayesian approach (Krueger et al., 2012). In the Bayesian approach, the subjective belief of an individual expert or the inter-subjective belief of several

experts about the value of parameters or plausibility of alternative model structures can be represented by prior probability distributions through expert elicitation techniques (Beer et al., 2013; Rinderknecht et al., 2014). These priors are then updated based on hard field measurement data. This approach has been used in many previous groundwater studies for parameter estimation or model selection and structural uncertainty analysis (e.g. Hassan et al., 2009). However, this approach has been criticized for a number of reasons, including: (1) it neglects the imprecision essentially embedded in expert provided soft data which may lead to biased result (Lele and Allen, 2006; Stein et al., 2013), (2) expressing expert knowledge in the form of probability distributions is often very difficult (Ross et al., 2009), and (3) it is bound to the incorporation of expert knowledge regarding model parameters and structures, and does not provide the means to include expert knowledge on other aspects of groundwater models such as state variables. To solve these problems, Rajabi and Ataie-Ashtiani (2016) proposed the use of fuzzy Bayesian inference based on MCMC, for incorporating expert knowledge in parameter estimation. Their method uses the power of fuzzy logic to provide a convenient framework for the representation of expert provided information regarding the various inputs to the Bayesian inference algorithm. Furthermore, it allows one to distinguishably model both uncertainty and imprecision in the fusion process.

4. Software and Codes

The MDI algorithms described in section 2 commonly involve large computational effort and hence their implementation inevitably requires computer programming. While many MDI efforts in the groundwater literature rely on codes that are developed for the specific problems at hand and are not made available to the public, some researchers/institutions have focused their efforts on developing generic codes for this purpose. These efforts have resulted in the development of

several MDI software since the 1990s, some of which are reviewed in **Table 2**. These software have essentially transformed the way MDI is conducted in groundwater studies, by paving the way for a mainstream change from trial-and-error techniques to automated MDI procedures. MDI software used in the groundwater literature include:

- 1. Model-dependent software that are developed for application in conjunction with specific groundwater models. An example is MODFLOWP (Hill, 1992),
- Model-independent generic software such as PEST (Doherty, 1994) PEST++ (Welter et al., 2015) and UCODE (Poeter and Hill, 1999),
- Code packages mostly in R (e.g. MCMC, see <u>http://www.stat.umn.edu/geyer/mcmc/</u>), python (e.g. pyNSMC (White et al., 2015), pyEMU (White et al., 2016) and pyMC (Patil et al., 2010)) and MATLAB (e.g. DREAM (Vrugt, 2016)).

The second and third categories can be used with any model, often the only requirement is that the input and output files of the model are numerical (ASCII or text only). These two categories are mostly community-supported and open-source, and are also being used in other engineering fields.

Most available codes and software are intended for parameter estimation, and the absolute majority employ batch frequentist weighted LSE with local derivative-based optimization methods (e.g. LM optimization). The widespread use of these algorithms can be attributed to their ease of coding and affordability of computations in low-dimensional problems. There are also several codes available that are based on MCMC algorithms for Bayesian parameter estimation, e.g. DREAM (in MATLAB), MCMC (in R), and UCODE MCMC (Lu et al., 2014). The reason for such focus on parameter estimation as compared to model identification or state estimation is twofold: (1) parameter estimation can often be performed by relatively simple

automated modifications to the model input files, while model structure or state updating requires a high level of interaction with the numerical model (Liu et al., 2012), and (2) currently the generation of model structure realizations is mostly based on expert insight rather than formal methods, making the development of automated model structure updating software difficult. Recent developments in MDI software for groundwater applications include developing parallel computing abilities (e.g. parallel PEST, as in Tang et al., 2010, BeoPEST, see Hunt et al., 2010, and MCMC UCODE, see Lu et al., 2014), enabling cloud computing (e.g. PEST.cloud, see https://pest.cloud/), incorporating advanced regularization (see

http://www.pesthomepage.org/Highly-parameterized_inversion.php about PEST), using state of the art global optimization schemes (see Finsterle, 2010, regarding iTOUGH2), and integrating MDI into risk-based environmental management optimization (e.g. ESTPP-OPT, see White et al., 2018). The PEST++ suite will likely play a key role in the future of groundwater MDI. One reason is that with the new, modular, nonintrusive, parallel run manager named PANTHER it is becoming easy for people to add programs to the PEST++ suite, and to run parallel computations on the cloud.

5. Discussion and Prognosis for Future Work

We defined MDI as a two way process between models and data, and reviewed recent advances in MDI methods and applications in the groundwater literature. The review shows that frequentist weighted LSE is still the most widely used method for MDI in groundwater scientific literature and professional practice, which is mostly due to the availability of open-source, userfriendly software, a multitude of case studies, and a number of well-established guidelines for its implementation. However aided by advances in computing power and data handling, there is a trend toward more extensive use of Monte Carlo-based methods such as MCMC, EnKF, SMC

and ES in the literature, and these methods are also gradually finding their way into professional practice. The classic KF became popular in the groundwater scientific literature in 1990s and early 2000s, but due to a number of limitations, failed to become a mainstream practical method and is now mostly replaced by the EnKF in the literature. The ExKF and VBMs, although popular in other fields (e.g. weather prediction and atmospheric sciences), have been given much less attention in the groundwater modeling community.

It is clear from this review that there are many tools and techniques for groundwater MDI, and this diversity is needed for supporting different MDI objectives, model and data types and computational constraints. It is important to understand that the success of groundwater MDI does not necessarily depend on providing more data or using more sophisticated models, but is mainly governed by their optimal synthesis, and properly addressing the associated uncertainties. This is an important reason for the significant interest in improving MDI methods in the literature. The continuing progress of data acquisition technologies and the evolution of models, means that the landscape of MDI in groundwater applications will continue to evolve in the future. Here we discuss a number of key issues that will likely form future directions: 1) Addressing computational challenges: Several key synthesis approaches used in MDI require a large number of model simulations. This is most notably true for methods that rely on MCS (such as MCMC, EnKF, SMC and ES) or stochastic/meta-heuristic optimization algorithms. These synthesis approaches become computationally very expensive when the computational demand of a single model run is substantially high. This has been an issue for the past several decades, and interestingly the enormous increase of computer processing speed in recent years has not solved this problem. The fact that computational demand remains to be an important issue to date and probably for the mid-term future, has mainly two reasons: first, recent advances

in groundwater simulation software have mostly focused on increasing model fidelity (i.e. improving the degree to which the model reproduces the behavior of the real-world system), resulting in continuous increase in their execution times (Carrera et al., 2005). Second, we are seeing a constant shift from theoretical research on MDI using synthetic toy problems with limited computational demand (e.g. Henry problem (Henry, 1964)), to real-world MDI applications involving computationally expensive groundwater models. Due to these reasons, the computational challenge must be somehow confronted within the groundwater community in order to facilitate the success of MDI. A review of literature shows that four main strategies are being pursued to address this problem: (1) parallelization and grid computing applied to either or both the model and the synthesis approach, for example see the parallel ParFlow model for simulating surface and subsurface flow (Kollet et al., 2010; Bürger et al., 2012, see https://parflow.org/), Xu et al. (2013) for parallel MCMC, Joseph and Guillaume (2013) for parallel EnKF, and Kurtz et al. (2017) for cloud-based modeling, (2) replacing the model with data-driven and physics-free, meta-models such as polynomial chaos expansion (PCE) (e.g. Laloy et al., 2013, Rajabi and Ataie-Ashtiani, 2016), (3) improving the computational efficiency of synthesis methods (e.g. Vrugt et al., 2009), and (4) employing cloud computing (see Hayley, 2017). Research on all four strategies is expected to continue in the future, and it is expected that cloud computing will play an increasingly important role in solving the computational challenges of MDI in groundwater.

2) Accounting for Local Heterogeneities: Every technique currently available for the estimation of heterogeneous groundwater model parameters from data, employs some level of averaging or smoothing through parameterization and regularization. The consequence is that the results of parameter estimation may be locally quite flawed, and this reduces the accuracy of model

predictions that are sensitive to these local values. Hence much work is being currently done and will continue in the future on several fronts to alleviate these problems. This includes developing methods for: (1) use of by-product information of regularized parameter estimation (such as spatial covariance structure of the estimated field) to identify misrepresented local details (Moore and Doherty, 2006), (2) optimizing the choice of the limited number of parameters that are used for the reconstruction of the heterogeneous field (e.g. Jung et al., 2011), and (3) adapting synthesis methods for the incorporation of more local details and hence more parameters in the estimation process (e.g. Tonkin and Doherty, 2009). The third approach inevitably includes developing strategies for solving the resulting computational challenges such as instability, non-convergence, solution non-uniqueness, and solution non-optimality.

3) *Real-time MDI*: Cases for the use of automated field sensors that rely on communication technologies such as WSNs, have recently emerged in the groundwater literature. These technologies allow for the monitoring of groundwater systems at much finer spatial and temporal resolutions than traditional manual sampling and analysis methods. But the resulting data is prone to sensor and WSN faults (e.g. stuck readings, out of range errors, abrupt shifts, abnormal noise, etc., see Szewczyk et al., 2004), and hence requires the integration of fault detection methods in groundwater MDI techniques especially when sensor readings are fed automatically into models. An example of studies addressing this issue is Barnhart et al. (2010), which employs a data reduction algorithm to identify and remove potentially faulty salinity data, and then uses the remaining data for groundwater model parameter estimation by PEST. Progress made in real-time groundwater monitoring technologies have facilitated the development of real-time or quasi real-time decision support systems (DSSs). The traditional and still widely used approach in groundwater DSSs is to employ offline rules that optimizes the controls with the objective of

minimizing the cost of operations (e.g. Rajabi and Ketabchi, 2017). But the emerging real-time alternative, also known as model predictive control (MPC), constantly revises and optimizes controls based on the feedback from the system (Liu et al., 2012). The ability to rapidly assimilate real-time data and provide answers to imminent questions is gradually becoming an important part of groundwater modeling (Langevin and Panday, 2012). An example reference is Drumheller et al. (2017). The study uses hydraulic head and electrical conductivity data from a distributed sensor network for quasi real-time model calibration and then performs simulation-optimization with the aim of controlling aquifer recharge and recovery operations in a laboratory setup. Hendricks Franssen et al. (2011) use EnKF for combining online observations of hydraulic head with numerical models, for the real-time characterization of groundwater flow in a real-world urban aquifer. Their computational tool became operational at the Water Works Zurich in 2009.

4) Developing codes and software for MDI: Despite the hugely valuable efforts by the groundwater and other research communities in developing software for MDI, major gaps still remain. These gaps, which underpin future research and collaboration opportunities, include: (1) developing community-supported, open-source data assimilations tools based on methods such as EnKF, SMC etc. Development of such tools (for instance the "parallel data assimilation framework", see http://pdaf.awi.de/trac/wiki), has helped other fields such as numerical weather prediction, atmospheric sciences and more recently hydrology. Work on such tools has already started in the groundwater modeling community. An example is PESTPP-IES, which is an implementation of the iterative ES LM algorithm of Chen and Oliver (2013) within the framework of PEST/PEST++ model interface protocols (see https://github.com/dwelter/pestpp). (2) Creating computer codes for groundwater model discrimination and structural uncertainty

analysis. Progress in this regard depends greatly on developing algorithms that can automate this process and reduce its dependence on subjective expert opinion. (3) Developing codes for formal fusion of different types of groundwater related data (e.g. geophysical and remote sensing data, expert knowledge, etc.).

5) *Bridging the gap between research and professional practice*: Lastly, the growing researchpractice gap has already been observed more generally in groundwater science (Simmons et al., 2012). We also acknowledge that there is an existing gap between MDI practices in the research community and those in consulting, industry and government; and this gap seems to be growing. This is evident from the fact that many of the more recent MDI methods are rarely applied outside of research settings. Closing this gap calls for more user-friendly software, guidelines for MDI method selection and application, worked examples and case studies, education and training.

Acknowledgement

The authors wish to thank the Editor-in-Chief, Professor Geoff Syme and two anonymous reviewers for their valuable comments which helped to improve the final manuscript. Behzad Ataie-Ashtiani and Craig T. Simmons acknowledge support from the National Centre for Groundwater Research and Training, Australia. Behzad Ataie-Ashtiani also appreciates the support of the Research Office of the Sharif University of Technology, Iran.

References

Abbaspour, K. C., Schulin, R., & Van Genuchten, M. T. (2001). Estimating unsaturated soil hydraulic parameters using ant colony optimization. *Advances in water resources*, *24*(8), 827-841.

Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018). Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. *Advances in Water Resources*, *111*, 192-204.

Almeida, A. S., & Frykman, P. (1994). Geostatistical modeling of chalk reservoir properties in the Dan Field, Danish North Sea.

Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. I. (2003). An introduction to MCMC for machine learning. *Machine learning*, *50*(1-2), 5-43.

Alcolea, A., Carrera, J., & Medina, A. (2006). Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Advances in Water Resources*, *29*(11), 1678-1689.

Ataie-Ashtiani, B., Rajabi, M. M., & Ketabchi, H. (2013). Inverse modelling for freshwater lens in small islands: Kish Island, Persian Gulf. *Hydrological Processes*, *27*(19), 2759-2773.

Bailey, R. T., & Baù, D. (2010). Ensemble smoother assimilation of hydraulic head and return flow data to estimate hydraulic conductivity distribution. *Water Resources Research*, *46*(12).

Bailey, R. T., & Baù, D. (2012). Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother. *Hydrology and Earth System Sciences*, *16*(2), 287.

Barnhart, K., Urteaga, I., Han, Q., Jayasumana, A., & Illangasekare, T. (2010). On integrating groundwater transport models with wireless sensor networks. *Groundwater*, *48*(5), 771-780.

Bayer, P., & Finkel, M. (2004). Evolutionary algorithms for the optimization of advective control of contaminated aquifer zones. *Water Resources Research*, 40(6).

Bayer, P., & Finkel, M. (2007). Optimization of concentration control by evolution strategies: Formulation, application, and assessment of remedial solutions. *Water resources research*, *43*(2).

Beer, M., Ferson, S., & Kreinovich, V. (2013). Imprecise probabilities in engineering analyses. *Mechanical systems and signal processing*, *37*(1-2), 4-29.

Becker, M. W. (2006). Potential for satellite remote sensing of ground water. *Groundwater*, 44(2), 306-318.

Bernardo, J. M., & Smith, A. F. (2001). Bayesian theory.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in water resources*, 16(1), 41-51.

Beven, K., & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, *6*(3), 279-298.

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of hydrology*, *249*(1-4), 11-29.

Beven, K. (2010). Environmental modelling: An uncertain future?. CRC Press.

Beven, K. J (2018). On hypothesis testing in hydrology: Why falsification of models is still a really good idea. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1278.

Beven, K., & Young, P. (2013). A guide to good practice in modeling semantics for authors and referees. *Water Resources Research*, *49*(8), 5092-5098.

Box, G. E., & Tiao, G. C. (2011). *Bayesian inference in statistical analysis* (Vol. 40). John Wiley & Sons.

Bredehoeft, J. (2005). The conceptualization model problem—surprise. *Hydrogeology journal*, *13*(1), 37-46.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., & Tans, P. (2005). An improved Kalman Smoother for atmospheric inversions. *Atmospheric Chemistry and Physics*, *5*(10), 2691-2702.

Brunner, P., Bauer, P., Eugster, M., & Kinzelbach, W. (2004). Using remote sensing to regionalize local precipitation recharge rates obtained from the chloride method. *Journal of Hydrology*, *294*(4), 241-250.

Brunner, P., Doherty, J., & Simmons, C. T. (2012). Uncertainty assessment and implications for data acquisition in support of integrated hydrologic models. *Water Resources Research*, 48(7).

Brunner, P., Therrien, R., Renard, P., Simmons, C. T., & Franssen, H. J. H. (2017). Advances in understanding river-groundwater interactions. *Reviews of Geophysics*, 55(3), 818-854.

Buchanan, S., & Triantafilis, J. (2009). Mapping water table depth using geophysical and environmental variables. *Groundwater*, 47(1), 80-96.

Burgers, G., Jan van Leeuwen, P., & Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly weather review*, *126*(6), 1719-1724.

BüRger, C. M., Kollet, S., Schumacher, J., & BöSel, D. (2012). Short note: Introduction of a web service for cloud computing with the integrated hydrologic simulation platform ParFlow. *Computers & Geosciences*, *48*, 334-336.

Camporese, M., Cassiani, G., Deiana, R., & Salandin, P. (2011). Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. *Water Resources Research*, *47*(12).

Carrera, J., & Neuman, S. P. (1986a). Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research*, *22*(2), 199-210.

Carrera, J., & Neuman, S. P. (1986b). Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms. *Water Resources Research*, *22*(2), 211-227.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology journal*, *13*(1), 206-222.

Carrera, J., Hidalgo, J. J., Slooten, L. J., & Vázquez-Suñé, E. (2010). Computational and conceptual issues in the calibration of seawater intrusion models. *Hydrogeology journal*, *18*(1), 131-145.

Catania, F., & Paladino, O. (2009). Optimal sampling for the estimation of dispersion parameters in soil columns using an iterative genetic algorithm. *Environmental Modelling & Software*, *24*(1), 115-123.

Chang, S. Y., Chowhan, T., & Latif, S. (2012). State and parameter estimation with an SIR particle filter in a three-dimensional groundwater pollutant transport model. *Journal of Environmental Engineering*, *138*(11), 1114-1121.

Chang, S. Y., & Sayemuzzaman, M. (2014). Using Unscented Kalman Filter in Subsurface Contaminant Transport Models. *Journal of Environmental Informatics*, 23(1).

Chang, H., Liao, Q., & Zhang, D. (2017). Surrogate model based iterative ensemble smoother for subsurface flow data assimilation. *Advances in Water Resources*, *100*, 96-108.

Chen, Z. (2003). Bayesian filtering: From Kalman filters to particle filters, and beyond. *Statistics*, *182*(1), 1-69.

Chen, Y., & Oliver, D. S. (2013). Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, *17*(4), 689-703.

Chen, Y., & Zhang, D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, *29*(8), 1107-1122.

Cirpka, O. A., Bürger, C. M., Nowak, W., & Finkel, M. (2004). Uncertainty and data worth analysis for the hydraulic design of funnel-and-gate systems in heterogeneous aquifers. *Water resources research*, 40(11).

Cox, H. (1964). On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on automatic control*, 9(1), 5-12.

Crestani, E., Camporese, M., Baú, D., & Salandin, P. (2013). Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. *Hydrology and Earth System Sciences*, *17*(4), 1517.

Cui, T., Dudley Ward, N., & Kaipio, J. (2013). Characterization of Parameters for a Spatially Heterogenous Aquifer from Pumping Test Data. *Journal of Hydrologic Engineering*, *19*(6), 1203-1213.

Daum, F., & Huang, J. (2003, March). Curse of dimensionality and particle filters. In *Aerospace Conference, 2003. Proceedings. 2003 IEEE* (Vol. 4, pp. 4_1979-4_1993). IEEE.

Dausman, A. M., Doherty, J., Langevin, C. D., & Sukop, M. C. (2010). Quantifying data worth toward reducing predictive uncertainty. *Groundwater*, 48(5), 729-740.

de Barros, F. P., & Dentz, M. (2016). Pictures of blockscale transport: Effective versus ensemble dispersion and its uncertainty. *Advances in Water Resources*, *91*, 11-22.

de Marsily, G. (1978). De l'identification des systèmes hydrogéologiques (Doctoral dissertation).

Dessai, S., & Hulme, M. (2007). Assessing the robustness of adaptation decisions to climate change uncertainties: a case study on water resources management in the East of England. *Global Environmental Change*, *17*(1), 59-72.

Djuric, P. M., Kotecha, J. H., Zhang, J., Huang, Y., Ghirmai, T., Bugallo, M. F., & Miguez, J. (2003). Particle filtering. *IEEE signal processing magazine*, *20*(5), 19-38.

Doherty, J. (1994). PEST: a unique computer program for model-independent parameter optimisation. *Water Down Under 94: Groundwater/Surface Hydrology Common Interest Papers; Preprints of Papers*, 551.

Doherty, J., & Simmons, C. T. (2013). Groundwater modelling in decision support: reflections on a unified conceptual framework. *Hydrogeology Journal*, 21(7), 1531-1537.

Dorigo, M., and T. Stützle (2004). Ant Colony Optimization, MIT Press, Cambridge, MA.

Doucet, A., De Freitas, N., & Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice* (pp. 3-14). Springer, New York, NY.

Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, 45-97.

Drécourt, J. P., Madsen, H., & Rosbjerg, D. (2006). Calibration framework for a Kalman filter applied to a groundwater model. *Advances in water resources*, *29*(5), 719-734.

Drumheller, Z., Smits, K. M., Illangasekare, T. H., Regnery, J., Lee, J., & Kitanidis, P. K. (2017). Optimal Decision Making Algorithm for Managed Aquifer Recharge and Recovery Operation Using Near Real-Time Data: Benchtop Scale Laboratory Demonstration. *Groundwater Monitoring & Remediation*, *37*(1), 27-41.

Eigbe, U., Beck, M. B., Wheater, H. S., & Hirano, F. (1998). Kalman filtering in groundwater flow modelling: problems and prospects. *Stochastic Hydrology and Hydraulics*, *12*(1), 15-32.

El Gharamti, M., Hoteit, I., & Valstar, J. (2013). Dual states estimation of a subsurface flowtransport coupled model using ensemble Kalman filtering. *Advances in water resources*, *60*, 75-88.

Erdal, D., & Cirpka, O. A. (2017). Preconditioning an ensemble Kalman filter for groundwater flow using environmental-tracer observations. *Journal of hydrology*, *545*, 42-54.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, *99*(C5), 10143-10162.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, *53*(4), 343-367.

Evensen, G. (2009). *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media.

Fasbender, D. (2008). Bayesian data fusion in environmental sciences: Theory and applications. *Thèse Présentée en Vue de l'Obtention du Grade de Docteur en Sciences, Louvain-la-Neuve: Université Catholique de Louvain*, 191

Fasbender, D., Peeters, L., Bogaert, P., & Dassargues, A. (2008). Bayesian data fusion applied to water table spatial mapping. *Water Resources Research*, *44*(12).

Finsterle, S. (2015). Practical notes on local data-worth analysis. *Water Resources Research*, *51*(12), 9904-9924.

Fernàndez-Garcia, D., Llerar-Meza, G., & Gómez-Hernández, J. J. (2009). Upscaling transport with mass transfer models: Mean behavior and propagation of uncertainty. *Water resources research*, *45*(10).

Feyen, L., Beven, K. J., De Smedt, F., & Freer, J. (2001). Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: conditioning on head observations. *Water Resources Research*, *37*(3), 625-638.

Feyen, L., & Gorelick, S. M. (2005). Framework to evaluate the worth of hydraulic conductivity data for optimal groundwater resources management in ecologically sensitive areas. *Water Resources Research*, *41*(3).

Finsterle, S. (1993). ITOUGH2 User's Guide, Report No. *LBL-34582, Lawrence Berkeley National Laboratory, Berkeley, Calif.*

Finsterle, S. A. (2010). *iTOUGH2 Universal optimization using the PEST Protocol* (No. LBNL-3698E). Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US).

Finsterle, S. (2015). Practical notes on local data-worth analysis. *Water Resources Research*, *51*(12), 9904-9924.

Freer, J., Beven, K., & Ambroise, B. (1996). Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research*, *32*(7), 2161-2173.

Freeze, R. A., Massmann, J., Smith, L., Sperling, T., & James, B. (1990). Hydrogeological decision analysis: 1. A framework. *Groundwater*, *28*(5), 738-766.

Freeze, R. A., James, B., Massmann, J., Sperling, T., & Smith, L. (1992). Hydrogeological decision analysis: 4. The concept of data worth and its use in the development of site investigation strategies. *Groundwater*, *30*(4), 574-588.

Fu, J., & Gómez-Hernández, J. J. (2009). Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking Markov chain Monte Carlo method. *Journal of Hydrology*, *364*(3), 328-341.

Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman and Hall/CRC.

Gelb, A. (Ed.). (1974). Applied optimal estimation. MIT press.

Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, *85*(410), 398-409.

Gelhar, L. W., Welty, C., & Rehfeldt, K. R. (1992). A critical review of data on field-scale dispersion in aquifers. *Water resources research*, *28*(7), 1955-1974.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian data analysis. Chapman and Hall/CRC.

Ghil, M., & Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography. In *Advances in geophysics* (Vol. 33, pp. 141-266). Elsevier.

Godoy, V. A., Zuquette, L. V., & Gómez-Hernández, J. J. (2018). Stochastic analysis of threedimensional hydraulic conductivity upscaling in a heterogeneous tropical soil. *Computers and Geotechnics*, 100, 174-187.

Gómez-Hernández, J. (1991). A stochastic approach to the simulation of block conductivity fields conditional upon data measured at a smaller scale. *Doctoral Dissertation, Stanford University*.

Gómez-Hernández, J. J., & Journel, A. G. (1993). Joint sequential simulation of multigaussian fields. In *Geostatistics Troia* '92 (pp. 85-94). Springer, Dordrecht.

Gómez-Hernánez, J. J., Sahuquillo, A., & Capilla, J. (1997). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory. *Journal of Hydrology*, *203*(1-4), 162-174.

Gordon, N. J., Salmond, D. J., & Smith, A. F. (1993, April). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)* (Vol. 140, No. 2, pp. 107-113). IET Digital Library.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., & Ye, M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, 48(8).

Haddad, O. B., Tabari, M. M. R., Fallah-Mehdipour, E., & Mariño, M. A. (2013). Groundwater model calibration by meta-heuristic algorithms. *Water resources management*, 27(7), 2515-2529.

Hayley, K. (2017). The present state and future application of cloud computing for numerical groundwater modeling. *Groundwater*, *55*(5), 678-682.

Han, X., & Li, X. (2008). An evaluation of the nonlinear/non-Gaussian filters for the sequential data assimilation. *Remote Sensing of Environment*, *112*(4), 1434-1449.

Han, M., Zhao, C., Feng, G., & Shi, F. (2015). Bayesian inference of the groundwater depth threshold in a vegetation dynamic model: A case study, lower reach, Tarim River. *Quaternary International*, *380*, 207-215.

Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, *9*(2), 159-195.

Hassan, A. E., Bekhit, H. M., & Chapman, J. B. (2008). Uncertainty assessment of a stochastic groundwater flow model using GLUE analysis. *Journal of Hydrology*, *362*(1-2), 89-109.

Hassan, A. E., Bekhit, H. M., & Chapman, J. B. (2009). Using Markov Chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model. *Environmental Modelling & Software*, *24*(6), 749-763.

Hayley, K. (2017). The present state and future application of cloud computing for numerical groundwater modeling. *Groundwater*, *55*(5), 678-682.

Hendricks Franssen, H. J., & Kinzelbach, W. (2008). Real-time groundwater flow modeling with the ensemble Kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resources Research*, *44*(9).

Hendricks Franssen, H. J., & Kinzelbach, W. (2009). Ensemble Kalman filtering versus sequential self-calibration for inverse modelling of dynamic groundwater flow systems. *Journal of Hydrology*, *365*(3-4), 261-274.

Hendricks Franssen, H. J., Gómez-Hernández, J. J., Capilla, J. E., & Sahuquillo, A. (1999). Joint simulation of transmissivity and storativity fields conditional to steady-state and transient hydraulic head data. *Advances in water resources*, 23(1), 1-13.

Hendricks Franssen, H. J., Stauffer, F., & Kinzelbach, W. (2004). Joint estimation of transmissivities and recharges—Application: Stochastic characterization of well capture zones. *Journal of hydrology*, *294*(1-3), 87-102.

Hendricks Franssen, H. J., Brunner, P., Makobo, P., & Kinzelbach, W. (2008). Equally likely inverse solutions to a groundwater flow problem including pattern information from remote sensing images. *Water resources research*, *44*(1).

Hendricks Franssen, H. H., Alcolea, A., Riva, M., Bakr, M., Van der Wiel, N., Stauffer, F., & Guadagnini, A. (2009). A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources*, *32*(6), 851-872.

Hendricks Franssen, H. J., Kaiser, H. P., Kuhlmann, U., Bauser, G., Stauffer, F., Müller, R., & Kinzelbach, W. (2011). Operational real-time modeling with ensemble Kalman filter of variably saturated subsurface flow including stream-aquifer interaction and parameter updating. *Water resources research*, *47*(2).

Henry, H.R. (1964). Effects of Dispersion on Salt Encroachment in Coastal Aquifers. USGS Water-Supply Paper 1613-C. Sea Water Coastal Aquifers, C71-84.

Heywood, C. E., & Yager, R. M. (2003). Simulated ground-water flow in the Hueco Bolson, an alluvial-basin aquifer system near El Paso, Texas (No. 2002-4108).

Heywood, C. E., Kahle, S. C., Olsen, T. D., Patterson, J. D., & Burns, E. (2016). *Simulation of groundwater storage changes in the eastern Pasco Basin, Washington* (No. 2016-5026). US Geological Survey.

Hill, M. C. (1992). A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional ground-water flow model using nonlinear regression (No. 91-484, pp. 1-358). US Geological Survey.

Hill, M. C., & Tiedeman, C. R. (2006). *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*. John Wiley & Sons.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical science*, 382-401.

Höge, M., Wöhling, T., & Nowak, W. (2018). A primer for model selection: The decisive role of model complexity. *Water Resources Research*, *54*(3), 1688-1715.

Højberg, A. L., & Refsgaard, J. C. (2005). Model uncertainty-parameter uncertainty versus conceptual models. *Water Science and Technology*, *52*(6), 177-186.

Holland, J. H. (1975). Adaptation in natural and artificial systems. An introductory analysis with application to biology, control, and artificial intelligence. *Ann Arbor, MI: University of Michigan Press*, 439-444.

Hooke, R., & Jeeves, T. A. (1961). Direct Search Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM)*, 8(2), 212-229.

Hosseini, M., & Kerachian, R. (2017). A data fusion-based methodology for optimal redesign of groundwater monitoring networks. *Journal of Hydrology*, *552*, 267-282.

Hsu, N. S., & Yeh, W. W. G. (1989). Optimum experimental design for parameter identification in groundwater hydrology. *Water Resources Research*, *25*(5), 1025-1040.

Huber, E., Hendricks-Franssen, H. J., Kaiser, H. P., & Stauffer, F. (2011). The role of prior model calibration on predictions with ensemble Kalman filter. *Groundwater*, 49(6), 845-858.

Hunt, R. J., Luchette, J., Schreuder, W. A., Rumbaugh, J. O., Doherty, J., Tonkin, M. J., & Rumbaugh, D. B. (2010). Using a cloud to replenish parched groundwater modeling efforts. *Groundwater*, *48*(3), 360-365.

Huyer, W., & Neumaier, A. (1999). Global optimization by multilevel coordinate search. *Journal of Global Optimization*, *14*(4), 331-355.

Huyer, W., & Neumaier, A. (2008). SNOBFIT--stable noisy optimization by branch and fit. *ACM Transactions on Mathematical Software (TOMS)*, *35*(2), 9.

Irsa, J., & Zhang, Y. (2012). A direct method of parameter estimation for steady state flow in heterogeneous aquifers with unknown boundary conditions. *Water Resources Research*, 48(9).

Jensen, J. B. (2003). *Parameter and uncertainty estimation in groundwater modelling* (Doctoral dissertation, The Hydraulics and Coastal Engineering Group, Dept. of Civil Engineering, Aalborg University).

James, A. L., & Oldenburg, C. M. (1997). Linear and Monte Carlo uncertainty analysis for subsurface contaminant transport simulation. *Water Resources Research*, *33*(11), 2495-2508.

Joseph, J. F., & Guillaume, J. H. (2013). Using a parallelized MCMC algorithm in R to identify appropriate likelihood functions for SWAT. *Environmental modelling & software*, *46*, 292-298.

Jung, Y., Ranjithan, R. S., & Mahinthakumar, G. (2011). Subsurface characterization using a D-optimality based pilot point method. *Journal of Hydroinformatics*, *13*(4), 775-793.

Kabir, S. F., Appiah Assumaning, G., & Chang, S. Y. (2017). Efficiency of using 4DVar, 3DVar and EnKF data assimilation methods in groundwater contaminant transport modelling. *European Journal of Environmental and Civil Engineering*, 1-17.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1), 35-45.

Kashyap, R. L. (1982). Optimal choice of AR and MA parts in autoregressive moving average models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), 99-104.

Keating, E. H., Doherty, J., Vrugt, J. A., & Kang, Q. (2010). Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality. *Water Resources Research*, *46*(10).

Keenan, T. F., Carbone, M. S., Reichstein, M., & Richardson, A. D. (2011). The model-data fusion pitfall: assuming certainty in an uncertain world. Oecologia, 167(3), 587.

Kennedy J, & Eberhart R (1995) Particle swarm optimization. In: Proceeding of International Conference on Neural Networks, Perth, Australia. IEEE Service Center, Piscataway, NJ, pp 1942–1948.

Kerrou, J., Renard, P., Franssen, H. J. H., & Lunati, I. (2008). Issues in characterizing heterogeneity and connectivity in non-multiGaussian media. *Advances in Water Resources*, *31*(1), 147-159.

Ketabchi, H., & Ataie-Ashtiani, B. (2015a). Coastal groundwater optimization—advances, challenges, and practical solutions. *Hydrogeology Journal*, 23(6), 1129-1154.

Ketabchi, H., & Ataie-Ashtiani, B. (2015b). Assessment of a parallel evolutionary optimization approach for efficient management of coastal aquifers. *Environmental Modelling & Software*, 74, 21-38.

Ketabchi, H., & Ataie-Ashtiani, B. (2015c). Evolutionary algorithms for the optimal management of coastal groundwater: a comparative study toward future challenges. *Journal of Hydrology*, *520*, 193-213.

Kikuchi, C. P., Ferré, T. P. A., & Vrugt, J. A. (2015). On the optimal design of experiments for conceptual and predictive discrimination of hydrologic system models. *Water Resources Research*, *51*(6), 4454-4481.

Kikuchi, C. (2017). Toward Increased Use of Data Worth Analyses in Groundwater Studies. *Groundwater*, *55*(5), 670-673.

Khader, A. I., & McKee, M. (2014). Use of a relevance vector machine for groundwater quality monitoring network design under uncertainty. *Environmental modelling & software*, *57*, 115-126.

Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, *14*(1), 28-44.

Kitagawa, G. (1987). Non-gaussian state—space modeling of nonstationary time series. *Journal of the American statistical association*, 82(400), 1032-1041.

Kollet, S. J., Maxwell, R. M., Woodward, C. S., Smith, S., Vanderborght, J., Vereecken, H., & Simmer, C. (2010). Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources. *Water Resources Research*, *46*(4).

Kollat, J. B., Reed, P. M., & Maxwell, R. M. (2011). Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics. *Water Resources Research*, *47*(2).

Konikow, L. F., & Bredehoeft, J. D. (1992). Ground-water models cannot be validated. *Advances in water resources*, *15*(1), 75-83.

Kowalsky, M. B., Finsterle, S., & Rubin, Y. (2004). Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. *Advances in Water Resources*, *27*(6), 583-599.

Kowalsky, M. B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., ... & Gee, G. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resources Research*, *41*(11).

Kowalsky, M. B., Chen, J., & Hubbard, S. S. (2006). Joint inversion of geophysical and hydrological data for improved subsurface characterization. *The Leading Edge*, *25*(6), 730-734.

Krueger, T., Page, T., Hubacek, K., Smith, L., & Hiscock, K. (2012). The role of expert opinion in environmental modelling. *Environmental Modelling & Software*, *36*, 4-18.

Kurtz, W., Lapin, A., Schilling, O. S., Tang, Q., Schiller, E., Braun, T., ... & Franssen, H. J. H. (2017). Integrating hydrological modelling, data assimilation and cloud computing for real-time management of water resources. *Environmental modelling & software*, *93*, 418-435.

Lahoz, W. A., Errera, Q., Swinbank, R., & Fonteyn, D. (2007). Data assimilation of stratospheric constituents: a review. *Atmospheric Chemistry and Physics*, 7(22), 5745-5773.

Laloy, E., & Vrugt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research*, *48*(1).

Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., & Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. *Water Resources Research*, *49*(5), 2664-2682.

Lambot, S., Javaux, M., Hupet, F., & Vanclooster, M. (2002). A global multilevel coordinate search procedure for estimating the unsaturated soil hydraulic properties. *Water Resources Research*, *38*(11).

Lambot, S., Rhebergen, J., Van den Bosch, I., Slob, E. C., & Vanclooster, M. (2004). Measuring the soil water content profile of a sandy soil with an off-ground monostatic ground penetrating radar. *Vadose Zone Journal*, *3*(4), 1063-1071.

Leaf, A. T. (2017). Using models to identify the best data: An example from northern Wisconsin. *Groundwater*, *55*(5), 641-645.

Leng, C. H., & Yeh, H. D. (2003). Aquifer parameter identification using the extended Kalman filter. *Water resources research*, *39*(3).

Langevin, C. D., & Panday, S. (2012). Future of groundwater modeling. *Groundwater*, 50(3), 334-339.

Leisenring, M., & Moradkhani, H. (2011). Snow water equivalent prediction using Bayesian data assimilation methods. *Stochastic Environmental Research and Risk Assessment*, *25*(2), 253-270.

Lele, S. R., & Allen, K. L. (2006). On using expert opinion in ecological analyses: a frequentist approach. *Environmetrics*, *17*(7), 683-704.

Li, X., & Tsai, F. T. C. (2009). Bayesian model averaging for groundwater head prediction and uncertainty analysis using multimodel and multimethod. *Water resources research*, *45*(9).

Li, L., Zhou, H., & Gómez-Hernández, J. J. (2011). Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media. *Advances in water resources*, *34*(4), 478-489.

Li, L., Zhou, H., Hendricks Franssen, H. J., & Gómez-Hernández, J. J. (2012). Modeling transient groundwater flow by coupling ensemble Kalman filtering and upscaling. *Water Resources Research*, *48*(1).

Linde, N., Ginsbourger, D., Irving, J., Nobile, F., & Doucet, A. (2017). On uncertainty quantification in hydrogeology and hydrogeophysics. *Advances in Water Resources*, *110*, 166-181.

Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*(7).

Liu, G., Chen, Y., & Zhang, D. (2008). Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. *Advances in Water Resources*, *31*(7), 975-986.

Liu, H. J., Hsu, N. S., & Lee, T. H. (2009). Simultaneous identification of parameter, initial condition, and boundary condition in groundwater modelling. *Hydrological processes*, *23*(16), 2358-2367.

Liu, Y., Weerts, A., Clark, M., Hendricks Franssen, H. J., Kumar, S., Moradkhani, H., ... & Van Velzen, N. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities.

Lu, D., Ye, M., Hill, M. C., Poeter, E. P., & Curtis, G. P. (2014). A computer program for uncertainty analysis integrating regression and Bayesian methods. *Environmental Modelling & Software*, *60*, 45-56.

Lu, D., Ye, M., & Curtis, G. P. (2015). Maximum likelihood Bayesian model averaging and its predictive analysis for groundwater reactive transport models. *Journal of Hydrology*, *529*, 1859-1873.

Lu, Z., & Zhang, D. (2003). On importance sampling Monte Carlo approach to uncertainty analysis for flow and transport in porous media. *Advances in Water Resources*, *26*(11), 1177-1188.

Ma, Y. Z., & La Pointe, P. R. (Eds.). (2011). Uncertainty Analysis and Reservoir Modeling: Developing and Managing Assets in an Uncertain World, AAPG Memoir 96 (Vol. 96). AAPG.

Montanari, A., Shoemaker, C. A., & van de Giesen, N. (2009). Introduction to special section on Uncertainty Assessment in Surface and Subsurface Hydrology: An overview of issues and challenges. *Water Resources Research*, *45*(12).
Mantovan, P., & Todini, E. (2006). Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of hydrology*, *330*(1-2), 368-381.

Mantoglou, A. (2003). Estimation of heterogeneous aquifer parameters from piezometric data using ridge functions and neural networks. *Stochastic Environmental Research and Risk Assessment*, *17*(5), 339-352.

McLaughlin, D. B. (1976). Application of Kalman filtering to groundwater basin modeling and prediction. *Real-Time Forecasting/Control of Water Resource Systems, IIASA Proc. Ser*, 109-123.

McLaughlin, D., & Townley, L. R. (1996). A reassessment of the groundwater inverse problem. *Water Resources Research*, *32*(5), 1131-1161.

Medina, A., & Carrera, J. (1996). Coupled estimation of flow and solute transport parameters. Water Resources Research, 32(10), 3063-3076.

Michalak, A. M. (2008). A Gibbs sampler for inequality-constrained geostatistical interpolation and inverse modeling. *Water Resources Research*, 44(9).

Miller, R. N., Ghil, M., & Gauthiez, F. (1994). Advanced data assimilation in strongly nonlinear dynamical systems. *Journal of the atmospheric sciences*, *51*(8), 1037-1056.

Moore, C., & Doherty, J. (2006). The cost of uniqueness in groundwater model calibration. *Advances in Water Resources*, *29*(4), 605-623.

Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in water resources*, *28*(2), 135-147.

Moradkhani, H., DeChant, C. M., & Sorooshian, S. (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resources Research*, *48*(12).

Morse, B. S., Pohll, G., Huntington, J., & Rodriguez Castillo, R. (2003). Stochastic capture zone analysis of an arsenic-contaminated well using the generalized likelihood uncertainty estimator (GLUE) methodology. *Water resources research*, *39*(6).

Mugunthan, P., Shoemaker, C. A., & Regis, R. G. (2005). Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models. *Water Resources Research*, *41*(11).

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, 7(4), 308-313.

Neuman, S. P. (1973). Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. *Water Resources Research*, *9*(4), 1006-1021.

Neuman, S. P. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment*, *17*(5), 291-305.

Neuman, S. P., Xue, L., Ye, M., & Lu, D. (2012). Bayesian analysis of data-worth considering model and parameter uncertainties. *Advances in Water Resources*, *36*, 75-85.

Ng, G. H. C., McLaughlin, D., Entekhabi, D., & Scanlon, B. (2009). Using data assimilation to identify diffuse recharge mechanisms from chemical and physical data in the unsaturated zone. *Water resources research*, *45*(9).

Nisbet, E. G., Dlugokencky, E. J., & Bousquet, P. (2014). Methane on the rise—again. *Science*, *343*(6170), 493-495.

Norberg, T., & Rosén, L. (2006). Calculating the optimal number of contaminant samples by means of data worth analysis. *Environmetrics*, *17*(7), 705-719.

Norby, R. J., De Kauwe, M. G., Domingues, T. F., Duursma, R. A., Ellsworth, D. S., Goll, D. S., ... & Pavlick, R. (2016). Model–data synthesis for the next generation of forest free-air CO2 enrichment (FACE) experiments. *New Phytologist*, 209(1), 17-28.

Nordqvist, R. (2000). Comparison of parameter estimation design criteria using a solute transport model with matrix diffusion. *Groundwater*, *38*(6), 827-835.

Nowak, W. (2010). Measures of parameter uncertainty in geostatistical estimation and geostatistical optimal design. *Mathematical Geosciences*, *42*(2), 199-221.

Nowak, W., & Cirpka, O. A. (2004). A modified Levenberg–Marquardt algorithm for quasilinear geostatistical inversing. *Advances in water resources*, *27*(7), 737-750.

O'Hagan, A. (2012). Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. *Environmental Modelling & Software*, *36*, 35-48.

Oliver, D. S., Cunha, L. B., & Reynolds, A. C. (1997). Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology*, *29*(1), 61-91.

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, *263*(5147), 641-646.

Patil, A., Huard, D., & Fonnesbeck, C. J. (2010). PyMC: Bayesian stochastic modelling in Python. *Journal of statistical software*, *35*(4), 1.

Pau, G. S. H., Finsterle, S., & Zhang, Y. (2016). Fast high-resolution prediction of multi-phase flow in fractured formations. *Advances in Water Resources*, *88*, 80-85.

Pappenberger, F., & Beven, K. J. (2006). Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water resources research*, *42*(5).

Poeter, E. P., & Hill, M. C. (1999). UCODE, a computer code for universal inverse modeling1. *Computers & Geosciences*, *25*(4), 457-462.

Pollock, D., & Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by electrical resistivity tomography. *Water Resources Research*, *48*(1).

Ponzini, G., & Lozej, A. (1982). Identification of aquifer transmissivities: the comparison model method. *Water Resources Research*, *18*(3), 597-622.

Porter, D. W., Gibbs, B. P., Jones, W. F., Huyakorn, P. S., Hamm, L. L., & Flach, G. P. (2000). Data fusion modeling for groundwater systems. *Journal of Contaminant Hydrology*, *42*(2), 303-335.

Peeters, L., Fasbender, D., Batelaan, O., & Dassargues, A. (2010). Bayesian data fusion for water table interpolation: incorporating a hydrogeological conceptual model in kriging. *Water Resources Research*, *46*(8).

Peng, C., Guiot, J., Wu, H., Jiang, H., & Luo, Y. (2011). Integrating models with data in ecology and palaeoecology: advances towards a model–data fusion approach. *Ecology Letters*, *14*(5), 522-536.

Pham, H. V., & Tsai, F. T. C. (2016). Optimal observation network design for conceptual model discrimination and uncertainty reduction. *Water Resources Research*, *52*(2), 1245-1264.

Qian, S. S., Stow, C. A., & Borsuk, M. E. (2003). On monte carlo methods for Bayesian inference. *Ecological Modelling*, *159*(2-3), 269-277.

Rajabi, M. M., & Ataie-Ashtiani, B. (2014). Sampling efficiency in Monte Carlo based uncertainty propagation strategies: application in seawater intrusion simulations. *Advances in Water Resources*, 67, 46-64.

Rajabi, M. M., Ataie-Ashtiani, B., & Janssen, H. (2015a). Efficiency enhancement of optimized Latin hypercube sampling strategies: application to Monte Carlo uncertainty analysis and meta-modeling. *Advances in water resources*, *76*, 127-139.

Rajabi, M. M., Ataie-Ashtiani, B., & Simmons, C. T. (2015b). Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. *Journal of Hydrology*, *520*, 101-122.

Rajabi, M. M., & Ataie-Ashtiani, B. (2016). Efficient fuzzy Bayesian inference algorithms for incorporating expert knowledge in parameter estimation. *Journal of Hydrology*, *536*, 255-272.

Rajabi, M. M., & Ketabchi, H. (2017). Uncertainty-based simulation-optimization using Gaussian process emulation: Application to coastal groundwater management. *Journal of Hydrology*, *555*, 518-534.

Rajib, A. I., Assumaning, G. A., Chang, S. Y., & Addai, E. B. (2017). Use of Multiple Data Assimilation Techniques in Groundwater Contaminant Transport Modeling. *Water Environment Research*, *89*(11), 1952-1960.

Rauch, H. (1963). Solutions to the linear smoothing problem. *IEEE Transactions on Automatic Control*, 8(4), 371-372.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., ... & Schmullius, C. C. (2005). Model–data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology*, *11*(3), 378-397.

Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., Li, D., Inverarity, G. W., ... & Payne, T. J. (2007). The Met Office global four-dimensional variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, *133*(623), 347-362.

Rayner, P., Michalak, A. M., & Chevallier, F. (2016). Fundamentals of data assimilation. *Geosci. Model Dev.* Discuss, 1-21.

Refsgaard, J. C., Van der Sluijs, J. P., Brown, J., & Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources*, *29*(11), 1586-1597.

Refsgaard, J. C., Højberg, A. L., Møller, I., Hansen, M., & Søndergaard, V. (2010). Groundwater modeling in integrated water resources management—Visions for 2020. *Groundwater*, *48*(5), 633-648.

Refsgaard, J. C., Christensen, S., Sonnenborg, T. O., Seifert, D., Højberg, A. L., & Troldborg, L. (2012). Review of strategies for handling geological uncertainty in groundwater flow and transport modeling. *Advances in Water Resources*, *36*, 36-50.

Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in water resources*, 31(11), 1411-1418.

Reichle, R. H., Crow, W. T., & Keppenne, C. L. (2008). An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water resources research*, *44*(3).

Rinderknecht, S. L., Albert, C., Borsuk, M. E., Schuwirth, N., Künsch, H. R., & Reichert, P. (2014). The effect of ambiguous prior knowledge on Bayesian model parameter inference and prediction. *Environmental modelling & software*, *62*, 300-315.

Ristic, B., Arulampalam, S., & Gordon, N. (2003). *Beyond the Kalman filter: Particle filters for tracking applications*. Artech house.

Rojas, R., Feyen, L., Batelaan, O., & Dassargues, A. (2010a). On the value of conditioning data to reduce conceptual model uncertainty in groundwater modeling. *Water Resources Research*, *46*(8).

Rojas, R., Kahunde, S., Peeters, L., Batelaan, O., Feyen, L., & Dassargues, A. (2010b). Application of a multimodel approach to account for conceptual model and scenario uncertainties in groundwater modelling. *Journal of Hydrology*, *394*(3-4), 416-435.

Romanowicz, R., Beven, K. J., & Tawn, J. (1994). Evaluation of predictive uncertainty in nonlinear hydrological models using a Bayesian approach. *Statistics for the Environment*, *2*, 297-317.

Rios, L. M., & Sahinidis, N. V. (2013). Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, *56*(3), 1247-1293.

Ritzel, B. J., Eheart, J. W., & Ranjithan, S. (1994). Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. *Water Resources Research*, *30*(5), 1589-1603.

Ross, J. L., Ozbek, M. M., & Pinder, G. F. (2009). Aleatoric and epistemic uncertainty in groundwater flow and transport simulation. *Water Resources Research*, *45*(12).

Sage, A. P., & Melsa, J. L. (1971). *Estimation theory with applications to communications and control*. SOUTHERN METHODIST UNIV DALLAS TEX INFORMATION AND CONTROL SCIENCES CENTER.

Samuel, M. P., & Jha, M. K. (2003). Estimation of aquifer parameters from pumping test data by genetic algorithm optimization technique. *Journal of irrigation and drainage engineering*, *129*(5), 348-359.

Sanchez-Vila, X., Guadagnini, A., & Carrera, J. (2006). Representative hydraulic conductivities in saturated groundwater flow. *Reviews of Geophysics*, *44*(3).

Sanford, W. E., Pope, J. P., & Nelms, D. L. (2009). *Simulation of groundwater-level and salinity changes in the Eastern Shore, Virginia*. Denver, CO: US Geological Survey.

Särkkä, S. (2013). Bayesian filtering and smoothing (Vol. 3). Cambridge University Press.

Schöniger, A., Nowak, W., & Hendricks Franssen, H. J. (2012). Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography. *Water Resources Research*, *48*(4).

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Sepúlveda, N., & Doherty, J. (2015). Uncertainty Analysis of a Groundwater Flow Model in East-Central Florida. *Groundwater*, *53*(3), 464-474.

Siade, A. J., Hall, J., & Karelse, R. N. (2017). A Practical, Robust Methodology for Acquiring New Observation Data Using Computationally Expensive Groundwater Models. *Water Resources Research*, *53*(11), 9860-9882.

Simmons, C. T., Hunt, R. J., & Cook, P. G. (2012). Using every tool in the toolbox. *Groundwater*, 50(3), 323-323.

Singh, A., & Minsker, B. S. (2008). Uncertainty-based multiobjective optimization of groundwater remediation design. *Water resources research*, 44(2).

Smith, L., & Schwartz, F. W. (1981). Mass transport: 2. Analysis of uncertainty in prediction. *Water Resources Research*, *17*(2), 351-369.

Smith, P. J., Beven, K. J., & Tawn, J. A. (2008). Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. *Water resources research*, *44*(1).

Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, *136*(12), 4629-4640.

Sovinsky, V. E. (2017). Comparing Groundwater Model Calibration Approaches: Does an Optimized Model Better Reflect Reality When Both Flow and Transport Observations are Applied?.

Spielgelhalter, D., Thomas, A., & Best, N. (2003). WinBUGS version 1.4. Bayesian inference using Gibbs sampling. *Biostatistics Unit, Institute for Public Health. Cambridge, United Kingdom.*

Stein, M., Beer, M., & Kreinovich, V. (2013). Bayesian approach for inconsistent information. *Information sciences*, *245*, 96-111.

Stone, N. (2011). *Gaussian process emulators for uncertainty analysis in groundwater flow* (Doctoral dissertation, University of Nottingham).

Sun, A. Y., Painter, S. L., & Wittmeyer, G. W. (2006). A robust approach for iterative contaminant source location and release history recovery. *Journal of contaminant hydrology*, *88*(3-4), 181-196.

Szalkai, A.R., P. Nagy, T. Tullner, et al. 2007. Organisation of ground water data management in the participating countries. WP3 report from the eWater project (Multilingual cross-border access to groundwater databases). <u>http://www.eWater.eu</u>.

Szewczyk, R., Polastre, J., Mainwaring, A., & Culler, D. (2004, January). Lessons from a sensor network expedition. In *European Workshop on Wireless Sensor Networks* (pp. 307-322). Springer, Berlin, Heidelberg.

Tang, G., D'Azevedo, E. F., Zhang, F., Parker, J. C., Watson, D. B., & Jardine, P. M. (2010). Application of a hybrid MPI/OpenMP approach for parallel groundwater model calibration using multi-core computers. *Computers & Geosciences*, *36*(11), 1451-1460.

Thiemann, M., Trosset, M., Gupta, H., & Sorooshian, S. (2001). Bayesian recursive parameter estimation for hydrologic models. *Water Resources Research*, 37(10), 2521-2535.

Tierney, L., & Mira, A. (1999). Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in medicine*, *18*(1718), 2507-2515.

Tonkin, M., & Doherty, J. (2009). Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques. *Water Resources Research*, *45*(12).

Triki, C., Zekri, S., Al-Maktoumi, A., & Fallahnia, M. (2017). An Artificial Intelligence Approach for the Stochastic Management of Coastal Aquifers. *Water Resources Management*, *31*(15), 4925-4939.

Troldborg, L., Refsgaard, J. C., Jensen, K. H., & Engesgaard, P. (2007). The importance of alternative conceptual models for simulation of concentrations in a multi-aquifer system. *Hydrogeology Journal*, *15*(5), 843-860.

Tsai, F. T. C., Sun, N. Z., & Yeh, W. W. G. (2003). Global-local optimization for parameter structure identification in three-dimensional groundwater modeling. *Water resources research*, *39*(2).

Tsai, F. T. C., & Li, X. (2008). Inverse groundwater modeling for hydraulic conductivity estimation using Bayesian model averaging and variance window. *Water Resources Research*, *44*(9).

Tsai, F. T. C. (2010). Bayesian model averaging assessment on groundwater management under model structure uncertainty. *Stochastic Environmental Research and Risk Assessment*, *24*(6), 845-861.

van der Spek, J. E., & Bakker, M. (2017). The influence of the length of the calibration period and observation frequency on predictive uncertainty in time series modeling of groundwater dynamics. *Water Resources Research*, *53*(3), 2294-2311.

Van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, *137*(12), 4089-4114.

Van Leeuwen, P. J., & Evensen, G. (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, *124*(12), 2898-2913.

van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2016). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic bulletin & review*, 1-12.

Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A., & Vesselinov, V. V. (2008). Inverse Modeling of Subsurface Flow and Transport Properties: A Review with New Developments All rights reserved. *Vadose Zone Journal*, *7*(2), 843-864.

Vrugt, J. A., Ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, *10*(3), 273-290.

Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, *75*, 273-316.

Wallis, I., Moore, C., Post, V., Wolf, L., Martens, E., & Prommer, H. (2014). Using predictive uncertainty analysis to optimise tracer test design and data acquisition. *Journal of hydrology*, *515*, 191-204.

Wang, Y. P., Trudinger, C. M., & Enting, I. G. (2009). A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. *Agricultural and Forest Meteorology*, *149*(11), 1829-1842.

Waters, P., Greenbaum, D., Smart, P. L., & Osmaston, H. (1990). Applications of remote sensing to groundwater hydrology. *Remote Sensing Reviews*, 4(2), 223-264.

Welter, D. E., White, J. T., Hunt, R. J., & Doherty, J. E. (2015). *Approaches in highly* parameterized inversion—*PEST++* Version 3, a Parameter ESTimation and uncertainty analysis software suite optimized for large environmental models (No. 7-C12). US Geological Survey.

Wheater, H. S., Tompkins, J. A., Van Leeuwen, M., & Butler, A. P. (2000). Uncertainty in groundwater flow and transport modelling—a stochastic analysis of well-protection zones. *Hydrological Processes*, *14*(11-12), 2019-2029.

Whitaker, J. S., & Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, *130*(7), 1913-1924.

White, J. T. (2018). A model-independent iterative ensemble smoother for efficient historymatching and uncertainty quantification in very high dimensions. *Environmental Modelling & Software*.

White, J., & Brakefield, L. K. (2015, December). pyNSMC: A Python Module for Null-Space Monte Carlo Uncertainty Analysis. In *AGU Fall Meeting Abstracts*.

White, J. T., Fienen, M. N., & Doherty, J. E. (2016). A python framework for environmental model uncertainty analysis. *Environmental Modelling & Software*, *85*, 217-228.

White, J. T., Fienen, M. N., Barlow, P. M., & Welter, D. E. (2018). A tool for efficient, modelindependent management optimization under uncertainty. *Environmental Modelling & Software*, *100*, 213-22

Woodbury, A. D., Smith, L., & Dunbar, W. S. (1987). Simultaneous inversion of hydrogeologic and thermal data: 1. Theory and application using hydraulic head data. *Water Resources Research*, *23*(8), 1586-1606.

Wöhling, T., Geiges, A., & Nowak, W. (2016). Optimal Design of Multitype Groundwater Monitoring Networks Using Easily Accessible Tools. *Groundwater*, *54*(6), 861-870.

Wu, J. C., Lu, L., & Tang, T. (2011). Bayesian analysis for uncertainty and risk in a groundwater numerical model's predictions. *Human and Ecological Risk Assessment: An International Journal*, *17*(6), 1310-1331.

Wu, J., & Zeng, X. (2013). Review of the uncertainty analysis of groundwater numerical simulation. *Chinese Science Bulletin*, *58*(25), 3044-3052.

Xu, T., Gómez-Hernández, J. J., Li, L., & Zhou, H. (2013). Parallelized ensemble Kalman filter for hydraulic conductivity characterization. *Computers & geosciences*, *52*, 42-49.

Xu, T., & Gómez-Hernández, J. J. (2016a). Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. *Water Resources Research*, 52(8), 6587-6595.

Xu, T., & Gómez-Hernández, J. J. (2016b). Characterization of non-Gaussian conductivities and porosities with hydraulic heads, solute concentrations, and water temperatures. *Water Resources Research*, 52(8), 6111-6136.

Xu, T., & Gómez-Hernández, J. J. (2018). Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. *Advances in Water Resources*, *112*, 106-123.

Xue, L., Zhang, D., Guadagnini, A., & Neuman, S. P. (2014). Multimodel Bayesian analysis of groundwater data worth. *Water Resources Research*, *50*(11), 8481-8496.

Ye, M., Neuman, S. P., & Meyer, P. D. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research*, 40(5).

Ye, M., Neuman, S. P., Meyer, P. D., & Pohlmann, K. (2005). Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff. *Water Resources Research*, *41*(12).

Ye, M., Pohlmann, K. F., Chapman, J. B., Pohll, G. M., & Reeves, D. M. (2010). A model-averaging method for assessing groundwater conceptual model uncertainty. *Groundwater*, *48*(5), 716-728.

Yeh, W. G. (1986). Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research*, *22*(2), 95-108.

Yeh, H. D., & Huang, Y. C. (2005). Parameter estimation for leaky aquifers using the extended Kalman filter, and considering model and data measurement uncertainties. *Journal of Hydrology*, *302*(1-4), 28-45.

Yeh, T. C. J., Lee, C. H., Hsu, K. C., & Tan, Y. C. (2007). Fusion of active and passive hydrologic and geophysical tomographic surveys: The future of subsurface characterization. *Subsurface Hydrology: Data Integration for Properties and Processes*, 109-120.

Zell, W. O., Culver, T. B., & Sanford, W. E. (2018). Prediction uncertainty and data worth assessment for groundwater transport times in an agricultural catchment. *Journal of Hydrology*.

Zhang, J., Li, W., Zeng, L., & Wu, L. (2016). An adaptive Gaussian process-based method for efficient Bayesian experimental design in groundwater contaminant source identification problems. *Water Resources Research*, *52*(8), 5971-5984.

Zhou, Y., McLaughlin, D., & Entekhabi, D. (2006). Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Monthly Weather Review*, *134*(8), 2128-2142.

Zhou, H., Li, L., & Gómez-Hernández, J. J. (2010). Three-dimensional hydraulic conductivity upscaling in groundwater modeling. *Computers & Geosciences*, *36*(10), 1224-1235.

Zhou, H., Gómez-Hernández, J. J., & Li, L. (2012). A pattern-search-based inverse method. *Water Resources Research*, *48*(3).

Zhou, H., Gómez-Hernández, J. J., & Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends. *Advances in Water Resources*, *63*, 22-37.

Highlights:

- Discussion are provided on how groundwater numerical models and data interact.
- Frequentist and Bayesian methods are reviewed and classified.
- Eight application areas of groundwater model-data interaction are reviewed.
- Discussions on the state of the art in software and codes are presented.
- Recent trends and issues that will likely form future directions are discussed.

Table 1. Review of common data types relevant to groundwater MDI

Classification	Energyles	Sources of data			
Classification	Examples	Hard data	Soft data		
Model	Subsurface characteristics, such as hydraulic conductivity, storativity, porosity, etc.	Slug tests, pumping tests, tracer techniques, permeameter test, laboratory analysis of field samples	Geophysical surveys, expert knowledge		
Parameters	Contaminant transport and transformation characteristics, such as dispersivity, sorption factor, chamical reaction rates ato	Tracer techniques	Expert knowledge		
External Forcing	Surface recharge, lateral inflow, river-aquifer interactions, extraction/injection rates, etc.	Lysimeter, tracer techniques, water balance methods, flowmeters	Remote sensing, expert knowledge		
Model Structure	Site geometry, geological formations, etc.	Borehole investigations	Geophysical surveys, remote sensing, expert knowledge		
	Hydraulic head, water table, tidal induced water table fluctuations	Measurement tapes, pressure sensors			
State variables	Chemical concentration	Laboratory analysis of field samples, portable field instruments	Geophysical surveys, remote sensing		
	Temperature	Portable field instruments	Ũ		
	Travel time	Tracer techniques			
	Discharge	Flowmeters			

Table 2. Review of some popular MDI codes in groundwater applications

Code	Key algorithms	Important features	Program ming language	Applications
PEST (Doherty, 1994)	Weighted LSE, GML optimization, Tikhonov (and other) regularization algorithms	MI, OS, Includes several utility software for, e.g. pilot points parameterization, pre/post processing and global, derivate-free optimization	Fortran	DWA (e.g. Wallis et al., 2014; Wöhling et al., 2016), PE & UA (e.g. Barnhart et al., 2010; Ataie-Ashtiani et al., 2013; Drumheller et al., 2017)

PEST++		MLOS object-		
(Welter et al., 2015) and various codes	Weighted LSE, GML	oriented, Integrated TCP/IP parallel run management		PE & UA (Heywood et al.,
within its framework, e.g. PESTPP- IES, ESTPP- OPT, etc.	other) regularization algorithms		C++	2016; Sovinsky, 2017; White et al., 2018)
- ,		MI, OS, Includes		
UCODE (Poeter and Hill, 1999)	Weighted LSE, Gauss– Newton optimization	MMA program which supports the use of multiple alternative models	Perl, Fortran	PE & UA (Sanford et al., 2009; Rojas et al., 2010b)
MODFLOWP (Hill, 1992)	Weighted LSE, modified Gauss-Newton or conjugate- direction optimization	MD (for MODFLOW)	Fortran	PE (Heywood and Yager, 2003)
iTOUGH2	Weighted LSE and other robust estimators (L ₁ , Huber Cauchy and	Initially intended as a MD code (for TOUGH2) but can		PE & UA (James and Oldenburg, 1997; Pau et al.,
(Finsterle, 1993)	Andrews), LM, Downhill Simplex, Simulated Annealing and Grid Search optimization	now be used with any model	Fortran	2016), DWA (Finsterle, 2015). See: http://esd1.lbl.gov/iTOUGH 2/
				PE & UA (Huber et al., 2011), SPE (Hendricks
EnKF3d	EnKF	MI, OS	C	2009; Hendricks Franssen et al., 2011)
WinBUGS		ML OC Includes		
OpenBUGS (Spielgelhalte r et al., 2003)	MCMC based on Gibs sampler	spatial models (GeoBUGS)	Pascal	PE & UA (Stone, 2011)
	MCMC based on			
DREAM (Vrugt, 2016)	Adaptive Metropolis algorithm	MI, OS, Can support BMA	MATLAB	PE & UA (Keating et al., 2010)
	MCMC based on random walk Metropolis algorithm			
MCMC	simulated tempering, and morphometric random walk Metropolis	MI, OS	R	PE & UA (Han et al., 2015)
pyEMU	Linear first-order, second-	MI, OS, Scalable to		
(White et a., 2016)	moment (FOSM) and non- linear uncertainty analyses	parameterized inverse problems	Python	DWA (Zell et al., 2018)
		mi, OS, includes methods includes		
pyMC (Patil	several types of MCMC algorithms (e.g. Metropolis-	summarizing	Dython	PE & UA (van der Spek and
et al., 2010)	Hastings, Gibs and Random- walk Metropolis-Hastings)	output, plotting, goodness-of-fit and convergence	ryuloli	Bakker, 2017)
		diagnostics		

MI: Model-independent MD: Model-dependent OS: open-source Acctinition GML: Gauss-Marquardt-Levenberg DWA: Data worth analysis





Fig. 2. Schematic representation of a groundwater flow and contaminant transport model

RCC

