Solute Tracer Tomography

Field Implementation and Parameter Estimation using the Ensemble Kalman Filter

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.Sc. Eduardo Emilio Sánchez León
aus Mexiko Stadt, Mexiko

Tübingen

2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	21.02.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr.–Ing. Olaf A. Cirpka
2. Berichterstatter:	Prof. Dr. Peter Dietrich

To Rosa, Laures, Margos,

Gogdo, Burus and Hofs...

Acknowledgments

I thank my supervisors Olaf Cirpka and Carsten Leven for the patience, guidance and support throughout the whole work. A very special gratitude to Claus Haslauer and Luise Hofmann for the comments, insights and proofreading of the thesis. I also thank those students involved in the different activities related to the work.

Finally, I thank the National Council of Research and Technology of Mexico (CONACYT) for assigning me a *Human Resources Development Scholarship* to support the PhD studies in Germany.

Abstract

The performance of groundwater flow and solute transport models depends, to a large extent, on the resolution at which aquifer heterogeneity is resolved. Large datasets are needed to estimate aquifer parameters at a high resolution, but their size is usually limited in real-world applications. The evolution of modern measurement techniques including better, smaller and affordable sensors have fostered the development of field tests with a tomographic layout, where the system is stressed in different directions. Recent numerical studies show the advantages of using large datasets of different type for describing small-scale features of aquifer properties. While hydraulic tomography, where multiple pumping tests are performed, has been repeatedly applied at field scale, applications of tracer tomography, where multiple tracer tests are performed, lag behind due to technical limitations.

This study pursues to narrow the gap between numerical studies and field applications, showing the potential of hydraulic and tracer tomography for high-resolution aquifer characterization. In contrast to the few reported applications of tracer tomography, in which heat was used as a tracer, the experimental setup developed in this work was designed to use fluorescein as a conservative tracer, and was applied in the shallow alluvial aquifer at the Hydrogeological Research Site Lauswiesen, Germany.

The experimental results demonstrate that solute-tracer tomography can be efficiently applied at the field scale, if a nested-cell forced-gradient flow field is generated prior to tracer injection. Field data were analyzed with the Ensemble Kalman Filter, coupled to a three dimensional groundwater flow and dual-domain transport model to estimate spatially distributed flow and solute transport parameters. The efficiency of the filter allows the description of aquifer heterogeneity at a high resolution while keeping reasonable computational costs. The filter was tested with a synthetic study based on a two-dimensional model resembling the hydrogeological features and well facilities at the field site. The filter settings with the best performance were applied for the estimation of aquifer parameters based on real data. Results of the synthetic study show that parameters estimated with the Ensemble Kalman Filter applied to hydraulic data already contain the main features of the reference field. However, with the inclusion of concentration data the spatial structure of the parameter fields is accentuated, their uncertainty is considerably reduced, and flow and transport model predictions are improved. While the standard update scheme of the Ensemble Kalman Filter is applicable to hydraulic head, it leads to mass balance errors during assimilation of concentration data. Therefore, a restart scheme was applied where the steady-flow model is reinitialized after each parameter update, and transport is simulated from the initial time until the next available measurement time-step.

The estimation of parameters based on drawdown curves measured during the field tracer tomography shows the potential of the Ensemble Kalman Filter in adjusting model parameters to improve groundwater flow simulations. When using the concentration data, the spatial structure of hydraulic conductivity was accentuated and the associated variance reduced. Transport simulations were largely affected by numerical dispersion, and by estimating effective transport parameters such as porosity and dispersivity as uniform values, rather than considering them as spatial fields. The Monte Carlo approach of the Ensemble Kalman Filter imposes limitations in the number of parameters that can be estimated, therefore efficient methods to optimize the grid resolution and reduce matrices dimensions are required.

Zusammenfassung

Die Modellierung der Grundwasserströmung und des Stofftransports hängt in hohem Maße von der Auflösung der Heterogeneität des Grundwasserleiters ab. Um die Parameter des Grundwasserleiters bestimmen zu können, sind große Datenmenge erforderlich, die jedoch in Feldanwendungen meist nicht in erforderlicher Menge zur Verfügung stehen. Die Entwicklung moderner Messtechniken insbesondere besserer, kleinerer und kostengünstigerer Sensoren ermöglichten die Entwicklung tomographischer Untersuchungsmethoden, in denen der Untergrund in unterschiedlicher Richtung angeregt wird. Aktuelle numerische Studien belegen die Vorteile große Datenmengen unterschiedlichen Typs zu nutzen, um kleinskalige Eigenschaften eines Grundwasserleiters zu beschreiben. Während die hydraulische Tomographie, bei der mehrere Pumpversuche durchgeführt werden, in Feldversuchen wiederholt angewendet wurde, ist der Einsatz der Tracertomographie, bei der mehrere Markierversuche durchgeführt werden, hauptsächlich auf Grund technischer Einschränkungen noch sehr eingeschränkt.

Das Ziel der vorliegenden Arbeit ist es, die Kluft zwischen numerischen Studien und Feldanwendungen zu verringern und das Potential von hydraulischer und Tracertomographie für die hochauflösende Charakterisierung von Grundwasserleitern darzustellen. Im Gegensatz zu den wenigen bisher durchgeführten Anwendungen der Tracertomographie mit Wärme als Tracer, wurde in dieser Arbeit Fluorescein (Uranin) als konservativer Tracer eingesetzt. Die Feldversuche wurden in dem flachen alluvialen Grundwasserleiter des Hydrogeologischen Forschungsgeländes Lauswiesen in Tübingen durchgeführt. Die experimentellen Ergebnisse zeigen, dass die Tracertomographie wirksam in Feldversuchen eingesetzt werden kann, wenn vor der Tracerinjektion ein Strömungsfeld erzeugt wurde, bei dem Grundwasserströmung mit zwei Brunnenpaaren erzwungen wird. Die Geländedaten wurden mittels eines Ensemble-Kalman Filters und mit einem dreidimensionalen Grundwasserströmungs- und Transportmodell mit Doppelporositätsansatz analysiert, um die räumliche Verteilung der hydraulischen Leitfähigkeit sowie Transportparameter zu bestimmen. Die Effizienz des Filters erlaubt die hoch auflösende Beschreibung der Heterogeneität des Grundwasserleiters mit angemessenem Rechenaufwand. Der Filter wurde mittels einer synthetischen Studie, basierend auf einem zweidimensionalen Modell, das sich an den hydrogeologischen Eigenschaften und den Brunnenanlagen des Forschungsgeländes orientiert, getestet. Die Filtereinstellungen mit den besten Ergebnissen wurden für die Bestimmung der Parameter des Grundwasserleiters anhand der Felddaten eingesetzt.

Die synthetische Studie zeigt, dass die Parameter, die aus den Piezometerhöhendaten mittels des Ensemble-Kalman Filters abgeleitet wurden, schon die hauptsächlichen Eigenschaften des Referenzfeldes enthalten. Die Berücksichtigung der Konzentrationsdaten verbessert jedoch die Schätzung der räumlichen Struktur der Parameter und auch die Prognosefähigkeit der Grundwasserströmungs- und Stofftransportmodelle. Während das Standardupdate des Ensemble-Kalman Filter hinreichend für die Assimilierung von Piezometerhöhendaten ist, führt es aber bei der Invertierung von Konzentrationsdaten zu Fehlern in der Massenbilanz. Deshalb wird empfohlen einen Neustart-Filter anzuwenden, bei dem das stationäre Strömungsmodell nach jedem Parameterupdate reinititalisiert wird und der Transport vom

Anfangszeitpunkt bis zum nächsten verfügbaren Messzeitpunkt simuliert wird.
Die Parameterbestimmung auf Grundlage tomographisch ermittelter Absenkungskurven zeigt das Potential des Ensemble-Kalman Filters, Simulationen der Grundwasserströmung mittels angepasster Modellparameter zu verbessern. Obwohl unter Einbeziehung der Konzentrationsdaten die räumliche Struktur der Durchlässigkeit verbessert und die damit verbundene Varianz reduziert wurde, wurden die Transportsimulationen größtenteils durch numerische Dispersion beeinflusst. Eine weitere Beeinträchtigung ergibt sich daraus, dass die Transportparameter als konstante Parameter und nicht als räumlich variable Funktionen geschätzt wurden. Der Monte-Carlo Ansatz des Ensemble-Kalman Filters begrenzt die Anzahl der zu bestimmenden Parameter, sodass effiziente Methoden zur Optimierung der Gitterauflösung und Reduzierung der Matrizendimension benötigt werden.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Water resources are indispensable for a growing society undergoing agricultural development, industrial expansion and urbanization. A considerable portion of the water is provided by groundwater, posing major management challenges. A better understanding of underlying processes of groundwater flow and solute transport is essential for meaningful predictions and the consequent application of responsible management policies.

The current understanding of groundwater flow and solute transport in the subsurface has evolved considerably over the last decades. Yet, detailed aquifer characterization remains a challenge for both scientists and practitioners, in part due to two main issues: (i) data scarcity, and (ii) the large computational resources required to resolve the spatial variation of real aquifer properties at a scale of relevance for management. A great effort has been put into the development of field methods for collecting robust datasets informative of aquifer heterogeneity. Their application in real problems, however, has been hindered by the high costs required for implementation.

Modern field methods, assisted with more precise and affordable sensors, facilitate the collection of reliable data during hydrogeological field investigations. Alongside, the ever-growing computing capacity allows constructing larger numerical models that can solve with higher accuracy the governing equations describing subsurface flow and mass transport.

Relevant progress has also been achieved in parameter estimation techniques. For the automation of the parameter estimation process, it is common to use inverse models, in which model parameters are derived from model state observations (e.g., hydraulic heads). Model inversion would most likely require to perform many model simulations, imposing prohibitive computational costs if a highly discretized numerical model is being considered.

Modern simulators yield an optimal computing-time that scales with $O(n \cdot log \cdot n)$, where n

is the number of unknowns (e.g. model parameters). A traditional conjugate gradient solver scales with $O(n^2)$, while defining the system of equations scales with $O(n)$. In the pursuit of optimizing the computational burden and increasing resolution of estimated parameter distributions, many model inversion methods have been developed (see below). Large datasets become then essential for a stable, meaningful, informative, and reliable parameter estimation, which comes at elevated costs and inhibits their application to real problems.

For a broader perspective of the efforts oriented to circumvent data-scarcity and computational burden problems, this chapter is dedicated to describe the state-of-the-art of the field methods for aquifer characterization and model inversion for parameter estimation. The chapter contains a statement about the motivation and objectives of this work and concludes with a brief outline of the thesis.

The main subject of this work is the implementation of solute tracer tomography as field-scale site-investigation method, and the integration of the Ensemble Kalman Filter to estimate the aquifer parameters. This chapter is *only* a brief introduction to the two topics, stressing their role in modern studies related to aquifer characterization. The theory behind the Ensemble Kalman Filter is given in Chapter 2, and a more elaborated description of hydraulic and tracer tomographic methods is presented in Chapter 3.

1.1 Hydrogeological Characterization

Flow of water and transport of solutes in aquifers is mainly controlled by subsurface heterogeneity. Detailed knowledge of the spatial distribution of the aquifer properties (e.g. hydraulic conductivity K, porosity n, dispersivity D) is key for a reliable assessment of groundwater flow and solute transport in the subsurface.

Direct measurements of aquifer parameters can be obtained with different methods. Hydraulic conductivity can be deduced from grain size analysis (Bear 1972; Wolf et al. 1991; Rosas et al. 2014), or measured by permeameter tests (Wolf et al. 1991; Butler et al. 2007) or centrifugation (Nimmo and Mello, 1991). These methods require the collection of samples that are usually prone to disturbances, affecting the natural conditions of the material. Furthermore, they are only representative of a very small (point-like) support volume.

A common practice among researchers and practitioners is to estimate effective hydraulic parameters by well-testing. This approach involves hydraulic stimulation of the aquifer and observing its response. The cause-effect relationship between stimuli and response is usually defined by mathematical formulations that describe the physical processes involved. Pumping (e.g., Butler 2009) and tracer testing (e.g., Ptak et al. 2004) are among the most popular

well-testing techniques. Hydraulic conductivity (or transmissivity) and storativity are typically estimated from hydraulic head changes monitored during hydraulic tests, while porosity and dispersivities can be obtained from tracer test data. Traditional methods for the analysis of pumping and tracer tests often involve matching a type-curve with a variety of analytical and semianalytical solutions of the groundwater flow (Theis 1935; Cooper and Jacob 1946; Neuman 1975) and advection-dispersion equations (Sauty 1980; Neville et al. 2000). Those solutions are usually derived assuming homogeneity within the aquifer system, and provide limited information on the spatial variation of aquifer parameters (e.g., Butler Jr. and Liu 1993; Welty and Gelhar 1994; Sánchez-Vila et al. 1999; Yeh and Liu 2000; Tiedeman and Hsieh 2004; Leven and Dietrich 2006; Li et al. 2007; Yeh and Zhu 2007; Butler 2009).

Borehole flowmeter tests and direct-push based techniques (slug, injection logging and permeameter) are informative of variations of hydraulic conductivity (Crisman et al. 2001; Butler et al. 2007; Dietrich et al. 2008; Dietrich and Leven 2009; Lessoff et al. 2010; Alexander et al. 2011; Bianchi et al. 2011; Basiricò et al. 2015) but as for direct measurements, they are representative of a very small support volume and therefore provide mainly point information. While the ease and fast measurement acquisition rates of these techniques allow performing multiple tests within a relatively short time period, the information obtained would most likely not be sufficient to resolve the spatial structure at the scale needed to guarantee accurate groundwater flow and solute transport simulations.

A combination of geophysical and hydrogeological methods can be used to enhance the resolution of the field characterization (Linde et al. 2006; Vereecken et al. 2006; Kirsch and Yaramanci 2009; Schwede et al. 2012; Doro et al. 2013; Camporese et al. 2015; Pollock and Cirpka 2008, 2010, 2012), however it is difficult to estimate a subsurface property directly from geophysical parameters (e.g., Behroozmand et al. 2015; Binley et al. 2016; Osterman et al. 2016). Even if a relationship between a geophysical and a hydrogeological parameter can be established, it is often limited to empirical formulations and their validity may be constrained to the site under investigation.

A more sophisticated approach to estimate subsurface parameters is inverse modeling. In its broadest sense inverse modeling is the process of building up a mathematical model that describes a natural system, and finding the model parameters that best reproduce the available observations. Estimating the distribution of aquifer properties at a high resolution usually implies having a large number of parameters, most of the times leading to an ill-posed problem, i.e. the solution is not unique or it is not a continuous function of the data (e.g., Zhou et al. 2014; Hansen 1994), if not enough informative data is available. Many options have been proposed to balance the amount of available data with the number of parameters being estimated. Some suggest reducing the number of parameters by introducing zones or

pilot points, adding regularization terms or including spatial correlation between parameters (de Marsily et al. 1984; Cooley 2000; Laveneu and de Marsily 2001; Carrera et al. 2005; Doherty et al. 2010). However, the necessity of more non-redundant field data to support model inversion will not be removed.

To overcome data insufficiency, field methods with a tomographic layout have gained the attention of the scientific community over the last two decades. The idea behind a hydrogeological tomographic survey is to sequentially stimulate the aquifer at multiple isolated sections, and measure the corresponding response at many observation points. The observations are then sequentially or jointly inverted to estimate hydraulic parameters.

The most common application of a hydrogeological tomographic method is called hydraulic tomography. This method involves a sequential stimulation of the aquifer by injecting and/or extracting water at multiple intervals of an aquifer, and measuring the hydraulic pressure changes at many locations distributed horizontally and (ideally) vertically throughout the aquifer. In one of the first works of hydraulic tomography, Gottlieb and Dietrich (1995) applied a standard least-squares method to estimate a spatially-distributed hydraulic conductivity field, based on pressure data from a synthetic 2D hydraulic tomography experiment. Since then, hydraulic tomography has been under constant development and improvement, and many examples of its application can be found in the literature (see Chapter 3).

A major drawback of hydraulic tomography is the diffusive propagation of the pressure signal, limiting the information contained in pressure data. Recent studies suggest to integrate data of different types during model inversion, e.g. hydraulic pressure and tracer concentrations. Differences in the sensitivity pattern of each type of data improve the resolution and reduce the uncertainty of the estimated parameter fields (e.g. Illman et al. 2010; Yeh and Zhu 2007; Schwede et al. 2014).

Tracer tomography has been suggested as a field method to obtain large tracer datasets and help to improve not only the estimation of aquifer parameters such as hydraulic conductivity, but also of important transport parameters such as porosity and dispervities. In analogy to hydraulic tomography, tracer tomography involves a series of tracer tests, with the tracer injection constrained to a specific portion of the aquifer. The tomographic layout is achieved by shifting the injection interval for each test, and monitoring the tracer plume at different locations and depths (Vasco and Datta-Gupta 1999; Yeh and Zhu 2007; Brauchler et al. 2013a; Doro et al. 2015).

With larger and varied datasets, e.g. from (combined) hydraulic and tracer tomography experiments, appears the need of model inversion schemes that can use the information contained in the data with an affordable computational cost. Different methods to estimate parameters

have been proposed (see the comprehensive reviews by Hendricks Franssen et al. 2009, Zhou et al. 2014, and Kitanidis 2015).

Gradient-based inverse methods require the computation of the sensitivity of all observations with respect to all parameters for each iteration, imposing a restrictive computational burden if a highly discretized model is used. For a stable inversion with gradient-based methods, the number of parameters has to be smaller than the number of independent observations, otherwise the problem is ill-posed. To alleviate the computational burden, in some applications the domain is divided into several zones, however this strongly controls the shape of the solution. Another alternative is to assume that the parameters are continuous spatial fields and regularize the inversion, e.g. by adding a smoothness term to the objective function (e.g., Neuman 1973; Moore and Doherty 2006; Hunt et al. 2007). In the pilot-point method, parameter values are estimated at a limited number of points and interpolated everywhere else within the domain by Kriging (e.g., de Marsily et al. 1984; RamaRao et al. 1995; Doherty 2003).

Geostatistical inversion techniques consider the parameters as spatially correlated random variables. Prior knowledge is included using the covariance function of the parameters, and conditioned on dependent measurements such as hydraulic heads (Kitanidis 1995; Gómez-Hernánez et al. 1997; Zimmerman et al. 1998; Cirpka and Kitanidis 2000; Nowak and Cirpka 2006). The sensitivities are obtained by solving as many adjoint problems as there are measurements, increasing the computational costs when transient processes or data from tomographic tests are being considered. An alternative to reduce the number of adjoint problems is to aggregate data by considering temporal moments of the time series (Cirpka and Kitanidis 2000; Li et al. 2005; Pollock and Cirpka 2008; Schwede et al. 2014).

Sequential data assimilation and Kalman filter methods have been increasingly used in groundwater applications (e.g., Chen and Zhang 2006; Hendricks Franssen and Kinzelbach 2008; Crestani et al. 2013; Panzeri et al. 2013; Tong et al. 2013; Panzeri et al. 2014; Camporese et al. 2015; Erdal and Cirpka 2016). Data assimilation is an iterative procedure in which noisy observations and simulation results at the current state of a system are merged. Model states and/or parameters are corrected to minimize the differences between modeled and real observations. Model simulations are then run forward in time until new observations are available and the errors are evaluated again, defining a forecast/correction cycle. A particularly popular data assimilation method in hydrogeology is the Ensemble Kalman Filter (EnKF) (Evensen, 1994). In its original application, the EnKF continuously updates the model states with incoming data, mitigating the restrictions of linear updating.

Formally, the conditioning in both the linearized geostatistical approach and the EnKF is identical. However, in the EnKF the uncertainty is propagated through Monte Carlo simula-

tions rather than by linearized uncertainties. The auto- and cross-covariances between model parameters and measurements are obtained from the ensemble, avoiding costly calculations of sensitivities. As the EnKF works with an ensemble of updated stochastic realizations, it provides parameter uncertainty estimates without any additional computational costs. To tune the filter capabilities, many adaptations have been made to the original formulation. Hendricks Franssen and Kinzelbach (2008) and Tong et al. (2012) updated not only model states, as in the original derivation of the EnKF, but also model parameters, and Nowak (2009) developed a modified EnKF that updates only model parameters. To estimate hydraulic conductivity fields, Tong et al. (2012), Xu et al. (2013), and Panzeri et al. (2015) used hydraulic-head observations, while Crestani et al. (2013) and Tong et al. (2013) used concentration data.

A limitation of the EnKF is the implicit assumption of multi-Gaussian relations among all variables considered, implying the hydraulic parameters being multi-Gaussian random functions, and the relationship between parameters and states being linear. To reduce the impact of non-Gaussianity, Zhou et al. (2011) and Schöniger et al. (2012) transformed the model states and parameters such that their marginal distributions become Gaussian. Although the performance of the EnKF is improved, the univariate transformations applied by Zhou et al. (2011) and Schöniger et al. (2012) do not affect the non-Gaussian statistical dependence between parameters and states, and multi-Gaussianity is not ensured.

For the implementation of the EnKF, an initial ensemble of parameter fields has to be generated. Erdal and Cirpka (2016) studied the impact of prior information included in the generation of the initial ensemble, i.e. by conditioning the random fields of model parameters to assumed distribution patterns. They showed that joint estimation of groundwater recharge and hydraulic conductivity fields using only hydraulic-head data, can lead to conductivity-to-recharge aliasing if wrong prior information is included in the parameter fields. It is therefore recommended to use the EnKF for parameter estimation in applications where some prior information about aquifer parameters is available.

1.2 Motivation and Objectives

The value of tracer tomography as a high-resolution field characterization method has been demonstrated with several numerical (e.g., Vasco et al. 1997; Schwede et al. 2014) and laboratory (e.g., Illman et al. 2010; Brauchler et al. 2013a) studies. However, the technical complexities and associated costs have shadowed its application in real aquifers. Already Vasco and Datta-Gupta (1999) estimated parameter fields based on tracer datasets from multiple tracer tests, however the experiments were not performed in a truly tomographic sequence. To

the author's knowledge, only two field applications of tracer tomography have been reported (Doro et al. 2015 and Somogyvári and Bayer 2017).

This work is motivated by the studies of Schwede et al. (2014) and Doro et al. (2015). The former showed with a synthetic example that integrating heat-tracer to hydraulic data from tomographic experiments, improves the resolution of the estimated hydraulic conductivity fields and reduces parameter uncertainties. The latter presented a field method to perform tracer tomography using heat as a tracer, illustrating its viability with a field experiment performed at the Lauswiesen Hydrogeological Research Site.

To cope with the high diffusivity of heat, dampening temperature signals within short distances from an injection point, a strong input signal is needed. Injecting water at high temperatures may lead to problems that include (*i*) changes of fluid density and viscosity, and hence distortion of the flow field and many transport processes, and (*ii*) negative biogeochemical effects on the aquifer system. It is also known that heat and solute tracers provide different type of information about the aquifer than heat. Irvine et al. (2015) showed that a solute tracer provides more insights about flow rate variability caused by aquifer heterogeneity. In contrast, the smoothed temperature signals caused by the additional transport of heat by conduction, may be more informative of mean groundwater velocities.

Different applications of the EnKF for the assimilation of hydraulic and tracer test data can be found in the literature. However, the combination of the filter with data from hydrogeological tomographic experiments has not been reported.

This work contributes to the development of tracer tomography by presenting an alternative to the recent investigations focused on using heat as a tracer, and by using the EnKF to estimate spatially distributed aquifer parameters based on the large datasets recorded with hydraulic and tracer tomography experiments. The main objectives are:

A. Adapt the experimental design presented by Doro et al. (2015) to the use of a solute tracer instead of heat. The main reason to avoid using heat as a tracer is the technical difficulty associated to the injection of warm water at a constant temperature over long time periods. Instabilities in the input signal propagate into the observed thermal breakthrough curves. To remove these effects, the application of complicated and computationally intensive signal processing tools would be required (Cirpka et al., 2007). The experimental designed proposed in this work is fully compatible with the hydraulic tomography test setup implemented by Sanchez-León et al. (2016) as well as the experimental design of Doro et al. (2015), and allows the inclusion of additional types of tracer (e.g., salt) with minimum additional effort. The proposed methodology for solute tracer tomography was applied at the Lauswiesen Hydrogeological Research Site.

B. Assess the potential, advantages, disadvantages and limitations of the ensemble Kalman filter for the estimation of hydraulic parameters based on hydraulic and tracer tomography data. The performance of the EnKF was evaluated using a synthetic scenario, with a 2-D groundwater flow-and-transport model that simulates a solute tracer tomography test. To facilitate the application of the EnKF to the real data, the model resembles the hydrogeological setting and well facilities at the Lauswiesen Hydrogeological Research Site.

C. The estimation of spatially distributed hydraulic conductivity fields based on the data collected during the field experiments. The quality of the estimation was evaluated by means of the predictive capacity using independent pumping and tracer tests.

1.3 Thesis Outline

This work is organized as follows: Chapter 2 reviews the underlying theory and governing equations describing groundwater flow, solute transport and parameter estimation using the EnKF. Chapter 3 explains the concept behind hydraulic and tracer tomographic tests, their development, evolution and applications since they were first suggested as a field method. Chapter 4 presents the experimental design adopted in this work. This chapter includes a description of the field hydrogeological setup, the well facilities and field instrumentation available to perform tracer-tomography experiments at Lauswiesen. Chapter 5 discusses the methodologies applied to denoise and pre-process the dataset obtained from the solute tracer tomography. Afterwards, the processed dataset is presented and described. The final part of Chapter 5 is dedicated to the theory and application of type-curve matching and temporal moment methods for the estimation of effective parameters from the processed data. Chapter 6 focuses on a synthetic study of a solute tracer tomography. Synthetic data were used to assess the performance of the EnKF during the estimation of hydraulic conductivity fields. Chapter 7 contains a description of the 3-D numerical model used to simulate the tomographic experiments, and the results of the parameter estimation using the EnKF and the real data. This work ends with conclusions, a discussion of the limitations of the proposed methodology, and recommendations to improve the applications of solute tracer tomography at a field scale.

Chapter 2

Theory

2.1 Groundwater Flow

Following the law of mass conservation, transient groundwater flow in a porous media can be described by the *groundwater flow equation* (Eq. 2.1):

$$S_o \frac{\partial h}{\partial t} - \nabla \cdot (\mathbf{K} \nabla h) = W_o \tag{2.1}$$

with initial boundary conditions (Eq. 2.2a), Dirichlet boundaries at the in- and outflow of the domain Γ_{in} and Γ_{out} (Eqs. 2.2b & 2.2c), and Neumann boundaries Γ_{no} along the other boundaries of the domain (Eq. 2.2d) defined as:

$$h = h_o \;\; at \;\; t = t_o \; \forall x \tag{2.2a}$$

$$h = h_{in} \;\; at \;\; \Gamma_{in} \tag{2.2b}$$

$$h = h_{out} \;\; at \;\; \Gamma_{out} \tag{2.2c}$$

$$\hat{\mathbf{n}} \cdot (-\mathbf{K} \nabla h) = q_{fix}(x) \;\; at \;\; \Gamma_{no} \tag{2.2d}$$

where S_o is the specific storage coefficient, h is the time- and location-dependent hydraulic head, t denotes time, t_o refers to the initial time, $\mathbf{K}$ is the spatially distributed hydraulic conductivity, W_o are volumetric sources or sinks (e.g., injection/extraction wells), h_o is the fixed hydraulic head along Γ_{in} and Γ_{out}, and $\hat{\mathbf{n}}$ is the unit vector normal to Γ_{no}. The values of h_{in}, h_{out} and h_o equal the hydraulic head distribution prior to any aquifer stress. Neumann boundaries Γ_{no} are set where the flux crossing a model boundary is known, having as a particular case the no-flux condition, in which $q_{fix} = 0$. The fixed normal volumetric flux q_{fix} follows Darcy's law (Eq. 2.3):

$$\mathbf{q}(x) = -\mathbf{K} \nabla h \tag{2.3}$$

As stated in Eq. 2.1, the transient behavior of groundwater flow subject to external stresses provides information about aquifer storativity.

2.2 Solute Transport

In this work a steady-state flow field is considered a prerequisite for solute transport, i.e. the partial derivative $\partial h/\partial t$ in Eq. 2.1 equals zero. Assuming a conservative compound (i.e. no sorption, reaction and/or decay) introduced via an inflow boundary $\Gamma_{in,c}$, its transport in groundwater can be described by the Advection-Dispersion Equation (ADE) (Eq. 2.4)

$$n\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{v} - n\mathbf{D}\nabla c) = 0 \tag{2.4}$$

with initial (Eq. 2.5a) and boundary conditions (Eq. 2.5b):

$$c = 0 \quad at \quad t = 0 \; \forall x \tag{2.5a}$$

$$c(t,x) = c_o(t,x) \quad at \quad \Gamma_{in,c} \tag{2.5b}$$

where c is the solute concentration, n is the porosity, $\mathbf{D}$ is the dispersion tensor, and c_o is a known time- and location-dependent concentration along the inflow boundary $\Gamma_{in,c}$. The seepage velocity $\mathbf{v}$ is defined as:

$$\mathbf{v} = \frac{\mathbf{q}}{n} \tag{2.6}$$

and according to Scheidegger (1961), the dispersion tensor $\mathbf{D}$ in Eq. 2.4 can be parameterized as follows:

$$\mathbf{D} = \frac{v_i v_j}{\|\mathbf{v}\|}(\alpha_l - \alpha_t) + \delta_{ij}(D_e + \alpha_t\|\mathbf{v}\|) \tag{2.7}$$

where v_i is the ith component of the velocity vector, α_l and α_t are the longitudinal and transverse dispersivities, D_e is the pore diffusion coefficient, $\|\mathbf{v}\|$ is the absolute value of seepage velocity, and δ_{ij} is the Kronecker delta which is unity for $i = j$ and zero otherwise.

In many real applications it is common to observe a non-Fickian behavior of solute transport, e.g., a dual-porosity domain characterized by early breakthrough resulting from strong advective transport in the mobile region and long tailing caused by the slow diffusive mass transfer between the mobile and immobile regions. This is the case for the Hydrogeological Research Site Lauswiesen, with heavily tailed breakthrough curves reported in previous investigations (see Riva et al. 2008; Sanchez-León et al. 2016). To account for unresolved small-scale variability in hydraulic parameters and to better simulate non-Gaussian breakthrough curves, it is a common approach to further parameterize the transport processes by attributing dual-domain properties to the aquifer (Feehley et al. 2000; Liu et al. 2007), and a mass-transfer

between mobile and immobile regions (Coats and Smith 1964; Sardin et al. 1991):

$$n_m \frac{\partial c_m}{\partial t} + n_{im} \frac{\partial c_{im}}{\partial t} + \nabla \cdot (\mathbf{q} c_m - n_m \mathbf{D} \nabla c_m) = 0 \tag{2.8a}$$

$$\frac{\partial c_{im}}{\partial t} = \lambda_{mt} (c_m - c_{im}) \tag{2.8b}$$

subject to the following initial and boundary conditions:

$$c_m = c_{im} = 0 \quad at \quad t = 0 \; \forall x \tag{2.9a}$$

$$c_m = c_o \quad at \quad \Gamma_{in,c} \tag{2.9b}$$

where c_m andc_{im} refer to the tracer concentration in the mobile and immobile domains, n_m and n_{im} are the corresponding porosities, and λ_{mt} is the first-order solute exchange coefficient between the mobile and immobile zones. If the mobile porosity equals the total porosity and the immobile porosity is set to zero, the model described by Eqs. 2.8a and 2.8b simplifies to the classic advection-dispersion equation (Eq. 2.4).

2.3 Parameter Estimation using the Ensemble Kalman Filter

The Kalman filter was first introduced to describe a recursive solution to the linear filtering problem (Kalman, 1960). The filter is used to recursively estimate the state of a process, minimizing the mean of the squared error. The set of equations needed to implement it assume a linear model and a linear measurement operator. In groundwater and solute transport processes, however, nonlinearities are present, rendering the Kalman filter, in its original form, suboptimal for model inversion. Several modifications to the filter have been suggested for its application in nonlinear dynamics. One approach known as the Extended Kalman Filter, consists of a linearization of the problem about the current model state, approximating the covariance matrices by discarding moments of third and higher order of the Taylor series (closure assumption) (e.g., Evensen 1992; Evensen 2009). However, the assumption that higher order terms are negligible in Bayesian updating may not be valid in all applications. An additional limitation of the extended Kalman filter is the definition of the best estimate from a single model realization, which lacks statistical interpretation (Evensen, 2009).

In another approach known as the ensemble Kalman filter (Evensen, 2003), the covariance matrices are approximated by performing statistics on an ensemble of model evaluations. The latter is the version of the Kalman filter used in this work, and its mathematical formulation for updating **only** model parameters is presented in the following (see also Nowak 2009; Schöniger et al. 2012).

Let's assume a true nonlinear system described by the following stochastic equation (**forecast equation**):

$$\mathbf{y}_t^{sim} = \mathbf{f}(\mathbf{p}_{t-1}) + \mathbf{q}_t \tag{2.10}$$

where $\mathbf{f}$ is a nonlinear model that simulates (or predict) the model states $\mathbf{y}^{sim}$ (e.g., drawdown or concentrations), $\mathbf{p}$ is a vector of model parameters (e.g., hydraulic conductivity), t is time, and $\mathbf{q}$ represents model uncertainties. In this work it is assumed that the accuracy of the model prediction is affected only by parameter uncertainties, i.e. model uncertainties are not considered, hence the term $\mathbf{q}_t$ in Eq. 2.10 is neglected altogether. Examples of the inclusion of model uncertainties in the EnKF are presented by e.g. Madsen and Canizares (1999), Rasmussen et al. (2016) and Erdal et al. (2014).

Adding an ensemble index i, where $i = 1, 2, \ldots, N_{ens}$ and N_{ens} is the total size of the ensemble, and neglecting model uncertainties, the forecast equation can be rewritten as:

$$\mathbf{y}_t^{sim,i} = \mathbf{f}(\mathbf{p}_{t-1}^{i}) \tag{2.11}$$

To ensure compatibility with the real observations $\mathbf{y}^{obs}$, measurement noise $\boldsymbol{\epsilon}$ is added to the simulated observations $\mathbf{y}_t^{sim,i}$ (Burgers et al., 1998). $\boldsymbol{\epsilon}$ is assumed to be white (uncorrelated) and drawn from a normal probability distribution with zero mean and covariance matrix (diagonal) of measurement errors $\mathbf{R}$:

$$p(\boldsymbol{\epsilon}) \sim N(0, \mathbf{R}) \tag{2.12}$$

The residuals, i.e. the deviations of the model states from real observations, can be computed for each ensemble member as:

$$\Delta\mathbf{y}^i = \mathbf{y}_t^{obs} - (\mathbf{y}_t^{sim,i} - \boldsymbol{\epsilon}^i) \tag{2.13}$$

Dropping out the time index t for clarity, the ensemble estimate of the cross-covariance matrix $\mathbf{Q}_{py}$ ($N_{par}xN_{obs}$) between model parameters and states is defined by Eq. 2.14:

$$\mathbf{Q}_{py} = E\left[(\mathbf{P}_u - E[\mathbf{P}_u])(\mathbf{Y}^{sim} - E[\mathbf{Y}^{sim}])^T\right] \tag{2.14}$$

and the auto-covariance of model states $\mathbf{Q}_{yy}$ ($N_{obs}xN_{obs}$) is estimated from the ensemble as:

$$\mathbf{Q}_{yy} = E\left[(\mathbf{Y}^{sim} - E[\mathbf{Y}^{sim}])(\mathbf{Y}^{sim} - E[\mathbf{Y}^{sim}])^T\right] \tag{2.15}$$

with the matrix of parameters $\mathbf{P}$ ($N_{par}xN_{ens}$) containing all parameters of the model for all ensemble members, and the matrix of model states $\mathbf{Y}^{sim}$ ($N_{obs}xN_{ens}$) composed of the vectors of modeled states of all ensemble members. The subscript u in Eq. 2.14 refers to

the unconditioned parameters contained in $\mathbf{P}$. The cross- and auto-covariance matrices (Eqs. 2.14 and 2.15, respectively) can be implemented using the sample mean of the ensemble:

$$\mathbf{Q}_{py} \approx \frac{1}{N_{ens}-1}(\mathbf{P}-\bar{\mathbf{P}})(\mathbf{Y}^{sim}-\bar{\mathbf{Y}}^{sim})^T \tag{2.16}$$

$$\mathbf{Q}_{yy} \approx \frac{1}{N_{ens}-1}(\mathbf{Y}^{sim}-\bar{\mathbf{Y}}^{sim})(\mathbf{Y}^{sim}-\bar{\mathbf{Y}}^{sim})^T \tag{2.17}$$

where $\bar{\mathbf{P}}$ and $\bar{\mathbf{Y}}^{sim}$ are the ensemble mean of the matrix of parameters and model states, respectively. Eqs. 2.16 and 2.17 imply that the larger the ensemble size, the better the approximations of the covariance matrices are. In the linear case, the EnKF for parameter estimation converges to the result of the classical Kalman filter with increasing ensemble size.

With the error matrices $\mathbf{Q}_{yy}$ and $\mathbf{Q}_{py}$ defined, the EnKF updates the parameters of each ensemble member, according to the **analysis** (or update) **equation** (see e.g. Sorenson 1970 for a complete derivation):

$$\mathbf{p}_c^i = \mathbf{p}_u^i + \mathbf{Q}_{py}(\mathbf{Q}_{yy}+\mathbf{R})^{-1}(\mathbf{y}^{obs}-(\mathbf{y}^{sim,i}-\boldsymbol{\epsilon}^i)) \tag{2.18}$$

Eq. 2.18 states that the posterior (or conditioned) parameter estimate is a linear combination of the prior (or unconditioned) estimate plus a weighted difference between the actual measurement and the associated model prediction. The iterative process consists in integrating the model states in time (by applying Eq. 2.10) using the updated model parameters. The new model states simulated one step forward in time, are compared with the corresponding real observations and the error covariances reevaluated. The iterative process finalizes once all available data are assimilated.

Fig. 2.1 describes the update procedure of the EnKF as implemented in this work. The horizontal axis is the time at which data are available. The vertical axis represents the number of parameter updates performed. The red rectangle represents the ensemble generation as a prior step. The black arrows indicate the solution of the EnKF as a function of time, updated at every time that a new measurement is incorporated. The blue lines represent the introduction of measurements for analysis, and the grey dotted arrows represent the filter update.

With the weighting factor $\mathbf{K}_g = \mathbf{Q}_{py}(\mathbf{Q}_{yy}+\mathbf{R})^{-1}$ (known as the Kalman gain) Eq. 2.18 can be finally rewritten as follows:

$$\mathbf{p}_c^i = \mathbf{p}_u^i + \mathbf{K_g}\Big(\mathbf{y}^{obs}-(\mathbf{y}^{sim,i}-\boldsymbol{\epsilon}^i)\Big) \tag{2.19}$$

The Kalman gain states that the closer the measurement error covariance $\mathbf{R}$ approaches zero, the higher the confidence in the observation data is and the less trusted the model predictions are. For multi-Gaussian state variables and a linear physical model, the Kalman filter determines

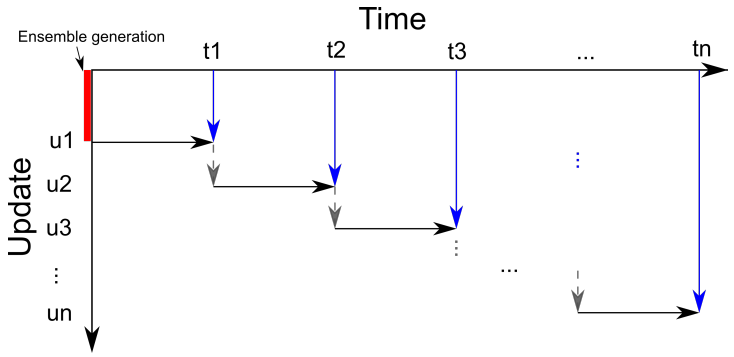


Figure 2.1: Description of the update procedure of the ensemble Kalman filter. Horizontal axis represents time and vertical axis the filter updates (only parameters in this work). Red square: ensemble generation; black line: model run forward in time; blue lines: introduction of measurement; gray lines: filter update (after Evensen 2009).

the parameter set maximizing the posterior probability density. In this context the ensemble Kalman filter generates the exact posterior distribution of parameters in the limit of an infinite ensemble size (Burgers et al., 1998).

The posterior covariance matrix of parameters (or error covariance matrix) can be computed from the updated parameters as:

$$\mathbf{Q}_{pp} = E\left[(\mathbf{P}_c - E[\mathbf{P}_c])(\mathbf{P}_c - E[\mathbf{P}_c])^T\right] \approx \frac{1}{N_{ens}-1}(\mathbf{P}_c - \bar{\mathbf{P}}_c)(\mathbf{P}_p - \bar{\mathbf{P}}_p)^T \tag{2.20}$$

The EnKF with the update scheme described in Fig. 2.1 has been successfully applied to assimilate hydraulic head data (Chen and Zhang 2006; Bailey and Baù 2010; Hendricks Franssen et al. 2011; Zhou et al. 2011; Bailey and Baù 2012; Schöniger et al. 2012; Tong et al. 2012). Other authors have used it to estimate parameters based on concentration data (e.g. Liu et al. 2008; Li et al. 2012). In this work, it was found that the classical update scheme of the EnKF leads to mass balance errors. To honor mass conservation and consistency between hydraulic parameters and concentrations, the steady flow model was reinitialized after each parameter update and transport was simulated from the initial time until the next available measurement time step. This version of the filter is known as the restart EnKF (Wen and Chen 2006; Hendricks Franssen and Kinzelbach 2008; Crestani et al. 2013; Tong et al. 2013). The updating scheme of the restart EnKF is described in Fig. 2.2. The computational effort of this scheme scales with the number of measurement time steps squared.

As mentioned before, an underlying assumption in the EnKF is that all variables involved (i.e. parameters, model states and observations), are random functions following a multivariate Gaussian distribution, and hence are fully characterized by their mean and covariance matrix. This assumption is justified for groundwater flow and solute transport applications only under

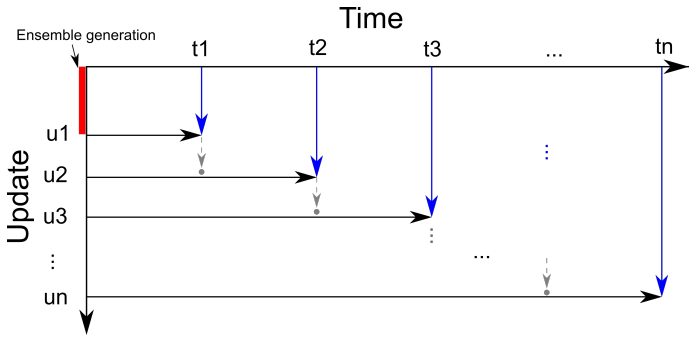


Figure 2.2: Description of the update procedure of the restart ensemble Kalman filter. Horizontal axis represents time and vertical axis the filter updates (only parameters in this work). Red square: ensemble generation; black line: model run forward in time; blue lines: introduction of measurement; gray lines: filter update; gray circles: exact time step at which the update of parameters occur (after Evensen 2009).

idealized boundary conditions (see Schwede et al. 2008). To improve the performance of the EnKF, two tuning features were adopted: (*i*) normal score transformation (Zhou et al. 2011; Schöniger et al. 2012), and (*ii*) a damping factor applied to $\mathbf{K}_g$, reducing the update of each parameter vector in the ensemble (e.g., Hendricks Franssen and Kinzelbach 2008; Erdal and Cirpka 2016).

2.3.1 Tuning the EnKF: Normal score transformation

Normal score transformation is a method designed to map a population with an arbitrary distribution to a standard normal distribution. This method consists in ranking the values of the variables to be transformed from lowest to highest and matching these ranks to equivalent ranks generated from a normal distribution. In this work, the transformation function was approximated using a Gaussian kernel. A description of its implementation is provided at the end of this chapter.

Normal score transformation is a monotonically increasing univariate transformation improving the shape of the probability density function at a marginal level, therefore multi-Gaussianity cannot be ensured. As highlighted by Schöniger et al. (2012) and Crestani et al. (2013), normal score transformations have to be used with caution since far from improving the performance of the filter, they can deteriorate the cross correlation structure between model parameters and states. This phenomena was observed in this work when the transformation was applied to concentration data, hence it was only used to transform hydraulic head data, where a considerable improvement of the filter was detected. After transforming the model states and the corresponding measurements, the transformed variables were included in the update equation

(Eq. 2.18) following Schöniger et al. (2012):

$$\mathbf{p}_c^i = \mathbf{p}_u^i + \mathbf{Q}_{\mathbf{p}\widehat{\mathbf{y}}}\left(\widehat{\mathbf{Q}_{\mathbf{yy}} + \mathbf{R}}\right)^{-1}\left(\widehat{\mathbf{y}^{obs}} - (\widehat{\mathbf{y}^{sim,i} - \boldsymbol{\epsilon}^i})\right) \tag{2.21}$$

where the symbol $\widehat{\ }$ represents transformed variables. An individual transformation function was built for each measurement location (Béal et al., 2010) and it was applied to the perturbed modeled states ($\mathbf{y}^{sim,i} - \boldsymbol{\epsilon}^i$) to ensure compatibility with real observations in the transformed space. After updating each parameter vector of the ensemble, back transformation is not required if only parameters are being updated (Schöniger et al., 2012), since they are computed by additional model runs targeting the new time step.

2.3.2 Tuning the EnKF: Damping the update

An important decision to make when implementing the EnKF is the number of realizations of the ensemble. The larger the ensemble, the better are the approximation of the covariance matrices (Eqs. 2.17 and 2.16), and the more stable the filter becomes. If a model with a large number of parameters is being used (e.g., finely discretized grid), it might be prohibitive to generate an ensemble with a large number of realizations and the so-called filter inbreeding problem may affect the updating process. Filter inbreeding is the process of increasingly underestimating the ensemble variance over time. Hendricks Franssen and Kinzelbach (2008) suggested to minimize this effect by including a factor β to dampen the update, which is also known as relaxation. The latter authors found that this factor improves the stability of the filter at each updating step. The factor β has also been used by Erdal and Cirpka (2016) with promising results. This factor is a number between 0 and 1 that directly affects the Kalman gain (Eq. 2.22):

$$\mathbf{p}_c^i = \mathbf{p}_u^i + \beta\mathbf{K}_g\left(\mathbf{y}^{obs} - (\mathbf{y}^{sim,i} - \boldsymbol{\epsilon}^i)\right) \tag{2.22}$$

There is not a clear method to define an optimal value for β, and it has to be chosen heuristically. In this work several values of β were tested using a synthetic case, and the value with the best performance was selected for parameter estimation with the real data.

2.3.3 Implementation of the Normal Score Transformation

Normal score transformation is a rank-based method designed to render a population with an arbitrary distribution to a standard normal distribution. The procedure implemented in this work consists of two main steps: (i) evaluating the empirical cumulative distribution function (CDF) based on the ranks of both the dataset and the standard Normal distribution

($\mu = 0, \sigma = 1$), and (ii) interpolating the empirical CDF to construct a continuous function and allow transformation of values not included in the initial dataset.

The vector of standard ranks (or standard scores) $\bar{z}$ is calculated as follows:

$$z_i = i - \left(\frac{1}{2} \cdot \frac{1}{n}\right) \tag{2.23}$$

where i is the corresponding rank and n the number of elements in the dataset. The Gaussian-transformed values $\bar{z}_{tr}$ can be determined by means of a simple rank transformation, defining then the empirical transformation function. The range of the transformed dataset is defined by $\bar{z}_{tr,min} = \frac{1}{2n}$ and $\bar{z}_{tr,max} = 1 - \frac{1}{2n}$. Finally, the continuous function is obtained by linear interpolation of the empirical transformation function. If values that have not been included in the initial dataset are to be transformed and fall outside the range defined by the dataset, a linear extrapolation needs to be performed. Since values are transformed linearly, this extrapolation is equivalent to assume that the tails follow those of the Gaussian distribution (Schöniger et al., 2012).

To overcome the non-Gaussianity of the observed and simulated hydraulic heads and concentration data, a transformation function was constructed for each measurement location, using the simulated values of all ensemble members. According to Eq. 2.21, if normal score transformations are included in the update step of the EnKF, the update of each parameter vector is based on the residuals in transformed space (i.e., differences between transformed real and simulated measurements). As Schöniger et al. (2012) suggests, the transformation function was based on the already perturbed simulated measurements and the same function was used to transform the corresponding real measurement. This ensures compatibility between simulated data and real observations in the transformed space.

Fig. 2.3 is an example of the normal score transformation applied to 500 (ensemble size) simulated concentration data points, at an arbitrary time step and measurement location. Although concentration data were not transformed during the application of the EnKF, this example was selected to show the performance of the normal score transformation with a clearly positive skewed dataset (Fig. 2.3-upper left). Non-Gaussianity is evident with the fit of a normal and exponential distributions to the histogram of the original data. The empirical transformation function is presented in Fig. 2.3 (upper right), and the histogram of the transformed data set with the corresponding fit of a normal distribution is shown in Fig. 2.3 (lower left). The probability plot in Fig. 2.3 (lower right) shows that a straight line fits perfectly to the data, confirming a Gaussian distribution of the transformed data.

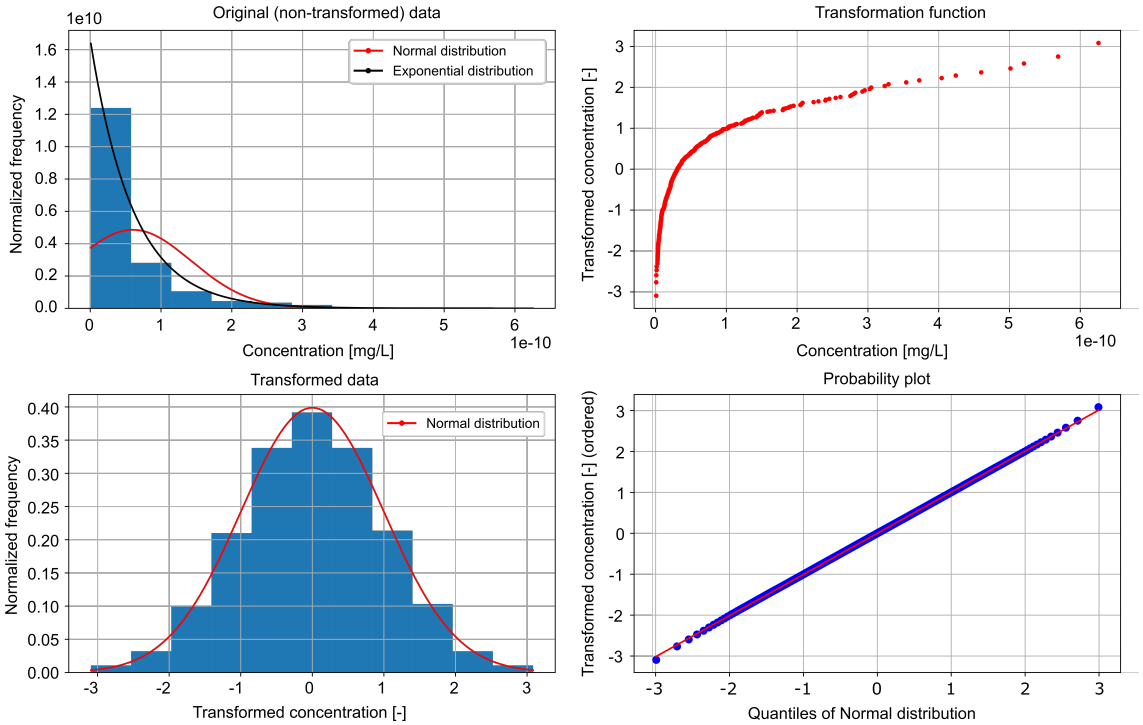


Figure 2.3: Normal score transformation applied to a positive (left) skewed dataset. Upper left: histogram of the original data, red line is a fit of the normal distribution and black line is a fit of the exponential distribution; Upper right: empirical transformation function; Lower left: histogram of the transformed data; Lower right: probability plot for the transformed variable, red line is a fit of a Normal probability function.

Chapter 3

Tomographic Aquifer Tests

The limitations of conventional hydraulic tests in providing enough details about subsurface heterogeneity have motivated the development of more elaborated field techniques such as hydraulic and tracer tomography. The main attribute of tomographic aquifer testing is the possibility of collecting datasets informative of different portions (and probably of different processes) of the subsurface, that may reveal details of material heterogeneity and increase the resolution in the estimation of the spatially distributed hydraulic parameters.

This work is focused on solute tracer tomography, and a truly hydraulic tomography experiment was not pursued during the field tests. However, the assessment of the EnKF with a synthetic scenario was based on the simulation of two different hydraulic flow fields, and consequently two different synthetic hydraulic datasets were used. During the real tests it was intended to generate the same hydraulic field prior to each individual tracer test, however, small variations in the injection and extraction rates were unavoidable, producing slightly different hydraulic responses. The different hydraulic fields were also considered independently during parameter estimation, mimicking in a sense a hydraulic tomography dataset.

3.1 Hydraulic Tomography

Hydraulic tomography is in the most simplified terms a series of hydraulic tests, where the aquifer is stressed either by injection and/or extraction of water at specific points, and the hydraulic head responses are monitored at several other ones (e.g., Gottlieb and Dietrich 1995; Straface et al. 2006; Berg and Illman 2013; Yeh et al. 2013). The tomographic layout consists in changing the stress locations while the hydraulic responses are still monitored at multiple points (Fig. 3.1). With the sequential stimulation of the aquifer, hydraulic tomography allows

collecting datasets from a larger investigation volume and from different portions of the aquifer.

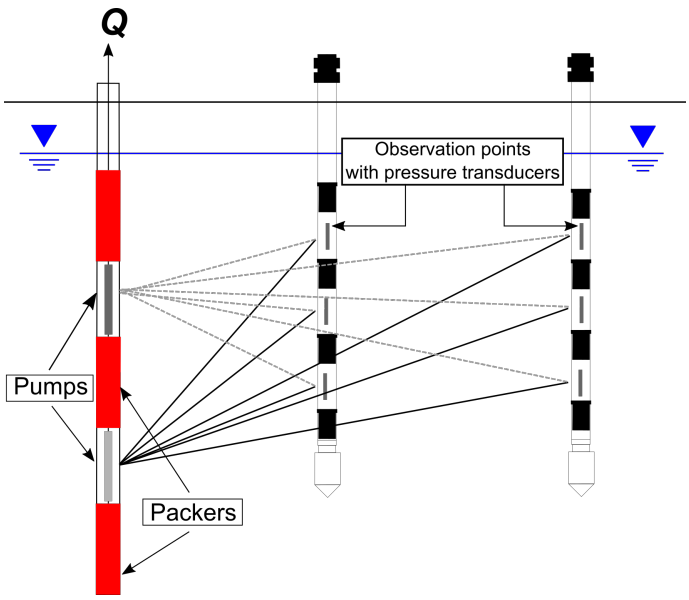


Figure 3.1: Experimental setup of a hydraulic tomography test.

Fig. 3.1 shows a schematic view of a hydraulic tomography test with two (independent) extraction intervals and two observation wells with measurement points at different depths. The different sections within the pumping well are usually generated with inflatable packers.

Hydraulic tomography has been applied in a large number of numerical, laboratory, and field experiments (e.g. Yeh and Liu 2000; Liu et al. 2002; Brauchler et al. 2003; Bohling et al. 2003; Zhu and Yeh 2006; Bohling et al. 2007; Illman et al. 2007; Illman et al. 2009; Xiang et al. 2009; Yin and Illman 2009; Berg and Illman 2011; Cardiff and Barrash 2011; Hu et al. 2011; Cardiff et al. 2012; Berg and Illman 2013; Brauchler et al. 2013b; Cardiff et al. 2013; Jiménez et al. 2013; Illman 2014; Lee and Kitanidis 2014; Berg and Illman 2015; D.L. et al. 2015; Sanchez-León et al. 2016), and different inversion schemes have been proposed to estimate aquifer parameters from the collected data. Vasco et al. (1999), Brauchler et al. (2003) and Brauchler et al. (2007) used a travel time-based inversion scheme to obtain distributions of hydraulic diffusivity (transmissivity/storativity). Berg and Illman (2011) and Liu et al. (2002) inverted transient hydraulic tomographic data with the sequential successive linear estimator developed by Zhu and Yeh (2005). Cardiff and Barrash (2011), Cardiff et al. (2012) and Schwede et al. (2014) used the quasi-linear geostatistical method of Kitanidis (1995).

While hydraulic tomography can be used to assess aquifer heterogeneity at a higher resolution than traditional tests, the resolution of the estimated hydraulic parameter fields are

limited by the diffusive nature of groundwater flow. Sensitivities of head measurements drop rapidly within short distance to the involved wells (e.g., Leven and Dietrich 2006; Bohling 2009; Bohling and Butler 2010; Schwede et al. 2014). Sanchez-León et al. (2016) applied a gradient based scheme for the estimation of a single 3-D hydraulic conductivity field and showed that K-fields derived from hydraulic head data alone (even from tomographic experiments) most probably will never resolve all aquifer heterogeneity relevant for solute transport.

3.2 Solute Tracer Tomography

The importance of tracer data for aquifer characterization lies in the close relation between solute transport and the variations in hydraulic properties. Vasco et al. (1997) showed with a sensitivity analysis of synthetic data, that a higher parameter resolution and acceptable uncertainties can be obtained with tracer rather than pressure data, and hence that tracer tests are better suited for estimating small-scale variability in the subsurface material.

Some studies have addressed the applicability of multi-level and multi-well tracer tests, and their utility for characterizing small-scale variations of subsurface properties. Vasco and Datta-Gupta (1999) inferred the variability in permeability from breakthrough curves recorded with several multilevel samplers and extraction wells. Dietrich et al. (2005) used multiple observation points to monitor several flow and transport experiments at a laboratory and field-block scales. Ptak et al. (2004) presented a comprehensive guideline for tracer testing using depth integrated and multilevel setups. Riva et al. (2008) showed the importance of multilevel measurements to recover relevant features required for accurate three-dimensional transport simulations. In a more recent study, Wagner et al. (2014) performed a thermal tracer test in a highly conductive aquifer and used various thermal breakthrough curves to define preferential flow paths. In a similar work, Klepikova et al. (2016) analyzed multiple heat tracer breakthrough curves to estimate the three-dimensional hydraulic conductivity field. All these studies represent the early stages of tracer tomography, which as for hydraulic tomography requires the sequential stimulation of different portions of an aquifer by injecting a tracer in the subsurface (Fig. 3.2).

It was not until the work of Yeh and Zhu (2007) that a true tomographic sequence for tracer testing was suggested. They showed with a numerical example the value of inverting partitioning-tracer tomographic measurements together with hydraulic data, to map DNAPL distributions in the subsurface. Zhu et al. (2009) analyzed tracer data from a synthetic tracer tomography with temporal moments of tracer concentration. On a step forward, Illman et al. (2010) adopted the methodology to estimate hydraulic conductivity fields and DNAPL satu-

ration distributions in a sandbox experiment. In another laboratory example, Brauchler et al. (2013a) performed a tracer tomography test in a sandstone block using helium gas as tracer, and reconstruct a three-dimensional velocity distribution in the block by analyzing multiple breakthrough curves with a line integral approach (Vasco and Datta-Gupta, 1999). With yet another numerical exercise, Schwede et al. (2014) showed that differences in sensitivity patterns of arrival times between thermal signals and head measurements provide additional information of aquifer heterogeneity, demonstrating the advantages of combined hydraulic and tracer tomographic experiments for parameter estimation.

In a recent investigation, Doro et al. (2015) presented the first experimental design to perform a field tracer tomography and evaluated its applicability at the Hydrogeological Research Site Lauswiesen, in Germany. In another field application, Somogyvári and Bayer (2017) obtained a hydraulic-conductivity tomogram based on tracer data collected during a field application of tracer tomography at the Widen field site, Switzerland. Both studies used heat as tracer.

Fig. 3.2 shows the general concept of a tracer tomography. Several sections are generated at a selected injection well (blocks in red). A pumping well extracts water and when a steady state flow field is achieved, tracer is injected in one of the sections of the injection well. The plume distribution is monitored at several observation wells. The tracer test is repeated with changed injection section, and the concentration distribution is monitored again.

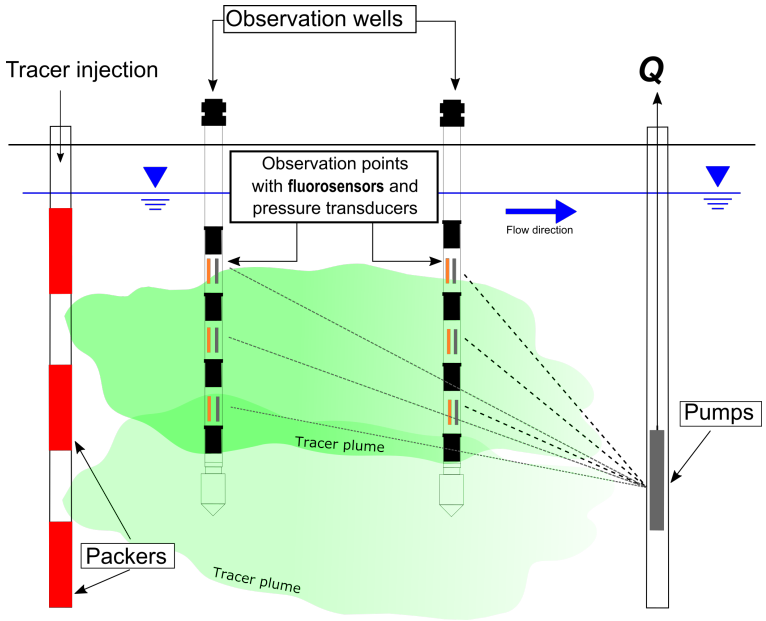


Figure 3.2: Schematic view of a solute tracer tomography with two different injection sections. The multilevel observation points contain both a fluorosensor and a pressure transducer.

Chapter 4

Experimental Setup

4.1 Hydrogeological Research Site Lauswiesen

As a proof-of-concept the experimental method for solute tracer tomography was applied at the Hydrogeological Research Site Lauswiesen. The site is located in a flat area of the Neckar Valley, east of Tübingen, Germany (Fig. 4.1).

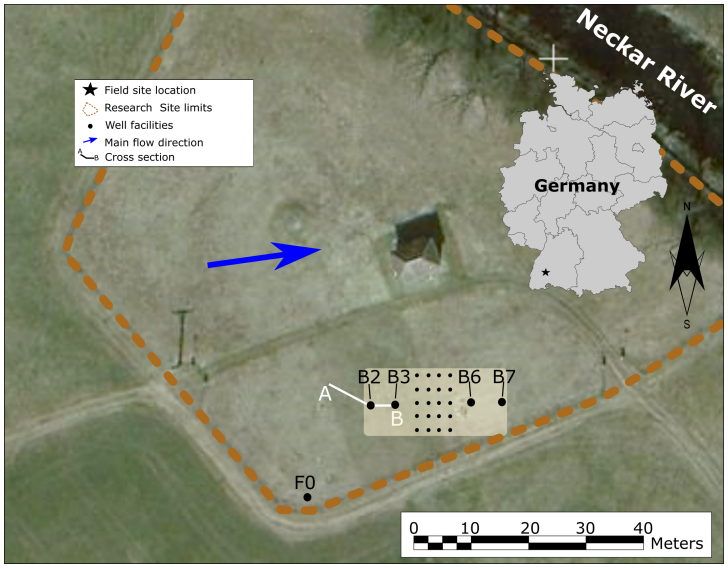


Figure 4.1: Location of the Hydrogeological Research Site Lauswiesen with the well field used in this study. Blue line: main groundwater flow direction; White line: location of the cross section presented in Fig. 4.2; Black dots: distribution of the wells used in this study. Well F0 was used to provide water during the experiments.

Kleinert (1976) suggested that the sediments of the region were deposited and reworked in a fluvial environment during the Holocene and Heinz (2001) proposed the Quaternary glaciated areas of the Black Forest as the source of water.

The aquifer is overlain by about $2\,\mathrm{m}$ of silt, and underlain by low-conductive claystones of the upper Triassic (*Bunte Mergel*) formation. The aquifer itself consists of $8\,\mathrm{m}$ to $9\,\mathrm{m}$ of gravel with small amounts of sand ($\sim 10\,\%$) and fines ($<10\,\%$) (see Fig. 4.2). The depth to groundwater is approximately $4\,\mathrm{m}$, defining an unconfined system with an ambient hydraulic gradient between $0.2\,\%$ to $0.3\,\%$ to the northeast, and variations of up to $15°$ induced by water level changes of the Neckar River.

A large number of investigations have been dedicated to the estimation of hydraulic parameters and characterization of the subsurface heterogeneity at Lauswiesen (Sack-Kühner 1996; Heinz 2001; Ptak et al. 2004; Hoffmann and Dietrich 2004; Rein et al. 2004; Riva et al. 2006; Riva et al. 2008; Li 2008; Schneidewind 2008; Lessoff et al. 2010; Händel and Dietrich 2012; Wagner et al. 2014; Doro et al. 2015; Sanchez-León et al. 2016, among others). In general, these studies agree on a mean hydraulic conductivity value of $3.0 \times 10^{-3}\,\mathrm{m\,s^{-1}}$, which is a typical value for sediments with high content of gravel (Sack-Kühner 1996; Riva et al. 2006; Schneidewind 2008; Lessoff et al. 2010; Wagner et al. 2014; Panzeri et al. 2015).

Based on an extensive direct-push injection-logging and direct-push slug-test field campaign, Lessoff et al. (2010) showed that the aquifer could be divided into two different zones, a highly conductive and relatively homogeneous upper zone, and a less conductive and more heterogeneous lower zone. This subdivision of the aquifer is further supported by Händel and Dietrich (2012) after fitting multi-level breakthrough curves using a two-layer model. With an inversion of a 3-D dual-domain flow and transport model, Sanchez-León et al. (2016) estimated a mean porosity value of about $26\,\%$. This is a more suitable porosity value for sandy gravel aquifers than the value of $9.8\,\%$ obtained by Riva et al. (2006) after fitting simulated breakthrough curves to multilevel–multi-tracer data.

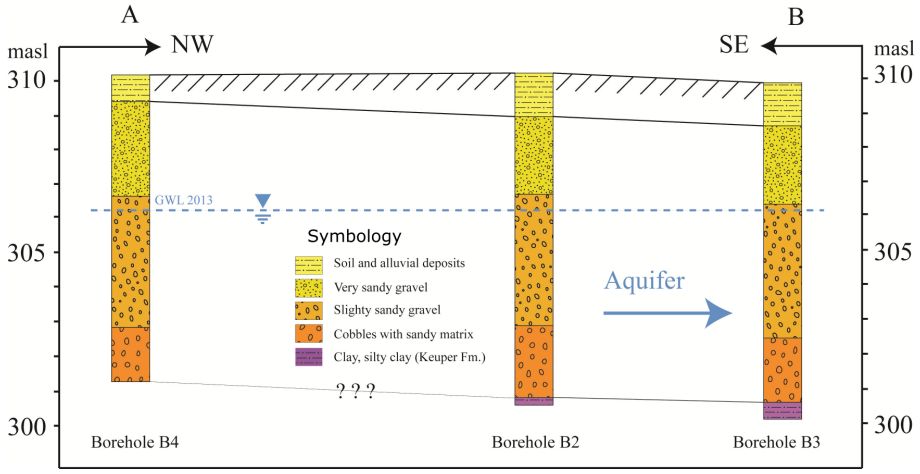


Figure 4.2: Lithological section describing the main (hydro-) geological units, based on bore log information from wells B2, B3 and B4 (after Sack-Kühner 1996 and Sanchez-León et al. 2016)

4.2 Well Inventory

The wells used in this study cover an area of about $30\,\mathrm{m}\times10\,\mathrm{m}$. The well field includes four fully screened wells (*B-wells*) with a diameter of $15\,\mathrm{cm}$. These B-wells are aligned with the main natural groundwater flow direction (see Fig. 4.1), and were used to inject/extract water and tracer during the tracer tomography test. Between wells B3 and B7, a total of 20 observation wells are distributed in a 5×4 regular grid (Fig. 4.3).

The wells B1-B5 were installed during a field campaign already in 1994. Wells B6, B7 and the observation wells were recently installed using direct push, as part of a field campaign focused on the development of tracer tomography at a field scale (see Doro et al. 2015 for more details).

Observation wells cmt1 to cmt4 (Fig. 4.3, plan view) are multilevel wells (Einarson and Cherry, 2002) with seven depth-discrete observation ports each, embedded in a high-density polyethylene (HDPE) multichannel pipe (Fig. 4.5A). The vertical distribution of the individual ports of the cmt wells is shown in Fig. 4.4.

Individual channels within the tube have an inner diameter of $2\,\mathrm{cm}$. The size of the individual channels allowed the installation of pressure and tracer sensors and a groundwater sampling hose (Fig. 4.5B and 4.5C). One channel of each cmt-well is occupied by equipment installed permanently for additional field tests. All other observation wells (*w-wells*) are fully screened

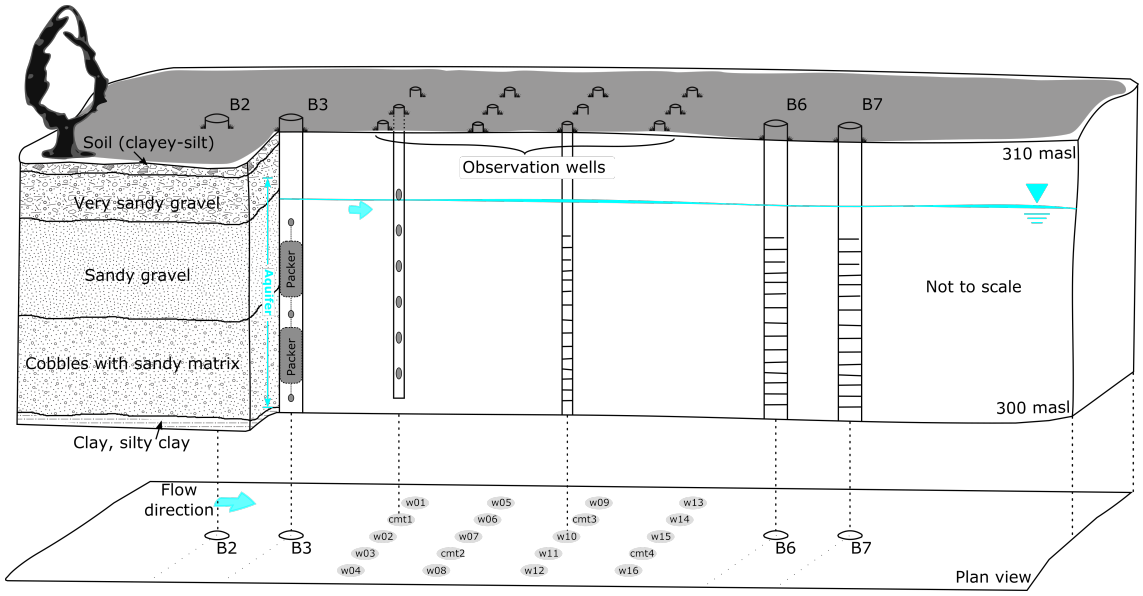


Figure 4.3: Schematic description of the well facilities used in this study. Cross section: distribution of the main geological units, details of the different types of observation wells available at the field site, and the multilevel injection system at well B3. Plan view: horizontal distribution of the four B-wells and observation wells.

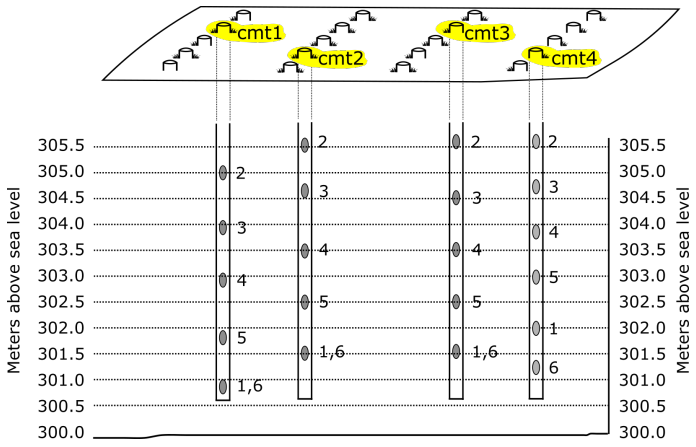


Figure 4.4: Vertical distribution (not to scale) of each port of the four multilevel wells available at the area under investigation. Units are in meters above sea level.

and vary between $1.9\,\mathrm{cm}$ to $2.5\,\mathrm{cm}$ in diameter. Overall, a total of 40 observation points were available for monitoring hydraulic pressure and tracer concentration. As Riva et al. (2008) suggested, the possibility to retrieve both depth-integrated and multi-level measurements may provide valuable information to represent many of the important features required to model three-dimensional transport processes.

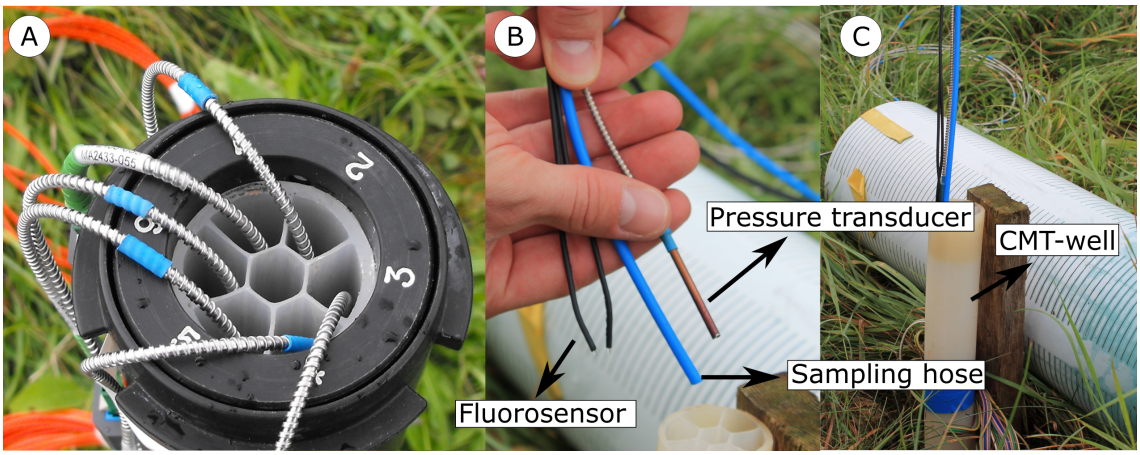


Figure 4.5: Details of observation well cmt1. A: the seven different channels within the HDPE tube; B: thickness of the fiber optic pressure transducer, fluorosensor and groundwater sampling hoses used during the field experiments; C: fiber optic pressure transducer, fluorosensor and groundwater sampling hoses used during the field experiments introduced in a single channel of the cmt well.

4.3 Adopted Flow Field

All tracer tests discussed in this work were performed with a forced-gradient flow field in a nested-cell setup. This four-well system was suggested by Luo et al. (2006b) to protect an in-situ remediation system from changes in the regional flow. It was implemented in a numerical study by Schwede et al. (2014) within the framework of a synthetic tomographic tracer test. The general idea of the nested-cell flow system is to create an inner cell nested within an outer flow field, using two extraction and two injection wells. A stable inner artificial ambient flow field is achieved with high extraction and injection rates at the outer well pair. The inner well pair is then used to create an internal nested flow field with lower injection/extraction rates.

Figure 4.6 shows an example of the nested-cell flow field simulated with the numerical model used in this work. Two Dirichlet boundaries were applied at the boundaries perpendicular to the natural flow direction, and two Neumann boundaries at the rest of the model boundaries. The model contains the location of the B- and w-wells, and similar injection and extraction rates to those applied during the real tests. The outer well pair B2-B7 was used to inject and extract water at high rates, and the inner well pair B3-B6 to create the internal nested flow with lower rates. The generated flow field was kept stable until completion of the individual tracer tests. To enhance tracer recovery rates during the tracer tests, an asymmetrical setup was adopted by applying higher extraction than injection rates in both well pairs (Ptak et al., 2004).

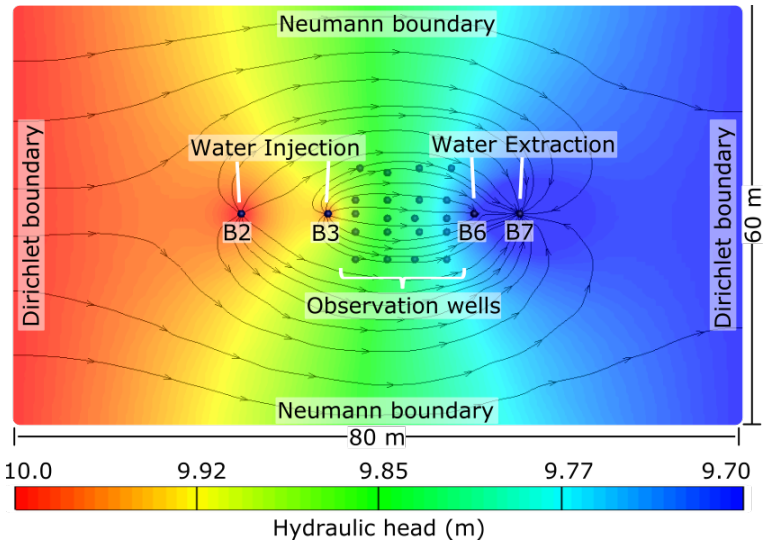


Figure 4.6: Expected distribution of hydraulic heads in an asymmetric nested-cell steady-state flow field. Observation and B-wells correspond to the actual location of the wells at the Hydrogeological Research Site Lauswiesen.

Benefits of this flow field design are: (*i*) a well-focused flow field, (*ii*) reduction of test duration due to the relatively high hydraulic gradients, (*iii*) lower impact of fluctuating boundary conditions, (*iv*) higher tracer recovery rates, and (*v*) minimized leakage of tracer into the environment, any tracer bypassing the inner extraction well would ideally be collected by the outer extraction well. Strongly advective dominated transport however, may result in underestimation of longitudinal dispersivities that characterize solute dispersion under natural flow conditions (Tiedeman and Hsieh, 2004).

A multilevel injection system was installed at well B3, generating three independent injection sections. Throughout all tests, water was injected simultaneously at the three sections, while tracer injection was restricted to a single section per test. To account for vertical variations of aquifer transmissivity and to enforce horizontal flow, water injection rates were adjusted for each level such that the injection rate in each injection segment was approximately proportional to the transmissivity of that section. Information about the vertical variations of hydraulic conductivity at the field site was determined with previous borehole flowmeter tests (Riva et al., 2006). The injected water was supplied by well F0, located $\sim 40\,\mathrm{m}$ upstream of the area under investigation (see Fig. 4.1). No influence of this pumping well was observed in the initial model simulations, neither during the field tests. To prevent tracer re-entering the system, pumped water was never reinjected. Instead it was released more than $\sim 40\,\mathrm{m}$ downstream, at the bank of the Neckar river.

4.4 Solute Tracer Tomography

The tracer experiments were performed using Na-fluorescein ($C_{20}H_{10}Na_2O_5$) as tracer. It is considered a nearly conservative compound under the slightly alkaline pH conditions prevailing at the site (Käss, 1964) and has the advantage of being detected at extremely low concentrations ($\sim \mathrm{ng\,l^{-1}}$) (Naurath et al. 2011; Weidner et al. 2011).

The tomographic experiment itself consisted of a series of tracer tests with the injection of fluorescein at one of the three isolated sections of well B3 (Fig. 4.8). The tomographic layout was achieved by repeating the experiments after shifting the tracer injection interval, until the three isolated sections were covered.

The multilevel injection system at the (inner) injection well B3, was established by two inflatable packers between the sections (customized sewage pipe test bladder). The device has a metal cap at the top of each inflatable packer, with several connection lines through which sensors for monitoring hydraulic pressure and tracer concentration can be placed (Fig. 4.7). All connections within the packer system are watertight to prevent leakage. The main purpose of the multilevel injection system was to constrain the injection of the dye tracer to a certain section of the aquifer. To achieve this, a bypass that delivers water to the individual sections of injection well B3 was attached to each line. For the tracer injection, the corresponding valve was connected to a surface water pump. The pump was connected to a container filled with water, where $1\,\mathrm{g}$ to $2\,\mathrm{g}$ of dye tracer were added.

Only after a quasi-steady-state flow field was reached, the surface pump was turned on and the corresponding valve opened to start the tracer injection at the selected depth. Ambient, tracer-free water was constantly added until all tracer was washed out of the container and the well casing. Injection times lasted less than one minute. In comparison with the overall test time it was assumed a point-like (*Dirac*) injection for the numerical simulations. The steady-state nested-cell flow field was maintained throughout the entire duration of the tracer test. For the tomographic layout, the surface pump was shifted to a different valve to inject the tracer at a different interval. The process was repeated until the tracer was injected in all three sections. Prior to a subsequent experiment, the system was left under natural conditions for at least one day, providing the remnant tracer with more time to flush out of the domain.

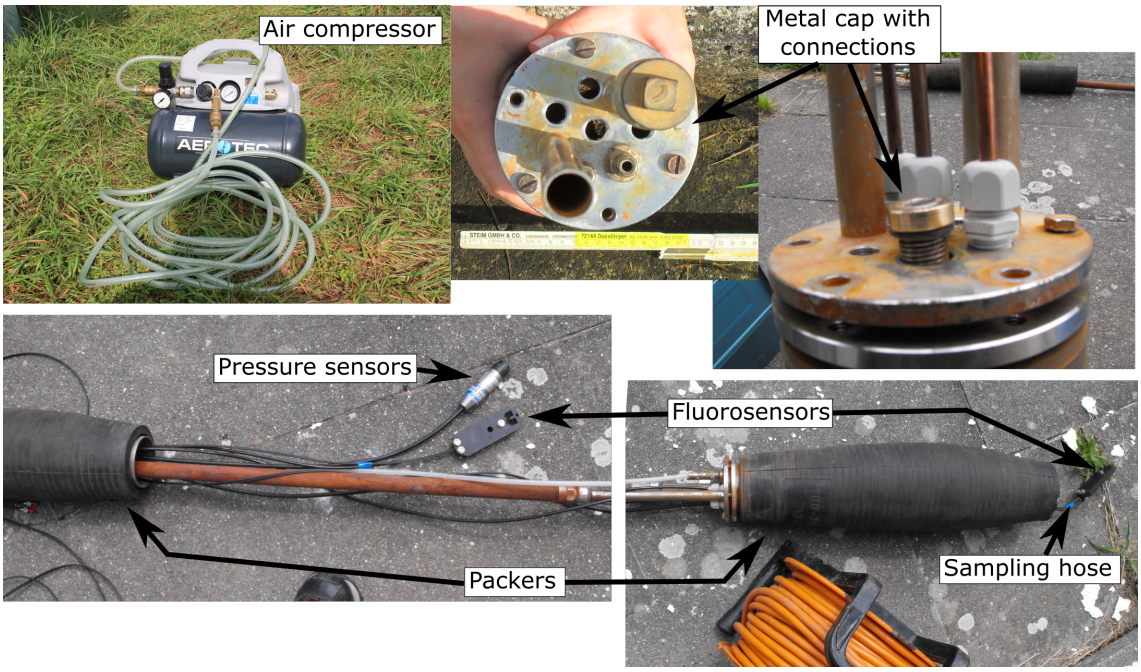


Figure 4.7: Details of the packer system installed in well B3. The system consists of an air compressor to inflate the two packers, a metal cap on top of each packer, sensors to monitor head changes and concentration and sampling hose.

Tracer breakthrough was monitored and recorded at the injection, extraction and many observation wells (see Section 4.5). Two tests per injection interval were performed, shifting both the pressure transducers and fluorosensors among the available observation wells. For clarity, each tracer test is labeled according to the injection location. The two tests with tracer injection at the top level are referred as Tests 1a and 1b. Tests 2a-2b and tests 3a-3b correspond to the experiments with tracer injection at the middle and bottom sections, respectively. Table 4.1 summarizes each test setup (i.e., injection/extraction rates, tracer injection location, mass of tracer injected). Small differences in the injection and extraction rates between tests "a" and "b" of each injection level, as well as the mass of dye tracer injected were unavoidable and had to be taken into account in the numerical simulations.

Table 4.1: Summary of the experimental setup adopted for each individual tracer test of the solute tracer tomography experiment. Top, Mid, Bot refer to the top, middle and bottom section generated with the multilevel injection system at well B3. Positive flow rates represent injection of water, while negative flow rates represent extraction rates.

	Inj.depth (m)	Tracer	Applied Flow Rates $(\mathrm{l s^{-1}})$					
Test	(Level)	Mass(gr)	B2	B3Top	B3Mid	B3Bot	B6	B7
1a	4.0-5.0 (Top)	1.10	5.2	1.6	0.9	0.6	-2.0	-9.0
1b	4.0-5.0 (Top)	1.15	4.8	2.0	0.87	0.6	-2.4	-9.0
2a	5.5-6.5 (Mid)	1.10	5.2	1.8	0.9	0.6	-2.5	-9.0
2b	5.5-6.5 (Mid)	1.20	4.9	2.1	0.81	0.6	-2.0	-9.0
3a	7.0-8.0 (Bot)	1.10	5.2	1.8	0.9	0.6	-2.5	-9.0
3b	7.0-8.0 (Bot)	1.15	4.9	2.2	0.82	0.6	-2.0	-8.7

4.5 Online Monitoring

Fig. 4.8 is a schematic representation of the monitoring system of hydraulic pressure and tracer concentrations that was installed in the available wells.

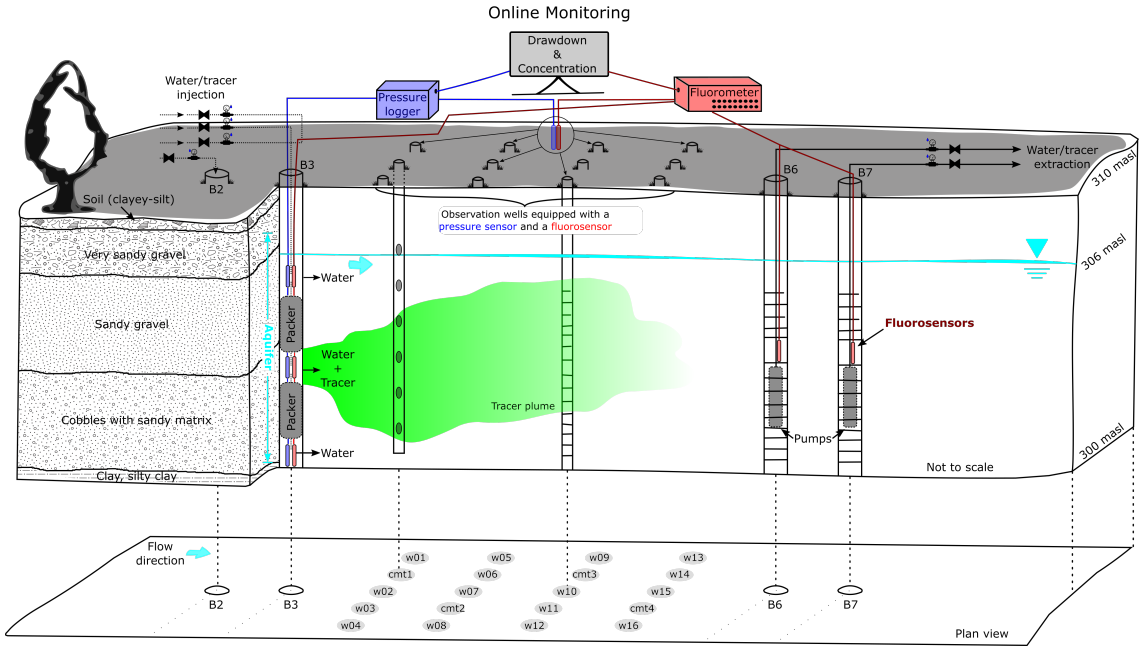


Figure 4.8: Schematic description of the available wells, the injection of tracer (green) at a specific isolated interval, and the online monitoring system installed at the field site. Cross section: main geological units, different types of observation wells, multilevel injection system at well B3, tracer injection at the middle section of B3, injection/extraction of water into/from B-wells. Plan view: horizontal distribution of the wells. The fiber optic pressure transducers and the data logger used to monitor hydraulic heads are represented in blue. The nineteen channel field fluorometer and the corresponding fluorosensors are represented in red. Notice the inclusion of both types of sensors at injection well B3, and fluorosensors at extraction wells B6 and B7.

4.5.1 Hydraulic Measurements

Before each test, natural water levels were measured manually at all B-wells and randomly selected w-wells using a standard E-tape. This data was used to generate potentiometric maps and estimate mean natural hydraulic gradients, needed to define the initial and boundary flow conditions in the numerical simulations of the individual tests.

Hydraulic pressure responses to water injection and extraction were measured with fiber

optic pressure transducers (FISO Technologies Inc., Quebec, QC, Canada, Sensor: FOP-MIV, signal conditions: FPI-HR) (Fig. 4.9), with an accuracy of $\pm 1\,\mathrm{mm}$ and a sampling frequency of $2.5\,\mathrm{Hz}$. These pressure sensors and their respective data logger are schematically represented in blue in Fig. 4.8. Real-time monitoring of hydraulic pressure changes helped to identify when a stable flow field was achieved, which was required prior to any tracer injection. Hydraulic heads were monitored throughout the entire tracer test to detect any disturbance of the flow field. A maximum of 14 fiber optic pressure transducers were distributed among the observation wells. Pressure changes at each level generated in well B3 were monitored using pressure sensors with a sampling frequency of $1\,\mathrm{Hz}$, and recorded with a GL820 GRAPHTEC logger (GRAPHTEC Corporation, Yokohama, Japan). These sensors were needed to properly regulate the injection flow rates, and monitor the exact starting time of water injection. No pressure sensors were placed at extraction wells B6 and B7 for two main reasons: (i) the exact starting time of pumping was recorded manually, and (ii) the noise in the recorded signal produced by the pumps was significant, leading to unreliable measurements.

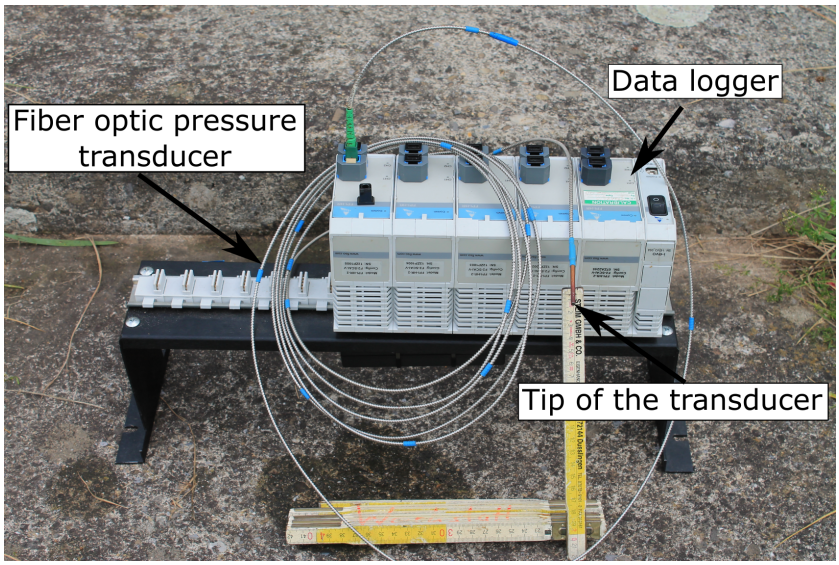


Figure 4.9: Fiber optic pressure transducer used to monitor hydraulic head changes during the tracer tomography experiment.

4.5.2 Tracer Measurements

Tracer concentration measurements were taken with fiber optic fluorosensors, connected to a 19-channel field fluorometer (Hermes Messtechnik, Stuttgart, Germany) (Fig. 4.10), with a sampling frequency of $1\,\mathrm{Hz}$. The field fluorometer measures fluorescence intensity as a response to a stimulation with an LED-light source. Light is transmitted back and forth

between the observation point and the detector through fiber optics. The fluorometer is capable of transforming signal intensity to actual values of concentration with an internal calibration that can be adjusted for the specific conditions of the field experiments. However, previous laboratory work showed that observed signals are highly susceptible to the length of the optic fiber and the color and material of the well casing, affecting concentration values. To avoid propagation of these effects into concentration data, it was decided to retrieve only signal intensity (i.e., signal in mV). To scale the intensity to concentration values, groundwater samples were collected and taken to the laboratory for fluorescence analysis. A detailed description of the sampling system is presented in Section 4.6.

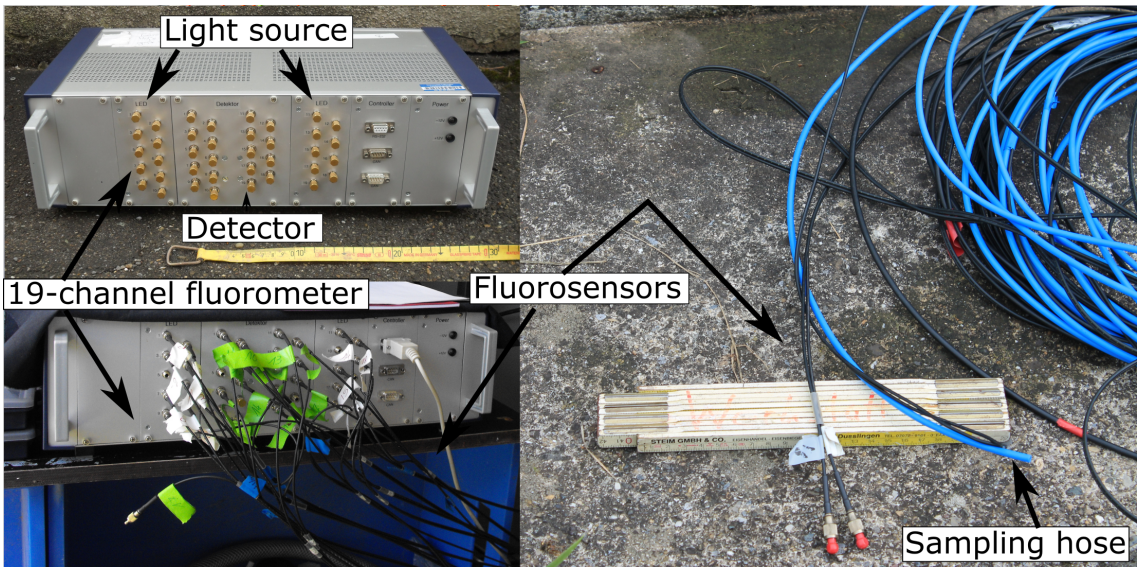


Figure 4.10: Field fluorometer and fluorosensors used to monitor concentrations during the tracer tomography experiment. Left: 19-channel field fluorometer; Right: fluorosensors with groundwater sampling hose.

The exact time of tracer injection was monitored with a fluorosensor installed at each section of the multilevel injection system. To estimate the total mass of tracer recovered, a fluorosensor was placed at extraction well B6. Any tracer bypassing well B6 was monitored with one more fluorosensor installed at the outer extraction well B7. All additional fluorosensors were distributed between the observation wells. The small diameter of both the fiber optic pressure transducers and fluorosensors facilitated logging pressure changes and tracer breakthrough at the same observation point, even in the small diameter multilevel cmt-wells (see Fig. 4.5).

4.6 Water Sampling and Laboratory Analysis

A major implementation in the method was a groundwater sampling system coupled to the online tracer monitoring device. Groundwater samples were needed to scale the signal intensity (in mV) of the recorded breakthrough curves to actual concentration values (in $\mathrm{mg\,l^{-1}}$). The sampling system consisted of $3\,\mathrm{mm}$ diameter tubes attached to each fluorosensor (Fig. 4.10, right) and connected to a high-precision 12-channel peristaltic pump (Ismatec IPC, Cole-Parmer, Germany; Fig. 4.11). To ensure sampling the water parcel that corresponds to the real-time observations, the tip of each sampling tube was fixed to the endpoint of every optic fiber. Two pumps were needed to measure all observation points. The sampling system worked with very low flow rates (max. $10\,\mathrm{ml\,min^{-1}}$) and its effect was not detected by the fiber optic pressure transducers, suggesting that the flow field remained undisturbed throughout the individual tracer experiments.

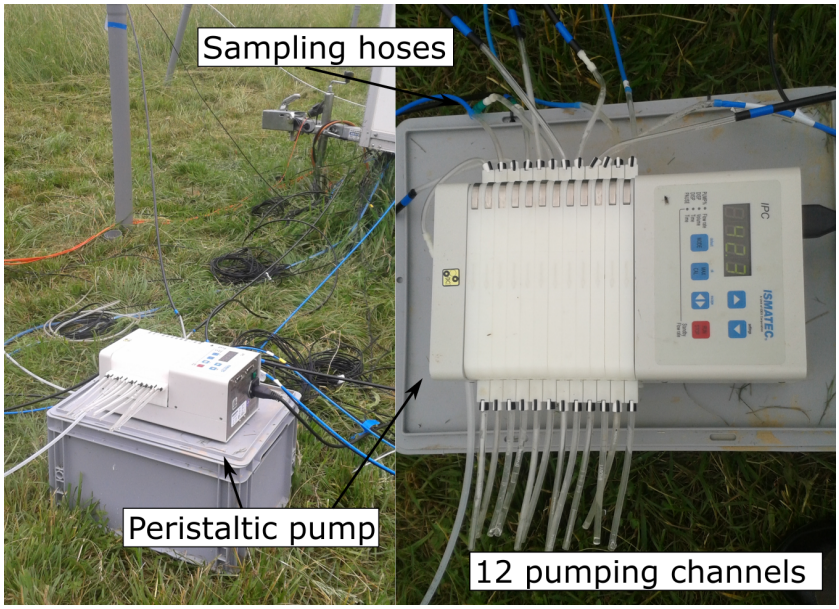


Figure 4.11: Peristaltic pump used to sample water at the observation points.

Extraction well B6 and the three intervals of the inner injection well B3 were monitored and sampled during all six tracer tests. During the second test series (tests 1b, 2b and 3b) the fluorosensor and sampling hose placed at extraction well B7 had technical problems, therefore breakthrough curves and water samples from this well were retrieved only for tests 1a, 2a and 3a. The additional sampling hoses were distributed among all observation wells. As heavy-tailed breakthrough curves have been previously reported at the Research Site Lauswiesen (e.g., Riva et al. 2006; Doro et al. 2015; Sanchez-León et al. 2016), it was decided to collect

at least five samples per observed breakthrough curve, targeting the first arrival, front, peak, early- and late-time tailing concentrations. The exact moment to take the samples was defined by the real-time monitoring system and the time-delay estimated for the water to arrive to the sampling vials. This time-shift was caused by the combination of low-flow rates and the long sampling hoses. Samples were collected in $8\,\mathrm{ml}$ brown vials and taken to the laboratory for fluorescence analysis. To assess background concentration values and verify if tracer from a previous test was still present in the system, a sample was taken at the inner extraction well B3 previous to tracer injection.

At the laboratory, the samples were analyzed with the emission-excitation matrix technique, using a HORIBA Fluoromax-4 spectrofluorometer (Horiba, Ltd. Kyoto, Japan) with a xenon lamp as a light source. This technique is based on the relative intensity of fluorescence of a compound. From previous laboratory analysis of fluorescein dissolved in groundwater samples collected at Lauswiesen, a fluorescent peak has been identified at an absorption (excitation) wavelength of $460\,\mathrm{nm}$, and a fluorescent emission wavelength of $521\,\mathrm{nm}$. These values were incorporated in the settings of the spectrofluorometer.

The spectrofluorometer measures fluorescence intensity in *Counts Per Second* (CPS), and concentration values were obtained with a calibration curve. This calibration curve was based on a stock solution containing the same fluorescent tracer used during the field tests (with a known concentration) and several dilutions of it. The different dilutions were measured to obtain intensity values that correspond to each controlled concentration, defining an intensity-concentration curve. The next step was to fit a parametric function (usually linear) to the curve, and use this function to transform all intensity values measured in the groundwater samples. To avoid extrapolation the calibration curve should cover the entire range of concentrations expected in the field samples. From previous tracer field experiments at the field site, performed with a similar injected mass and investigation distances, it was expected to have tracer concentrations in the range of $\mu\mathrm{g\,l^{-1}}$ and to not exceed a few $\mathrm{mg\,l^{-1}}$. To guarantee a full coverage of the calibration curve, stock solutions of $1\,\mathrm{g\,l^{-1}}$ and $5\,\mathrm{mg\,l^{-1}}$, its 1:10, 1:100, 1:1000 main dilutions, and three secondary dilutions in between each of them were generated. To remove background values, the intensity obtained for water sampled prior to each tracer injection (blank) were subtracted from all further measurements.

Although a linear behavior is expected in the calibration curve, a non-linear curve was observed (Fig. 4.12, left). This was attributed to the wide range of concentrations selected, which was needed to include the samples with the highest concentrations. However, $\sim 98\,\%$ of the samples showed concentration values below $1\,\mathrm{mg\,L^{-1}}$, falling in the linear range of the calibration curve. A linear, cubic and polynomial fit to the calibration points is presented in Fig. 4.12 (left). It was observed that linearity was improved when the data was log-transformed and therefore it was decided to perform a linear fitting in log space (Fig. 4.12, right).

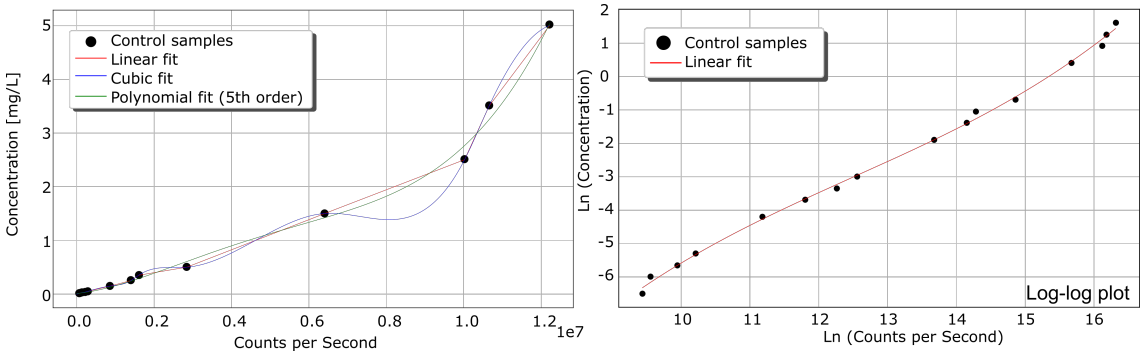


Figure 4.12: Concentration as a function of counts per second. Left: comparison of the linear, cubic, and polynomial fit to the data; Right: linear fit in log-space.

The concentration values for all water samples were used to scale the recorded breakthrough curve signals (in mV) to actual values of concentration. Details of the methodology applied to scale the breakthrough curves are provided in Chapter 5.

Chapter 5

Data Processing

Six individual tracer tests were performed during the solute tracer tomography experiment reported in this work (see Table 4.1). The breakthrough curves of each tracer test are accompanied by a set of hydrographs (later drawdown) and fluorescence laboratory measurements of corresponding groundwater samples. The full raw dataset is stored in a HDF5 database that contains all relevant information as metadata (e.g. well location, units, acquisition rates, injection/extraction rates, injected tracer mass). The database is provided as electronic appendix (Appendix: Data/Raw), including a brief description of the content and structure of the file.

To increase the signal-to-noise ratio of the raw data, detect and remove background values and trends introduced by the sensors or correct signal shifts, a python-based time-series processing module named PyProc, was developed (see Appendix: Data/PyProc). PyProc has a command-line interface and is capable of loading and plotting data directly from the HDF5 database, and from ASCII files with a broad variety of formats, but is specially compatible with the files generated by the sensors used in this work. All modifications made to the raw data were automatically stored using a logging feature that records sufficient information to guarantee reproducibility of the entire data processing workflow.

The first part of this chapter presents the theory behind some of the tools implemented in PyProc, and their application to the hydraulic and tracer data collected during the field experiments. Afterwards, the complete processed dataset is presented graphically together with a brief description of the main characteristics. The final part of the chapter focuses on the analysis of the processed drawdown and breakthrough curves with available analytical solutions and temporal moments. The processed data is provided in ASCII format as an electronic appendix (Appendix: Data/Processed).

5.1 Time Series Processing Methods

Examples of two raw hydrographs and two raw breakthrough curves recorded during test 3b (tracer injected at the bottom section of well B3) are shown in Fig. 5.1. The two hydraulic head changes were measured at observation wells w14 and cmt3-6, and the two breakthrough curves were recorded at observation well w14 and pumping well B6. For clarity, Fig. 5.1 also contains the relative location of the wells at the field site.

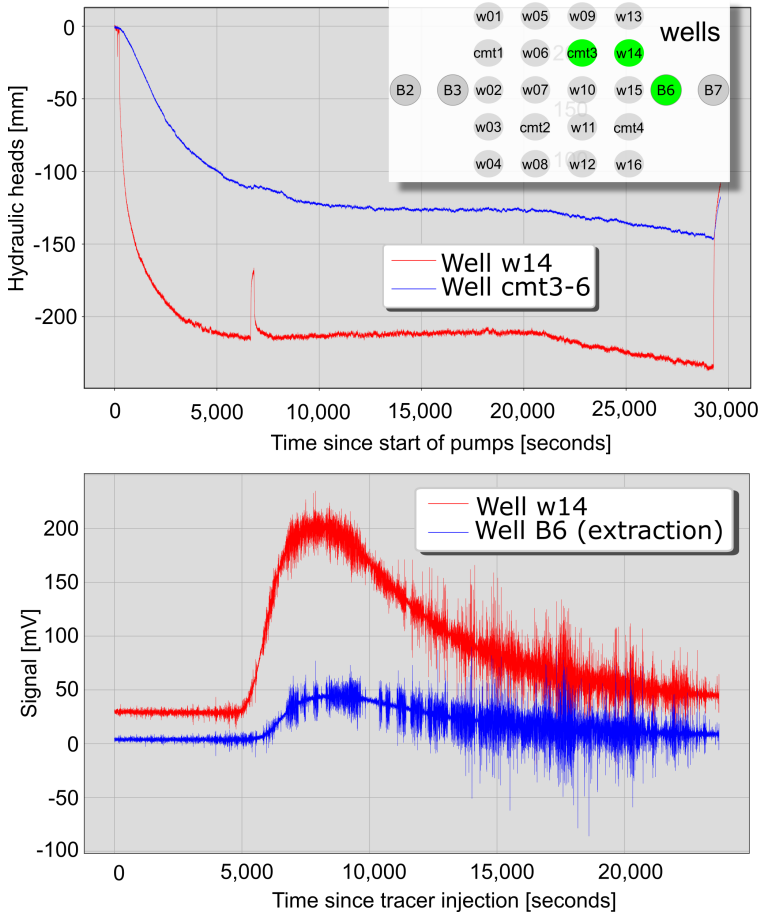


Figure 5.1: Example of real measurements obtained during the tracer tomography experiment performed at the Lauswiesen research site. Curves recorded during test 3b: tracer injected at bottom section of well B3. Top: hydraulic head response measured at two different observation wells; Bottom: two breakthrough curves (in mV); Additionally, well distribution at the field site: green circles show the corresponding location of the hydraulic head and breakthrough curves plotted in the figure.

These curves were selected to illustrate common perturbations observed in the signals. The two hydrographs in Fig. 5.1 (top) show the hydraulic response observed at wells w14 (blue line) and cmt3-6 (red line). They present a short-time disturbance (intensified in well w14) at about $6500\,\mathrm{s}$ to $7000\,\mathrm{s}$ after start recording. This disturbance was related to electrical problems that affected the pumps. A linear trend appears at both curves at $\sim 18\,000\,\mathrm{s}$ after initiating the test. These linear trends are often observed in records of tests with long execution times and are mainly attributed to sensor drifts .

The two breakthrough curves in Fig. 5.1 (bottom) were selected due to the relatively large measurement noise but obvious response to the tracer observed at wells w14 and B6. The breakthrough curve recorded at extraction well B6 was specially selected to exemplify the application of temporal-moment analysis to breakthrough curves.

The first processing step consisted in **trimming** the datasets according to the exact beginning of the hydraulic and tracer tests. Data recorded before the starting of the tests, were used to obtain an estimate of the measurement noise of each sensor.From the hydrographs it can be observed that a steady flow was achieved after $\sim 8000\,\mathrm{s}$ of pumping and injecting water. The flow field was kept constant for $\sim 22\,000\,\mathrm{s}$, corresponding to the time needed for the tracer test, and afterwards the injection/extraction of water was stopped. In this work the recovery phase of the hydraulic field was not used, and therefore the hydrographs were also trimmed to the time where steady state was reached. Additionally, the measured values of hydraulic head were converted to drawdown.

The next step was to apply basic tools for the **correction** of shifts, **detection/removal** of (*obvious*) outliers, **removal** of linear trends, and **interpolation** for missing values. The application of these tools is straightforward and therefore they are not further explained. The resulting curves are shown in Fig. 5.2. The hydrographs were converted into drawdown curves and the jump in the signal was removed and replaced by points generated by linear interpolation. The breakthrough curves were trimmed according to the exact time of tracer injection and the background values registered by the sensors were subtracted.

After these basic processing steps, more elaborated functions to reduce measurement noise such as smoother functions, automated outlier detection algorithms and optimization algorithms to fit parametric functions to the data were applied. In the next part of this chapter, the theory behind and application of the following processing methods is presented:

- Smoother functions: Windows
- Smoother functions: Kalman filter
- Outlier detection: Modified Z-scores
- Fit parametric functions: Generalized Inverse Gaussian

- Breakthrough signal scaling: from voltage to concentration units
- Breakthrough tail completion

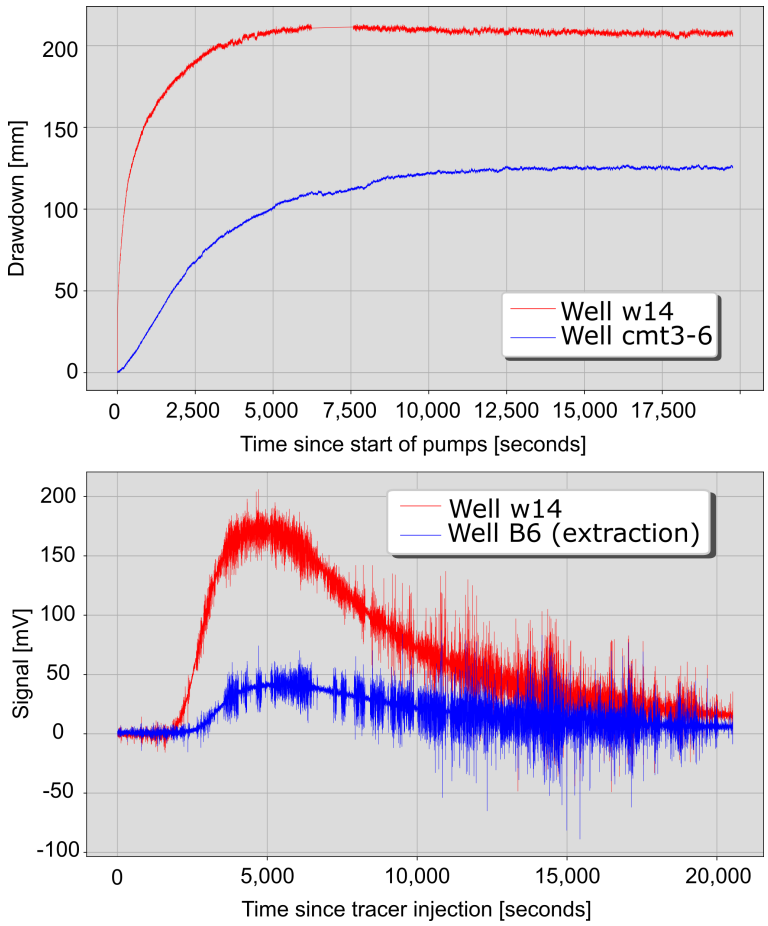


Figure 5.2: Curves from Fig. 5.1 after the initial processing steps: trimming the series to the beginning of pumping and tracer injection , shifting, obvious outlier/jumps removal. Left: additional transformation of hydraulic heads to drawdown; Right: background values removed from breakthrough curves signal.

5.1.1 Smoother Functions: Windows

A window function is a mathematical function with a defined shape and length, composed of factors (or weights) that usually range between 0 and 1. To smooth the original time series, this function is convolved, i.e. overlapped and shifted over time, with the noisy data. Window functions help to reduce the variance of a dataset in exchange for a reduction of resolution. In this work the most common windows used were the Hanning, Hamming, Gaussian and Exponential windows. Their shape is defined as follows (see Fig.5.3):

$$Hanning: \quad w(n) = 0.5 - 0.5\cos\left(\frac{2\pi n}{L-1}\right) \quad with\ 0 \leq n \leq L-1 \tag{5.1a}$$

$$Hamming: \quad w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{L-1}\right) \quad with\ 0 \leq n \leq L-1 \tag{5.1b}$$

$$Gaussian: \quad w(n) = \exp^{-0.5\left|\frac{n}{\sigma}\right|^{2p_{sh}}} \tag{5.1c}$$

$$Exponential: \quad w(n) = \exp^{-\left|n-\frac{L-1}{2}\right|/\tau} \tag{5.1d}$$

where $w(n)$ is the factor assigned to the n sample, L is the length of the window, σ is the standard deviation of the function, p_{sh} is a shape parameter, and τ defines the decay in the exponential function.

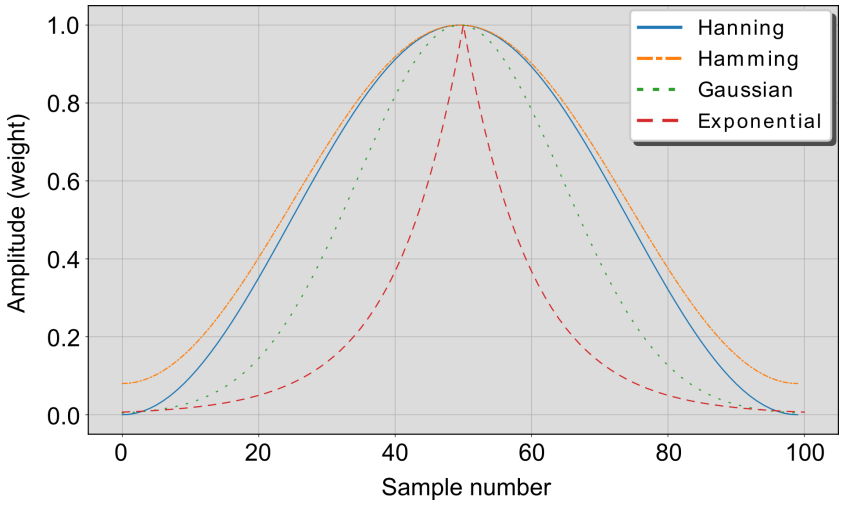


Figure 5.3: Shape of different windows used to smooth the time series recorded during the tomographic field experiments.

The shapes of the smoothing windows described by Eqs. 5.1a - 5.1d are plotted in Fig. 5.3. These shapes do not have large variations, however each window treats the tails with a different approach. A Hanning window for example, assigns a value of zero to the values at the limits of the window, while the Gaussian and exponential distribution approach zero only

asymptotically. An advantage of the Gaussian and exponential windows is the possibility of adjusting not only the length of the window, but also their input parameters. For the Gaussian window the input parameter is the shape parameter p_{sh}, with $p_{sh} = 1$ for a shape identical to a normal distribution, and $p_{sh} = 0.5$ for a Laplace distribution shape. In the exponential window the shape is defined by the decay rate τ.

The effect of the four different windows in the drawdown curve of well w14 is shown in Fig. 5.4. To compare their performance, a length $L = 10$ was defined for all windows. All extra parameters (p_{sh}, σ, τ) were set to 1 for the Gaussian and exponential windows. Fig. 5.4A shows an overview of the entire record. In a more detailed inspection, Fig. 5.4B shows the original and smoothed signals for the first $50\,\mathrm{s}$ of the pumping test, and Fig. 5.4C shows the same signals but at a later time ($\sim 4300\,\mathrm{s}$ to $4600\,\mathrm{s}$). From Eqs. 5.1a & 5.1b it is not surprising that the processed signal presents a similar behavior for the Hanning and Hamming windows. Although these two windows outperformed the Gaussian window at later times, the latter produces better results at early times. An important advantage of the Gaussian window is that the shape can be adjusted to be consistent with the measurement noise of each series, which can be measured directly from the time series. The exponential window is outperformed by all other windows throughout the entire record. This can be explained by the shape of the window, which decreases exponentially from the center of the function, resulting in a strong damping effect during convolution and a loss of most of the early time behavior. Additionally, the exponential window has a comparably large spectral density at high frequencies, whereas all other window functions are low-pass filters.

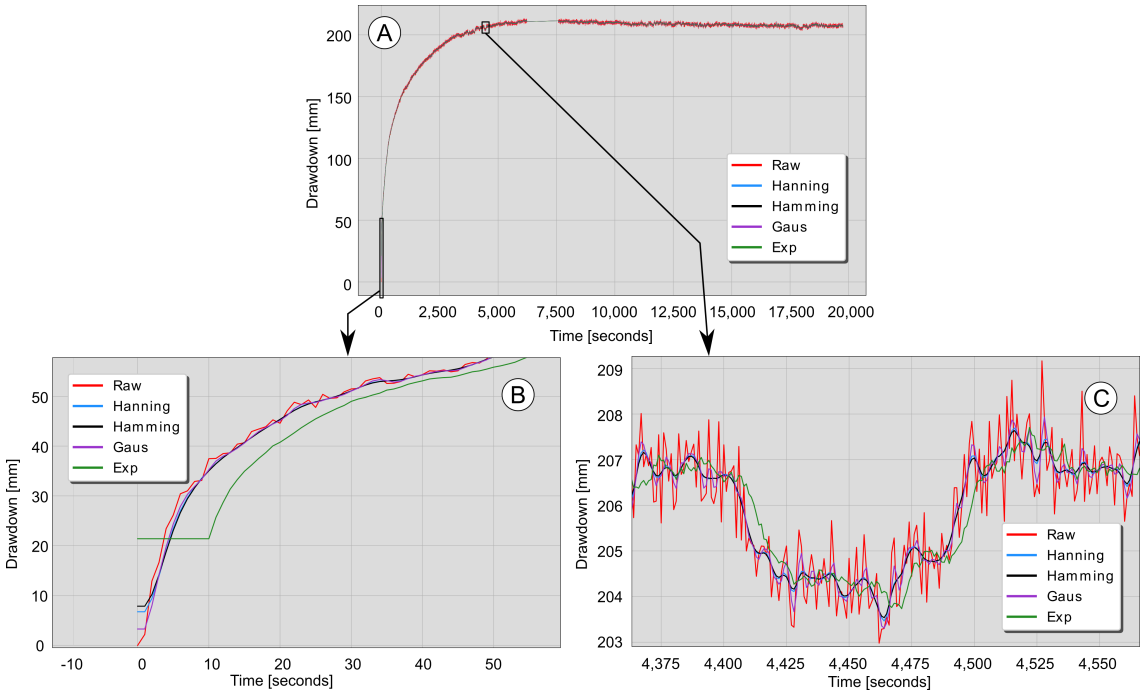


Figure 5.4: Application of the different types of window function to drawdown.

5.1.2 Smoother Functions: Kalman Filter

Equations 2.10 - 2.20 in Section 2.3 describe a version of the Kalman filter adapted for the estimation of model states and parameters with an ensemble approach. The classical Kalman filter, however, was a development of estimation theory for the estimation of a process from uncertain measurements (Kalman, 1960), using a probabilistic version of the least-squares method. In this work, the original version was applied as a smoother function to process many of the measured field curves. Obviously, both the classical Kalman filter and the ensemble based approach share the same underlying theory. Hence, in the following only a short description highlighting the main differences and assumptions for its implementation is presented.

Rewriting the forecast equation (Eq. 2.10) to describe the state x of a linear system at a discrete time t:

$$x_t = \mathbf{F}x_{t-1} + q_t \tag{5.2}$$

and describe the available (noisy) measurements $\mathbf{y}^{obs} \in \mathfrak{R}^n$ of that process as:

$$y_t^{obs} = \mathbf{H}x_t + \epsilon_t \tag{5.3}$$

In this version of the filter, matrix $\mathbf{F}$ defines the system's process and relates the state estimate (head or concentration) at time $t-1$ to the estimate at the current time step t. $\mathbf{H}$ relates the current estimate x_t to measurement y_t^{obs}. As before, variables q_t and ϵ_t represent the process and measurement noise, respectively. The error in the process is assumed constant and with a very low value (e.g., $q = 1 \times 10^{-4}$). ϵ_t is obtained from the measurement noise estimated directly from the data. The process in Eq. 5.2 does not change and the state is defined by the shape of the time-series, therefore matrix $\mathbf{F}$ equals one, and the noisy measurements are of the state directly so matrix $\mathbf{H}$ equals unity and $x_t = y_t$. y_t^{obs} then becomes (Welch and Bishop, 2006):

$$y_t^{obs} = y_t + \epsilon_t \tag{5.4}$$

where y_t is the (unknown) real state of the process. In the discrete Kalman filter, the prior variance Q_t^- at the current time step is calculated from the posterior variance at the previous time step (Q_{t-1}) plus the error assumed in the process:

$$Q_t^- = Q_{t-1} + q \tag{5.5}$$

The posterior estimate $\hat{y}_t$ can then be computed as a linear combination of the prior estimate $\hat{y}_t^-$ plus a weighted difference between the noisy measurement y_t^{obs} and the prediction $(\hat{y}_t^- + \epsilon_t)$ of that measurement:

$$\hat{y}_t = \hat{y}_t^- + k_t^g(y_t^{obs} - (\hat{y}_t^- + \epsilon_t)) \tag{5.6}$$

The weight k_t^g corresponds to a scalar form of the Kalman gain (Eq. 2.19). As stated before, k_t^g is derived to minimize the posterior variance Q_t (see Kalman 1960; Brown and Hwang 1997). Following the procedure presented so far, k_t^g is defined as:

$$k_t^g = Q_t^-(Q_t^- + \epsilon_t)^{-1} \tag{5.7}$$

Finally, the posterior variance Q_t is updated via:

$$Q_t = (1 - k_t^g)Q_t^- \tag{5.8}$$

After each time and measurement update, the process is repeated with the previous posterior estimates used to predict the new prior estimate (Welch and Bishop, 2006) (Fig. 5.5).

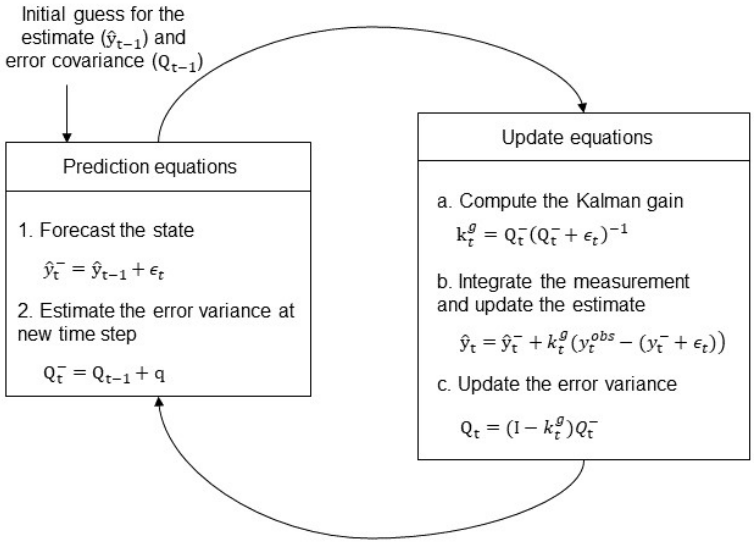


Figure 5.5: Recursive scheme of the Kalman filter, together with the equations for the prediction and update steps. Modified from Welch and Bishop (2006).

The first value of the original data series was used as the initial guess of the estimated state $\hat{y}_{t-1}$, and an initial value of 1 was assigned to Q_{t-1}. As Welch and Bishop (2006) demonstrated, the initial Q_{t-1} is not critical, and the filter would eventually converge as long as $Q_{t-1} \neq 0$.

Previous to the application of the filter, the error measurement was calculated for each curve. A standard deviation $\sigma_{dr} = 1.3\,\mathrm{mm}$ was estimated for the drawdown curve measured at well w14, and a $\sigma_{bt} = 7.5\,\mathrm{mV}$ for the breakthrough curve of well B6. Figs. 5.6 and 5.7 show the effect of the Kalman filter with different standard deviations in the drawdown and breakthrough curves, respectively.

The details of Figs. 5.6B and 5.6C correspond to the beginning of the pumping test and a section of the curve with a high slope. The close-up plots of Figs. 5.7B and 5.7C focus on the first arrival and peak of the breakthrough curve. With a visual inspection, it is noticed that the performance of the filter is optimized when the standard deviation used in the filter is consistent with the measurement noise of the dataset. The higher the standard deviation, the less confidence is given to the data and a higher smoothing is applied to the series (see Eq. 5.7).

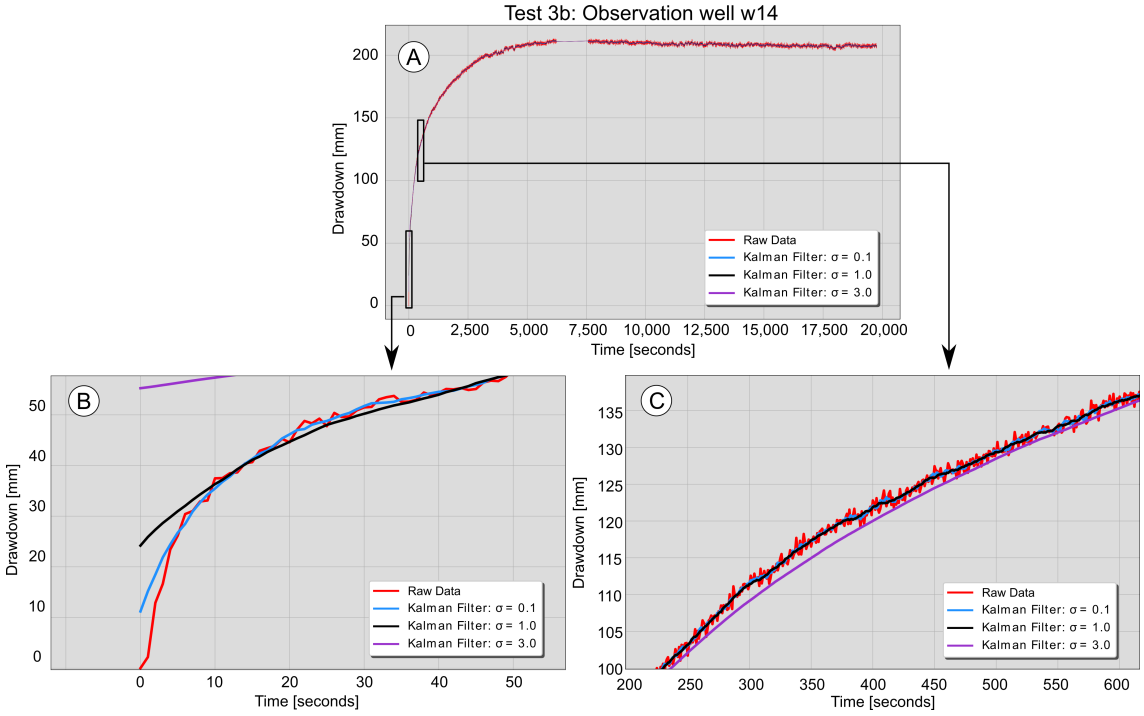


Figure 5.6: Application of the Kalman filter (with different standard deviations) to the drawdown curve (in mm) measured during field test 3b at observation well w14. A: Full record; B: close up to the first $50\,\mathrm{s}$ of the pumping test; C: close up to the time range $200\,\mathrm{s}$ to $600\,\mathrm{s}$, where high hydraulic head changes are observed.

The Kalman filter was specially useful for processing breakthrough curves. The signals measured with the 19-channel fluorometer usually contained high level of noise (Fig. 5.7A). Yet, the shape of the signal can be clearly differentiated. The main process is well identified by the filter even when the standard deviation is over- or underestimated. However, a closer look to the first arrival (5.7B) and peak (5.7C) reveals that main features of the signal, such as the first arrival, begin to distort if the standard deviation is overestimated. Conversely, if the standard deviation is underestimated the resulting smoothed signal still contains strong variations produced by the measurement noise.

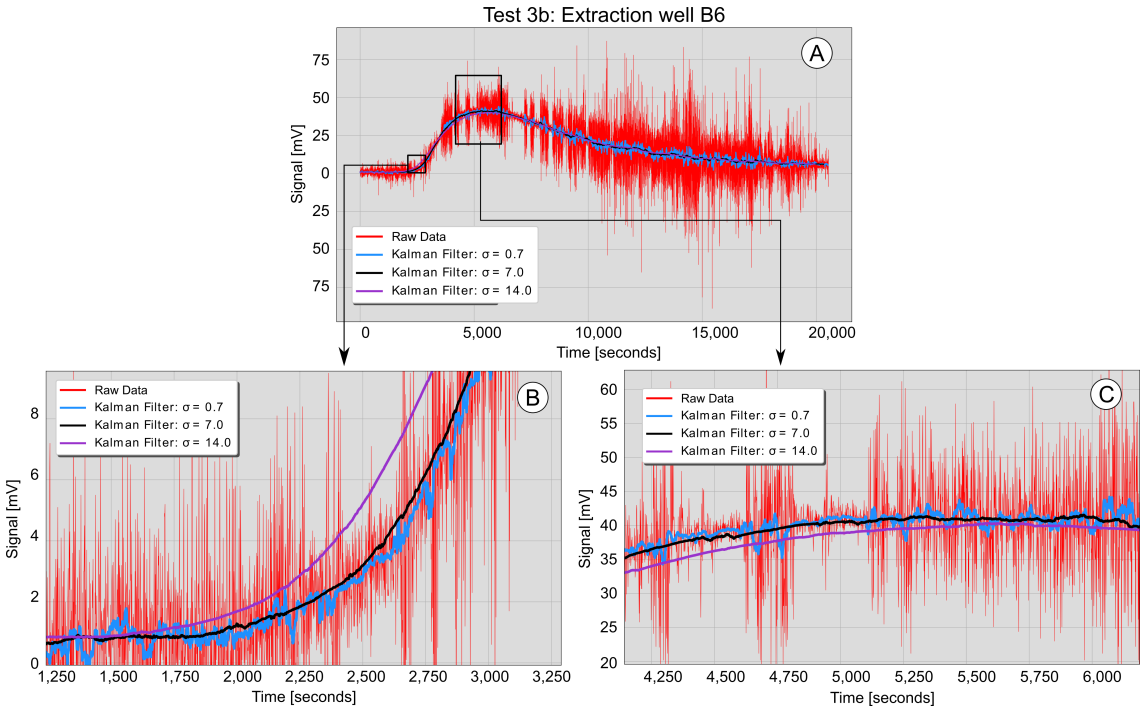


Figure 5.7: Application of the Kalman filter (with different standard deviations) to the breakthrough curve (in mV) measured during field test 3b at extraction well B6. A: Full record; B: close up to the first arrival of the tracer; C: close up to the peak of the breakthrough curve.

5.1.3 Outlier Detection: Modified Z-scores

The Z-scores (or standard score) method is an outlier test based on the relationship between a single data point and the mean and standard deviation of a group of points. An underlying assumption of the method is that the dataset follows an approximately normal distribution. This assumption is not valid for data collected during pumping and tracer tests, therefore a modified version of the method was implemented, in which standard scores are estimated and evaluated in a piecewise approach, rather than using the full dataset. If white noise is present in the dataset, this piecewise approach improves the normality of each sub-dataset, however it can generate considerably small intervals of data rendering the Z-scores method unstable (Iglewicz and Hoaglin, 1993).

To improve the stability, Iglewicz and Hoaglin (1993) suggested using a modified version to calculate the Z-scores z_i:

$$z_i = \frac{1}{1.482} \frac{x_i - \tilde{x}}{MAD(x)} \tag{5.9}$$

with $\tilde{x}$ representing the median of the current sub-dataset. MAD denotes the median absolute

deviation and is defined as:

$$MAD(x) = median(|x_i - \tilde{x}|) \tag{5.10}$$

where the bars $|\cdot|$ denote the absolute value of the argument. The denominator $1.482{*}MAD$ approximately equals the standard deviation of the subset. Iglewicz and Hoaglin (1993) defined as a potential outlier any modified Z-score higher than 3.5. In this work the threshold value of 3.5 was not always optimal, and therefore different values were evaluated. Fig. 5.8 shows an example of the modified Z-scores applied to the hydrograph recorded at observation well w14.

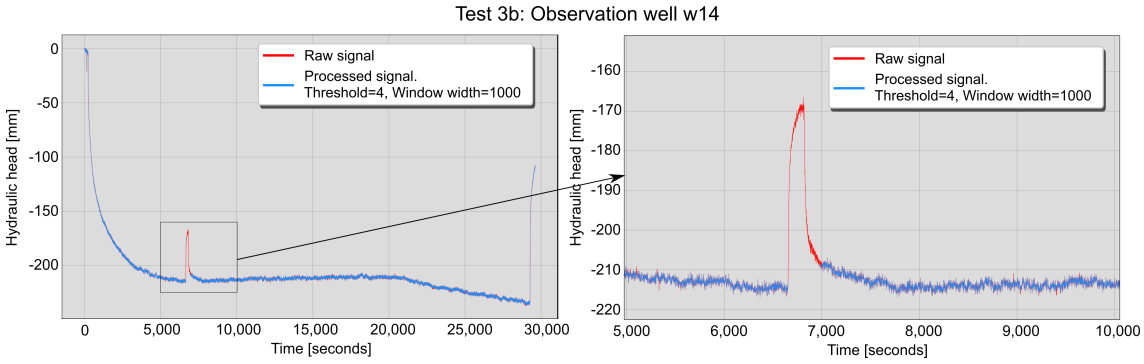


Figure 5.8: Modified Z-score method applied to the original hydrograph measured at well w14, during the field test 3b. Red lines: raw dataset; Blue lines: processed signal; Left: overview of the full record; Right: close up to the time range where a clear disturbance of the signal occurred.

As mentioned at the beginning of this chapter, the peak observed between $6500\,\mathrm{s}$ to $7000\,\mathrm{s}$ was attributed to a short-circuit that affected the performance of the pumps for $\sim 3\,\mathrm{min}$. To remove the effect of the pumping instabilities, the modified Z-score method was applied to the raw signal (red line) with a piecewise interval of 1000 data points and a threshold value of 4. The method was able to remove most of the perturbations in the signal (Fig. 5.8: blue line), however it fails in recognizing the late effects (between $7000\,\mathrm{s}$ to $8000\,\mathrm{s}$) of the short-circuit. The remained affected data was manually removed. The performance of the modified Z-score method is optimal for drastic changes in the signal (e.g., appearance of single outliers) and suboptimal for smooth transitions.

5.1.4 Fit Parametric Functions: Generalized Inverse Gaussian

The objective of fitting a parametric function to noisy data is to find a smooth curve, which is fully characterized by a few parameters, that follows the trend of the data. While a parametric model has the disadvantage that it predefines the general shape of the function, it has

the advantage that it can be used for extrapolation. In this work, it was observed that the generalized inverse Gaussian (GIG) distribution (Jørgensen, 1982) resembles the heavy tailed concentration breakthrough curves typically observed at the Hydrogeological Research Site Lauswiesen. The GIG distribution is defined as follows:

$$f(x) = \frac{(a/b)^{p/2}}{2K_b(\sqrt{ab})} x^{(p-1)} e^{-(ax+b/x)/2}, \ \forall x > 0 \tag{5.11}$$

where K_b is the modified Bessel function of the second kind and order b, a and b are non-negative parameters, and p is a scalar. For a proper fitting of the GIG distribution to the tracer data, a scaling factor (η) and a shift in the abscissa (Δx) and ordinate (Δy) were included, resulting in the following form:

$$f(x) = \eta \frac{(a/b)^{p/2}}{2K_b(\sqrt{ab})} (x - \Delta x)^{(p-1)} e^{-(a(x-\Delta x)+b/(x-\Delta x))/2} + \Delta y, \ \forall x > 0 \tag{5.12}$$

Parameters a and b regulate the concentration and scale of the densities. The scalar p has no statistical meaning. Δx and Δy allow shifts on both dimensions, accounting for delayed arrival times and background values in the observed signal, and η scales the distribution to the units of the data. These parameters have no physical meaning for solute transport. The fitting was performed using non-linear least squares with the Levenberg-Marquardt algorithm, and the acceptance criteria was based on the accuracy of reproducing key parts of the breakthrough curve such as first arrival, location/intensity of the peak and tailing. An initial guess for the parameters were assigned as follows:

- $a_o = 1/N_x$, with N as the number of data points,
- $b_o = 1/t_{end}$, with t_{end} being the last time step of the series,
- $p = -0.5$, defining an inverse Gaussian model (Efoevi Koudou and Ley, 2013), which is the analytical solution to the advection-dispersion equation,
- $\Delta y = y_{ini}$, with y_{ini} being the initial y value of the data series,
- $\eta = m_0$, with m_o as the 0^{th} temporal moment of the breakthrough curve,
- $\Delta x = \frac{m_1}{m_0} - \sqrt{m_{2c}/m_0}$, with m_1 and $m_2 c$ representing the first and second central temporal moments of the original curve (see below for the definition of temporal moments).

Fig. 5.9 shows the processed curve after fitting the GIG distribution to the noisy data measured at extraction well B6 during tracer test 3b. Despite the high level of noise in the raw signal (red line), the fitted curve (blue line) is a noise-free breakthrough curve with all relevant features of the original data.

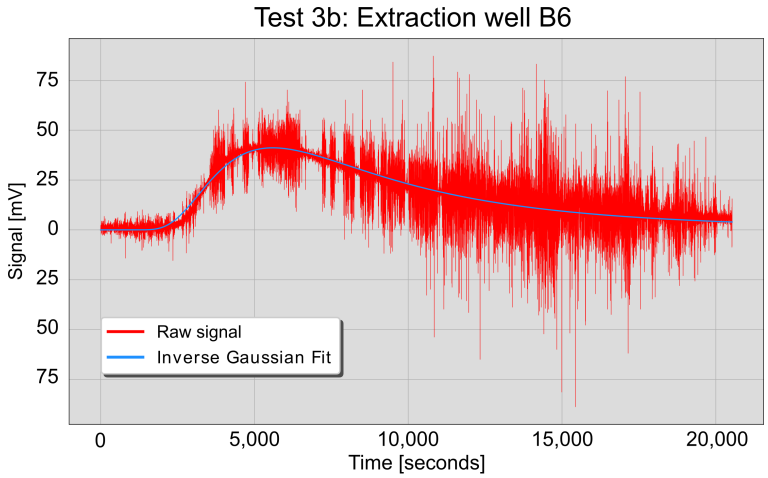


Figure 5.9: Fit of the generalized inverse Gaussian distribution to the breakthrough curve measured during test 3b at the extraction well B6.

5.1.5 Signal Scaling: from voltage to concentration units

Once the noise in the tracer data was reduced, the recorded breakthrough curves (originally measured in mV) were scaled to actual values of concentration, minimizing the L2-norm (or least squares) between the laboratory measurements obtained from the groundwater samples and the signal values at the corresponding times.

As an example, Fig. 5.10A shows the scaling process of the same breakthrough curve presented above. The concentration values of the groundwater samples measured at the laboratory (blue dots), are overlapped with the processed breakthrough curve signal (red line) at the corresponding time. The signal in mV was then scaled to concentration by minimizing the Euclidean distance (represented by the black arrows in Fig. 5.10A) between the laboratory concentrations and the raw signal. The scaled breakthrough curve is shown in Fig. 5.10B.

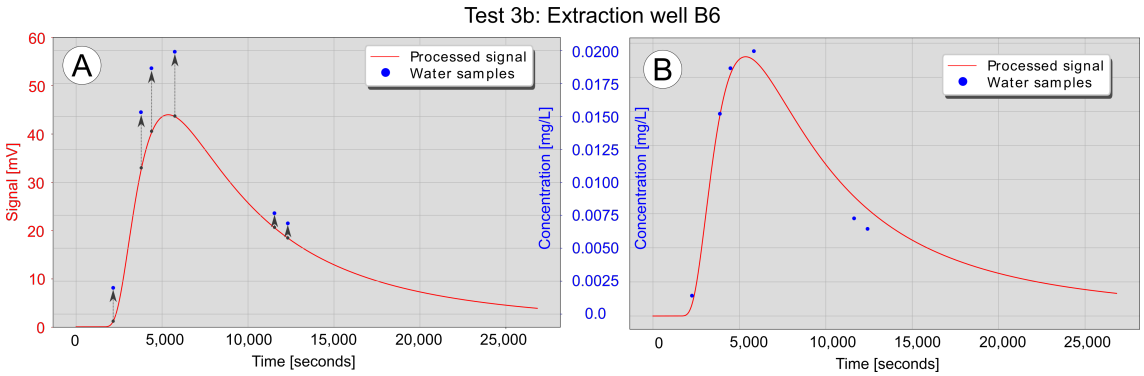


Figure 5.10: Visual representation of the scaling process of a breakthrough curve. Left: raw signal in red lines with the concentration values of the water samples. Black arrows represent the Euclidean distance between each concentration value and its corresponding value in the breakthrough curve. Right: breakthrough curve already scaled to $\mathrm{mg\,l^{-1}}$.

The residual distance between the laboratory concentrations and the breakthrough curves is attributed to the inherent errors introduced by sampling at low pumping rates, plus errors from the field fluorometer and laboratory measurements.

5.1.6 Breakthrough Tailing Completion

Effective transport parameters can be estimated by matching the temporal moments of breakthrough curves simulated with transport models to those of the real data. Classical methods to estimate effective parameters from temporal moments assume complete breakthrough curves. However, truncated records are usually retrieved from real tracer tests, which usually have a constrained execution time. Jawitz (2004) suggested evaluating complete moments of truncated breakthrough curves by completing the dataset with an exponential distribution, when the final portion of the curve exhibits a linear trend in a semilogarithmic scale. The exponential distribution function, including a shift on the abscissa, is defined as:

$$f(x) = a\exp(-bx) + \Delta x \tag{5.13}$$

where a is a scaling factor, b is the base of the function and Δx represents a shift factor. The fitting was performed using the same non-linear least squares algorithm used for fitting the generalized inverse Gaussian distribution to the tracer data.

Fitting an exponential tail is efficient and easy to implement, however, it could not be properly applied to all recorded datasets because the semilogarithmic trend was not clearly identified. In those cases, it was preferred to first fit a GIG distribution to the data, and

use the fitted function for extending the dataset. Fig. 5.11 shows a comparison of the two different approaches used to extend the breakthrough curve tail to a time value of $35\,000\,\mathrm{s}$.

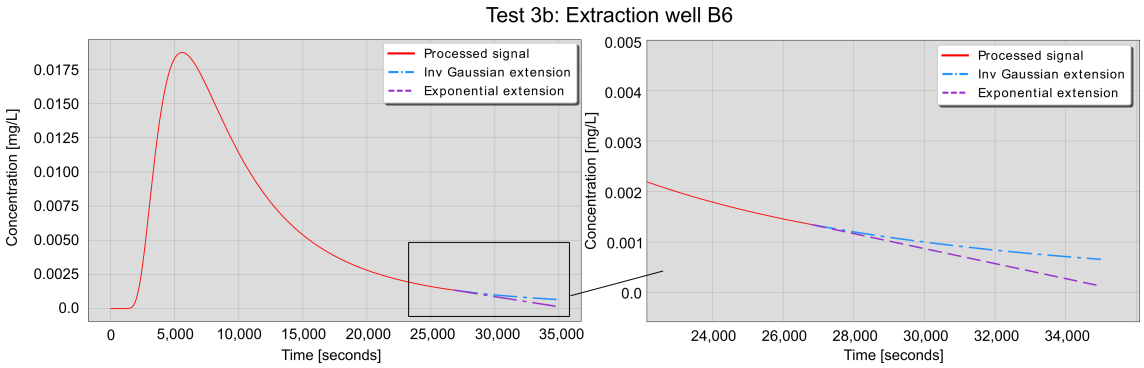


Figure 5.11: Tail completion of the breakthrough curve measured at extraction well B6 during test 3b. Left: full record of the processed breakthrough curve (red line), and the tail extension obtained with the GIG (blue dashed line) and exponential (purple dashed line) methods. Right: details of the shape of the different tails obtained with the two different methods.

In the example presented in Fig. 5.11, the extension with the GIG distribution (blue line) has a smooth slope that agrees with the late-time behavior of the original signal. In contrast, the extension with the exponential distribution (purple line) shows a higher slope, approaching the concentration value of zero more rapidly.

So far there is not a consensus within the research community on the proper way to complete breakthrough curves. Rather than completing the tail, which is inherently an extrapolation of the dataset, some have attempted to derive relationships between complete moment-equations and the truncated dataset (e.g., Jawitz 2004 and Luo et al. 2006a). To fail in considering breakthrough tailing may result in a poor quantitative assessment of remediation techniques where the volume of water to be treated needs to be estimated accurately (Neville et al., 2000).

The estimation of parameters with the EnKF was based in the transient behavior of both the hydraulic and (truncated) tracer datasets and not in their temporal moments, avoiding the additional uncertainties related to their calculation. The temporal moment analysis was implemented only for the assessment of the individual tracer tests, and to provide a general idea of the effective transport parameters at the field site.

5.2 Processed Drawdown and Breakthrough Curves

To obtain a final dataset that can be used for parameter estimation, characteristic time points were selected from the smoothed curves. A total of 27 time points were selected from each processed drawdown curve (for all tests), with more data points from the early-time responses than at late time. Early-time hydraulic changes are more sensitive to hydraulic parameters than those at late time (Leven and Dietrich, 2006). The number of characteristic points selected from the breakthrough curves varied according to the length of each test.

The hydraulic dataset for all six tests is presented in Fig. 5.12. Hydraulic head changes during test 2a could not be retrieved due to technical problems with the data logger. The head responses can be separated between those affected primarily by the injection wells and those mainly affected by the extraction wells. The curves with a negative drawdown correspond to the wells located closer to the injection wells. Conversely, the curves with only positive drawdown were recorded at the wells closer to the extraction points. The magnitude of the drawdown is related to the distance to the injection and extraction wells. The larger the drawdown, the closer the observation point was to the extraction wells.

In general, the hydraulic response to injection and extraction of water had a similar response throughout the five tests shown in Fig. 5.12. The small changes observed from test to test were attributed to the slightly different injection and extraction rates applied. During the field experiments, steady flow was ensured prior to injection of the tracer. For tests 1a and 1b, steady state was achieved after $5000\,\mathrm{s}$. In all other tests, steady-state conditions were not observed before $7000\,\mathrm{s}$. Multiple points taken from the steady-state portion of the drawdown curves would not provide additional information for parameter estimation, therefore the curves were trimmed at the first time point when a steady flow was observed at most observations wells.

Fig. 5.13 shows the breakthrough curves measured at extraction wells B6 and B7 during the six tracer tests of the tracer tomography experiment. Due to difficulties with the sampling channel placed at extraction well B7, breakthrough curves for the second series of tracer tests (tests 1b, 2b and 3b) could not be scaled. The tracer was detected in both wells, but concentrations observed in well B7 were by one order of magnitude smaller than those in well B6. This indicates that most of the tracer indeed arrived at well B6, and only a small portion leaked to well B7. Concentrations and peak arrival times observed at well B7 were more or less identical in the three tracer tests.

The peak of the breakthrough curves measured at well B6 during the second test series (tests 1b, 2b and 3b) appears earlier than in the first test series. This may be caused by

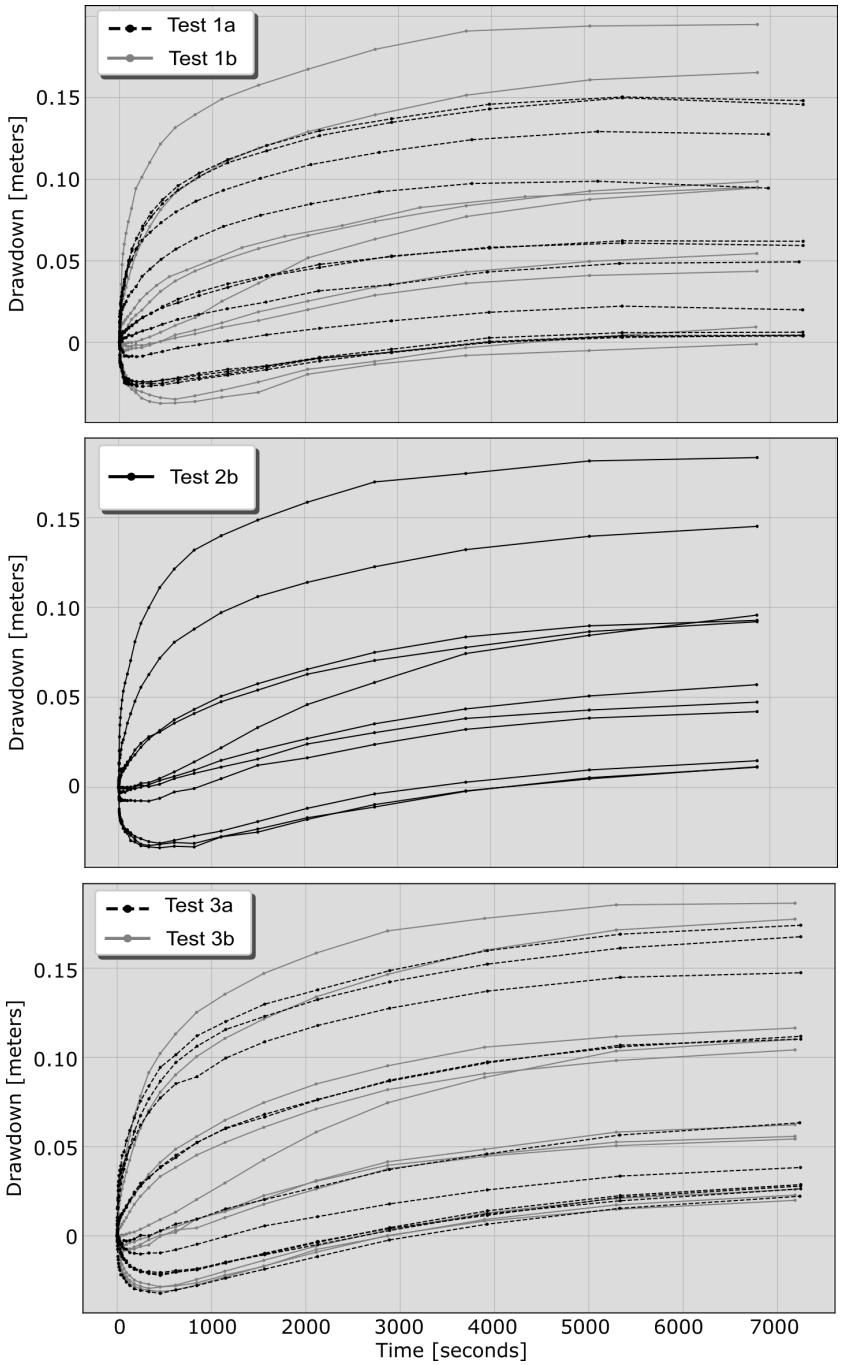


Figure 5.12: Drawdown curves corresponding to five out of the six individual tests of the tracer tomography. Tests 1a and 1b: tracer injection at the top section of well B3; Test 2b: tracer injection at the middle section of well B3; Tests 3a and 3b: tracer injection at the bottom section of well B3

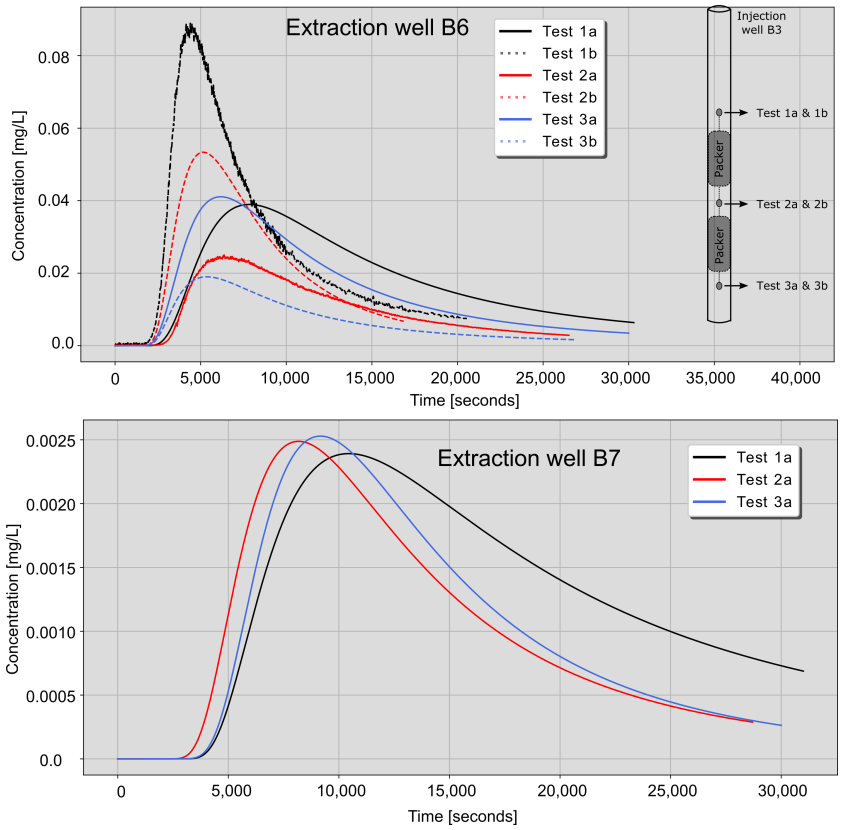


Figure 5.13: Breakthrough curves measured at extraction wells B6 (left) and B7 (right) during the six individual tests of the tracer tomography. Inset: corresponding injection section of each tracer test.

the higher mean natural hydraulic gradients observed during the second test series (Table 5.1). The slightly larger amount of tracer injected during test 1b is reflected in the high peak observed in well B6.

Table 5.1: Mean natural hydraulic gradient (I) based on manual measurements of groundwater levels. Measurements were obtained prior to start injecting and extracting water.

Test	Mean I %	Test	Mean I %
1a	0.34	1b	1.15
2a	0.17	2b	0.49
3a	0.20	3b	0.23

Fig. 5.14 presents the complete processed tracer dataset. The dataset consists of a total of 46 breakthrough curves of fluorescein measured during the six individual tracer tests of the

tracer tomography experiment. The breakthrough curves in Fig. 5.14 have been normalized by the injected mass. The three plots on the left correspond to the first test series, and the plots on the right are the breakthrough curves of the second test series. Each series consisted of three individual tests with tracer injection at the top, middle, and bottom sections of the multilevel injection system installed in well B3. The only curve still containing considerably large levels of noise was the breakthrough curve of well cmt1-1 from test 2b.

A visual inspection of all curves reveals that, independent of the tracer injection location, the tracer plume was distributed horizontally and vertically in the investigation area. Breakthrough curves were observed at the wells located at the corners of the observation well grid (e.g., in wells w4, w13 and w16) and also at different depths of the multilevel observation wells (e.g., in wells cmt1, cmt2 and cmt3).

In general, tracer concentrations decrease with a shift of the tracer injection towards the bottom of the aquifer (except for test 1a). The highest concentrations were measured at observation point cmt1-3. This observation point is located directly next to injection well B3, at a depth of $4\,\mathrm{m}$. The vertical location of this point coincides with the injection depth of the top section. High tracer concentrations were observed at this point also in tests 3a and 3b, indicating a large and fast vertical spread of the plume.

The tails of the breakthrough curves were accentuated in the tests with tracer injection at the lower sections of the aquifer. This might indicate a higher degree of heterogeneity and zones with lower hydraulic conductivity, which is in agreement with earlier descriptions of the aquifer at the site.

A quantitative assessment of the field curves was performed by fitting analytical solutions to the data, and the spatial distribution of hydraulic conductivity was estimated using this processed datatset with the EnKF.

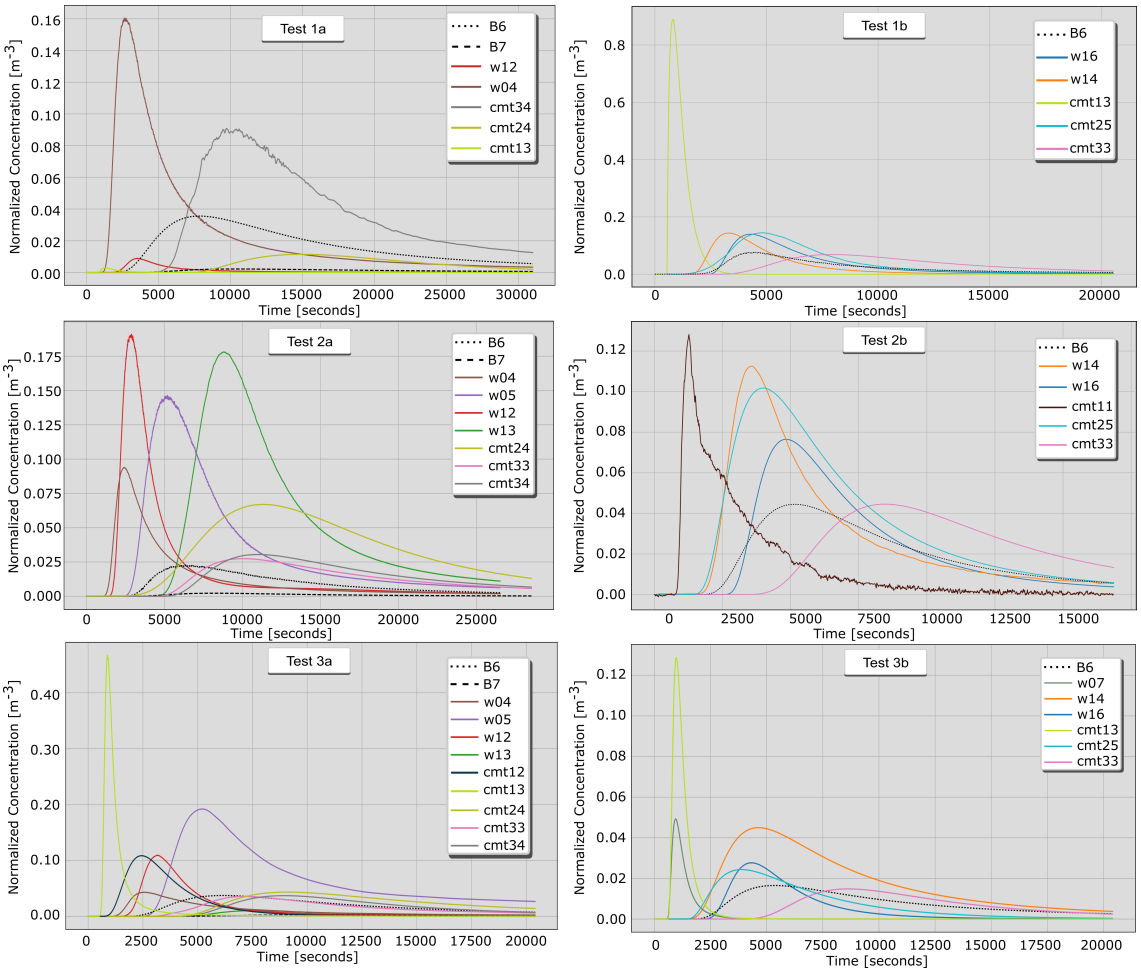


Figure 5.14: Full processed dataset of the tracer tomography test. Breakthrough curves are normalized by the injected mass. Left column: breakthrough curves corresponding to the first test series; Right: breakthrough curves of the second test series. Top row: tests 1a and 1b with tracer injection at the top section of well B3; Middle row: tests 2a and 2b with tracer injection at the middle section of well B3; Bottom row: tests 3a and 3b with tracer injection at the bottom section of well B3.

5.3 Drawdown and Breakthrough Curves Analysis

The processed dataset presented in the previous section was used for parameter estimation with the Ensemble Kalman Filter. To provide averaged hydraulic and transport parameter values, available analytical and semi-analytical solutions of Eqs. 2.1, 2.4 and 2.8 were fitted to the data. The method of temporal moments was used to analyze the breakthrough curves and also estimate effective aquifer parameters.

5.3.1 Drawdown Curves

To estimate effective hydraulic conductivity and storativity, the analytical solution to Eq. 2.1 after Theis (1935), here for multiple injection and extraction wells, was fitted to the processed drawdown curves.

$$s = \sum_{i=1}^{NW} \frac{Q_w^i}{4\pi T} W(U_i) \tag{5.14a}$$

$$U_i = \frac{r_i S}{4 T t_i} \tag{5.14b}$$

where s is the drawdown, NW is the number of wells, Q_w^i denotes the pumping rate of well i, with $Q_w^i > 0$ defining an infiltration well and $Q_w^i < 0$ an extraction well. $W(U_i)$ is the exponential integral function, known in hydrogeology as Theis' well function, T is the aquifer transmissivity, S is the storage coefficient, t is the elapsed time since the start of the test and r_i is the radial distance from the observation point to pumping well i.

The aquifer parameters T and S were optimized using the Nelder-Mead simplex algorithm (Nelder and Mead, 1965) as implemented in the *fminsearch* routine of Python. A constant aquifer thickness of $6\,\mathrm{m}$ was used to obtain hydraulic conductivity (K) and storativity (S_o) values from the depth averaged T and S.

As an example, results of the fitting applied to the hydraulic datasets measured during tests 1a and 3b are shown in Fig. 5.15. Table 5.2 summarizes the estimated values of hydraulic conductivity and storativity for all six tests.

The Theis solution, as implemented in this work (Eqs. 5.14a and 5.14b), is able to reproduce the hydraulic response to the injection and extraction of water, including the negative drawdown observed at the observation points closer to the injection wells (Fig. 5.15). The Theis model matches the steady-state values observed during the pumping tests (Fig. 5.15, left) but fails to reproduce early-time behavior (Fig. 5.15, right).

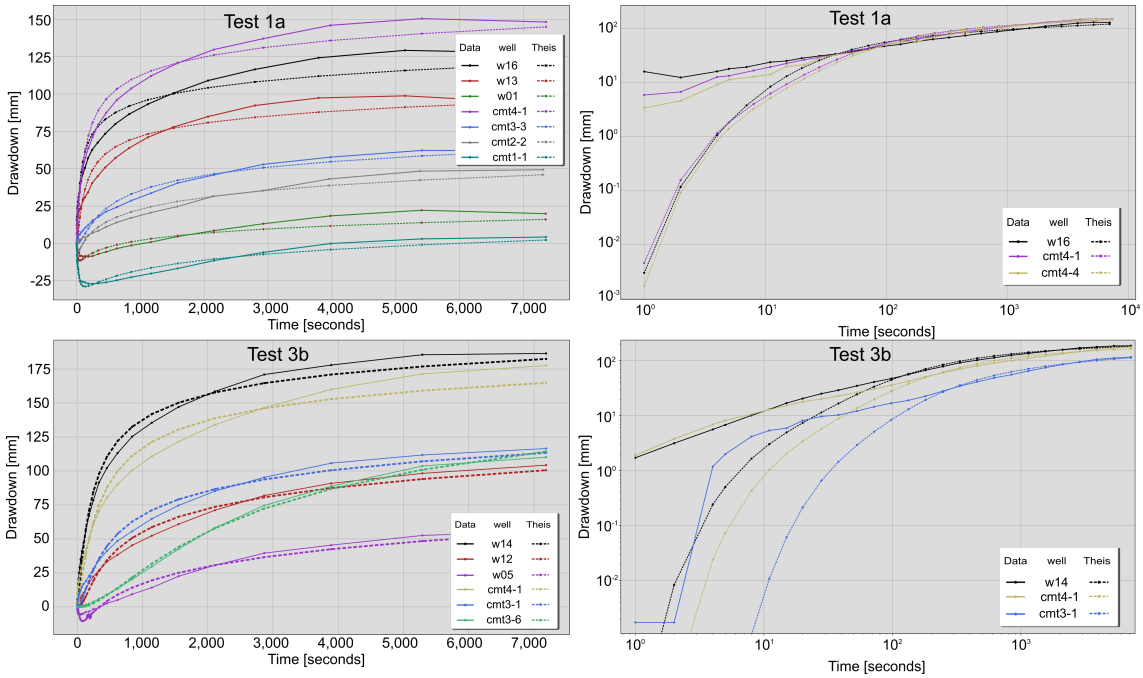


Figure 5.15: Type curve matching after Theis (1935) applied to selected drawdown curves of tests 1a and 3b. Left: time-drawdwon curves (solid lines) and their corresponding Theis curve (dashed line); Right: log-log plot of drawdown showing details of the large deviations of the fitted curves at early times.

Table 5.2: Mean (μ) and standard deviation (σ) of hydraulic conductivity (K) and storativity (S_o) values estimated with the Theis solution (Theis, 1935), for each test performed during the tomographic experiment. Global: μ and σ of hydraulic conductivity and storativity for the complete dataset.

	S_o (m^{-1})		K $(\mathrm{m\,s}^{-1})$	
Test	μ	σ	μ	σ
1a	7.3×10^{-3}	7.1×10^{-3}	3.9×10^{-3}	7.08×10^{-4}
1b	1.26×10^{-2}	1.40×10^{-2}	2.5×10^{-3}	7.97×10^{-4}
2a	–	–	–	–
2b	9.0×10^{-3}	1.29×10^{-2}	2.8×10^{-3}	1.375×10^{-3}
3a	5.8×10^{-3}	3.4×10^{-3}	4.1×10^{-3}	2.171×10^{-3}
3b	8.4×10^{-3}	9.7×10^{-3}	3.0×10^{-3}	2.469×10^{-3}
Global	8.66×10^{-3}	1.13×10^{-2}	3.2×10^{-3}	1.504×10^{-3}

Estimated effective hydraulic conductivities agree with previous research (Riva et al. 2006; Schneidewind 2008; Lessoff et al. 2010; Wagner et al. 2014), with a global mean value of

$3.2 \times 10^{-3}\,\mathrm{m\,s^{-1}}$. The observed drawdown is representative of the volume of the domain being investigated. This volume expands with time and therefore the hydraulic responses are increasingly averaged (Leven and Dietrich, 2006). This effect is evidenced in the similar slopes obtained at late times for all drawdown curves (Fig. 5.15, right) and small variations in the calculated hydraulic conductivities ($\sigma = 1.5 \times 10^{-3}\,\mathrm{m\,s^{-1}}$).

Estimated storativities have a global average value of $8.6 \times 10^{-3}\,\mathrm{m^{-1}}$ typical of confined systems. The large variation ($\sigma = 1.13 \times 10^{-2}\,\mathrm{m^{-1}}$) reflects the heterogeneity in hydraulic conductivity between pumping and observation wells (Meier et al., 1998), and is a result of the misfit of the Theis model to the early time data (Fig. 5.15, right), which contains the most information about the storage capacity of the aquifer. The inaccuracy in the estimation of aquifer storativity from pumping test data, is a direct consequence of the application of methods derived under the assumption of homogeneity (e.g., Theis solution) in heterogeneous systems (Meier et al. 1998; Leven and Dietrich 2006).

The effective hydraulic conductivity and storativity values obtained with type-curve matching were considered in the geostatistical model used to generate the random fields for the ensemble Kalman filter.

5.3.2 Breakthrough Curves: Analytical and Semi-analytical Solutions

Effective porosities and dispersivities were estimated by fitting a one-dimensional analytical solution of Eq. 2.4, to the breakthrough curves measured at extraction wells B6 and B7 (Sauty, 1980):

$$c(x,t) = \frac{M}{A n_{eff}\sqrt{4\pi\alpha_l v t}} \exp\left(-\frac{(x-vt)^2}{4\alpha_l v t}\right) \tag{5.15}$$

where c is the time- and location-dependent concentration, M is the injected tracer mass, n_{eff} is the effective porosity, α_l is the longitudinal dispersivity, A is the flow cross section, v refers to the average groundwater flow velocity, and x is the distance between the injection and the observation point.

After fitting the parameters α_l and n_{eff} with the same optimization algorithm used for the drawdown curves, mean arrival times (t_{mean}), average velocity (v), dispersion coefficient (D, here neglecting effective molecular diffusion) and hydraulic conductivity (K), were estimated according to:

$$t_{mean} = \frac{\pi x^2 b n_{eff}}{Q} \tag{5.16}$$

$$v = \frac{x}{t_{mean}} \tag{5.17}$$

$$D = \alpha_l v \tag{5.18}$$

$$K = \frac{v n_{eff}}{I} \tag{5.19}$$

where b denotes the aquifer thickness, Q is the extraction rate and I refers to the hydraulic gradient between injection and extraction points.

To assess the non-Fickian transport typically observed at the Hydrogeological Research Site Lauswiesen, the semi-analytical solution for one-dimensional transport derived by Neville et al. (2000) was fitted to the same breakthrough curves. This solution can represent one-dimensional advection and dispersion processes, combined with a dual-porosity domain and a first-order mass transfer reaction between the two regions. They derived the solution using the Laplace transform technique and obtained the final result via numerical inversion of the solution in Laplace space (see Neville et al. 2000 for the derivation). The Nelder-Mead simplex algorithm (Nelder and Mead, 1965) was used to fit the semi-analytical solution to the data. The parameters adjusted were the first-order mass transfer coefficient (λ_{mt}), total porosity (n_{tot}), hydrodynamic dispersion coefficient (D) and the proportion of mobile pore water (ϕ).

Figure 5.16 compares the field breakthrough curves (red lines) measured at well B6 during all six tests, with the 1-D analytical solution (blue dashed lines) and semi-analytical solution for the dual-domain version of the advection dispersion equation (grey dashed lines).

In all tests, the 1-D analytical solution shows an acceptable agreement with the measured breakthrough curve. However, it fails to capture the long tailing, and the estimated porosity below $1\,\%$ (see Table 5.3) is unrealistically low, indicating unresolved small-scale variability of aquifer heterogeneity (Sanchez-León et al., 2016).

Table 5.3 summarizes the transport parameters estimated with the 1-D analytical solution for the extraction well B6. The average longitudinal dispersivity of $2.75\,\mathrm{m}$ is considerably larger than previously reported values ($\alpha_l \approx 1.0\,\mathrm{m}$, e.g., Riva et al. 2008; Händel and Dietrich 2012; Wagner et al. 2014). This is most likely a consequence of the Fickian approach of the dispersion term in the classical advection-dispersion equation, which does not consider all transport processes caused by heterogenity (Händel and Dietrich, 2012). Overestimation of dispersivity is related to the systematic underestimation of the breakthrough peak observed in Fig. 5.16. The estimated t_{mean}, v, D and K are directly affected by α_l and n_{eff} yielding for example low values of K.

Table 5.4 shows the parameters estimated with the dual-domain semi-analytical solution. The solution acceptably fits the breakthrough curves of tests 1a, 1b and 3a, with porosity values ranging between $37\,\%$ to $45\,\%$ (see Table 5.4). The estimated values of λ_{mt}, n_{tot}, D

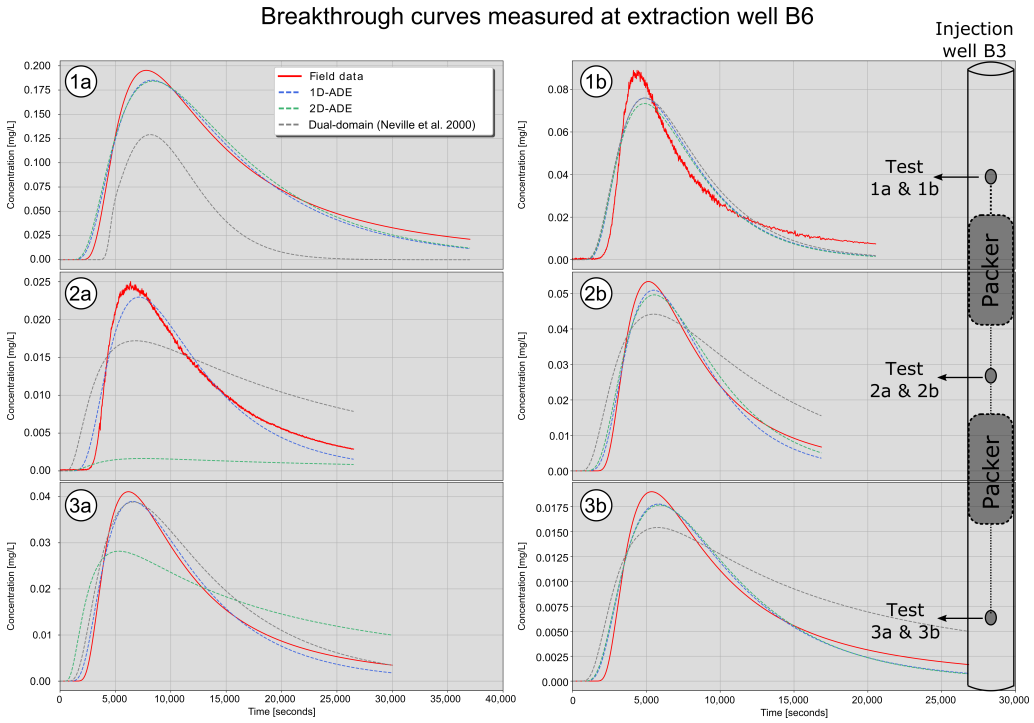


Figure 5.16: Breakthrough curves measured at extraction well B6 during the six tracer tests of the tracer tomography. Red solid lines: processed breakthrough curves. Blue dashed lines: 1D-analytical solution (Eq. 5.15); Gray dashed lines: semi-analytical solution of the dual-domain version of the advection dispersion equation (Neville et al., 2000). Left plots: first series of tests with injection at the top (test 1a), middle (test 2a) and bottom (test 3a) sections of injection well B3; Right plots: second series of tests (tests 1b, 2b and 3b). For clarity, a schematic representation of the injection well B3 is superimposed at the right side of the plots, showing the corresponding injection section for each test.

and ϕ for these two tests were further considered in the parameter estimation with the EnKF. In the other tests, the solution did not fit the data and therefore the estimated parameters are unreliable (Table 5.4 in red). The fact that it was not possible to fit the semi-analytical solution to the data may be due to the applied optimization scheme. The Nelder-Mead simplex algorithm is unstable in both high-dimensional problems and strongly non-linear functions, and therefore finding a global minimum value for the objective function was not guaranteed. In such cases, the solution strongly depends on the initial values of the fitting parameters.

As mentioned before, technical difficulties affected the measurements in well B7 during tests 1b, 2b and 3b. Therefore only breakthrough curves of tests 1a, 2a and 3a could be transformed into concentration. Fig. 5.17 compares these three field breakthrough curves with the corresponding fitting solutions, and Table 5.5 summarizes the parameter values estimated

Table 5.3: Transport parameters estimated for all tests and for extraction well B6, with the 1-D analytical solution (Eq. 5.15). α_l: longitudinal dispersivity; n_{eff}: effective porosity; t_{mean}: mean arrival time; v: seepage velocity; D: dispersion coefficient; K: hydraulic conductivity; μ: mean; σ: standard deviation.

	Well B6: 1-D analytical solution					
Test	α_l (m)	n_{eff} (%)	t_{mean} (s)	v ($\mathrm{m\,s^{-1}}$)	D ($\mathrm{m\,s^{-1}}$)	K ($\mathrm{m\,s^{-1}}$)
1a	3.41	0.65	15223	1.04×10^{-3}	3.56×10^{-3}	6.8×10^{-4}
1b	2.26	0.37	7415	2.14×10^{-3}	4.86×10^{-3}	7.9×10^{-4}
2a	2.66	0.60	11568	1.37×10^{-3}	3.66×10^{-3}	8.2×10^{-4}
2b	1.94	0.41	7876	2.01×10^{-3}	3.92×10^{-3}	8.2×10^{-4}
3a	3.1	1.5	11582	1.37×10^{-3}	4.36×10^{-3}	8.2×10^{-4}
3b	3.17	0.430	10261	1.54×10^{-3}	4.91×10^{-3}	6.6×10^{-4}
μ	2.75	0.66	10654	1.5×10^{-3}	4.2×10^{-3}	7.7×10^{-4}
σ	5.6×10^{-1}	0.432	2861	4.2×10^{-4}	5.9×10^{-4}	7.8×10^{-5}

Table 5.4: Transport parameters estimated for all tests and for extraction well B6, with the dual-domain semi-analytical solution (Neville et al., 2000). λ_{mt}: first-order mass transfer coefficeint; n_{tot}: total porosity; ϕ: proportion of mobile porosity with respect to total porosity; D: dispersion coefficient; μ: mean; σ: standard deviation. Cells in red: unreliable results.

	Well B6: Dual-domain semi-analytical solution			
Test	λ ($\mathrm{s^{-1}}$)	n_{tot} (%)	ϕ (%)	D ($\mathrm{m\,s^{-1}}$)
1a	4.20×10^{-12}	35.7	0.43	8.28×10^{-1}
1b	6.20×10^{-11}	44.9	78.0	6.17×10^{-3}
2a	4.45×10^{-12}	42.8	76.3	1.16×10^{-2}
2b	1.56×10^{-11}	37.8	81.6	9.53×10^{-3}
3a	1.39×10^{-11}	44.5	79.6	6.82×10^{-3}
3b	7.12×10^{-16}	12.3	74.1	7.31×10^{-3}
μ	2.67×10^{-11}	41.7	52.7	2.80×10^{-1}
σ	2.41×10^{-11}	4.2	35.1	3.67×10^{-1}

with the 1-D analytical solution. To be able to fit the 1-D analytical solution, the breakthrough curves of tests 1a and 2a were extended using the generalized inverse Gaussian distribution. It was not possible to properly fit the dual-domain semi-analytical solution to these breakthrough curves (Fig. 5.17), and therefore the optimized values are not reported.

Table 5.5: Transport parameters estimated for tests 1a (top), 2a (middle), and 3a (bottom) for extraction well B7, with the 1D analytical solution (Eq. 5.15).α_l: Longitudinal dispersivity; n_{eff}: effective porosity; t_{mean}: mean arrival time; v: seepage velocity; D: dispersion coefficient; K: hydraulic conductivity; μ: mean; σ: standard deviation.

	Well B7: 1-D analytical solution					
Test	α_l (m)	n_{eff} (%)	t_{mean} (s)	v ($\mathrm{m\,s^{-1}}$)	D ($\mathrm{m\,s^{-1}}$)	K ($\mathrm{m\,s^{-1}}$)
1a	5.13	2.4	22358	9.3×10^{-4}	4.79×10^{-3}	2.98×10^{-3}
2a	3.30	1.50	13760	1.51×10^{-3}	5.01×10^{-3}	3.18×10^{-3}
3a	2.36	1.48	13597	3.63×10^{-3}	3.63×10^{-3}	3.18×10^{-3}
μ	3.60	1.81	16572	1.3×10^{-3}	4.4×10^{-3}	3.11×10^{-3}
σ	1.40	0.55	5011	3.4×10^{-4}	7.4×10^{-4}	1.17×10^{-4}

The 1-D analytical solution has the same behavior as for well B6. The mean porosity of $1.8\,\%$ is one order of magnitude larger than in well B6, however it is still unrealistic. A direct consequence of larger porosities, is a mean value of hydraulic conductivity of $3.11 \times 10^{-3}\,\mathrm{m\,s^{-1}}$ that agrees with the drawdown data analysis. The mean longitudinal dispersivity estimated for well B7 is larger than the value obtained for well B6.

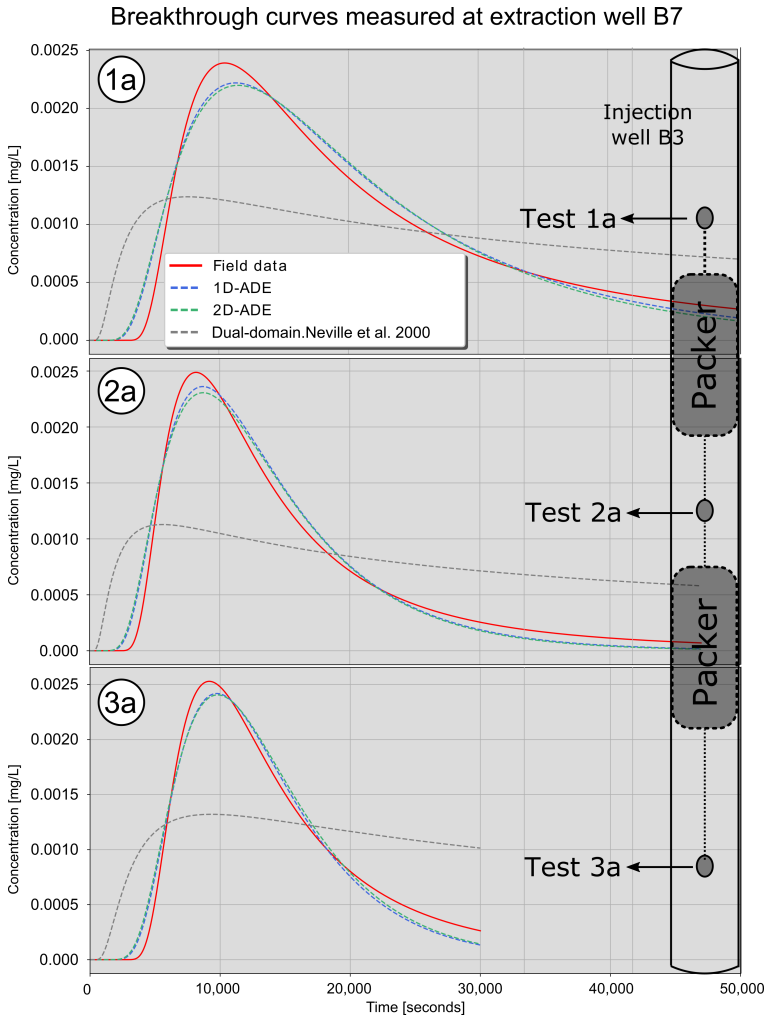


Figure 5.17: Breakthrough curves measured at extraction well B7 during the first series of tracer tests (top injection: test 1a, middle injection: test 2a, bottom injection: test 3a). Red solid lines: processed breakthrough curves. Blue dashed lines: 1-D analytical solution (Eq. 5.15); Gray dashed lines: semi-analytical solution of the dual-domain version of the advection dispersion equation (Neville et al., 2000). For clarity, a schematic representation of the injection well B3 is superimposed at the right side of the plots, showing the corresponding injection section for each test.

5.3.3 Breakthrough Curves: Temporal Moments

The method of temporal moments is another approach that can be used for the estimation of effective transport parameters. Applying this method to the breakthrough curves measured at the extraction wells, the total recovered tracer mass can be estimated, helping to assess the quality of the tracer tests.

For a breakthrough curve of a conservative tracer measured at location x, the *i-th* temporal (μ_i) and central (μ_i^c) moments are defined as:

$$\mu_i = \int_{t=0}^{\infty} t^i c(x,t)\,dt \tag{5.20}$$

$$\mu_i^{ce} = \int_{t=0}^{\infty} (t-\bar{t})^i c(x,t)\,dt \tag{5.21}$$

where t is time, c is the tracer concentration, and $\bar{t}$ is the mean breakthrough time and can be calculated by normalizing the first with the 0^{th} moment:

$$\bar{t} = \frac{\mu_1}{\mu_0} \tag{5.22}$$

with the 0^{th} temporal moment μ_0 representing the effective tracer mass (mass per discharge). Apparent dispersion coefficients can be estimated via the variance of the travel time distribution (σ_{tt}), defined by the second central moment (μ_2^{ce}) normalized with the zeroth moment (μ_0):

$$\sigma_{tt} = \frac{\mu_2^{ce}}{\mu_0} = \frac{2Dx}{v^3} \tag{5.23}$$

in which D is the effective dispersion coefficient, x denotes distance from the injection point, and $v = \frac{x}{\bar{t}}$ is the mean flow velocity. The temporal moments defined by Eqs. 5.20 and 5.21 assume complete breakthrough curves. However, truncated records are usually retrieved from real tracer tests, specially in a medium with strong non-Fickian transport behavior.

To account for the missing data, Luo et al. (2006a) derived a set of equations that relate the temporal moments of complete breakthrough curves to the moments from the truncated dataset. For a Dirac delta input, the zeroth, first and second truncated moments are defined as:

$$\mu_0^{tr} = \sum_{i=1}^{N} c(T - it_0) \tag{5.24}$$

$$\mu_1^{tr} = T\mu_0^{tr} - \sum_{i=1}^{N} \mu_0 (T - it_0) \tag{5.25}$$

$$\mu_2^{tr} = T\mu_0^{tr} - 2\sum_{i=1}^{N}(\mu_1(T - it_0) + it_0\mu_0(T - it_0)) \tag{5.26}$$

where c is the observed tracer concentration, t_0 is the tracer injection time, N is the largest integer value smaller than or equal to $\frac{t}{t_0}$, T is the truncation time, and μ_0 and μ_1 are the complete zeroth and first temporal moments (see Eq. 5.20). The second-central truncated moment can then be defined by Eq. 5.27:

$$m_2^{ce,tr} = \frac{\mu_2^{tr}}{\mu_0^{tr}} \tag{5.27}$$

and the normalized first-central and second-central truncated moments, corresponding to the mean ($\bar{t}$) and variance (σ_{tt}) of travel time are computed as:

$$\bar{t} = \frac{m_1^{ce,tr}}{\mu_0^{tr}} \tag{5.28}$$

$$\sigma_{tt} = \frac{m_2^{ce,tr}}{\mu_0^{tr}} = \frac{2Dx}{v^3} \tag{5.29}$$

Table 5.6 summarizes the temporal moment analysis (Eqs. 5.20 - 5.23) of the breakthrough curves recorded at the extraction wells during the six tests of the tracer tomography experiment. In the first test series (tests 1a, 2a and 3a), the shorter travel times correspond to the test with tracer injection at the middle section (2a). Although a dataset of hydraulic heads could not be retrieved for this test, higher mean velocities can be explained by the higher hydraulic gradients expected due to the larger injection and extraction rates applied (see Table 4.1).

In the second series (tests 1b, 2b and 3b) the pumping rates were more stable, and therefore similar hydraulic gradients were achieved in the three tests. Mean travel times calculated for extraction well B6 increase with depth, and might indicate the presence of less conductive material at lower depths of the aquifer. The longitudinal dispersivities and dispersion coefficients are similar to those estimated with the 1-D analytical solution, ranging between $1.5\,\mathrm{m}$ to $4.2\,\mathrm{m}$ and $2.7 \times 10^{-3}\,\mathrm{m^2\,s^{-1}}$ to $5 \times 10^{-3}\,\mathrm{m^2\,s^{-1}}$, respectively.

With the temporal moment analysis it was possible to estimate the mass of tracer recovered in the extraction wells. This estimation is an indication of the reliability of the recorded tracer data. The closer the value is to the injected mass, the more reliable are the conclusions based on the collected dataset. This value however, is highly affected by all the processing steps required to transform the raw and noisy breakthrough curve signals into a smooth time series in units of concentration.

Overestimation of the recovered mass during tests 1a, 1b, and 3a (see Table 5.6) is attributed to complications during water sampling, uncertainties in the laboratory measurements and errors during scaling the data to units of concentration, where a perfect fit between the breakthrough curves and the laboratory measurements could not be achieved. These results suggest that the sampling system should be further developed to improve the quality of the collected samples.

Tests 2a and 2b have acceptable estimations of recovered mass of $87\,\%$ and $62\,\%$, respectively. Tracer test 3b shows the lowest percentage of recovered mass ($32\,\%$), indicating that the majority of the tracer was lost into the system. Disturbances of the hydraulic field were not observed, and the reason for the low recovered mass could not be identified. A larger measurement uncertainty was assigned to this dataset when used for parameter estimation.

Table 5.6: Temporal moment analysis of the breakthrough curves measured at the extraction wells B6 and B7 during the six tests of the tracer tomography. x: distance to injection well B3; μ_0: o^{th} moment; μ_1: 1st moment; $\bar{t}$: mean travel time; μ_2^{ce}: 2nd central moment; v: mean velocity; D: dispersion coefficient; α_l: longitudinal dispersivity.

Test	Well	x (m)	Inj. Mass (g)	μ_0 ($\mathrm{kg\,m^{-3}\,s}$)	μ_1 ($\mathrm{kg\,m^{-3}\,s^2}$)	$\bar{t}$ (s)	μ_2^{ce} ($\mathrm{kg\,m^{-3}\,s^3}$)	v ($\mathrm{m\,s^{-1}}$)	D ($\mathrm{m^2\,s^{-1}}$)	α_l (m)	Recovered Mass (g)
1a	B6	15.9	1.09	0.5	8.63×10^3	14685	4.0×10^7	1.0×10^{-3}	2.7×10^{-3}	2.52	1.00
	B7	20.9	–	0.03	1.10×10^3	15939	4.4×10^7	9.5×10^{-4}	3.9×10^{-3}	4.16	0.3(118%)
1b	B6	15.9	1.15	0.53	3.99×10^3	7493	7.7×10^6	2.1×10^{-3}	4.3×10^{-3}	2.06	1.28
	B7	–	–	–	–	–	–	–	–	–	–(111%)
2a	B6	15.9	1.09	0.26	2.92×10^3	11032	7.7×10^6	1.4×10^{-3}	2.7×10^{-3}	1.91	0.66
	B7	20.9	–	0.03	4.72×10^2	14650	2.3×10^6	1.4×10^{-3}	5.1×10^{-3}	3.59	0.29(87%)
2b	B6	15.9	1.22	0.37	2.94×10^3	7753	4.3×10^6	2.0×10^{-3}	3.0×10^{-3}	1.50	0.76
	B7	–	–	–	–	–	–	–	–	–	–(62%)
3a	B6	15.9	1.10	0.46	5.23×10^3	11340	1.8×10^7	1.4×10^{-3}	3.4×10^{-3}	2.43	1.10
	B7	20.9	–	0.03	4.13×10^2	13559	1.0×10^6	1.5×10^{-3}	2.9×10^{-3}	1.93	0.27(124%)
3b	B6	15.9	1.15	0.18	1.85×10^3	10035	5.6×10^6	1.5×10^{-3}	3.8×10^{-3}	2.43	0.37
	B7	–	–	–	–	–	–	–	–	–	–(32%)

Fig. 5.18 shows in a plan view the mean travel time distribution (in hours) of the tracer plume calculated for tests 1b, 2a and 3a. It can be observed that the tracer plume travels faster when the tracer is injected in the upper section of the injection well, and slows down while the tracer injection is shifted towards the bottom of the aquifer. It is also observed that while the tracer spreads over the entire investigation area, it travels faster at the sides than in the center of the domain. This contradicts the general idea that the tracer would travel faster along the main line defined by the two injection and extraction wells. The rapid spread of the plume may be caused by aquifer heterogeneity and enhanced by the high gradients generated with the nested-cell flow setup.

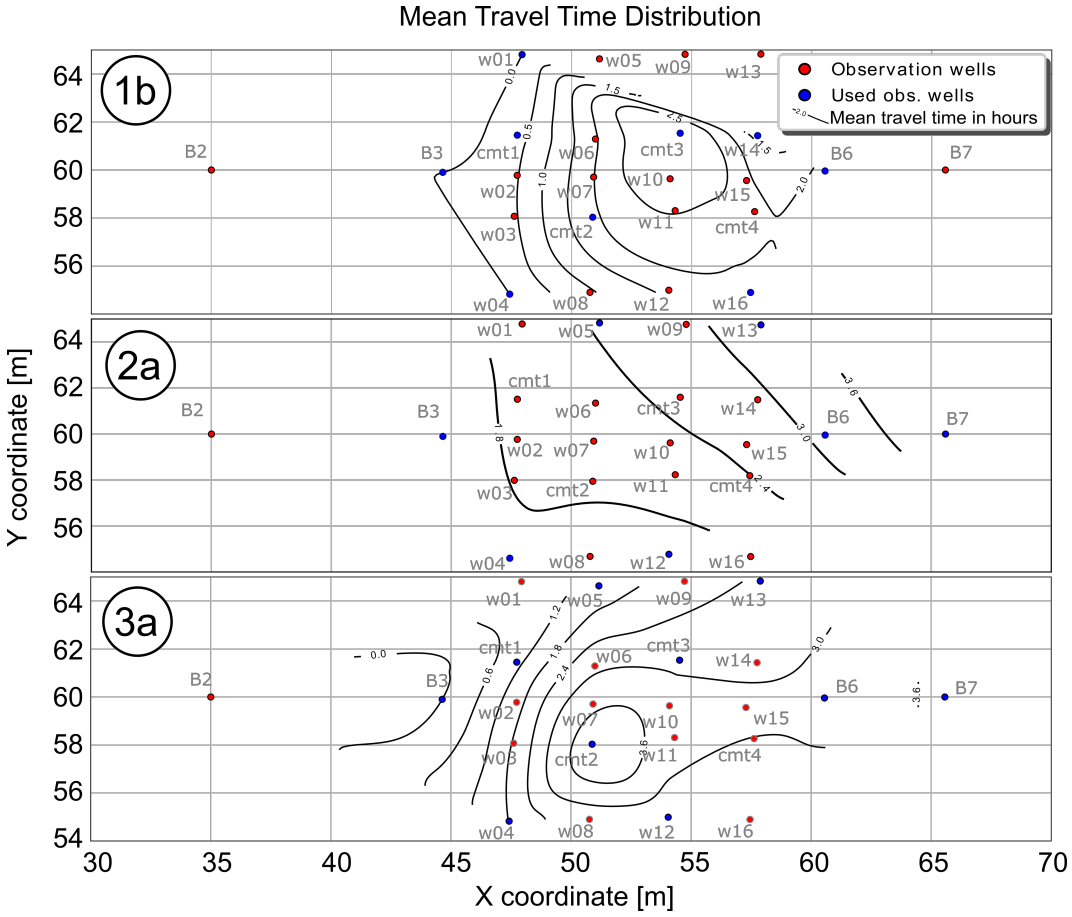


Figure 5.18: Mean travel time distribution (in hours) of the tracer plume, for tracer tests 1b (top), 2a (middle) and 3a (bottom). Estimation performed via temporal moments applied to the processed breakthrough curves without tail extension.

For completeness, a temporal moment analysis was also performed using the breakthrough curves extended with the exponential and generalized inverse Gaussian distribution functions (see subsection 5.1.6). Additionally, temporal moments were also estimated using the equa-

tions derived by Luo et al. (2006a) for truncated datasets. Estimated recovered mass and mean travel times for each breakthrough curve are compared in Tables 5.7 and 5.8, respectively.

Table 5.7: Comparison of the tracer recovered mass calculated with the original breakthrough curve and (*i*) temporal moment equations assuming a complete dataset, (*ii*) with truncated temporal moment equations, (*iii*) an extension of the breakthrough curve with the exponential distribution, and (*iv*) an extension of the breakthrough curve with the generalized inverse Gaussian distribution.

					Recovered Mass			
Test	Well	Inj. Mass (g)	Original t_{end} (s)	Extended t_{end} (s)	No extension (g)	Truncated (g)	Exp. tail (g)	GIG tail (g)
1a	B6	1.096	36999	40000	1.00	1.20	1.20	1.20
	B7	–	69999	70000	0.4	0.45	0.45	0.45
1b	B6	1.154	20548	25000	1.28	1.28	1.30	1.30
	B7	–	–	–	–	–	–	–
2a	B6	1.096	26507	40000	0.66	0.66	0.68	0.70
	B7	–	47000	47000	0.29	0.29	0.29	0.29
2b	B6	1.223	16858	25000	0.76	0.75	0.79	0.82
	B7	–	–	–	–	–	–	–
3a	B6	1.107	29999	40000	1.10	1.15	1.20	1.21
	B7	–	29999	40000	0.27	0.27	0.28	0.29
3b	B6	1.153	26824	35000	0.37	0.36	0.38	0.39
	B7	–	–	–	–	–	–	–

As expected, considering truncated temporal moments or tail extension increases the estimated values of mean travel times and mass recovered. This might be seen as an improvement if the estimations of recovered mass are below but close to the injected tracer mass (e.g. tests 2a and 2b), but even larger errors can be introduced if the recovered mass is already overestimated.

In general, mean travel times estimated with the truncated temporal moment equations do not vary from the original values. The effect of extending the breakthrough tail could produce a shift of more than $1000\,\mathrm{s}$.

Table 5.8: Comparison of the mean travel time calculated with the original breakthrough curve and (*i*) temporal moment equations assuming a complete dataset, (*ii*) with truncated temporal moment equations, (*iii*) an extension of the breakthrough curve with the exponential distribution, and (*iv*) an extension of the breakthrough curve with the generalized inverse Gaussian distribution.

				Mean travel time			
Test	Well	Original t_{end} (s)	Extended t_{end} (s)	No extension (s)	Truncated (s)	Exp. tail (s)	GIG tail (s)
1a	B6	36999	40000	14685	14684	14685	14685
	B7	69999	70000	21839	21838	21839	21839
1b	B6	20548	25000	7493	7492	7769	7811
	B7	–	–	–	–	–	–
2a	B6	26507	40000	11032	11031	14650	14650
	B7	47000	47000	14650	14649	14650	14650
2b	B6	16858	25000	7753	7752	8309	8715
	B7	–	–	–	–	–	–
3a	B6	29999	40000	11340	11339	12173	12450
	B7	29999	40000	13559	13558	14321	14657
3b	B6	26824	35000	10035	10034	10745	11329
	B7	–	–	–	–	–	–

Fig. 5.19 shows as an example the two processed breakthrough curves recorded during test 3b at extraction well B6 and observation well w14. Fig. 5.19 also shows the differences between mean travel times of the truncated temporal moment equations (green dashed lines) and the original temporal moments equations applied to the original dataset (red solid lines), the curve extended with an exponential distribution (magenta dashed lines) and the curve extended with the generalized inverse Gaussian distribution (blue dashed lines).

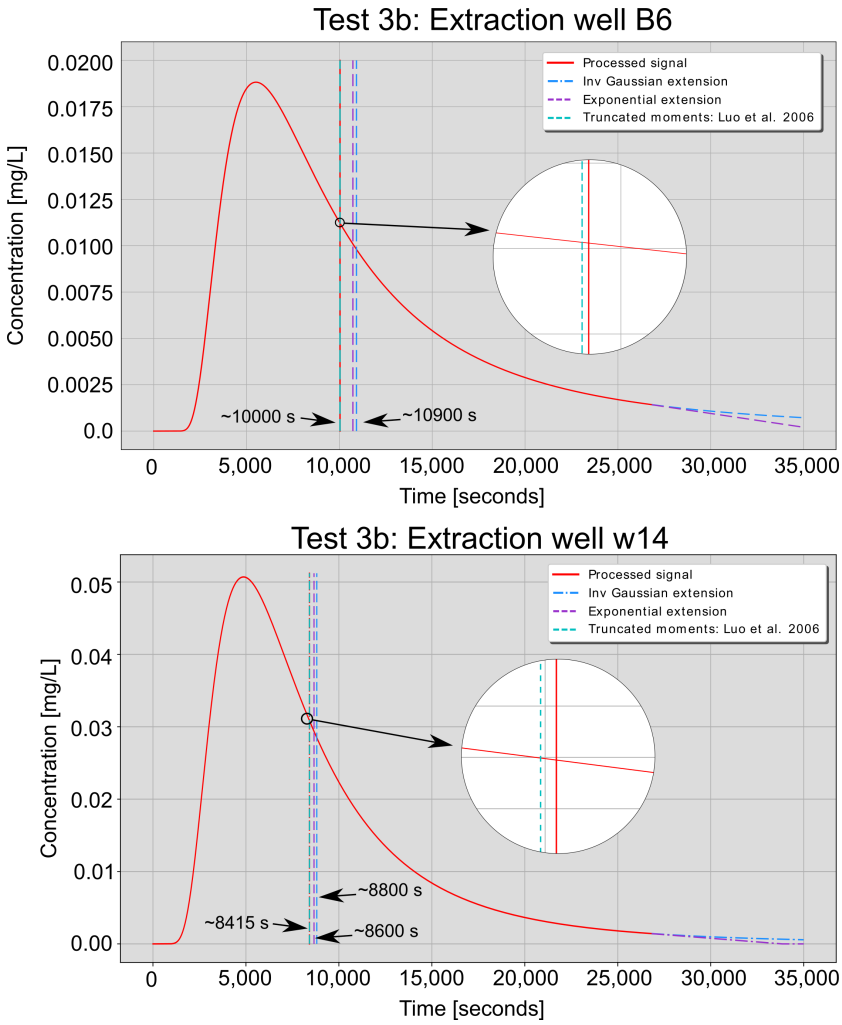


Figure 5.19: Original and extended breakthrough curves and different mean travel times estimated for each extension method. Circles with white background show details of the mean travel time variations according to the different extension models. The breakthrough curves were measured at wells B6 and w14 during test 3b (injection at the mid section).

Chapter 6

Parameter Estimation: Synthetic Case

The performance of the EnKF applied to tracer tomography data was evaluated with a synthetic test case. This process implied running the filter many times, and to alleviate the computational burden, the synthetic case consisted of a two-dimensional flow and transport model. The optimal filter setup was then implemented in the estimation of parameters using the real data and a fully three-dimensional model.

The two-dimensional model was used to simulate the tracer tomography test and produce synthetic hydraulic and tracer data. The model was based on the hydrogeological setup of the aquifer, and the experimental design of the real tests performed at the Hydrogeological Research Site Lauswiesen (see Figs. 4.1-4.3).

This chapter describes the numerical simulations, and presents and discusses the results of the parameter estimation using the EnKF with synthetic data. All HydroGeoSphere models and synthetic data are provided as an electronic appendix (Appendix: Models/synthetic).

6.1 Numerical Simulations

The 2-D numerical model was constructed using the pde-based hydrological model HydroGeoSphere (Aquanty, Waterloo, ON, Canada, Therrien and Sudicky 1996). HydroGeoSphere is capable of solving the groundwater flow equation (Eq. 2.1) and the dual-domain version of the advection-dispersion equation (Eq. 2.8). Early-stage simulations with this model were used to assess important factors for a proper design of field-scale tracer tests such as transport scale, boundary conditions, flow regime, test duration, types of tracer, types of observation points, monitoring network distribution and methods for data analysis.

The model grid consists of $161\,202$ nodes forming $80\,000$ regular block elements ($\Delta x = \Delta y = \Delta z = 0.25\,\mathrm{m}$) and covering an area of $100{\times}50\ \mathrm{m}^2$ (Fig. 6.1). A nested-cell flow field was simulated with two injection and two extraction wells. A total of 25 observation wells were distributed in the area between the inner injection and extraction points (Fig. 6.1).

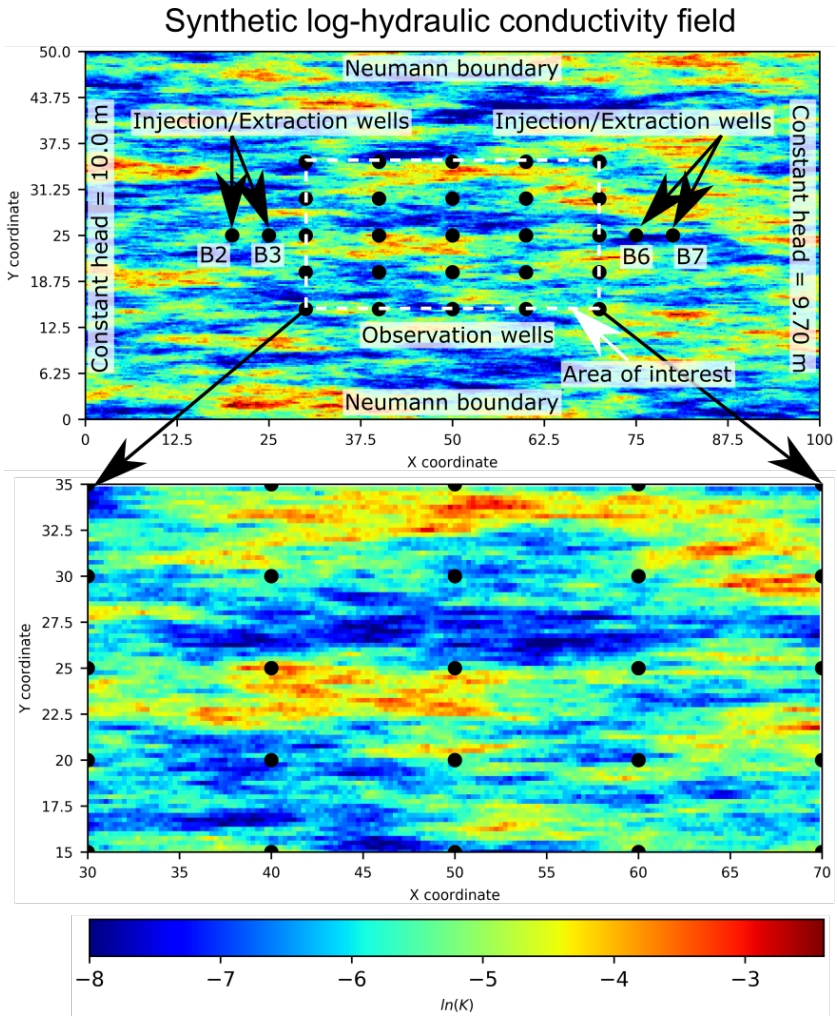


Figure 6.1: Two-dimensional model used to generate hydraulic and tracer data from a synthetic tracer tomography experiment. Black dots: injection, extraction and observation wells. White dotted line: area of interest for parameter estimation. Neumann boundaries (no-flux) at the top and bottom limits. Dirichlet boundaries (constant head) at the left and right limits.

The observation wells are consecutively numbered between 1 and 25, starting from left to right and bottom to top. To emphasize the similarities between the 2-D model and the available well facilities at Lauswiesen, the injection and extraction wells were also labeled B2, B3, B6 and B7. Several injection depths could not be defined in the 2-D model. Instead, the

tomographic layout was achieved by inverting the injection and extraction setup (inverted flow field) and changing the location of the tracer injection from well B3 to well B6.

Hydraulic and transport parameters were based on previous field investigations (see Section 4.1) and the analysis of the real tracer tomography dataset performed in this work (Section 5.3). To account for heterogeneity, fields of log-hydraulic conductivity (log-K) were generated by spectral methods (Dietrich and Newsam, 1993) and a geostatistical model based on parameter values within the ranges obtained by Lessoff et al. (2010) for Lauswiesen (exponential model, mean hydraulic conductivity $\mu_k = 3.0 \times 10^{-3}\,\mathrm{m\,s^{-1}}$, log-K variance $\sigma^2_{Ln(k)} = 1$, and longitudinal and transverse correlation lengths $l_l = 8\,\mathrm{m}$ and $l_t = 1\,\mathrm{m}$, respectively) (see Fig. 6.1). Storativity (S_o), total porosity (n_{tot}) and longitudinal dispersivity (α_l) were assumed homogeneous and equal to $3.5 \times 10^{-3}\,\mathrm{m^{-1}}$, $30\,\%$ and $1\,\mathrm{m}$, respectively. The dual-domain transport model was implemented with a mass transfer coefficient $\lambda_{mt} = 1 \times 10^{-9}\,\mathrm{s^{-1}}$, immobile porosity value of $n_{im} = 20\,\%$, and a porosity value for the mobile zone $n_m = 10\,\%$ (values listed in Table 6.1). Constant hydraulic heads of $10\,\mathrm{m}$ and $9.7\,\mathrm{m}$ were defined at the left and right limits of the domain, respectively, generating an ambient hydraulic gradient of $0.3\,\%$. Neumann boundaries (no-flux) were assigned to the upper and lower limits and an initial concentration of zero at the beginning of the simulations was defined all-over the model domain (Fig. 6.1).

Table 6.1: Parameters and aquifer properties used to generate the reference 2-D-model. μ_k: mean hydraulic conductivity; $\sigma^2_{Ln(k)}$: log-hydraulic conductivity variance; l_l: longitudinal correlation length; l_t: transverse correlation length; S_o: storativity; n_{im}: immobile porosity ; n_m: mobile porosity; n_{tot}: total porosity; α_l: longitudinal dispersivity; λ_{mt}: mass transfer coefficient between mobile and immobile zones.

Spatially distributed hydraulic conductivity					
μ_k ($\mathrm{m\,s^{-1}}$)	$\sigma^2_{Ln(k)}$	l_l (m)	l_t (m)		
3.0×10^{-3}	1	8	1		
Homogeneous parameters					
S_o ($\mathrm{m^{-1}}$)	n_{im} (%)	n_m (%)	n_{tot} (%)	α_l (m)	λ_{mt} ($\mathrm{s^{-1}}$)
3.5×10^{-3}	20	10	30	1	1×10^{-9}

Flow and transport simulations were performed sequentially. First, transient groundwater flow was simulated until steady flow was achieved. For transport, an initial steady-sate flow simulation was performed to obtain the groundwater flow velocity distribution, followed by the injection of tracer at the selected location (i.e. well B3 or B6). Table 6.2 contains the injection/extraction rates and the mass of tracer injected during each test of the synthetic

tracer tomography test. Test labels *1syn* and *2syn* refer to the full sequence of flow and transport simulations, whereas each individual simulation is further identified by an additional *Fl* and *Tr* for the corresponding flow or transport simulation, respectively.

Table 6.2: Settings of the groundwater flow and transport simulations used for the synthetic study. Negative flow rates denote extraction of water, whereas positive numbers refer to the injection of water.

	Injection	Tracer	Applied Flow Rates ($\mathrm{l\,s^{-1}}$)			
Test	well	Mass(g)	B2	B3	B6	B7
1syn	B3	100	9.0	6.0	-7.0	-10
2syn	B6	100	-10	-7.0	6.0	9.0

The synthetic hydraulic dataset consists of two sets of 25 drawdown curves recorded at 27 time steps (Fig. 6.2 top-row). Each set corresponds to one of the two different hydraulic tests simulated. In real experiments, the hydraulic response observed directly at the injection and extraction wells is affected by the injection and extraction pumps, so that these sets were excluded in both the real and synthetic datasets. The synthetic tracer data were composed of two sets of 25 breakthrough curves measured at the observation wells (Fig. 6.2 middle-row) and two additional curves recorded at the extraction wells. Each set of breakthrough curves corresponds to one of the two different tracer tests contributing to the synthetic tracer-tomography experiment. The tracer information collected at the extraction wells was used to estimate tracer recovery rates (Fig. 6.2 bottom-row).

The drawdown and breakthrough curves of Fig. 6.2 are grouped and colored according to the location (column) of the corresponding observation point. A maximum drawdown of $\sim 1\,\mathrm{m}$ was simulated in test 1syn-Fl, whereas in test 2syn-Fl a maximum drawdown of $\sim 1.5\,\mathrm{m}$ was computed. The stronger drawdowns during the second test are caused by the natural hydraulic gradient and the location downstream of the injection wells (wells B6 and B7). Negative drawdown was observed in locations close to the injection wells, indicating a rise of water levels due to a strong influence of water injection. Steady state was achieved faster in test 1syn-Fl, were water was injected in wells B2 and B3, , than in tets 2syn-Fl. The larger times needed to achieve steady state during test 2syn-Fl (with the inverted flow field) is an effect of the defined natural hydraulic gradient. For a faster steady flow, higher extraction rates would be needed, however it was decided to keep the extraction and injection rates close to the pumping capacity available at the field site.

Maximum tracer concentrations of $\sim 10\,\mathrm{mg\,l^{-1}}$ were obtained during both tracer tests of the synthetic tracer-tomography experiment. Slightly larger concentrations and shorter mean

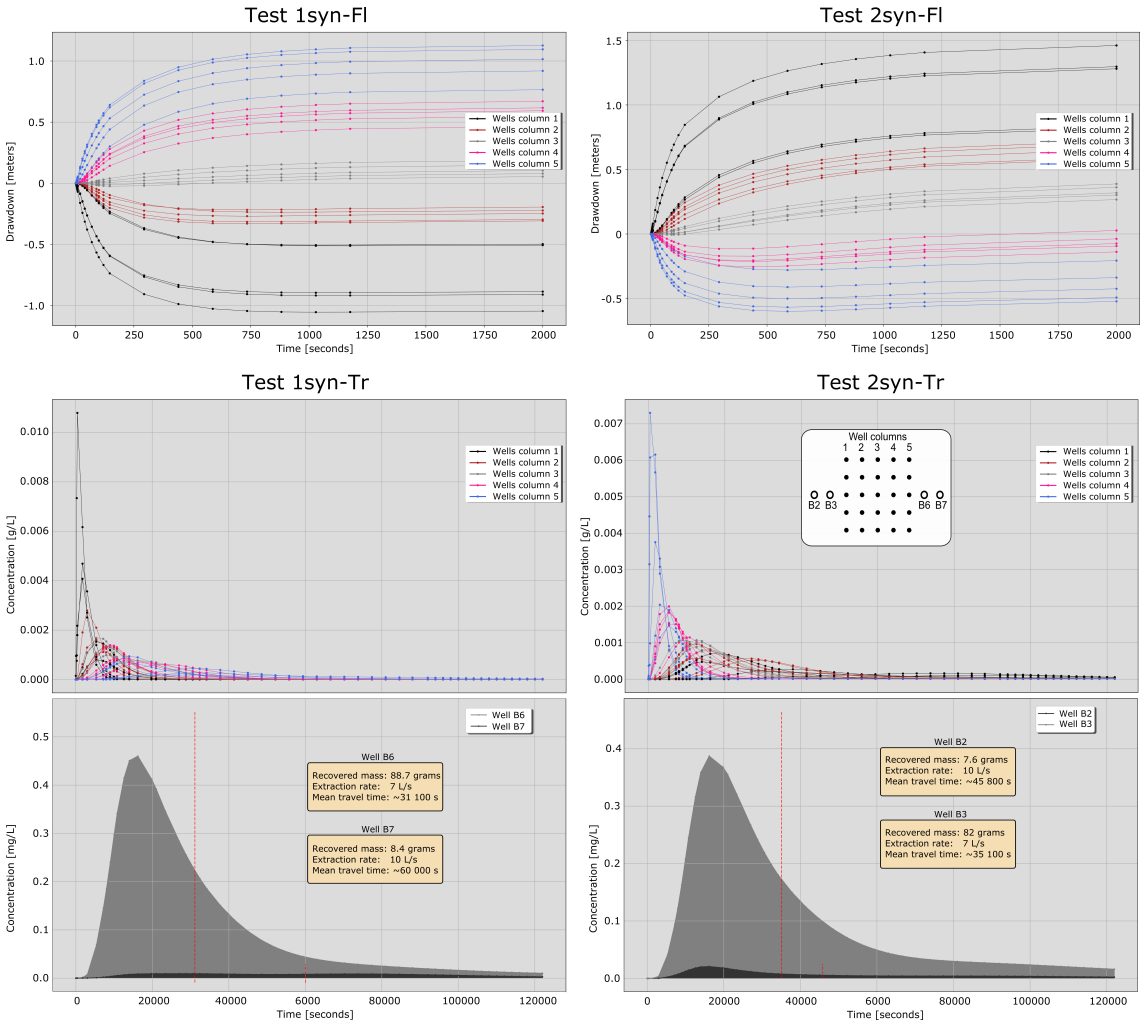


Figure 6.2: Dataset generated by flow and transport simulations in the synthetic tracer-tomography experiment. Top row: drawdown curves simulated during tests 1syn-Fl (left) and 2syn-Fl (right), at 25 observation points . Middle row: breakthrough curves simulated during the two different tracer tests that conformed the tracer tomography, at the same 25 observation points. Bottom row: breakthrough curves and estimated recovered tracer mass and mean tracer travel times of both extraction wells defined in each individual tracer test. Additionally, a small diagram shows the spatial distribution of all wells defined in the model together with the column numbers defined to group the observation points.

travel times were registered for tracer test 1syn-Tr (tracer injection at well B3). This was again caused by the natural hydraulic gradient and the relatively low pumping rates implemented in the upstream wells (B2 and B3) during test 2syn-Tr. In test 1syn-Tr a total tracer mass of $97\,\%$ was recovered, whereas only $90\,\%$ was recovered in test 2syn-Tr. The remain tracer was either released into the natural ambient flow or has not yet arrived at the extraction wells.

6.2 Parameter Estimation

Data assimilation and parameter estimation with the EnKF was sequentially performed using first, the hydraulic datasets of both tests of the synthetic tracer-tomography experiment (1syn-Fl followed by 2syn-Fl). The ensembles of parameters conditioned to hydraulic heads, were then used as initial parameter fields for the assimilation of transport data (1syn-Tr followed by 2syn-Tr). The sequence of the assimilation of synthetic data is schematically represented in Fig. 6.3.

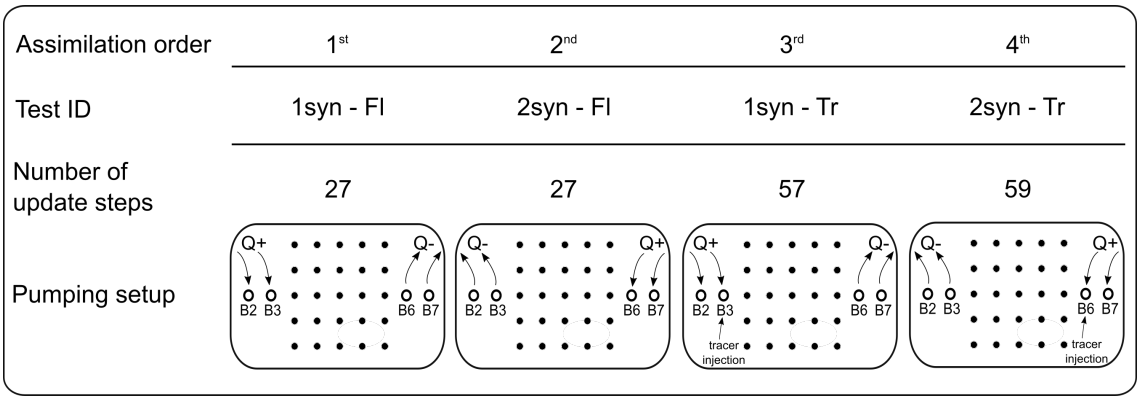


Figure 6.3: Sequence of the assimilation of data from the synthetic tracer tomography. Q^+: injection of water; Q^-: extraction of water. Notice the change in the sign of Q between flow simulations, and the change of tracer injection from well B3 (1syn-Tr) to well B6 (2syn-Tr).

An ensemble of 500 randomly sampled realizations was used in this work. Hendricks Franssen and Kinzelbach (2008) showed that this ensemble size was appropriate to reduce filter inbreeding problems (ensemble variance underestimation) in groundwater flow applications. Each realization of the ensemble was initialized using the same geostatistical parameters as in the reference model (see Table 6.1). To further alleviate filter inbreeding, a damping factor β was applied (see Eq. 2.22). Following Erdal and Cirpka (2016), a value of $\beta = 0.6$ is usually appropriate for groundwater problems. Due to the higher nonlinearity of the advection-dispersion equation, a stronger damping ($\beta = 0.1$) was applied during assimilation of transport data. The evaluation of the conditioned parameters was concentrated on the area covered by the 5×5 grid of observation wells (*area of interest* in Fig. 6.1).

The following settings of the EnKF were evaluated:

- Noise added to modeled states via the covariance matrix of measurement errors $\mathbf{R}$ (Eqs. 2.12 and 2.13).

- Application of normal score data transformation according to Eq. 2.21.
- Application of a damping factor β according to Eq. 2.22.
- Inclusion of additional (homogeneous) hydraulic and transport parameters (e.g., S_o, n_{im}, n_m, n_{tot}, α_l, λ_{mt}) in the update parameter vector (see e.g., Eq. 2.18).
- The update schemes of the classical EnKF (Fig. 2.2) and the Restart EnKF (Fig. 2.1) (only for transport simulations).

For the assimilation of transport data, cumulative breakthrough curves rather than normal concentration series were used. The monotonicity of cumulative breakthrough curves increases the sensitivity of the observations to the model parameters and improve the stability of the parameter update when the predicted concentrations are considerably off the real values.

The results of the scenarios defined with different filter settings were analyzed based on four criteria:

(i) Visual comparison between the ensemble mean of updated parameters and the reference field.

(ii) Visual comparison between the ensemble predictions of the model state (drawdown or concentration) and the synthetic dataset.

(iii) Root mean square error (RMSE) and mean relative error (MRE) of the simulated states.

(iv) Time-step evolution of the Nash-Sutcliffe (NS) coefficient of model efficiency.

The RMSE and absolute MRE of all model observations and update steps, were estimated according to:

$$RMSE = \frac{1}{U}\sum_{j=1}^{U}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\bar{y}_j^{mod,i} - y_j^{ref,i}\right)^2} \tag{6.1}$$

$$MRE = \frac{1}{U}\sum_{j=1}^{U}\frac{1}{N}\sum_{i=1}^{N}\left(\left|\frac{\bar{y}_j^{mod,i} - y_j^{ref,i}}{y_j^{ref,i}}\right|\right) \tag{6.2}$$

where U is the total number of update steps, N is the number of observations, $|\cdot|$ represents the absolute value, and $y_j^{ref,i}$ and $\bar{y}_j^{mod,i}$ are the ith reference and ensemble mean observations at time step j, respectively.

The Nash-Sutcliffe coefficient of model efficiency was calculated for each update step as:

$$NS = 1 - \frac{\sum_{i=1}^{N}\left(\bar{y}_i^{mod} - y_i^{ref}\right)^2}{\sum_{i=1}^{N}\left(y_i^{ref} - \frac{1}{N}\sum_{j=1}^{N} y_j^{ref}\right)} \tag{6.3}$$

A RMSE of zero and a NS-coefficient of one would imply a perfect match between model and observed (or synthetic) data.

In addition to the visual inspection, the updated ensemble of parameters was analyzed at each update step with the average absolute error (AAE) and the average ensemble standard deviation (AESD). According to Hendricks Franssen and Kinzelbach (2008), the AAE is the difference between the ensemble mean and the true value at a grid element, averaged over all grid elements. The AESD is the average of all standard deviations evaluated at all grid elements:

$$AAE_u = \frac{1}{G}\sum_{i=1}^{G}\left|(\bar{p}_{i,u}^{mod} - p_i^{ref}\right| \tag{6.4}$$

$$AESD_u = \frac{1}{GM}\sum_{j=1}^{M}\sum_{i=1}^{G}\left|p_{i,j,u}^{mod} - \bar{p}_{i,u}^{mod}\right| \tag{6.5}$$

where G is the number of grid elements, M is the number of realizations in the ensemble, p^{mod} and p^{ref} refer to the updated and referenced parameter value, the overbar represents the ensemble average and u indicates that the analysis is performed for each update step.

Inroduce all scenarios: which are important?

Test 1syn-FI

Test 1syn-FI was the first hydraulic test of the synthetic tracer tomography. The nested-cell flow field was defined by injecting water at the upstream wells B2 and B3, whereas the downstream wells B6 and B7 were defined as extraction points. For this synthetic test, 25 drawdown curves were generated, recording hydraulic changes at 27 different times (Figs. 6.2 and 6.3).

Table 6.3 is a summary of the six most relevant scenarios defined for parameter estimation with drawdown data from Test 1syn-FI.

For those scenarios in which S_o was included in the update, Table 6.4 shows the mean and variance of log-S_o (σ^2_{LnSo}) used to generate the initial ensemble of random realizations, and the fixed value assumed for those scenarios with no update of S_o. The variance of the natural logarithm of the corresponding value was used to avoid the generation of negative numbers.

Scenario `A1_flow` was a standard application of the EnKF, with no data transformation, no damping of the update, and an absolute (fixed) value of measurement noise added to the model states, whereas data transformation and a damping factor $\beta = 0.6$ where included in scenarios `B1_flow` and `C1_flow`. For scenario `A1_flow` only the spatially distributed log-K was updated (i.e., no update of aquifer storativity).

Table 6.3: Settings of the six scenarios used for the assimilation of drawdown data from test 1syn-Fl. ✓and × indicate whether the setting was activated or not; absolute: indicates if a fixed value for the measurement standard deviation was applied; relative: implies that an error percentage relative to the current measurement value was applied.

Test 1syn-Fl											
Scenario ID	Data transf.	Damping factor β		Measurement error σ_ϵ (m)		Scenario ID	Data transf.	Damping factor β		Measurement error σ_ϵ (m)	
A1_flow	×	×	1.0	absolute	5×10^{-3}	D1_flow	✓	✓	0.6	absolute	5×10^{-3}
B1_flow	✓	×	1.0	absolute	5×10^{-3}	E1_flow	✓	✓	0.6	absolute	5×10^{-3}
C1_flow	×	✓	0.6	absolute	5×10^{-3}	F1_flow	✓	✓	0.6	relative	10 %

Table 6.4: Effective transport parameters included in the update during assimilation of drawdown data of test 1syn-Fl. S_o: aquifer storativity; ✓and × indicate if the parameter was included in the update or not; μ and σ^2_{LnSo}: mean value and variance of log-S_o used to generate random values for the initial ensemble.

Test 1syn-Fl							
Scenario ID	S_o ($x10^{-3}$ m^{-1})			Scenario ID	S_o ($x10^{-3}$ m^{-1})		
		μ	σ^2_{LnSo}			μ	σ^2_{LnSo}
A1_flow	×	$S_o = 3.5$		D1_flow	×	$S_o = 0.35$	
B1_flow	×	$S_o = 3.5$		E1_flow	✓	3.5	0.3
C1_flow	✓	0.35	1.0	F1_flow	✓	3.5	0.3

Aquifer storativity was included in the update for scenario C1_flow, with an initial mean value $\bar{S}_o$ of $3.5 \times 10^{-4}\,\mathrm{m}^{-1}$, which is one order of magnitude smaller than the value used in the reference model. To assess the impact of uncertainties in the assumed value of S_o and its exclusion from the update, for scenario D1_flow an S_o value smaller than the reference was used for the ensemble, and kept constant throughout the parameter estimation. This scenario resembles a common situation of real applications, where uncertainties in effective aquifer storativity are usually large. The effect of measurement noise was evaluated with scenarios E1_flow and F1_flow. In the former, a fixed measurement error of $5 \times 10^{-3}\,\mathrm{m}$ was added to all modeled states at all update steps, whereas in the latter an error of $10\,\%$ relative to the reference measurement was applied.

Figs. 6.4 and 6.5 present the ensemble mean and associated variance of log-K at update steps 9, 18 and 27 (corresponding to a pumping time $t_9 = 9\,\mathrm{s}$, $t_{18} = 120\,\mathrm{s}$ and $t_{27} = 2000\,\mathrm{s}$, respectively), for all six scenarios. The initial ensemble of parameters ($t_o = 0\,\mathrm{s}$) was the same for all scenarios. The reference and the initial ensemble mean of log-K are displayed in the top

row of Fig. 6.4, and the variance associated in the top row of Fig. 6.5.

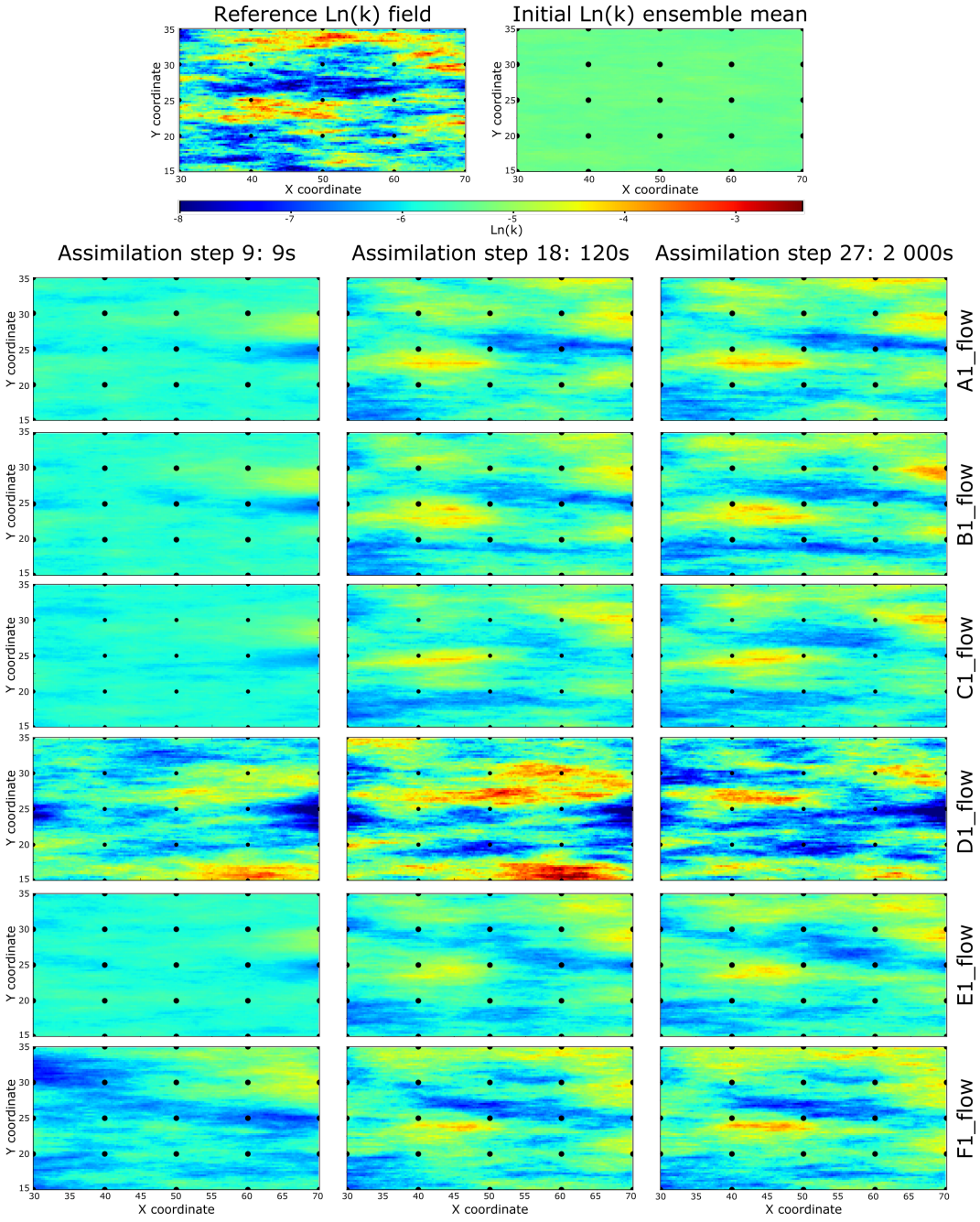


Figure 6.4: Ensemble mean log-hydraulic conductivity fields for the assimilation of drawdown data of test 1syn-FI with six different filter settings, after 9 (left row), 18 (central row) and 27 (right row) assimilation steps. Top: reference (left) and initial ensemble mean (right) log-hydraulic conductivity fields.

A visual inspection of the mean ensemble of parameters suggests that independent of the settings of the filter, the main spatial structure of the reference field are recovered (except for scenario `D1_flow`). The ensemble variance of log-K is considerably reduced after 27 update steps, from an initial mean value of 1 to a mean value of ~ 0.6. Except for scenario `D1_flow`, the AAE was reduced between $16\,\%$ to $25\,\%$ after 27update steps (Table 6.6). The best estimation of log-K was obtained for scenarios `B1_flow` and `F1_flow`, with an AAE reduction of $22\,\%$ and $25\,\%$, respectively.

Table 6.5 shows the mean (μ) and variance (σ^2_{LnSo}) of aquifer storativity obtained after 27 update steps. As expected, mean values of updated parameters remained close to the value used for the generation of the initial ensemble, and the uncertainty associated is reduced by $20\,\%$ and $37\,\%$ for scenarios `E1_flow` and `F1_flow`, respectively. Results from scenario `C1_flow`, in which a wrong initial mean value was used for the generation of the initial ensemble, the associated uncertainty remained constant throughout the 27 update steps.

Table 6.5: Updated mean (μ) and variance (σ^2_{LnSo}) of aquifer storativity (S_o) after the assimilation of data from test 1syn-Fl. S_o: aquifer storativity; × indicates that the parameter was not included in the update.

Test 1syn-Fl					
Scenario ID	S_o (m^{-1})		Scenario ID	S_o (m^{-1})	
	μ	σ^2_{LnSo}		μ	σ^2_{LnSo}
`A1_flow`	×		`D1_flow`	×	
`B1_flow`	×		`E1_flow`	3.3×10^{-3}	0.24
`C1_flow`	3.1×10^{-4}	1.01	`F1_flow`	3.3×10^{-3}	0.19
Reference value $S_o = 3.5 \times 10^{-3}\,\mathrm{m}^{-1}$					

The effect of wrong prior information given to S_o, and its exclusion from the updating process is evidenced in the results of scenario `D1_flow`. To compensate for the wrong S_o (note: $S_o^{ref} = 3.5 \times 10^{-3}\,\mathrm{m}^{-1}$), stronger perturbations of log-K are applied at every update step, even if a damping factor is included, leading to the abundance of extremes and larger deviations from the true hydraulic conductivity. This negative effect is not observed in scenario `C1_flow`, where also a wrong initial value of S_o= $3.5 \times 10^{-4}\,\mathrm{m}^{-1}$ was applied, but S_o was included in the update. Although the EnKF is not able to correct the mean value used to generate initial random realizations, allowing a higher degree of variation in S_o helped the update of log-K.

Stronger reductions of AAE and AESD of log-K were observed for scenarios `B1_flow` and `F1_flow`, whereas a considerable increment was observed for scenario `D1_flow` (Table 6.6). According to Hendricks Franssen and Kinzelbach (2008), ratios of AESD/AAE lower than 1

might be indicative of filter inbreeding. This problem was reduced for scenarios `E1_flow` and `F1_flow` with ratios closer to 1.

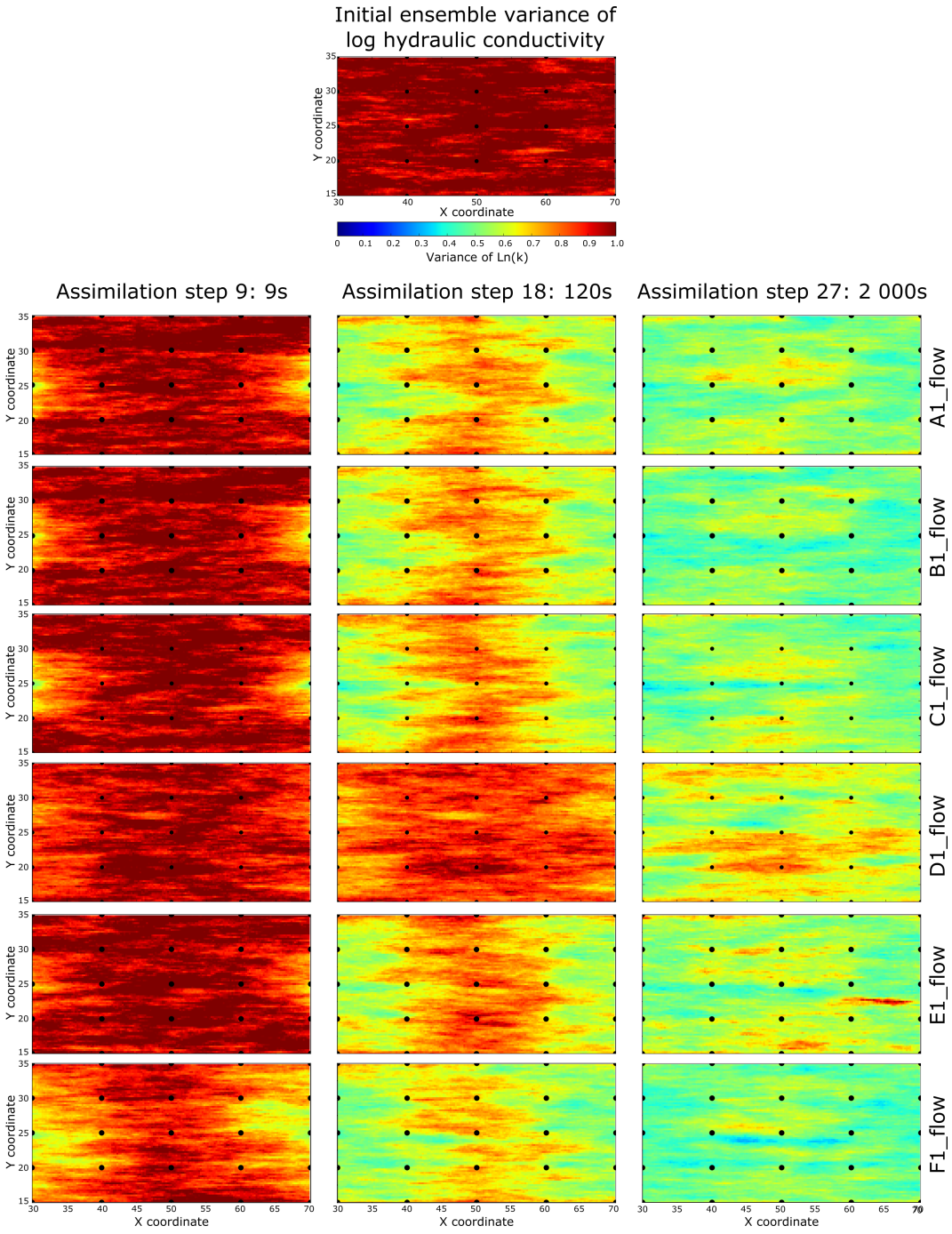


Figure 6.5: Ensemble variance of the log-hydraulic conductivity fields presented in Fig. 6.4. Top: initial ensemble variance of log-hydraulic conductivity.

The evolution of drawdown predictions as a function of the update step is plotted in Fig. 6.6 for four randomly selected observation points.

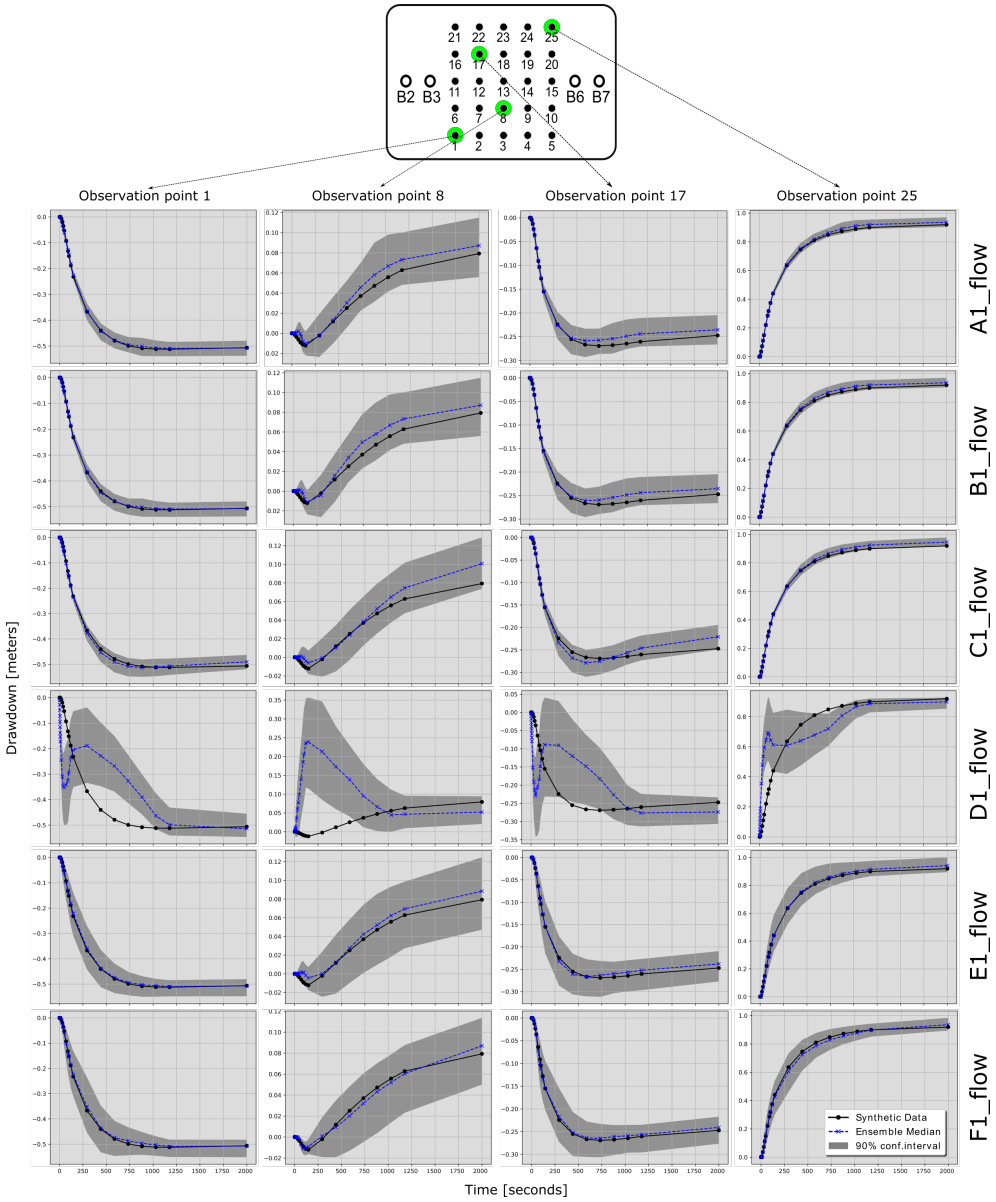


Figure 6.6: Ensemble predictions of drawdown as a function of the updating step, of four randomly selected observation points. The curves correspond to the six different scenarios generated to assimilate the drawdown data of test 1syn-FI. Black lines: reference drawdown simulated at each observation point; Gray area: 90 % confidence interval of the ensemble predictions; Blue lines: drawdown ensemble median.

The reference drawdown simulated at each observation point is plotted in black, the en-

semble median is plotted in blue and the $90\,\%$ confidence interval of the ensemble is described by the gray zone. It is noticeable that the steady-state value is well captured in all scenarios, illustrating the importance of the information contained in the transient records of drawdown. Although computationally more demanding, transient data provides valuable information about aquifer heterogeneity, that otherwise would be missed if data assimilation is based only on steady-state values. Furthermore, aquifer storativity can be incorporated in the estimation only if transient records are included.

The large errors in drawdown predictions at early times for scenario `D1_flow` confirm the negative effect of a wrong prior of aquifer storativity and its exclusion from parameter updating. At late times, steady state is achieved, which does not depend on the storativity at all. For all other scenarios, the model predictions are equally good, with a slight improvement for scenarios `B1_flow` and `F1_flow`, visible especially in the drawdown curves of observation points 8 and 17. This is confirmed by lower RMSE and MRE estimated for scenarios `B1_flow` and `F1_flow` (Table 6.6).

From the results of scenarios `B1_flow` and `C1_flow`, it appears that data transformation has a stronger effect than the inclusion of a damping factor. This is supported by a visual comparison between their respective updated ensemble mean and the reference field, and by smaller RMSE and MRE of the model predictions (Table 6.6). The inclusion of a damping factor becomes less relevant with increasing ensemble size, since numerical covariances are better approximated. Previous studies have shown that an ensemble of 500 realization is usually appropriate for groundwater flow problems (Hendricks Franssen and Kinzelbach 2008; Franssen and Kinzelbach 2009). A better performance of scenario `B1_flow` than scenario `E1_flow` (even when a damping factor was not included in the former) can be explained by the reduced relevance of the damping factor, and the fact that the true aquifer storativity was provided in scenario `B1_flow` and not updated, whereas for scenario `E1_flow` it was also updated.

The NS-coefficient estimated for the mean model predictions, as a function of the updating step, further confirms the relevance of the full transient records. Since the correct steady-state values are met in all scenarios, it is of no surprise that after the last assimilation step, the NS-coefficient is close to one in all cases, with one being a perfect match between modeled and reference values. The evolution of the NS-coefficient with respect to the update step reveals instabilities if the classical version of the EnKF (scenario `A1_flow`) is used (Fig. 6.7).

Table 6.6: Average absolute error (AAE), Average Ensemble Standard Deviation (AESD) and Total Variance of log-hydraulic conductivity at the end of the assimilation of the data from test 1syn-FI, and Root Mean Square Error (RMSE) and Mean Relative Error (MRE) of drawdown averaged over all assimilation steps.

	Test 1syn-FI				
	Model Parameters (Lnk)			Model States	
Scenario ID	AAE ($\log(\mathrm{m\,s^{-1}})$)	AESD ($\log(\mathrm{m\,s^{-1}})$)	AESD/AAE (–)	RMSE(ymod) (m)	MRE (ymod) (–)
`A1_flow`	0.66	0.57	0.86	7.85×10^{-3}	1.58
`B1_flow`	0.63	0.56	0.89	6.15×10^{-3}	1.27
`C1_flow`	0.64	0.56	0.87	7.09×10^{-3}	1.09
`D1_flow`	1.07	0.61	0.57	1.48×10^{-1}	231.7
`E1_flow`	0.64	0.61	0.95	6.50×10^{-3}	1.78
`F1_flow`	0.60	0.57	0.95	5.84×10^{-3}	1.24
Initial	0.76	0.8	1.05	–	–

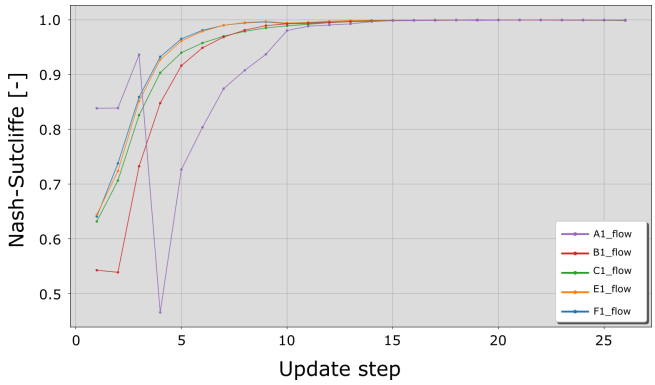


Figure 6.7: Evolution of the Nash-Sutcliffe coefficient of model efficiency, as a function of the update step. Note: values corresponding to scenario `D1_flow` were discarded due to the extremely low values, affecting the clarity of the plot.

Test 2syn-FI

Test 2syn-FI was the second hydraulic test of the synthetic tracer tomography. The inverted nested-cell flow field applied in this test was defined by interchanging the injection and extraction wells. Water was now injected at the downstream wells B6 and B7, whereas extraction of water was defined at the upstream wells B2 and B3. For this synthetic test, 25 drawdown curves were generated, recording the hydraulic changes at 27 different times (see Table 6.2 and Figs. 6.2 and 6.3).

Table 6.7 is a summary of the three scenarios defined for the estimation of parameters with drawdown data from test 2syn-FI. The conditioned parameters from scenarios `C1_flow`, `E1_flow` and `F1_flow` of test 1syn-FI, were used as initial ensemble of parameters for scenarios `C2_flow`, `E2_flow` and `F2_flow`, respectively. Data transformation and a damping factor of 0.6 were applied in all three scenarios.

Table 6.7: Description of the three scenarios defined for the assimilation of drawdown data of test 2syn-FI. ✓and × indicate if the setting was activated or not; μ: mean value; σ^2_{LnSo}: variance of log-aquifer storativity; absolute: indicates if a fixed value for the measurement standard deviation was applied; relative: implies that an error percentage relative to the current measurement value was applied.

Test 2syn-FI					
Scenario ID	Dat transf.	Damping factor β		Measurement error σ_ϵ (m)	
`C2_flow`	✓	✓	0.6	absolute	5×10^{-3}
`E2_flow`	✓	✓	0.6	absolute	5×10^{-3}
`F2_flow`	✓	✓	0.6	relative	$10\,\%$

Benefits of including aquifer storativity in the update (assumed homogeneous) where observed in the results of test 1syn-FI, therefore aquifer storativity was also included during parameter updating of test 2syn-FI, with initial random realizations generated with a mean value $\mu_{So} = 3.5 \times 10^{-3}\,\mathrm{m}^{-1}$ and σ^2_{LnSo} equal to 1.0 for scenario `C2_flow`, and $\sigma^2_{LnSo} = 0.3$ for scenarios `E2_flow` and `F2_flow` (Table 6.8).

The settings of scenarios `E2_flow` and `F2_flow` correspond to those applied in scenarios `E1_flow` and `F1_flow` of test 1syn-FI, and served to assess their performance once new data has been assimilated. For scenario `E2_flow`, a fixed error of $5 \times 10^{-3}\,\mathrm{m}$ was applied to all model outputs and update steps, whereas an error of $10\,\%$ relative to the current reference (synthetic) measurement was applied in scenario `F2_flow`.

Table 6.8: Effective transport parameters included in the update during assimilation of drawdown data from test 2syn-Fl. S_o: aquifer storativity; ✓and × indicate if the parameter was included in the update or not; μ and σ^2_{LnSo}: mean value and variance of log-S_o used to randomly generate parameter values for the initial ensemble.

Test 2syn-Fl			
Scenario ID	S_o ($x10^{-3}$ m^{-1})		
		μ	σ^2_{LnSo}
C1_flow	✓	3.5	1.0
E1_flow	✓	3.5	0.3
F1_flow	✓	3.5	0.3

Figs. 6.8 and 6.9 show the ensemble mean and associated variance of log-K after 9, 18 and 27 update steps, and the initial ensemble mean and initial ensemble variance of log-hydraulic conductivity used for each scenario.

A visual inspection of the updated mean ensemble of parameters (Fig. 6.8) reveals that some features of the reference field are accentuated after the new data was assimilated. However, it is noticeable that no big changes occur during the update, suggesting a stabilization of the parameters. Since reverting the flow field simply reverts the hydraulic response, this is an expected result. That is, there is no new information contained in the dataset of 2syn-Fl in comparison to that of 1syn-Fl. The top and central regions of the reference field (with relatively high and low log-K values, respectively) are better captured in scenarios `E2_flow` and `F2_flow`. A drastic reduction in ensemble variance occurred for scenario `F2_flow` (Fig. 6.8), confirmed by the low AAE of 0.485 ($\log(\mathrm{m\,s}^{-2})$) (Table 6.9), estimated after the last update. The $\sim 21\,\%$ reduction of the AAE with respect to the initial value, and a AESD/AAE ratio of 0.7 estimated for `F2_flow` might be indicative of filter inbreeding effects. The AAE and AESD values for scenarios `C2_flow` and `E2_flow` were reduced between $1\,\%$ to $14\,\%$. A fixed error measurement value added to the simulated states seems to provide more stable results in the assimilation of drawdown data.

Drawdown predictions as a function of the updating step are plotted in Fig. 6.10 for four exemplary observation points. The relatively large variance of log-storativity used in scenario `C2_flow` ($\sigma^2_{LnSo} = 1.0$) is reflected in the broader area covered by the $90\,\%$ confidence interval of the ensemble predictions (Fig. 6.10-top row) and the unstable evolution of the NS-coefficient shown in Fig. 6.11 (blue line). In applications with large uncertainties about aquifer storativity, the use of a large σ^2_{LnSo} might be justified by the acceptable drawdown predictions of the ensemble median (Fig. 6.10).

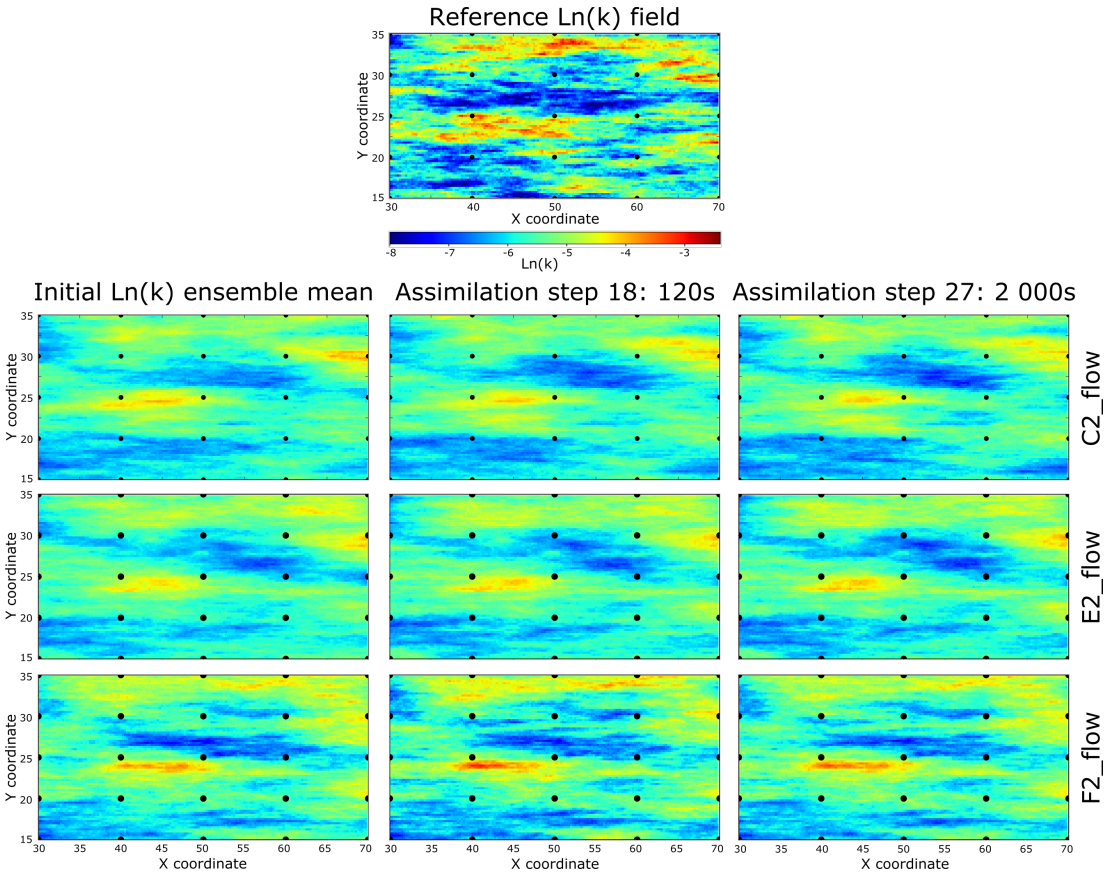


Figure 6.8: Ensemble mean log-hydraulic conductivity fields for the assimilation of drawdown data of test 2syn-FI after 9, 18 and 27assimilation steps. Additionally, the initial ensemble mean of log-hydraulic conductivity of each scenario is included.

Table 6.9: Average absolute error (AAE) and Average Ensemble Standard Deviation (AESD) of log-hydraulic conductivity at the end of the assimilation of data from test 2syn-FI, and Root Mean Square Error (RMSE) and Mean Relative Error (MRE) of drawdown averaged over all assimilation steps.

	Test 2syn-FI				
	Model Parameters (Lnk)			Model States	
Scenario ID	AAE ($\log(\mathrm{m\,s^{-1}})$)	AESD ($\log(\mathrm{m\,s^{-1}})$)	AESD/AAE (–)	RMSE(ymod) (m)	MRE (ymod) (–)
C2_flow	0.653	0.554	0.84	0.02	295.23
E2_flow	0.638	0.535	0.83	8.7×10^{-3}	0.664
F2_flow	0.665	0.470	0.71	7×10^{-3}	0.573

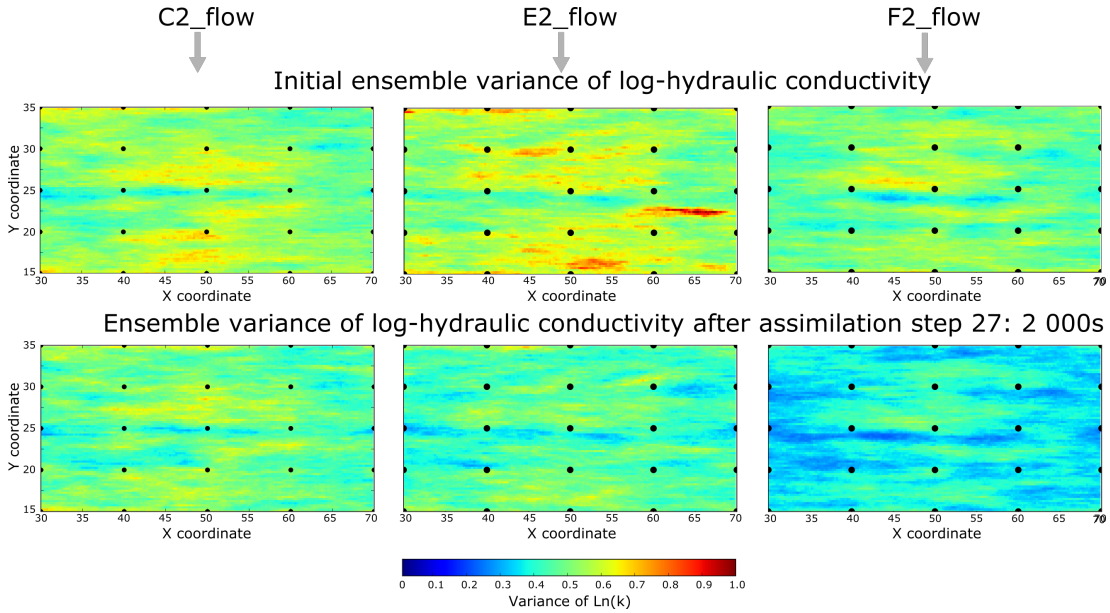


Figure 6.9: Top: Initial ensemble variance of the log-hydraulic conductivity fields used in scenarios `C2_flow`, `E2_flow` and `F2_flow`; Bottom: ensemble variance of log-hydraulic conductivity after 27 assimilation steps .

In general, drawdown predictions are improved for late times in all scenarios, explaining the similar behavior of the NS-coefficient (Fig. 6.11) to test 1syn-Fl. Appropriate drawdown predictions from scenarios `E2_flow` and `F2_flow` can be observed in Fig. 6.10 and verified by the stable evolution of the NS-coefficient (Fig. 6.11), and the low RMSE and MRE values averaged over all assimilation steps (Table 6.9).

From the results of the assimilation of hydraulic data, the following conclusions can be drawn:

(i) As expected, relatively small parameter updates were observed during the assimilation of the second hydraulic dataset (test 2syn-Fl). Reverting the flow field yield no additional information, and therefore, a stabilization of the parameters occurs. To gain new information, a linearly independent combination of pumping/extraction wells would be needed.

(ii) The inclusion of data transformation, a damping factor and the update of aquifer storativity in the settings of the EnKF, improves the performance of the parameter estimation (results from scenarios `C1_flow`, `E1_flow` and `F1_flow`).

(iii) The best performance was observed with low variance values of log-storativity ($\sigma^2_{LnSo} = 0.3$) (scenarios `E2_flow` and `F2_flow`), however larger values might be justified if no prior information about storativity is available.

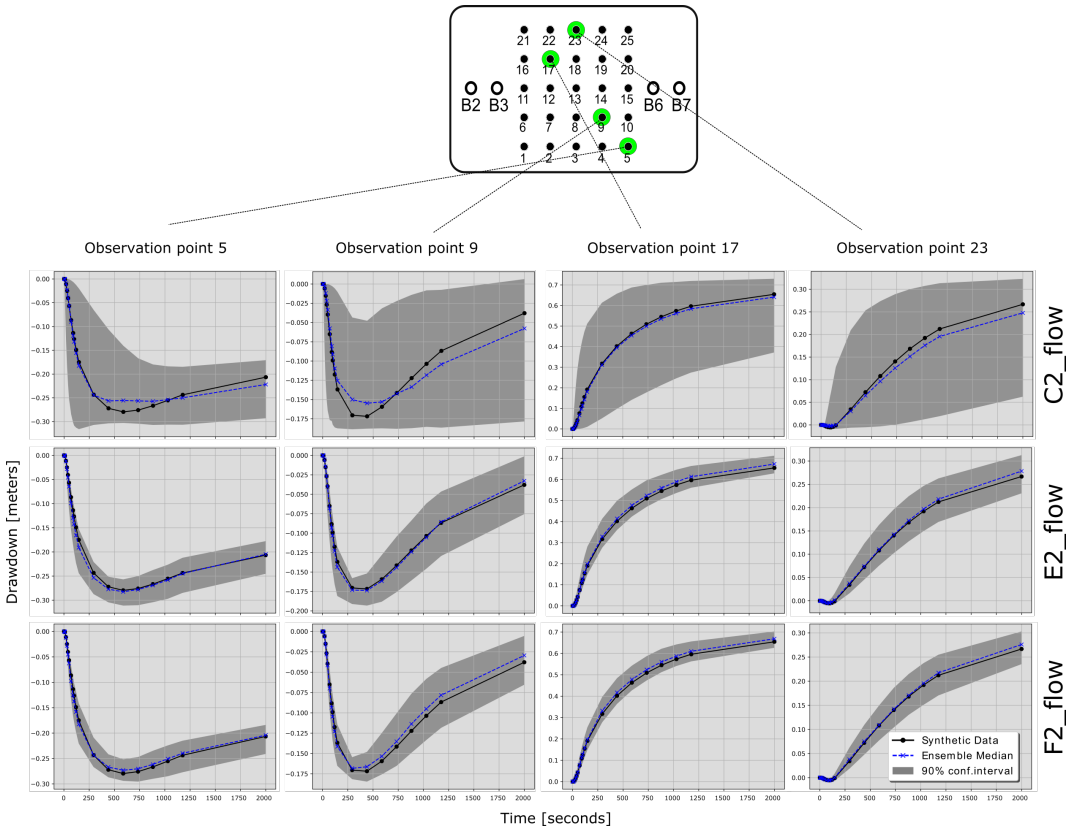


Figure 6.10: Ensemble predictions of drawdown as a function of the updating step, of four randomly selected observation points. The curves correspond to the three scenarios generated to assimilate drawdown data of test 2syn-FI. Black lines: reference drawdown simulated at each observation point; Gray area: $90\,\%$ confidence interval of the ensemble predictions; Blue lines: drawdown ensemble median.

(iv) Lower RMSE and MRE of drawdown estimated for scenario `F2_flow` indicate better model predictions, therefore the updated ensemble of parameters from this scenario was used as initial ensemble for the assimilation of synthetic tracer data from test 1syn-Tr.

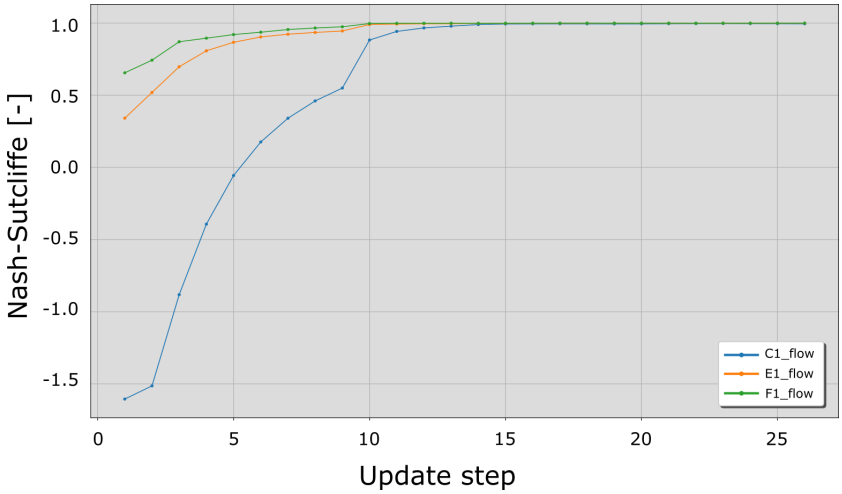


Figure 6.11: Evolution of the Nash-Sutcliffe coefficient of model efficiency as a function of the update step, for the assimilation of drawdown data of test 2syn-Fl.

Test 1syn-Tr

Test 1syn-Tr was the first tracer test of the synthetic tracer tomography, and was based on the steady flow from test 1syn-Fl. Tracer injection was performed at well B3, and extracted at wells B6 and B7. The data assimilation with data from test 1syn-Tr was evaluated with four different scenarios (Table 6.10).

Table 6.10: Four scenarios used for the assimilation of concentration data from test 1syn-Tr. Standard and restart EnKF refer to the update scheme applied. ✓and × indicate if the setting was activated or not; rel.: implies the application of an error percentage relative to the current reference measurement; max & min: maximum and minimum values of measurement error.

Test 1syn-Tr						
Scenario ID	Update scheme	Data transf.	Damping factor β			Measurement error σ_ϵ (g/l)
`A1_transport`	standard EnKF	✓	✓	0.1	rel.	$10\,\%$, max=5×10^{-3},min=5×10^{-6}
`B1_transport`	Restart EnKF	✓	✓	0.1	rel.	$10\,\%$, max=5×10^{-5},min=5×10^{-8}
`C1_transport`	Restart EnKF	✓	✓	0.1	rel.	$10\,\%$, max=5×10^{-3},min=5×10^{-6}
`D1_transport`	Restart EnKF	×	✓	0.1	rel.	$10\,\%$, max=5×10^{-3},min=5×10^{-6}

The updated ensemble of log-K from scenario `F2_flow` was used as initial ensemble, and the following effective transport parameters of a dual-domain system were included in the update: porosities of the mobile and immobile regions, n_m and n_{im}, respectively, the longitudinal dispersivity α_l, and the first-order mass transfer coefficient λ_{mt} between the mobile and immobile domains. The mean and standard deviation used to generate random realizations (with a Gaussian kernel) of effective transport parameters are presented in Table 6.11.

Table 6.11: Mean(μ) and variance (σ^2) used to generate random realizations of effective transport parameters. The parameters were generated at the beginning of the assimilation of tracer data from test 1syn-Tr. n_m: mobile zone porosity; n_{im}: immobile zone porosity; λ_{mt}: first-order mass transfer coefficient; α_l: longitudinal dispersivity. Variance values are expressed for the natural logarithm (Ln) of the corresponding parameter value.

Test 1syn-Tr							
Effective transport parameters included in the update							
n_m (%)		n_{im} (%)		α_l (m)		λ_{mt} (s^{-1})	
μ	$\sigma^2_{Ln(nm)}$	μ	$\sigma^2_{Ln(nim)}$	μ	$\sigma^2_{Ln(\alpha)}$	μ	$\sigma^2_{Ln(\lambda)}$
$10\,\%$	0.01	20	5×10^{-3}	1	0.3	1×10^{-9}	0.1

The classical EnKF was implemented in scenario `A1_transport`. As mentioned in Section 2.3, the update scheme of the classical EnKF leads to mass balance errors during parameter estimation with real-time assimilation of transient concentration data, mainly due to inconsistencies between the predefined steady flow field and the updated hydraulic parameters. Conversely, the update scheme of the restart EnKF (see Fig. 2.2) ensures mass conservation and consistency between hydraulic parameters, groundwater flow velocities and concentrations by re-initializing the flow field after each parameter update, and simulating transport from the initial time until the next available measurement. Hence, the restart EnKF was implemented in scenarios `B1_transport`, `C1_transport` and `D1_transport`. A damping factor $\beta = 0.1$ was applied in all scenarios, while data transformation was excluded only from scenario `D1_transport`. Due to the wide range of concentrations usually observed during tracer testing (e.g., from $\mathrm{ng\,l^{-1}}$ to $\mathrm{g\,l^{-1}}$), and difficulties with providing appropriate estimations of measurement uncertainties, an error of $10\,\%$ relative to the current reference measurement was added to the modeled states in all four scenarios. Minimum and maximum limits of the added noise were lower for scenario `B1_transport` than for all other scenarios (see Table 6.10).

Fig. 6.12 shows the ensemble mean of log-K obtained during assimilation of concentration data from test 1syn-Tr, at three different update steps. The reference log-K field is presented at the top of Fig. 6.12. The updated parameter fields estimated after ten assimilation steps show similar characteristics in all three scenarios (Fig. 6.12-left column), whereas larger differences can be observed after 30 update steps (Fig. 6.12-middle column).

The similarities between the ensemble mean of log-K after ten update steps, can be explained by the relatively low concentrations obtained at early update steps. Concentrations close to zero imply small residuals (Eq. 2.13), hence small parameter perturbations (see Eq. 2.19). With larger (cumulative) concentrations at later times and inaccurate predictions, the magnitude of the parameter perturbation potentially increases with each update step, even if a damping factor is applied.

The final ensemble mean of log-K for scenarios `A1_transport` and `B1_transport` deviate considerably from the reference field, and contain a large number of extreme values. Wrong updates of scenario `E1_transport` are attributed to the update scheme. After every update step, the standard EnKF integrates the model states from the current time to a forward time-step in which new observations are available. For transport simulations, this implies that the steady flow is not reinitialized, yielding inconsistencies between the predefined steady flow and the updated hydraulic parameters. Errors in the model predictions are then propagated forward in time, producing spurious covariances and larger residuals, and affecting the performance of the standard EnKF. For scenario `A1_transport`, no tracer mass was left in the system after 37 assimilation steps, therefore the transport model was not able to run forward in time.

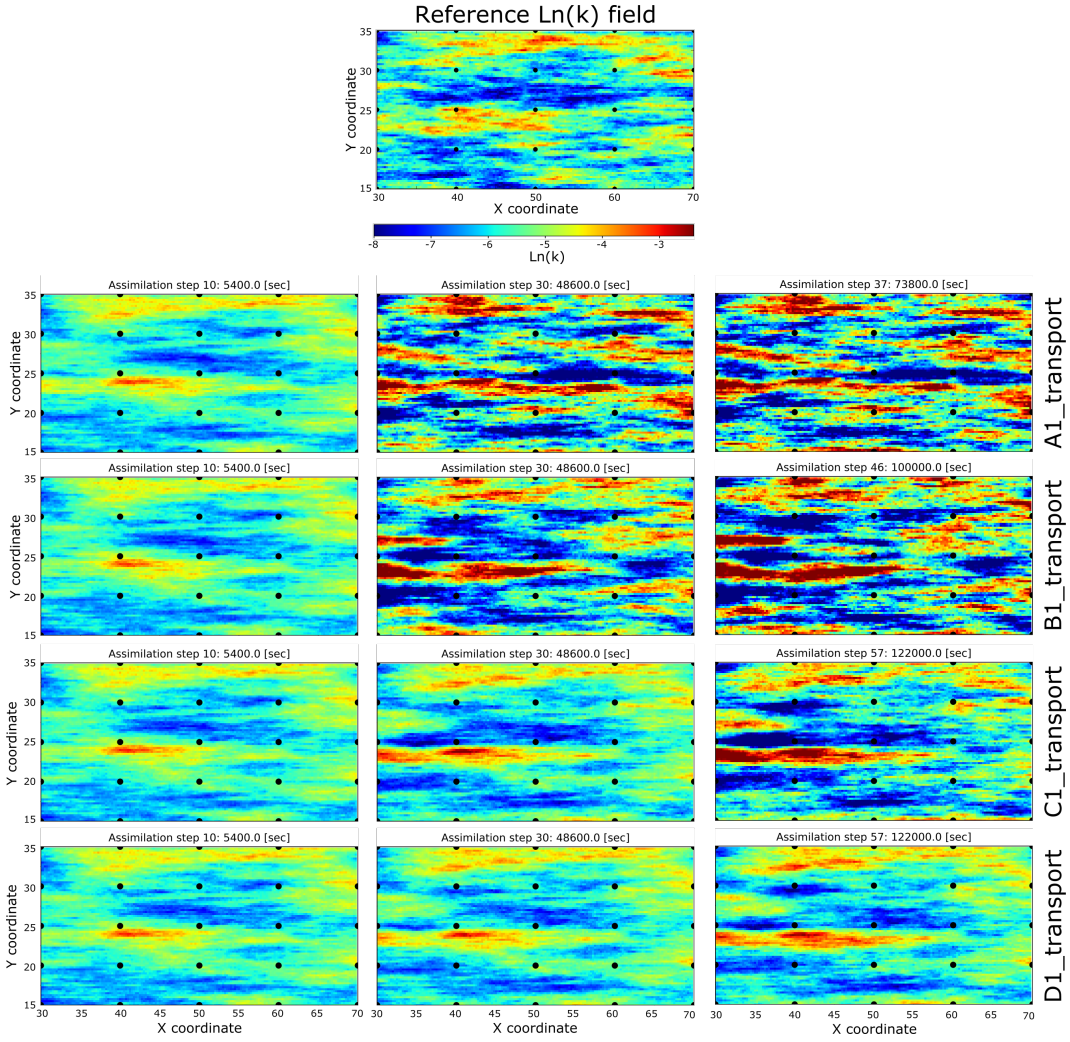


Figure 6.12: Ensemble mean of log-hydraulic conductivity for the assimilation of tracer data of test 1syn-Tr, after 10, 30 and the final assimilation step of each scenario. The initial ensemble of log-hydraulic conductivity correspond to the updated ensemble of scenario `F2_flow`, sequentially conditioned to hydraulic data of tests 1syn-Fl and 2syn-Fl.

The structure of the reference parameter field is better recovered in scenario `D1_transport`, suggesting better correlations of the estimated covariances and more stable updates if no data transformation is applied. As previous research indicates (e.g., Schöniger et al. 2012), rank-based transformations are not recommended for variables with high nonlinear dependence on log-hydraulic conductivity (as is the case of concentration). For those cases, data transformation might introduce additional artifacts that enhance deterioration of the estimated covariance correlations. The negative effect of data transformation applied to concentration data is evidenced in the drastic updates observed for scenarios `B1_transport` after 30 update steps, and

for `C1_transport` at the end of the update (57 update steps).

Table 6.12 shows the mean and variance of the effective parameters included in the update, after all data were assimilated. In general, the mean and variance of the updated parameters remained close to the initial values. The closest mean n_m value to the reference was observed for scenario `A1_transport`, however with a variance one order of magnitude larger than the the value used to generate the random realizations. All mean values of n_{im} were close to the reference value, and the same $\sigma^2_{Ln(nim)}$ was obtained for all cases. The best estimation of α_l was observed for scenario `D1_transport`, with a deviation of $22\,\%$ from the reference value. Considerable larger deviations to the reference value of α_l were obtained for all other scenarios. No relevant differences were observed in the estimation of λ_{mt} and its associated variance.

Table 6.12: Updated mean(μ) and variance (σ^2) of the effective parameters included in the update during assimilation of tracer data from test 1syn-Tr, after 57 update steps. n_m: mobile zone porosity; n_{im}: immobile zone porosity; λ_{mt}: first-order mass transfer coefficient; α_l: longitudinal dispersivity. Variances are expressed for the natural logarithm (Ln) of the corresponding parameter.

	Test 1syn-Tr							
Scenario	n_m (%)		n_{im} (%)		α_l (m)		λ_{mt} (s^{-1})	
ID	μ	$\sigma^2_{Ln(nm)}$	μ	$\sigma^2_{Ln(nim)}$	μ	$\sigma^2_{Ln(\alpha)}$	μ	$\sigma^2_{Ln(\lambda)}$
Initial	10	0.01	20	5×10^{-3}	1	0.3	1.0×10^{-9}	0.1
`A1_transport`	10.1	10×10^{-2}	19.8	5×10^{-3}	1.27	0.30	9.79×10^{-10}	0.089
`B1_transport`	10.9	9×10^{-3}	19.0	5×10^{-3}	1.38	0.25	9.93×10^{-10}	0.098
`C1_transport`	10.6	8×10^{-3}	20.0	5×10^{-3}	1.31	0.24	1.05×10^{-9}	0.085
`D1_transport`	10.5	8×10^{-3}	20.2	5×10^{-3}	1.22	0.24	1.06×10^{-9}	0.102

Despite different estimations of log-K obtained for each scenario, the ensemble variance was reduced to similar levels in all cases, confirmed by the AESD of log-hydraulic conductivity estimated after all update steps (Table 6.13). The lowest average absolute error (AAE) of log-K corresponds to scenario `D1_transport`. However, the AAE value of 0.75 $\log(\mathrm{m\,s^{-1}})$ represents a $\sim 12\,\%$ increment with respect to the AAE of the initial ensemble used for test 1syn-Tr ($\mathrm{AAE}_{ini} = 0.665\,\log(\mathrm{m\,s^{-1}})$). This increment is attributed to the high log-conductivity zone located between X=(30,50)m and Y=(20,25)m at the end of the assimilation for scenario `D1_transport`.

Errors in parameter estimation might be reduced if a smaller damping factor β is applied, reducing the parameter perturbations at every update step, and possibly also the filter inbreed-

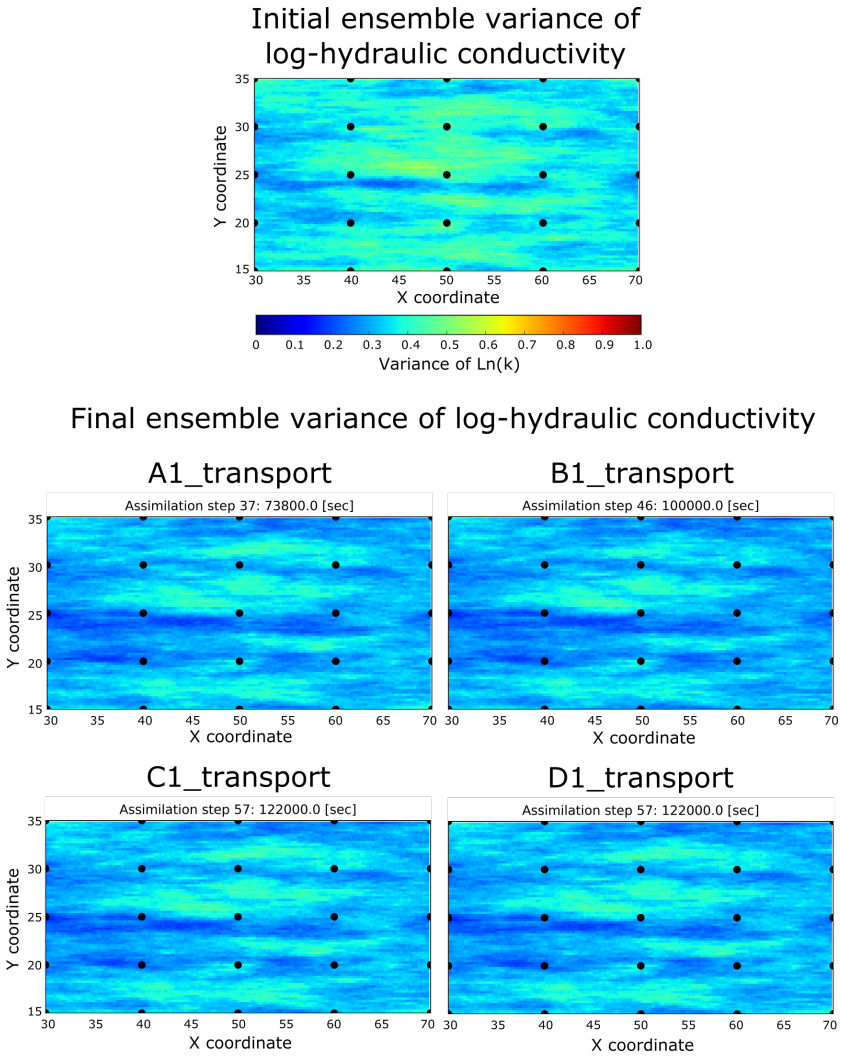


Figure 6.13: Top: Initial ensemble variance of the log-hydraulic conductivity fields (taken from the results of the assimilation of hydraulic data with scenario `F2_flow`). Bottom: ensemble variance of the log-hydraulic conductivity fields estimated after the last update step of paramters performed in each scenario.

ing effects evidenced with AESD/AAE ratios smaller than one (Table 6.13). In this work, however, rather than applying a stronger damping to the update, it was decided to use the dataset of the second tracer test (test 2syn-Tr), and assess the benefits of including additional information obtained during a tracer tomography experiment.

Fig. 6.14 shows the breakthrough curves obtained at 6 different observation points, with a complete simulation of the first tracer test (1syn-Tr). The parameters used for the simulation correspond to the updated ensemble of parameters after all update steps. Again, the best

Table 6.13: Average absolute error (AAE) and Average Ensemble Standard Deviation (AESD) of log-hydraulic conductivity at the end of the assimilation of data from test 1syn-Tr, and Root Mean Square Error (RMSE) and Nash-Sutcliffe coefficient (NS) of concentrations averaged over all assimilation steps.

	Test 1syn-Tr				
	Model Parameters (Lnk)			Model States	
Scenario ID	AAE ($\log(\mathrm{m\,s^{-1}})$)	AESD ($\log(\mathrm{m\,s^{-1}})$)	AESD/AAE (–)	RMSE(ymod) ($\mathrm{g\,l^{-1}}$)	NS-coeff. (–)
`A1_transport`	1.45	0.445	0.31	2.7×10^{-4}	−1.2
`B1_transport`	1.5	0.44	0.29	2.6×10^{-4}	−0.07
`C1_transport`	1.04	0.45	0.43	2.9×10^{-4}	0.63
`D1_transport`	0.75	0.45	0.6	1.4×10^{-4}	0.90

results are obtained for scenario `D1_transport`, confirmed by a RMSE between updated and reference breakthrough curves of $1.4 \times 10^{-4}\,\mathrm{g\,l^{-1}}$, and a NS-coefficent of 0.90 (see Table 6.13). Problems in simulating the breakthrough curves at wells 1 and 7 in scenario `D1_transport` are associated to the high log-K zone mentioned above. The updated ensemble of parameters for scenarios `A1_transport`, `B1_transport` and `C1_transport` fail to reproduce the breakthrough curves at almost all observation points, leading to higher RMSE values and lower NS-coefficeints (Table 6.13).

The lack of information outside the area of interest defined by the grid of observation wells, generates spurious parameters in the external region, affecting to a larger extent those observations points located at the boundaries (e.g. wells 1 to 5 and 21 to 25). The same effect is observed in the breakthrough curves simulated at the outer extraction well B7, located further away from the area under investigation. The updated ensemble of parameters of scenario `D1_transport` was used as initial ensemble of parameters for the assimilation of tracer data from the second tracer test of the synthetic tracer tomography.

From the results of the parameter estimation with concentration data from test 1syn-Tr, two main conclusions can be formulated:

(i) The update scheme of the classical EnKF is not well suited for the assimilation of transient records of tracer tests.

(ii) Normal-score transformations are not recommended for parameter estimation based on concentration data. The application of rank-based transformations to data with a strongly nonlinear dependency on log-hydraulic conductivity may lead to weak pseudolinearizations, introducing additional artifacts that corrupt the correlations of the numerical

covariances, and therefore affect the performance of the filter.

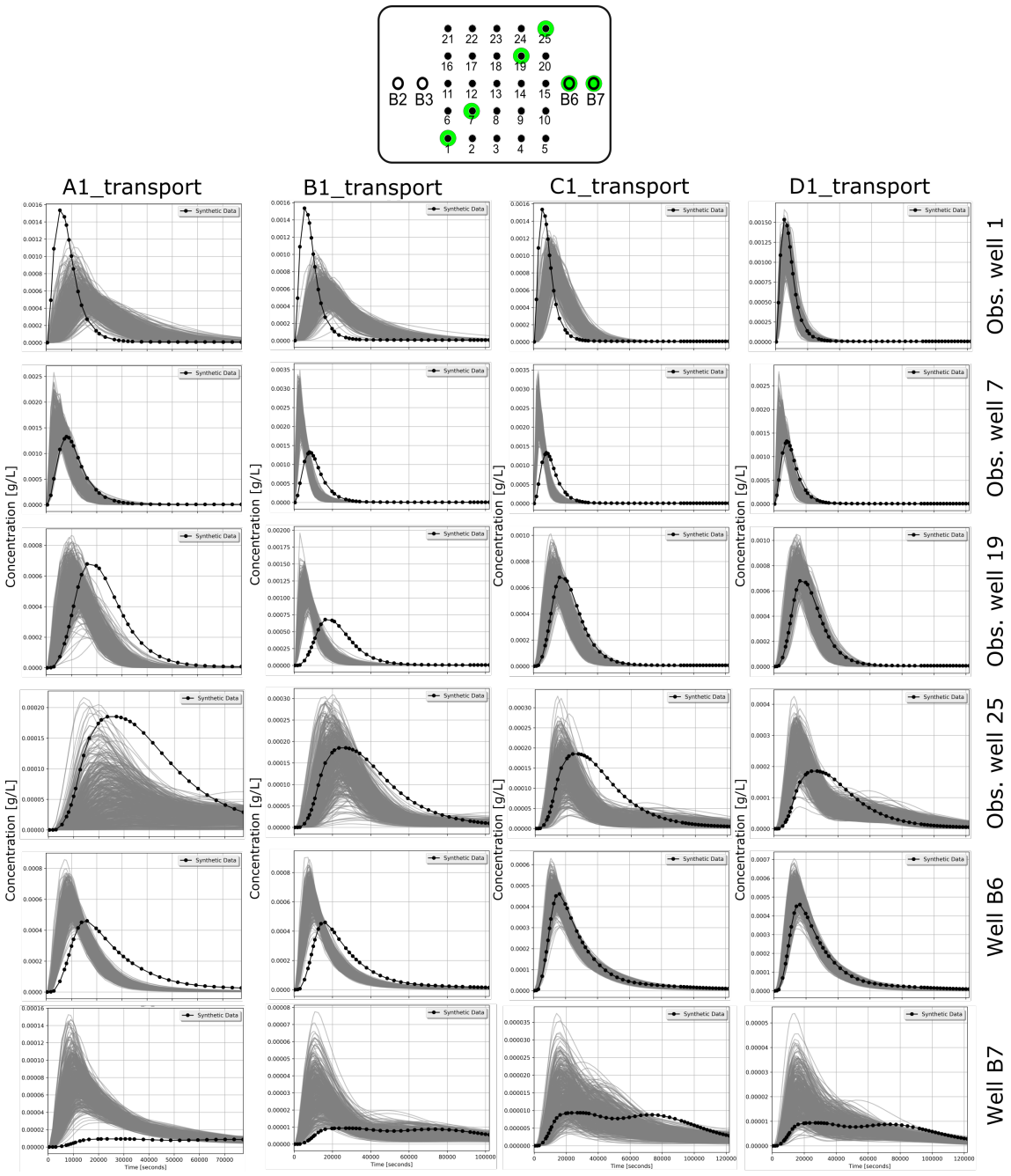


Figure 6.14: Breakthrough curves simulated after the final update step with tracer data from test 1syn-Tr. Gray lines: breakthrough curves of each individual realization of the updated ensemble; Black lines: reference breakthrough curves.

Test 2syn-Tr

Test 2syn-Tr is the second tracer test of the synthetic tracer tomography. The steady flow of this test corresponds to the reverse flow field generated in test 2syn-Fl. Tracer injection was performed at well B6, and extracted at wells B2 and B3.

The assimilation of data from test 2syn-Tr was performed with two different scenarios (Tables 6.14) and 59 update steps. Both scenarios used the updated ensemble from test 1syn-Tr scenario `D1_transport` as initial ensemble of parameters (mean and variance of effective parameters shown in Table 6.15). To follow the effect of rank-based transformations during the assimilation of transient concentration data, normal-score transformations were applied in scenario `A2_transport`, and excluded in scenario `B2_transport`. All additional settings correspond to those of scenario `D1_transport`.

Table 6.14: Settings of the two scenarios generated for the assimilation of concentration data from test 2syn-Tr. ✓ and × indicate if the setting was activated or not; rel.: implies the application of an error percentage relative to the current measurement value; max & min: maximum and minimum values allowed for the measurement error.

Test 2syn-Tr						
Scenario ID	Update scheme	Data transf.	Damping factor β			Measurement error σ_ϵ (g/l)
`A2_transport`	standard EnKF	✓	✓	0.1	rel.	$10\,\%$, max=5×10^{-3},min=5×10^{-6}
`B2_transport`	Restart EnKF	×	✓	0.1	rel.	$10\,\%$, max=5×10^{-3},min=5×10^{-6}

Table 6.15: Initial mean(μ) and variance (σ^2) of the effective parameters included in the update. Values taken from the results of test 1syn-Tr, scenario `D1_transport`. n_m: mobile zone porosity; n_{im}: immobile zone porosity; λ_{mt}: first-order mass transfer coefficient; α_l: longitudinal dispersivity. Variances are expressed for the natural logarithm (Ln) of the corresponding parameter.

Test 2syn-Tr Effective transport parameters included in the update (values taken from results of test 1syn-Tr scenario `D1_transport`)							
n_m (%)		n_{im} (%)		α_l (m)		λ_{mt} (s^{-1})	
μ	$\sigma^2_{Ln(nm)}$	μ	$\sigma^2_{Ln(nim)}$	μ	$\sigma^2_{Ln(\alpha)}$	μ	$\sigma^2_{Ln(\lambda)}$
10.5	8×10^{-3}	20.2	5×10^{-3}	1.22	0.24	1.06×10^{-9}	0.102

Fig. 6.15 shows the ensemble mean of log-hydraulic conductivity after three different update steps of the assimilation of concentration data from test 2syn-Tr. The reference log-K field is presented at the top of Fig. 6.15.

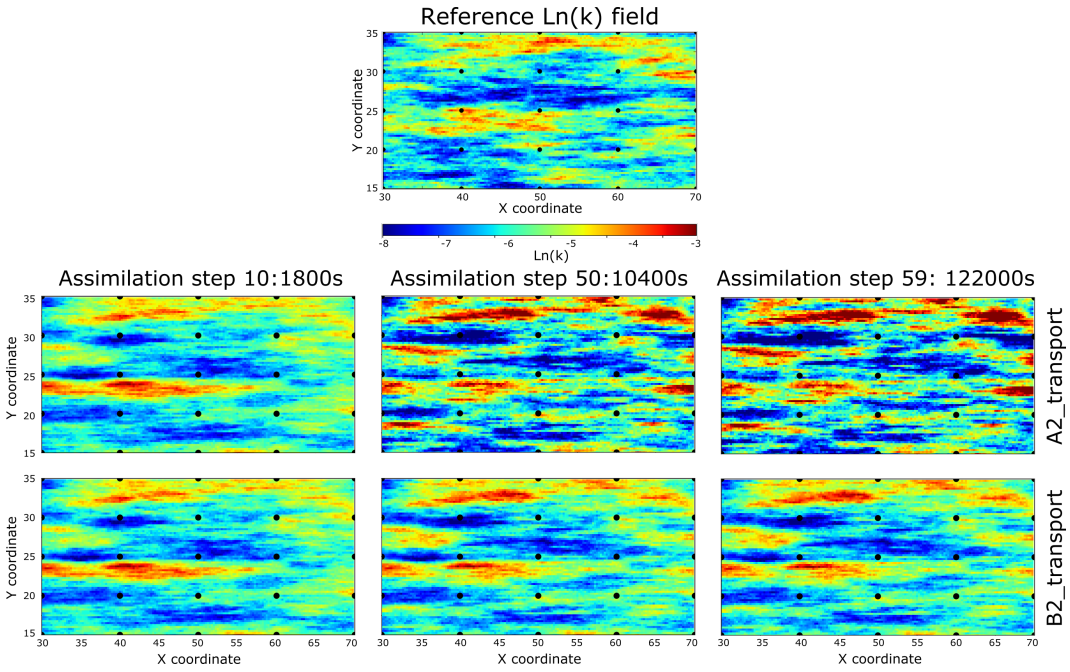


Figure 6.15: Ensemble mean log-hydraulic conductivity fields for the assimilation of concentration data from test 2syn-Tr after 10, 50 and 59 assimilation steps. The reference log-hydraulic conductivity field is shown at the top.

Despite different settings of the filter, the ensemble mean of log-hydraulic conductivity is similar in both scenarios after 10 update steps, most likely due to the small residuals estimated during the assimilation of very small concentrations. The negative effect of rank-based transformations is confirmed with a visual inspection of the updated ensemble mean of log-K for scenario `A2_transport`, after 50 and 59 update steps (see Fig. 6.15), and by the $60\,\%$ increment of the AAE with respect to the initial value of 0.75 $\log(\mathrm{m\,s^{-1}})$ (Table 6.16).

Better results were observed for scenario `B2_transport`, without normal-score transformation. A visual comparison suggests that details of the reference field are accentuated in the ensemble mean of log-hydraulic conductivity after 59 update steps. It appears that the Restart EnKF is able to correct some of the high log-hydraulic conductivity values previously observed between X=$(30,50)$ m and Y=$(20,25)$ m. The low conductivity zones at the center and bottom of the reference field are clearly present in the estimated ensemble mean. The better estimation of parameters for scenario `B2_transport` is confirmed by a $7\,\%$ reduction of the AAE of log-hydraulic conductivity with respect to the initial value (Table 6.16). Estimated

values of AESD also suggest a stabilization in the estimation of ensemble variability. The latter is also appreciated in Fig. 6.16 with no considerable changes of the ensemble variance, after 59 assimilation steps.

Table 6.16: Average absolute error (AAE) and Average Ensemble Standard Deviation (AESD) of log-hydraulic conductivity at the end of the assimilation of data from test 2syn-Tr, and Root Mean Square Error (RMSE) and Nash-Sutcliffe coefficient (NS) based on breakthrough curves simulated with the updated ensemble of parameters.

	Test 2syn-Tr				
	Model Parameters (Lnk)			Model States	
Scenario ID	AAE ($\log(\mathrm{m\,s^{-1}})$)	AESD ($\log(\mathrm{m\,s^{-1}})$)	AESD/AAE (–)	RMSE(ymod) ($\mathrm{g\,l^{-1}}$)	NS-coeff. (–)
`A2_transport`	1.2	0.42	0.35	1.56×10^{-4}	0.73
`B2_transport`	0.7	0.42	0.6	1.13×10^{-4}	0.90

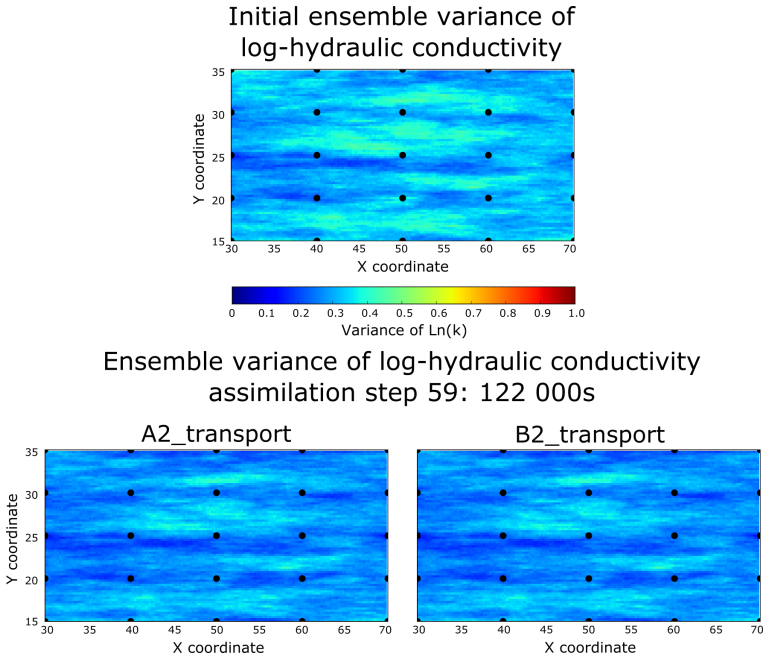


Figure 6.16: Top: Initial ensemble variance of the log-hydraulic conductivity fields used in scenarios `C2_flow`, `E2_flow` and `F2_flow`; Bottom: ensemble variance of log-hydraulic conductivity after 59 assimilation steps.

Table 6.17 shows the final mean and variances of the effective parameters included in the update. Although considerable variations between both scenarios were not observed, all mean effective parameter values from scenario `B2_transport` were closer to the initial and reference

values.

Table 6.17: Updated mean (μ) and variance (σ^2) of the effective parameters included in the update during assimilation of tracer data from test 2syn-Tr, after 59 update steps. n_m: mobile zone porosity; n_{im}: immobile zone porosity; λ_{mt}: first-order mass transfer coefficient; α_l: longitudinal dispersivity. Variances are for the natural logarithm (Ln) of the corresponding parameter.

	Test 2syn-Tr							
Scenario	n_m (%)		n_{im} (%)		α_l (m)		λ_{mt} (s^{-1})	
ID	μ	$\sigma^2_{Ln(nm)}$	μ	$\sigma^2_{Ln(nim)}$	μ	$\sigma^2_{Ln(\alpha)}$	μ	$\sigma^2_{Ln(\lambda)}$
Initial	10.5	8×10^{-3}	20.2	5×10^{-3}	1.22	0.24	1.06×10^{-9}	0.102
`A2_transport`	11.7	6×10^{-3}	19.9	5×10^{-3}	1.7	0.21	1.07×10^{-9}	0.102
`B2_transport`	10.4	6×10^{-3}	20.2	5×10^{-3}	1.3	0.20	1.07×10^{-9}	0.101

Fig. 6.17 shows the breakthrough curves obtained at 6 different observation points (including both extraction wells), with a complete simulation of test 2syn-Tr. The parameters used for the simulation correspond to the updated ensemble of parameters after 59 assimilation steps. In general, improved simulations can be observed for scenario `B2_transport`, without normal-score transformation. The estimated RMSE between updated and reference breakthrough curves equals $1.56 \times 10^{-4}\,\mathrm{g\,l^{-1}}$ for scenario `A2_transport`, with normal-score transformation. This implies an $\sim 11\,\%$ increment with respect to the RMSE obtained for test 1syn-Tr (scenario `D1_transport`). In contrast, the RMSE value of $1.13 \times 10^{-4}\,\mathrm{g\,l^{-1}}$ represents a $\sim 20\,\%$ reduction with respect to the same initial value. The NS-coefficient decreased for scenario `A2_transport` (with normal-score transformation) to a value of 0.73, whereas it was kept constant (0.90) for scenario `B2_transport`, without transformation (see Table 6.16). Problems in simulating the breakthrough curves at wells 1, B2 and B3 can be attributed to a larger effect of the spurious estimation of parameters outside the area of interest.

6.3 Chapter Summary

Chapter 6 described the application of the EnKF for the estimation of aquifer parameters using data from a synthetic tracer-tomography experiment. First, numerical simulations of a synthetic tracer-tomography test were performed to generate a reference dataset based on a known log-hydraulic conductivity field. The tomographic experiment consisted of two pumping and two tracer tests with a nested-cell flow regime. To estimate the spatial distribution of log-hydraulic conductivity and (homogeneous) aquifer storativity S_o, the EnKF was sequentially

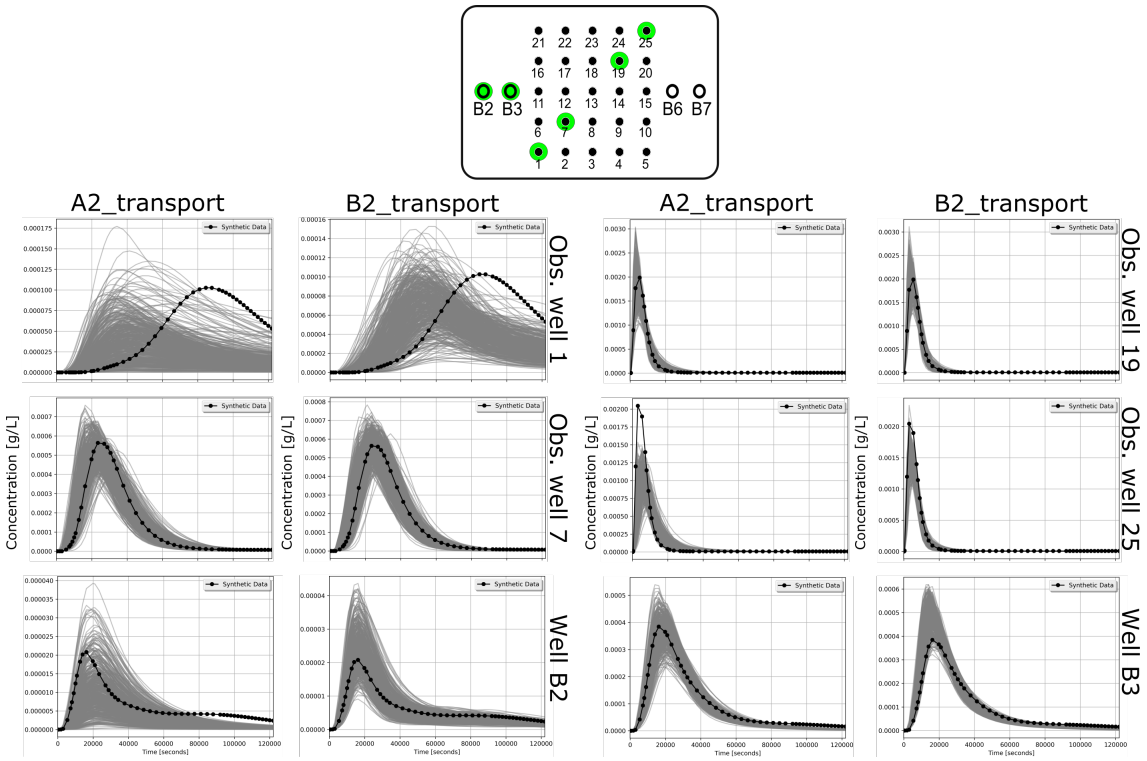


Figure 6.17: Breakthrough curves simulated after the final update step with tracer data from test 2syn-Tr. Model simulations correspond to those of test 2syn-Tr. Gray lines: breakthrough curves of each individual realization of the updated ensemble; Black lines: reference breakthrough curves.

applied to the two drawdown datasets. The ensemble of parameters conditioned to drawdown data was used as initial ensemble for the assimilation of concentration data. Homogeneous porosity of the mobile and immobile regions n_m and n_{im}, longitudinal dispersivity α_l, and a first-order mass transfer coefficient λ_{mt} were also updated.

Fig. 6.18 summarizes the evolution of the ensemble mean (left) and variance (right) of log-hydraulic conductivity throughout the assimilation of the synthetic tracer tomography data. The results correspond to those scenarios with the best performance observed after all update steps were finalized for each test. The reference log-hydraulic conductivity field is shown at the top of the figure, followed by the initial ensemble mean and ensemble variance of log-hydraulic conductivity. The corresponding absolute average error (AAE) estimated for each test, is also included in Fig. 6.18.

Estimated parameters based on hydraulic-head data already contain the main features of the reference field, and a considerable reduction in ensemble variance is achieved (Fig. 6.18, `F1_flow` and `F2_flow`). However, the fields are relatively smooth in comparison to the refer-

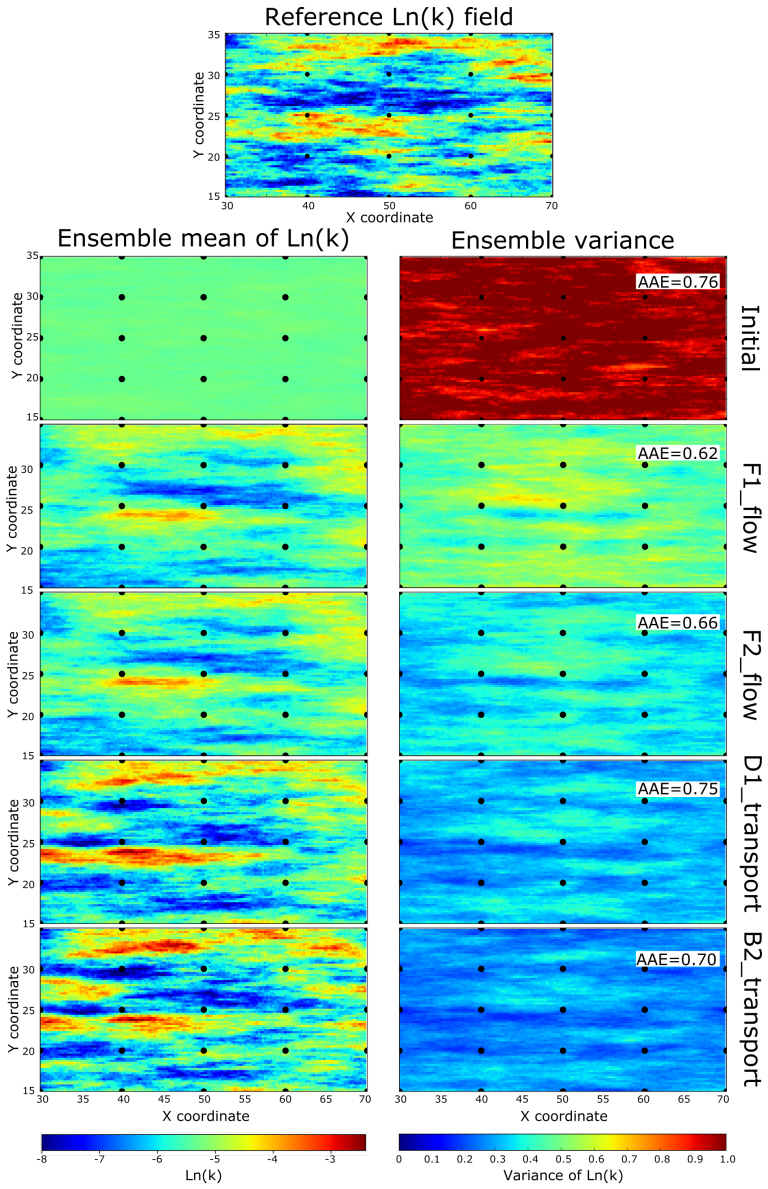


Figure 6.18: Evolution of the ensemble mean (left) and variance (right) of log-hydraulic conductivity during the sequential assimilation of data from the individual hydraulic and tracer tests of the synthetic tracer tomography. Top: reference log-hydraulic conductivity field.

ence. With the inclusion of concentration data, the spatial structure of the ensemble mean of log-K is accentuated and the ensemble variance of parameters considerably reduced. The average absolute error, which is a measure of the deviation from the reference field, is reduced by $\sim 20\,\%$ after drawdown data from both tests have been assimilated. The AAE increased (AAE=0.75) after the assimilation of data from the first tracer test (Test 1syn-Tr), mainly caused by the high log-conductivity zone located between X=(30,50) m and Y=(20,25) m.

However, the AAE of log-K is again reduced (AAE $= 0.7$) after the assimilation of concentration data from the second tracer test (test 2syn-Tr), suggesting an improvement in the estimation of parameters. From Fig. 6.18 (bottom left), it can be observed that the inclusion of concentration data from the second test helps correcting some of the extreme values of the high log-conductivity zone mentioned before, while the structure of the field is further accentuated.

To show that parameters conditioned to tracer data still produce acceptable results for groundwater flow, test 1syn-Fl was simulated using as input parameters the final ensemble mean. Fig. 6.19 compares the simulated drawdown to the reference curves at eight different observation points. After a visual inspection, it is concluded that the ensemble mean of parameters conditioned to tracer data is still capable of producing reasonable flow simulations. This is confirmed by a NS-coefficient and RMSE between all reference and simulated drawdowns (from test 1syn-Fl) of 0.9 and $0.025\,\mathrm{m}$, respectively. Difficulties in reaching the steady state values at observation wells 6 and 7 can be attributed to the high log-conductivity zone mentioned below. Discrepancies in drawdown simulations at well 17 are caused by the low log-hydraulic conductivity values surrounding the observation well, and which do not appear in the reference log-hydraulic conductivity field.

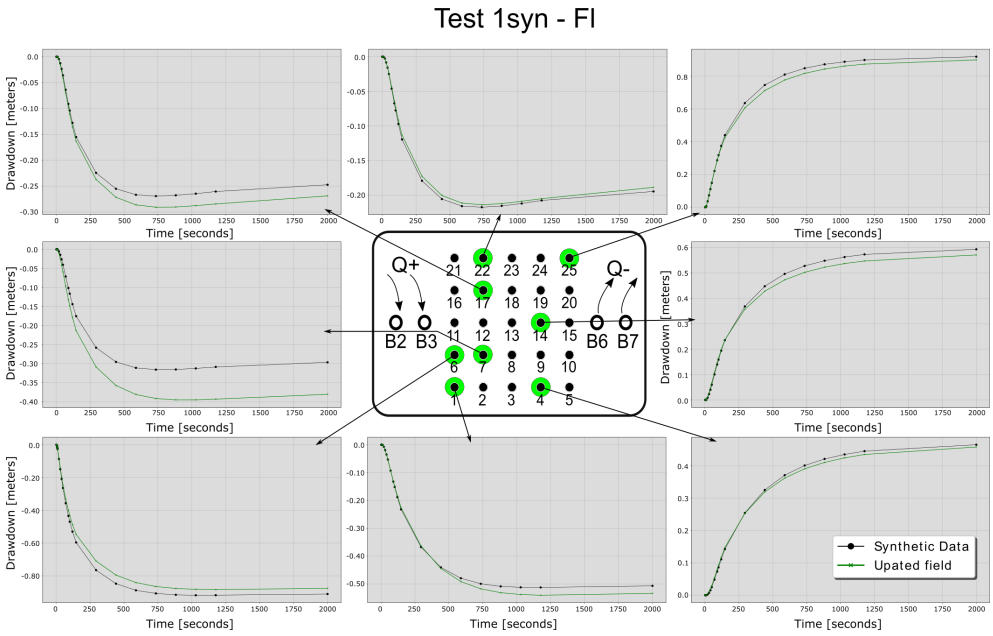


Figure 6.19: Drawdown curves at eight different observation points, for a simulation of test 1syn-Fl with the ensemble mean of parameters estimated after all data from the synthetic tracer tomography was assimilated.

The estimation of effective parameter values strongly depends on the initial (mean) value

used to generate the random realizations. This may affect the application of the EnKF in real applications where high uncertainties of parameters, such as storativity, porosity, and dispersivity are common.

The selection of the damping factor β to reduce parameter perturbations after each update step, was based on previous studies. Further research is required to define the optimal β for tracer tomography applications. Normal-score transformation improves the estimation of parameters based on drawdown data, but is not recommended for the assimilation of transient concentration records, mainly due to their strong nonlinear dependency on log-hydraulic conductivity.

Finally, it was found that the update scheme of the classical EnKF leads to mass balance errors and inconsistencies between hydraulic parameters and concentrations. The update scheme of the Restart EnKF allows to reinitialize the steady flow field and recalculate the plume distribution after each update step, being a suitable option for the assimilation of transient concentration data. Results of the sequential assimilation of drawdown and concentration data demonstrate the benefits of the information provided by tracer tomography for parameter estimation and hence, improved groundwater flow and solute transport simulations.

Chapter 7

Parameter Estimation: Real Data

This chapter presents the methodology to estimate the hydraulic parameters of the aquifer at the Hydrogeological Research Site Lauswiesen, inverting drawdown and breakthrough curves measured during the field tracer tomography experiment presented in this work. The processed field dataset was introduced in Section 5.2 and the design of the field experiments has been described in Chapter 4. Here, the methodology and results of the parameter estimation are presented and discussed. The chapter includes a comparison of the current results with previous studies performed at the field site. All HydroGeoSphere models discussed in this chapter are provided in an electronic appendix (Appendix: Models/real).

7.1 Numerical Simulations

The field tracer tomography experiment presented in this work was simulated with a three-dimensional numerical model using HydroGeoSphere. The model grid consisted of $420\,693$ nodes forming $384\,000$ regular block elements, distributed over 12 layers ($\Delta x = \Delta y = \Delta z = 0.5\,\mathrm{m}$). The bottom layer was set to $0\,\mathrm{m}$, corresponding to the bottom of the aquifer, and the top was defined as a confined layer at a constant elevation of $6\,\mathrm{m}$. Constant-head boundary conditions were defined at the left and right limits of the domain, whereas no-flux (Neumann) boundaries were defined at the front, back, bottom, and top faces of the model (Fig. 7.1). In all simulations, the constant value at the left was set to $10\,\mathrm{m}$, and the constant value at the right was adjusted according to the mean ambient hydraulic gradient estimated with manual water level measurements taken before each test. Although the hydrogeological setting of the aquifer under investigation corresponds to an unconfined system, previous investigations have shown that the confined assumption is acceptable for the time scales of the current

experiments (e.g., Sanchez-León et al. 2016). Dual-domain transport was assumed for the tracer-test simulations.

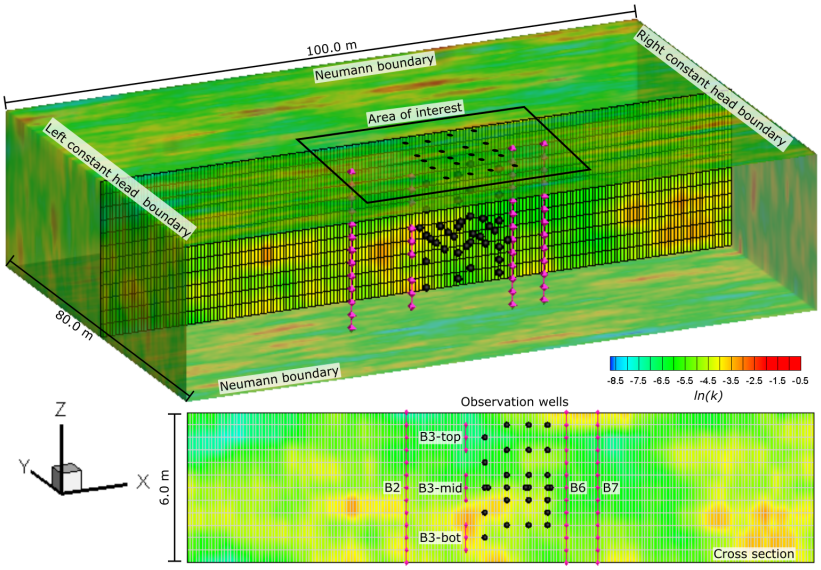


Figure 7.1: Three-dimensional model used to estimate hydraulic parameters of the aquifer at Lauswiesen. Top: 3D perspective of the model grid with an exemplary log-hydraulic conductivity field, the constant head and no-flux (Neumann) boundary conditions. Bottom: cross section with details of the distribution of extraction/injection (purple diamonds) and observation (black dots) wells. The black rectangle at the center refers to the area of interest for parameter estimation.

The extension of the domain was chosen such that pumping-induced effects were minimized at the boundaries, while keeping model run times acceptable. The information provided by the available wells is concentrated in the central region of the model, identified as *area of interest* in Fig. 7.1. This area is delimited by the XY points P1=$(34,54)$ m, P2=$(34,66)$ m, P3=$(66,54)$ m and P4=$(66,66)$ m, and covers the entire thickness of the model domain. The injection and extraction wells B2, B6 and B7 were defined as fully screened wells, covering the whole aquifer thickness. The three isolated sections generated in injection well B3 (Fig. 4.8 and 4.7) were implemented by creating three different wells identified as B3-top, B3-mid and B3-bot for the top, middle and bottom sections, respectively (see cross section in Fig. 7.1). All w-wells were implemented as observation wells covering the whole thickness of the aquifer, however, hydraulic head changes were monitored only at the middle. All observation ports of the multilevel wells (cmt-wells) were defined as point observations at the corresponding elevation.

As mentioned in Chapter 4, two tests were performed for each of the three tracer injection

sections of well B3. The idea behind the two tests per injection section was the integration of the measured data from both tests in a single set, obtaining a broader spatial coverage. However, small differences in the natural gradient and injection/extraction rates between each test-pair were strongly reflected in the data. Consequently, the measurements could not be treated as a single dataset and six groundwater flow and their corresponding transport models had to be constructed (note the missing processed dataset for hydraulic test 2a due to problems with the data-logger). The numerical models contained information about the mean natural gradient, injection and extraction rates, injected mass of tracer and active observation wells of the corresponding real tests. For clarity, the main settings of the individual tests are repeated in Table 7.1.

Table 7.1: Experimental setup adopted for each individual tracer test of the solute tracer tomography experiment. These settings were integrated in the 3-D models generated for parameter estimation with the EnKF. Top, Mid, Bot refer to the top, middle and bottom sections generated at well B3. Positive flow rates represent the injection of water, whereas negative flow rates represent extraction rates. Mean I refers to the mean natural hydraulic gradient based on manual measurements of groundwater levels.

	Inj.depth (m)	Tracer	Applied Flow Rates (ls^{-1})						Mean I
Test	(Level)	Mass(g)	B2	B3Top	B3Mid	B3Bot	B6	B7	(%)
1a	4.0-5.0 (Top)	1.10	5.2	1.6	0.9	0.6	-2.0	-9.0	0.34
1b	4.0-5.0 (Top)	1.15	4.8	2.0	0.87	0.6	-2.4	-9.0	1.15
2a	5.5-6.5 (Mid)	1.10	5.2	1.8	0.9	0.6	-2.5	-9.0	0.17
2b	5.5-6.5 (Mid)	1.20	4.9	2.1	0.81	0.6	-2.0	-9.0	0.49
3a	7.0-8.0 (Bot)	1.10	5.2	1.8	0.9	0.6	-2.5	-9.0	0.20
3b	7.0-8.0 (Bot)	1.15	4.9	2.2	0.82	0.6	-2.0	-8.7	0.23

Like in the synthetic case, flow and transport simulations were performed separately. First, transient groundwater flow was simulated with the settings of each real test (Table 7.1). For transport, the groundwater flow velocity distribution was calculated with an initial steady-flow simulation, followed by the injection of tracer at the corresponding location. In the following, flow and transport simulations of each test in Table 7.1, are differentiated by an extra Fl or Tr label, respectively (e.g. test 2a-Tr corresponds to the transport simulation of test 2a, in which the tracer was injected at the middle section of well B3).

Fig. 7.2 shows an example of a 3-D simulation of the hydraulic and tracer test 2a, with tracer injection at the middle section of well B3. Transient groundwater flow was simulated until a steady flow was achieved (Fig. 7.2 top) (or until the time of the last available field measurement of the test). Tracer test was then simulated independently (Fig. 7.2 bottom),

with a previous steady-flow simulation to calculate groundwater flow velocities.

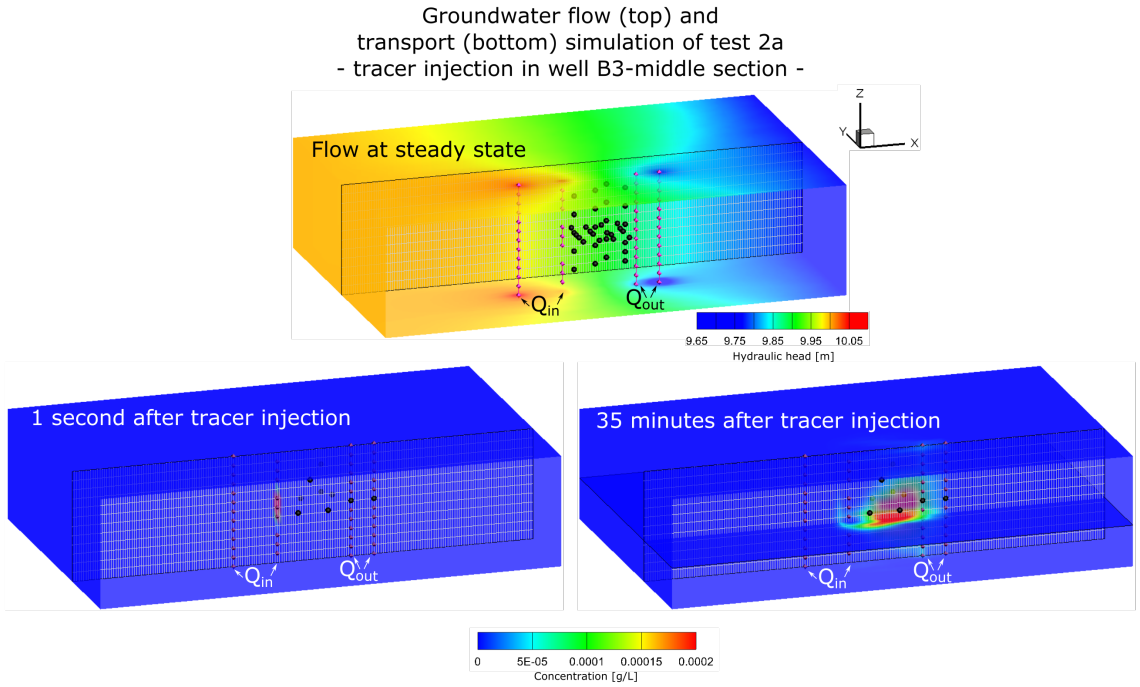


Figure 7.2: Example of a groundwater flow (top) and tracer test (bottom) simulation with the three-dimensional model constructed for the aquifer at Lauswiesen. The tracer test corresponds to test 2a-Tr, with tracer injection at the middle section of well B3. Q_{in} and Q_{out}: injection and extraction of water, respectively. Purple diamonds: injection and extraction wells; Black dots: observation wells.

7.2 Parameter Estimation

For the estimation of aquifer parameters based on real data, the EnKF settings were defined following the results of the synthetic case (as listed in Table 7.2). The classical update approach of the EnKF and data transformation were applied for the assimilation of drawdown data, whereas the restart EnKF scheme was implemented for concentration data without transformation.

Damping factors β of 0.6 and 0.1 were applied during the assimilation of drawdown and breakthrough curves, respectively. A fixed measurement error was added to the modeled drawdowns, and an error percentage relative to the current modeled concentration was implemented for the assimilation of transport data.

Table 7.2: Settings of the EnKF applied during sequential assimilation of real drawdown and concentration data. Standard and restart correspond to the type of update scheme implemented. ✓and × indicate if the setting was activated or not. *Absolute* and *relative* imply the application of a fixed measurement error or an error percentage relative to the current measurement value.

Model Type	Update steps	Update scheme	Data transf.	Damping factor β	Measurement error σ_ϵ (m)	
Flow	27	standard	✓	0.6	absolute	5×10^{-3} m
Transport	34&27	restart	×	0.1	relative	10 %

The spatially distributed log-hydraulic conductivity fields were generated with the same spectral methods applied in the synthetic study (Dietrich and Newsam, 1993). The geostatistical parameters used to generate the 3-D log-K fields for an ensemble of 500 realizations are shown in Table 7.3.

Table 7.3: Geostatistical parameters used to generate the initial ensemble of log-hydraulic conductivity fields in the 3-D model. μ_k: geometric mean of hydraulic conductivity; $\sigma^2_{Ln(k)}$: variance of log-hydraulic conductivity; l_l: longitudinal correlation length; l_t: transverse correlation length; l_v: vertical correlation length.

Spatially distributed hydraulic conductivity geostatistical model: exponential				
μ_k (m s^{-1})	$\sigma^2_{Ln(k)}$	l_l (m)	l_t (m)	l_v (m)
8.0×10^{-3}	1.0	10	1	0.5

During assimilation of drawdown data, a homogeneous aquifer storativity (S_o) was added to the ensemble of parameters to update. The homogeneous mobile and immobile porosities (n_m and n_{im}), mass transfer coefficient (λ_{mt}) and longitudinal dispersivity (α_l) were included in the update during assimilation of tracer concentrations. Table 7.4 shows the statistical parameters used to generate normal distributed random realizations of the ensemble. Initial mean values of (effective) aquifer parameters were chosen based on the analysis of the field drawdown and breakthrough curves presented in Section 5.3, and adjusted by evaluating the initial model simulations.

During adjustment, it was found that a larger S_o was needed to resemble the hydraulic responses observed in the field data. Disagreements with values of S_o obtained with type-curve analysis are attributed to the early-time misfit of the Theis model to the drawdown curves (Fig. 5.15), producing unreliable estimates of S_o. Total porosity was set to 30 %,

Table 7.4: Mean (μ) and variance (σ^2) applied to generate different realizations of additional aquifer parameters. S_o: storativity; n_m: mobile porosity; n_{im}: immobile porosity; α_l: longitudinal dispersivity; λ_{mt}: mass transfer coefficient between mobile and immobile zones; Ln: natural logarithm.

(Homogeneous) Parameters included in the update									
Flow		Transport							
S_o (m^{-1})		n_m (%)		n_{im} (%)		α_l (m)		λ_{mt} (s^{-1})	
μ	σ^2_{LnSo}	μ	$\sigma^2_{Ln(nm)}$	μ	$\sigma^2_{Ln(nim)}$	μ	$\sigma^2_{Ln(\alpha)}$	μ	$\sigma^2_{Ln(\lambda)}$
5.0×10^{-2}	0.3	10	0.2	20	0.2	1	0.5	1×10^{-8}	1.0

which is a value between the results from fitting the semi-analytical solution for the dual domain advection-dispersion-equation to the breakthrough curves observed at extraction well B6 (Table 5.4), and the value reported by Sanchez-León et al. (2016) from the calibration of a numerical model with hydraulic and tracer data. Finally, it was observed that the mass-transfer coefficient λ_{mt} obtained with the same semi-analytical solution ($2.6 \times 10^{-11}\,\mathrm{s}^{-1}$) was too small to produce relevant interactions between the mobile and immobile zones, and therefore it was set equal to $1 \times 10^{-8}\,\mathrm{s}^{-1}$ and a relatively large variability.

The initial ensemble mean of log-hydraulic conductivity and associated ensemble variance are presented in Figure 7.3, with a 3-D view of the numerical model containing the following:

(i) Delineation of the area of interest,

(ii) Location of the injection/extraction (purple diamonds) and active observation (black spheres) wells,

(iii) Position of three slices defined in the XY plane at a Z elevation of $1\,\mathrm{m}$, $3\,\mathrm{m}$ and $5\,\mathrm{m}$,

(iv) Position of three slices in the XZ plane at Y equals to $55\,\mathrm{m}$, $60\,\mathrm{m}$ and $65\,\mathrm{m}$.

(Note: The spatial coverage of the six slices corresponds only to the area of interest)

From the full dataset of the field tracer tomography experiment, drawdown data only from tests 3b-Fl, 3a-Fl and 1b-Fl, and breakthrough curves from tests 3b-Tr, 3a-Tr, 2b-Tr and 2a-Tr were used in the parameter estimation. Data from the additional tests were used to evaluate the prediction capabilities of the updated parameters. As for the synthetic study, concentration data were used for updating parameters only after drawdown data were assimilated. The assimilation order was the following:

Test 3b-Fl→Test 3a-Fl→Test 1b-Fl→Test 3b-Tr→Test 3a-Tr→Test 2b-Tr→Test 2a-Tr

Illman et al. (2008) showed that the order of assimilation may affect the final estimate,

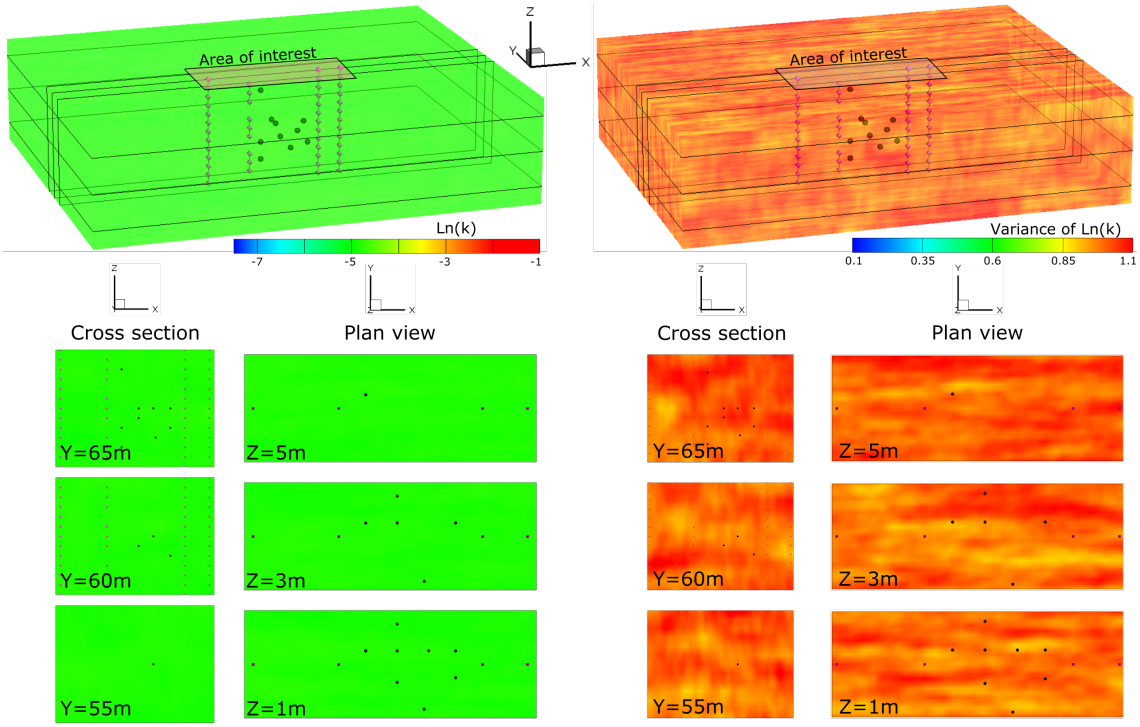


Figure 7.3: Initial ensemble mean (left) and variance (right) of log-hydraulic conductivity used for the estimation of aquifer parameters via assimilation of real data from the field tracer tomography experiment. Purple diamonds: injection/extraction wells; Black spheres: observation wells; Top: 3-D perspective of the numerical model; Bottom: XZ and XY slices with details of the spatial distribution of the ensemble mean and variance of log-K at the area of interest.

however, its effects were not assessed in this study. In the Results section of this chapter, the evolution of the parameter estimation process is presented in the same order of assimilation and with the same 3-D perspective view and slices used in Fig. 7.3, with the corresponding ensemble mean and variance of log-hydraulic conductivity.

7.3 Results

7.3.1 Sequential Assimilation of Drawdown Data

Figs. 7.4 and 7.5 show the evolution of the ensemble mean and associated variance of log-hydraulic conductivity, throughout the sequential assimilation of drawdown data from tests 3b-FI, 3a-FI and 1a-FI. In addition to the 3-D plots, the XZ (cross section) and XY (plan view) slices of the ensemble mean and variance of log-hydraulic conductivity are presented in Figs. 7.4 and 7.5. The area covered by the slices corresponds only to the area of interest delimited in the 3-D plot.

The spatial structure of log-hydraulic conductivity is gradually accentuated after the sequential assimilation of data from each test. Hydraulic conductivity values ($Lnk < -6$ or $K < 2.4 \times 10^{-3}\,\mathrm{m\,s^{-1}}$) that are smaller than the original ensemble mean ($\mu_k = 8.0 \times 10^{-3}\,\mathrm{m\,s^{-1}}$) dominate the upper portion of the area of interest, whereas values closer to the mean, or larger, prevail in the middle and lower regions ($Z \leq 3\,\mathrm{m}$). These findings disagree with previous studies, in which a more conductive upper layer overlaying a less conductive and more heterogeneous lower layer was found (e.g., Lessoff et al. 2010; Doro et al. 2015). However, similar results to those of Fig. 7.4 were found by Sanchez-León et al. (2016) who calibrated a three-dimensional model with drawdown data from a hydraulic-tomography experiment.

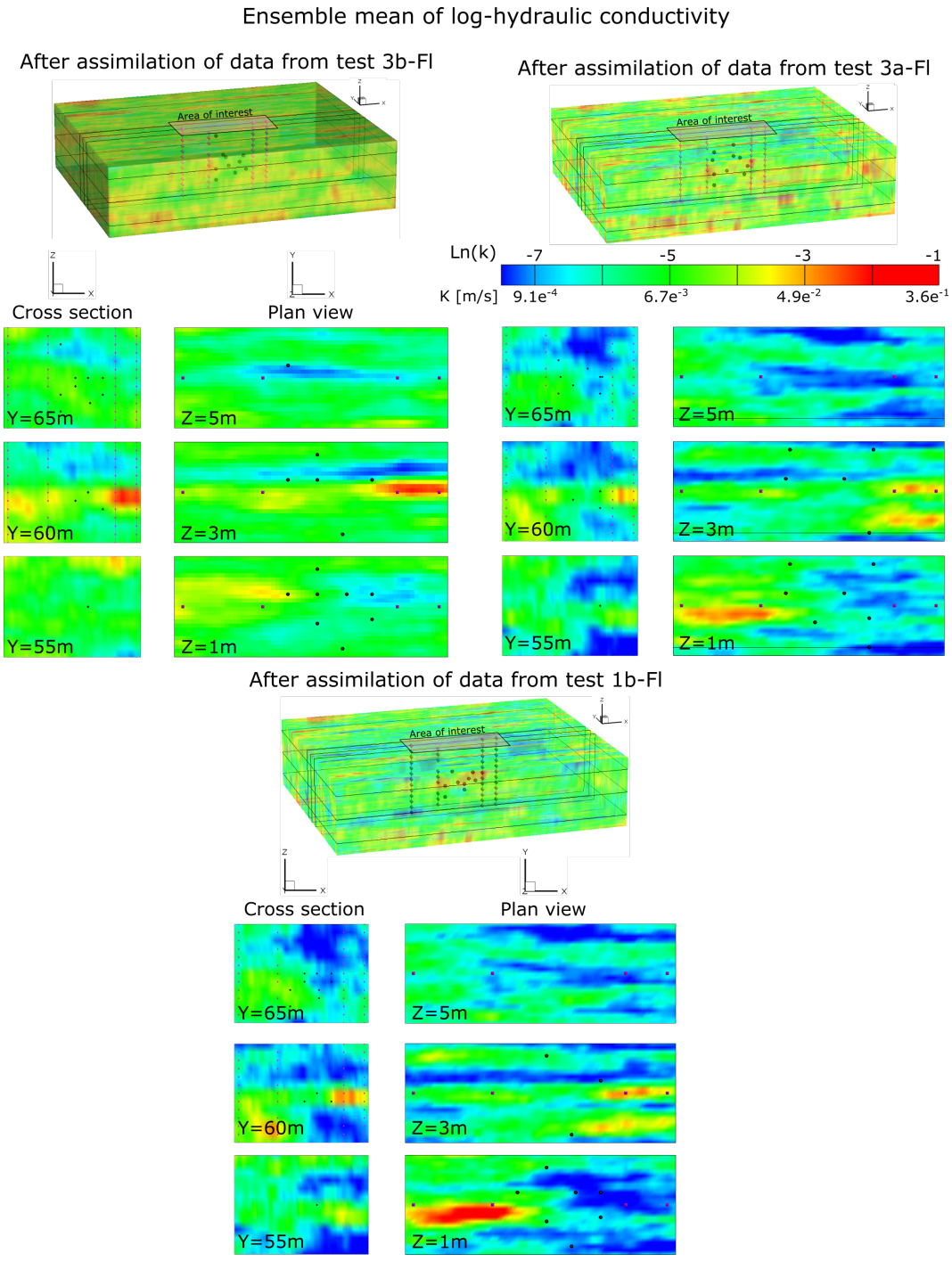


Figure 7.4: Ensemble mean of log-hydraulic conductivity estimated after assimilation of drawdown data from tests 3b-Fl, 3a-Fl and 1b-Fl. XZ (cross section) and XY (plan view) slices of the ensemble mean of log-K at the area of interest are also plotted for each test. Purple diamonds: injection/extraction wells; Black spheres: observation wells.

The high-K values developed in the vicinity of the extraction wells (middle-right area) and injection wells (bottom left) after assimilating data from tests 3b-FI and 1b-FI, respectively, are most likely an artifact caused by the lack of information outside the grid of observation wells. Relevant ensemble variance reductions were not observed after the assimilation of data from test 3b-FI (Fig. 7.5 left column). This is confirmed quantitatively by a reduction in the average ensemble standard deviation (AESD), for the entire model domain, of $\sim 1.3\,\%$ with respect to the initial value (Table 7.5). A similar behavior was observed for the ensemble variance of log-hydraulic conductivity (σ^2_{LnK}), with a reduction of only $3\,\%$ after all update steps for test 3b-FI were performed. Values of AESD and σ^2_{LnK} estimated for the entire model domain are largely affected by the spurious correlations developed throughout parameter updating in regions with no data. Therefore, AESD and σ^2_{LnK} values considering only those elements within the area of interest are also presented in Table 7.5. An AESD and σ^2_{LnK} reduction of $\sim 3.8\,\%$ and $\sim 7.1\,\%$, respectively, are achieved if only the area of interest is considered.

The largest ensemble variance reduction was observed after the assimilation of drawdown from test 3a-FI, in particular in the middle region of the area of interest (see Fig. 7.5). This is again qualitatively confirmed by a decrease in AESD in the entire model domain and area of interest of $\sim 2.5\,\%$ and $\sim 6.3\,\%$, with respect to the initial values. Considering only the area of interest, final AESD and σ^2_{LnK} reductions of $\sim 8.9\,\%$ and $\sim 16\,\%$, respectively, were achieved at the end of the assimilation of data from test 1b-FI.

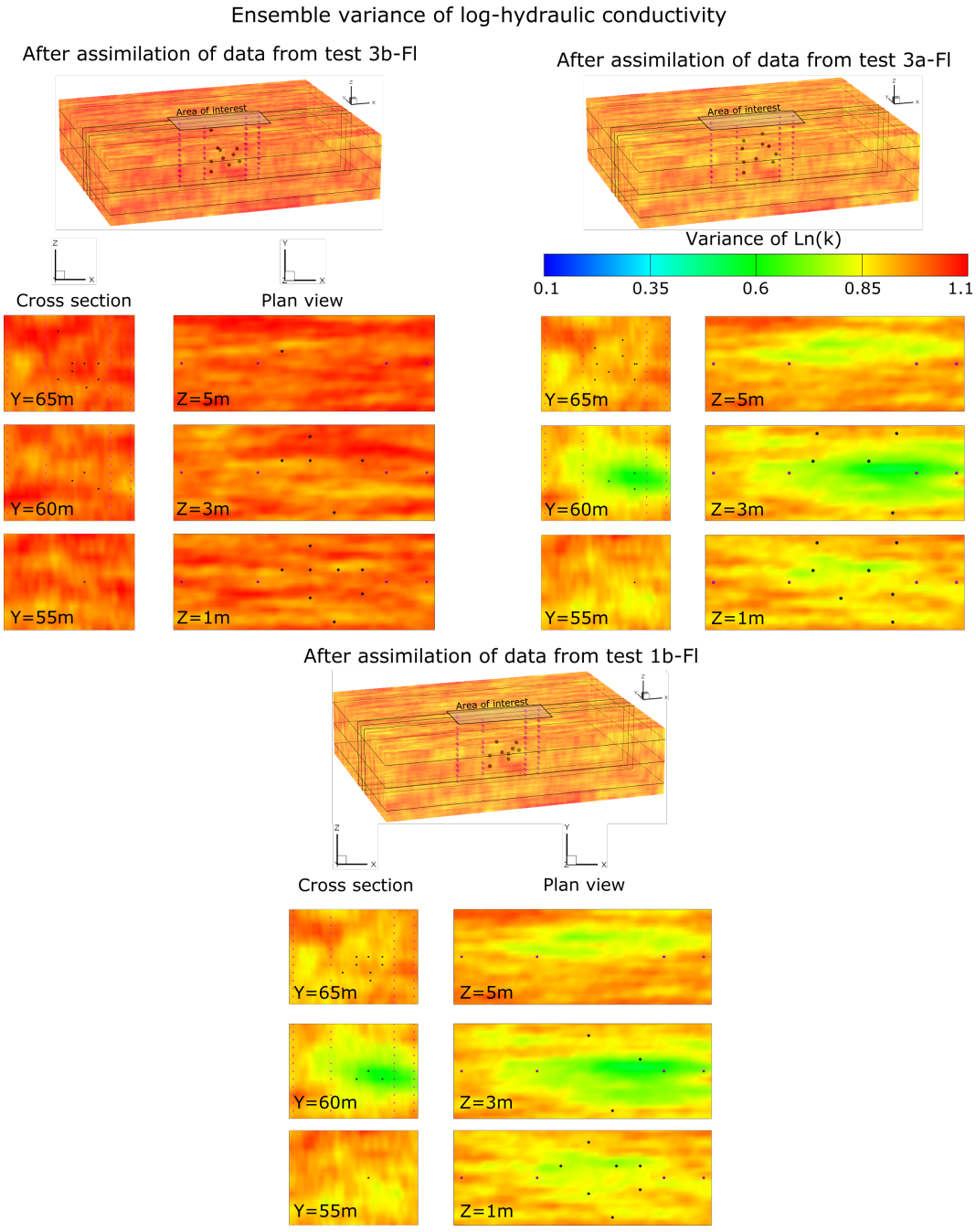


Figure 7.5: Ensemble variance of log-hydraulic conductivity estimated after assimilation of drawdown data from tests 3b-Fl, 3a-Fl and 1b-Fl. XZ (cross section) and XY (plan view) slices with the ensemble variance of log-K at the area of interest are also plotted for each test. Purple diamonds: injection/extraction wells; Black spheres: observation wells.

Table 7.5 lists the evolution of the mean hydraulic conductivity and aquifer storativity. The mean hydraulic conductivity (μ_k in $\mathrm{m\,s^{-1}}$) estimated at the end of the assimilation of drawdown data remains close to the initial value if all elements of the model are considered. For the area of interest, a shift towards a lower mean is observed, being consistent with effective hydraulic conductivity estimates from type-curve matching of measured drawdown curves. The ensemble mean of aquifer storativity is slightly increased, and is accompanied by a drastic reduction of σ^2_{LnSo} after assimilation of data from tests 3b-FI and 3a-FI. Final values of $S_o = 0.016\,\mathrm{m^{-1}}$ and $\sigma^2_{LnSo} = 0.26$ were obtained at the end of the assimilation of drawdown data.

Table 7.5: Evolution of the parameter estimation during assimilation of drawdown data from three different hydraulic tests of the field tracer tomography. μ: ensemble mean; AESD: average ensemble standard deviation of log-hydraulic conductivity; σ^2_{LnK}: variance of log-hydraulic conductivity; σ^2_{LnSo}: variance of log-storativity; S_o: aquifer storativity. Values of AESD and σ^2_{LnK} are reported for the entire model domain as well as for the area of interest. Reported values correspond to the last update step of each hydraulic test.

	Hydraulic conductivity						Storativity S_o	
Test	Entire domain			Area of interest			Homogeneous	
ID	μ_K $(\mathrm{m\,s^{-1}})$	σ^2_{LnK}	AESD LnK	μ_k $(\mathrm{m\,s^{-1}})$	σ^2_{LnK}	AESD Lnk	μ_{So} $(\mathrm{m^{-1}})$	σ^2_{LnSo}
Initial	8.0×10^{-3}	1	0.79	7.9×10^{-3}	0.99	0.79	5×10^{-2}	0.30
3b-FI	7.3×10^{-3}	0.97	0.78	2.7×10^{-3}	0.92	0.76	0.07	0.08
3a-FI	7.0×10^{-3}	0.95	0.77	2.6×10^{-3}	0.88	0.74	0.09	0.036
1b-FI	7.0×10^{-3}	0.92	0.76	1.8×10^{-3}	0.83	0.72	0.016	0.26

Fig. 7.6 shows the evolution of drawdown predictions as a function of the update step for five different observation points in all three tests. Field measurements are represented with black lines, whereas the ensemble median and $90\,\%$ confidence interval of the ensemble predictions are shown in blue lines and a gray zone, respectively.

Difficulties in reproducing the transient behavior of the field measurements from test 3b-FI can be observed in Fig. 7.6 (left). This may be attributed to the inclusion of the damping factor β during parameter updating, which reduces the perturbations of parameters and dampens their adjustment. However, results from the synthetic study showed that a damping factor provides stability during data assimilation and reduces filter inbreeding effects (see also Franssen and Kinzelbach 2009). Additional factors affecting the transient evolution of drawdown could be related to (i) unresolved variability of aquifer storativity due to the adoption of a single effective value for the entire domain, and (ii) uncertainties associated with the mean value used for generating effective aquifer storativities for all ensemble members.

The maximum drawdown observed at the observation points is well captured by the ensemble predictions for most of the observation wells. Larger disagreements are observed at observation points closest to the injection wells (e.g., wells w01, cmt1-1, cmt1-5), and may be caused by discrepancies between real and modeled injection rates. Instabilities during water injection at low rates increase the uncertainty of the flowmeter measurements registered at the multilevel injection well (well B3).

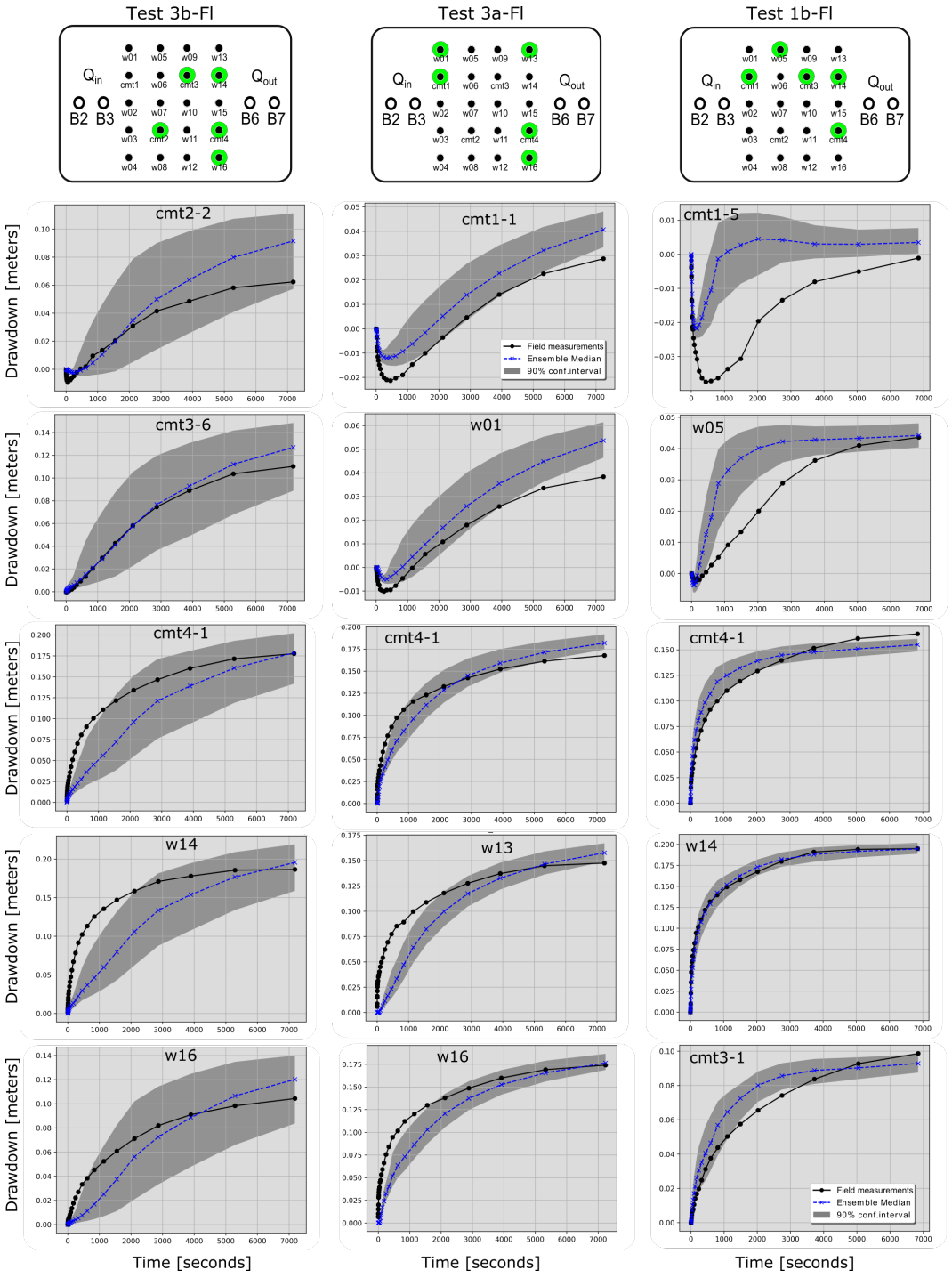


Figure 7.6: Drawdown predictions of tests 3b-Fl, 3a-Fl and 1b-Fl as a function of the updating step and for five different observation points. Black dots: field measurements; Gray area: 90 % confidence interval of the ensemble predictions; Blue lines: drawdown ensemble median.

The transient behavior is better captured in drawdown predictions of tests 3a-FI and 1b-FI, suggesting an improvement in the estimation of parameters. This is evidenced by the drawdown curves of well cmt4-1 (monitored during the three hydraulic tests), and well w16 (monitored during tests 3b-FI and 3a-FI). These findings agree with the lower parameter uncertainties observed in Fig. 7.5 also for tests 3a-FI and 1b-FI, and are qualitatively confirmed by a progressive increment in the NS-coefficient and a gradual reduction of the RMSE for all drawdown predictions as a function of the update step (see Table 7.6 and Fig.7.7).

Table 7.6: Evolution of the Nash-Sutcliffe coefficient (NS-coeff.) and root mean square error (RMSE) for drawdown, after the last update step of each of the three different hydraulic tests used for the estimation of parameters.

Test ID	NS-coeff. (-) (for drawdown)	RMSE (m)
3b-FI	0.85	2.4×10^{-2}
3a-FI	0.96	1.2×10^{-2}
1b-FI	0.98	8.8×10^{-3}

The evolution of the RMSE and the NS-coefficient as a function of the update step is plotted for all three tests in Fig. 7.7. The RMSE increases during the first update steps, and is followed by a considerable reduction after 15 update steps for test 3a-FI, and 19 update steps for tests 3b-FI and 1b-FI.

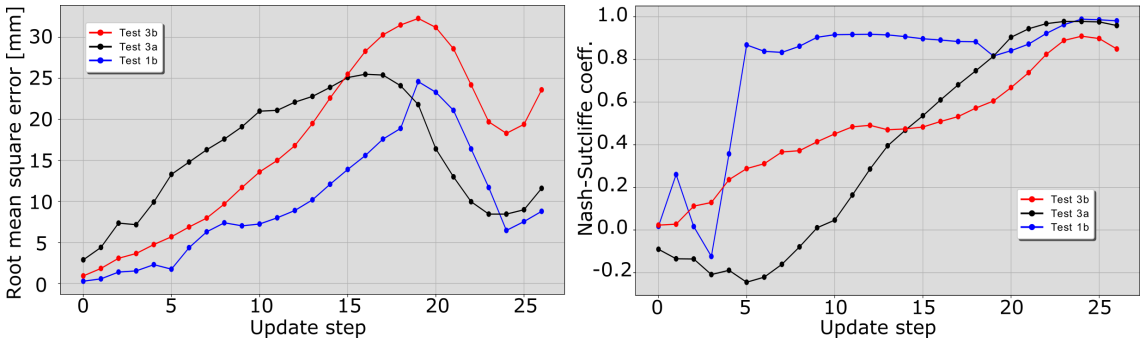


Figure 7.7: Evolution of the root mean square errors (left) and Nash-Sutcliffe coefficient of model efficiency (right), estimated for the drawdown predictions as a function of the updating step. Red lines: estimates for test 3b; Black lines: estimates for test 3a; Blue lines: estimates for test 1b. Note the scale in millimeters for the root mean square errors.

The RMSE estimated after the final update step of each test is consistently reduced (see Table 7.6). Simulated drawdowns at the end of the update deviate from field observations by $\leq 8.1\,\mathrm{mm}$ on average, a value close to the measurement error assumed during updating ($\sigma_\epsilon = 5.0\,\mathrm{mm}$). A similar improvement is observed for the NS-coefficient, with values approaching unity after the last update step of each test.

Fig. 7.8 is a comparison between all measured and simulated drawdowns for the two hydraulic tests not included for data assimilation and parameter estimation (tests 1a-FI and 2b-FI). These calibration plots evaluate the prediction capabilities of the estimated ensemble mean of parameters. A perfect prediction would imply that all points in the calibration plot lay on the 1:1-line (blue dashed line in Fig. 7.8), the RMSE would be zero, and the NS- and Pearson's correlation coefficients would be one. Acceptable predictions of the two additional tests are corroborated by a RMSE $\leq 13\,\mathrm{mm}$, and NS- and Pearson's correlation coefficients > 0.9.

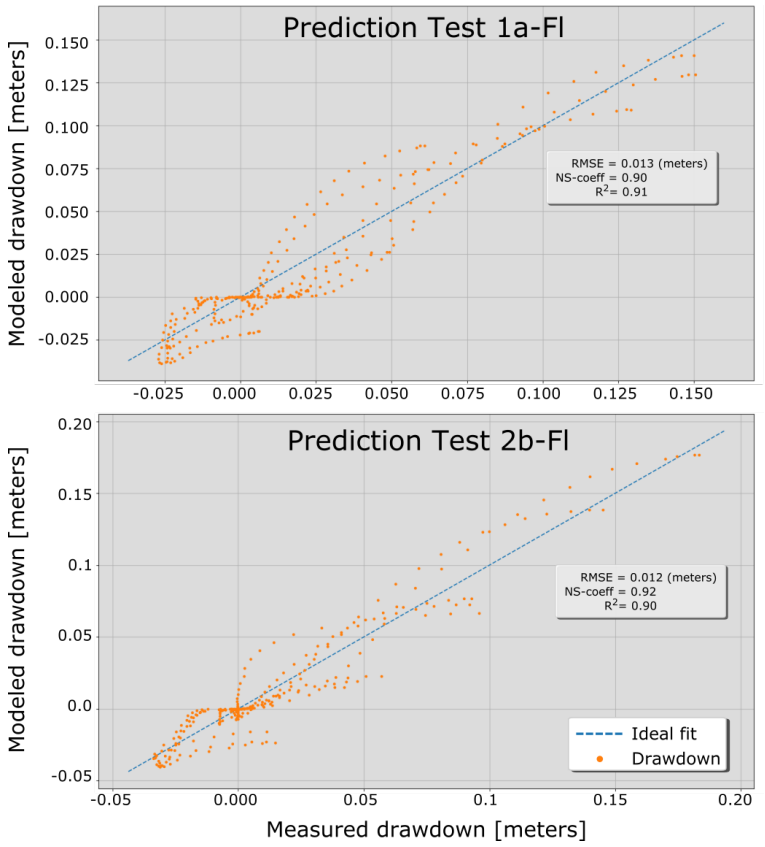


Figure 7.8: Calibration plot for model predictions of the two additional hydraulic tests (tests 1a-FI and 2b-FI) of the field tracer tomography. RMSE: root mean square error; NS-coeff.: Nash-Sutcliffe coefficient of model efficiency; R^2: Pearson's correlation coefficient.

7.3.2 Sequential Assimilation of Concentration Data

Parameter updating based on concentration data was performed in 34 update steps for tracer tests 3b-Tr and 3a-Tr, and 27 update steps for tracer tests 2b-Tr and 2a-Tr. As for the synthetic study, the assimilation of transport data was based on cumulative concentrations rather than normal tracer breakthrough curves. As mentioned before, monotonicity of cumulative breakthrough curves is expected to increase the stability of the parameter update when breakthrough times deviate considerably from the real values.

Figs. 7.9 and 7.10 illustrate the evolution of the ensemble mean and variance of log-hydraulic conductivity throughout the assimilation of breakthrough curves from the four different tracer tests, whereas Tables 7.7 and 7.8 contain a quantitative assessment of the hydraulic conductivity fields and additional effective aquifer parameters, respectively.

Table 7.7: Evolution of hydraulic conductivity during assimilation of concentration data from four different tracer tests. μ_k: geometricmean hydraulic conductivity; AESD: average ensemble standard deviation of log-hydraulic conductivity; σ^2_{LnK}: variance of log-hydraulic conductivity. Values are reported for both the entire model domain and the area of interest and for the last update step of each tracer test. Initial refers to the statistics of the parameters prior to any update step, and *after flow* refers to parameter statistics performed after drawdown data was assimilated.

	Hydraulic conductivity					
Test	Entire domain			Area of interest		
ID	μ_k (ms^{-1})	σ^2_{LnK}	AESD Lnk	μ_k (ms^{-1})	σ^2_{LnK}	AESD Lnk
Initial	8.00E-03	1	0.79	7.90×10^{-3}	0.99	0.79
After Flow	7.0×10^{-3}	0.92	0.76	1.8×10^{-3}	0.83	0.72
3b-Tr	7.2×10^{-3}	0.91	0.76	2.0×10^{-3}	0.81	0.72
3a-Tr	7.07×10^{-3}	0.88	0.75	1.69×10^{-3}	0.78	0.70
2b-Tr	7.34×10^{-3}	0.87	0.75	1.6×10^{-3}	0.76	0.69
2a-Tr	7.20×10^{-3}	0.86	0.74	1.59×10^{-3}	0.73	0.68

A visual inspection of Fig. 7.9 reveals that relatively low hydraulic conductivity values at the center of the area of interest shift towards higher values. While the same behavior is extended to the top of the aquifer, there is a general shift towards lower conductivity values at the lower portions of the aquifer. In comparison to the ensemble mean of log-hydraulic conductivity estimated with drawdown data (Fig. 7.4), parameter differences are accentuated at the lower portion of the area of interest, whereas the middle and top sections appear to converge towards more averaged values. This is evidenced in the XY slices at an elevation of

$1\,\mathrm{m}$ and $3\,\mathrm{m}$ in which the high conductivity zone close to the injection wells is reduced and constrained to the lower-left region of the area of interest, after concentration data from all tests have been assimilated.

Mean hydraulic conductivity values of $7.2 \times 10^{-3}\,\mathrm{m\,s^{-1}}$ and $1.6 \times 10^{-3}\,\mathrm{m\,s^{-1}}$ were estimated for the entire model domain and the area of interest, respectively. These values are similar to those estimated after the assimilation of drawdown data (see Table 7.7) suggesting a stabilization of the ensemble mean of hydraulic conductivity.

A visual inspection of Fig. 7.10 reveals a relevant reduction of the ensemble variance (σ^2_{LnK}). Lower parameter uncertainties were obtained at the central region of the area of interest. As expected, the breakthrough curves measured at the extraction wells provided the estimation with relevant information, helping to reduce parameter uncertainties at the surroundings of wells B6 and B7. The variance of log-hydraulic conductivity for the area of interest was reduced in average, up to $26\,\%$ with respect to the value prior to data assimilation, and the average AESD of log-hydraulic conductivity was reduced by $14\,\%$, also with respect to the initial value. These results suggest an improvement in the estimation of parameters after concentration data have been assimilated. However, variance reduction at places far away from the sensitive region (i.e., where data is available) may indicate a general underestimation of the ensemble variance. Causes for variance underestimation could be associated to the corruption of covariance matrices throughout the assimilation and the relatively few and large time steps used for the assimilation. Additional data can be included in the assimilation by reducing the length of the selected time steps, increasing the computational effort.

Table 7.8 shows the evolution of the mean and variance of the additional (effective) transport parameters included in the update. The ensemble mean of aquifer storativity sequentially decreased to a value of $7.9 \times 10^{-3}\,\mathrm{m^{-1}}$ and the associated variance increased $14\,\%$ with respect to the initial value used to generate random realizations of S_o for all ensemble members.

The ensemble mean of mobile-zone porosity (n_m) and longitudinal dispersivity (α_l) were shifted towards values smaller than the initial ensemble mean used to generate the corresponding random realizations, while the associated variances at the end of the updating were strongly reduced. The shift in mean values can be attributed to a wrong initial mean used to generate the random realizations and the high sensitivity of these two parameters, therefore requiring a stronger adjustment to achieve relevant improvements in model predictions. The ensemble mean of immobile-zone porosity (n_{im}) and first-order mass transfer coefficient (λ_{mt}) remained similar to the initial values throughout the assimilation, and the associated variance showed no relevant reductions.

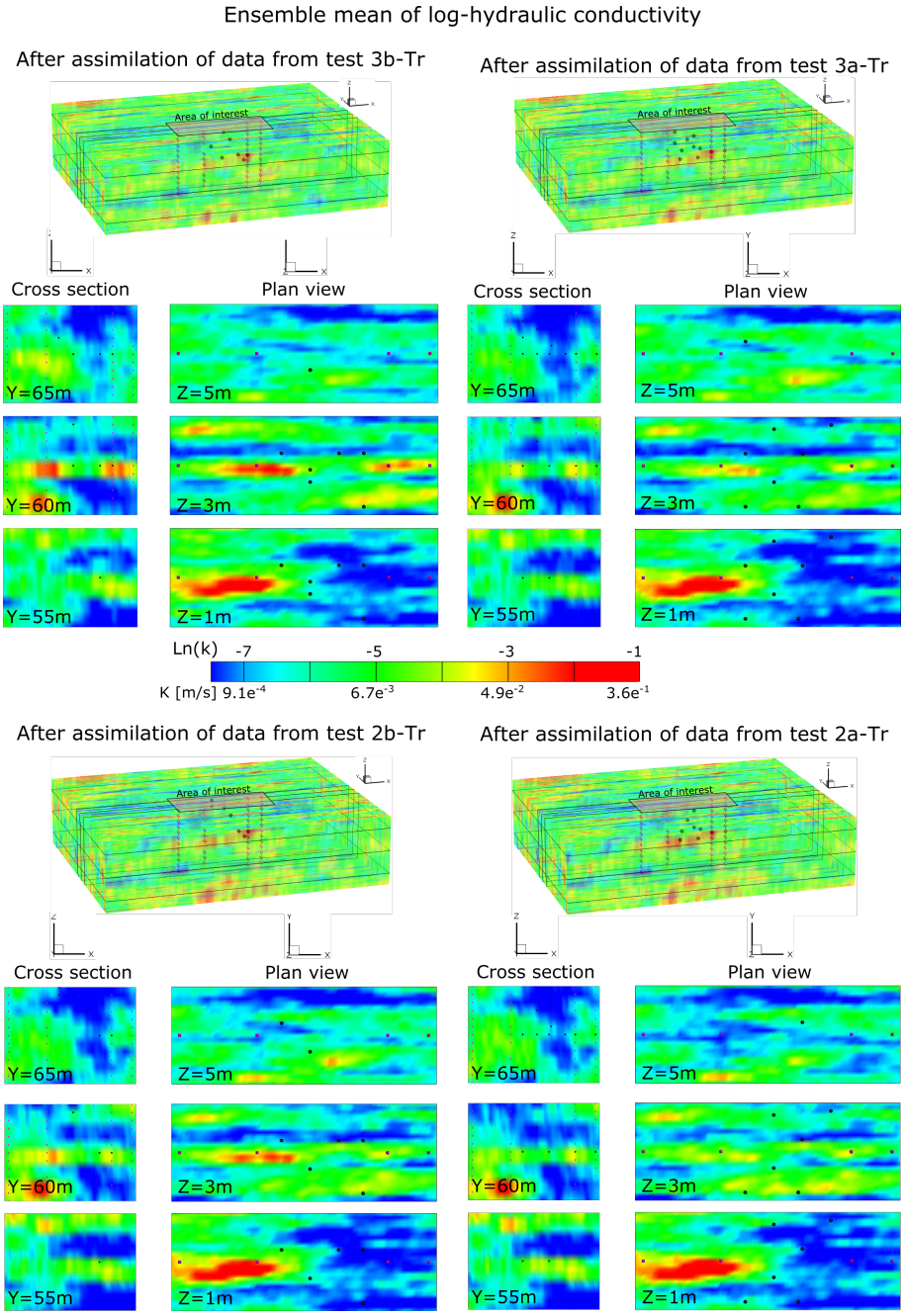


Figure 7.9: Ensemble mean of log-hydraulic conductivity estimated after assimilation of concentration data from tests 3b-Tr, 3a-Tr, 2b-Tr and 2a-Tr. XZ (cross section) and XY (plan view) slices with the ensemble mean of log-K at the area of interest are also plotted for each test. Purple diamonds: injection/extraction wells; Black spheres: observation wells.

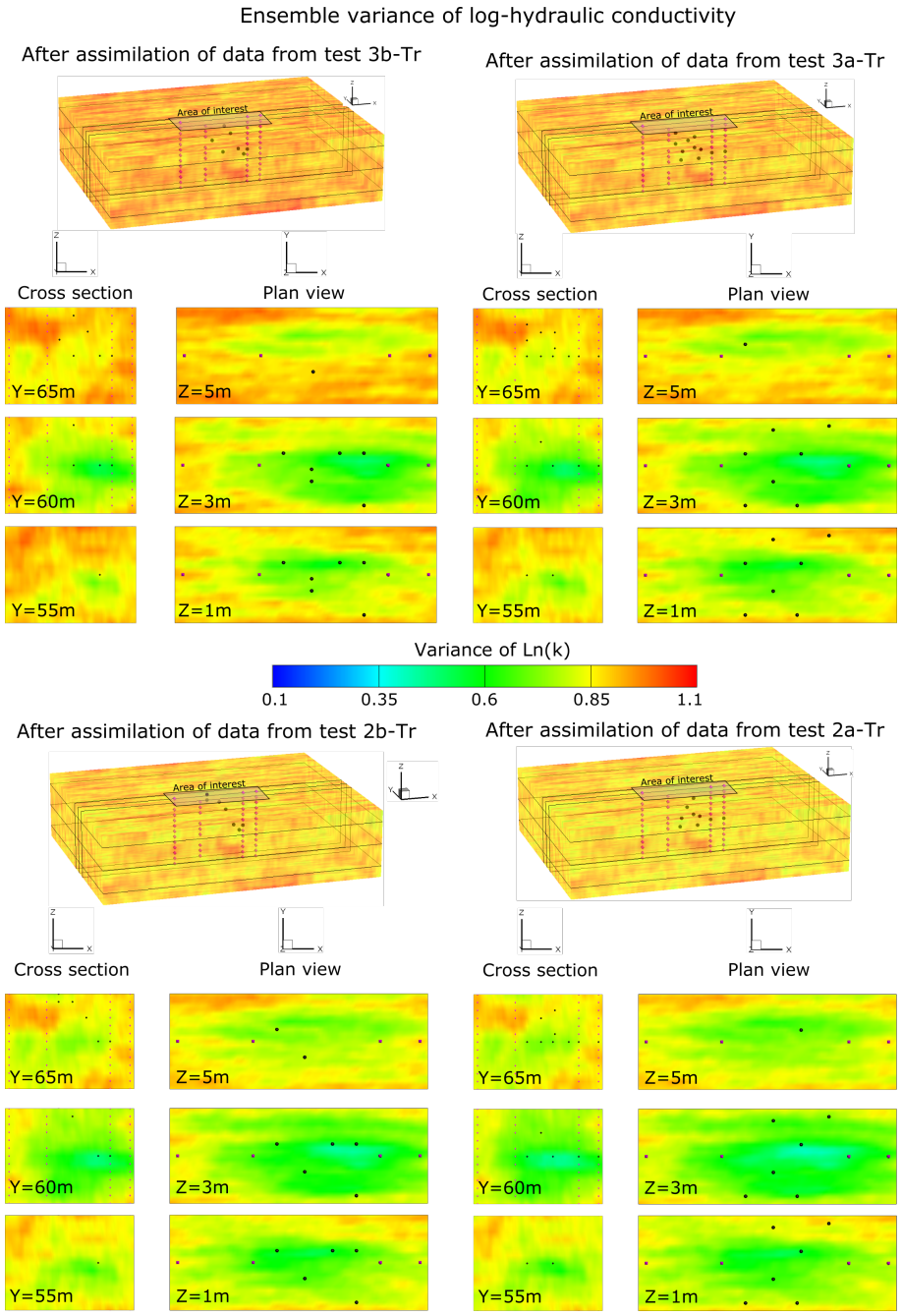


Figure 7.10: Ensemble variance of log-hydraulic conductivity estimated after assimilation of concentration data from tests 3b-Tr, 3a-Tr, 2b-Tr and 2a-Tr. XZ (cross section) and XY (plan view) slices with the ensemble variance of log-K at the area of interest are also plotted for each test. Purple diamonds: injection/extraction wells; Black spheres: observation wells.

Table 7.8: Evolution of additional (effective) aquifer parameters throughout the assimilation of concentration data from four different tracer tests of the field experiment. μ: ensemble mean of the corresponding parameter; σ^2 ensemble variance of the corresponding (log-) parameter; S_o: aquifer storativity; n_m: mobile zone porosity; n_{im}: immobile zone porosity; α_l: longitudinal dispersivity; λ_{mt}: first-order mass transfer coefficient.

	Effective transport parameters included in the update									
Test	S_o		n_m		n_{im}		α_l		λ_{mt}	
ID	μ (m^{-1})	σ^2_{LnSo}	μ (%)	$\sigma^2_{Ln(nm)}$	μ (%)	$\sigma^2_{L(nim)}$	μ (m)	$\sigma^2_{Ln(\alpha)}$	μ (s^{-1})	$\sigma^2_{Ln(\lambda)}$
Initial	0.05	0.3	10	0.20	20	0.2	1.0	0.4	9.96×10^{-9}	1.0
After Flow	1.6×10^{-2}	0.26	–	–	–	–	–	–	–	–
3b-Tr	1.7×10^{-2}	0.26	8.7	0.17	20.2	0.19	0.97	0.35	1.03×10^{-8}	0.977
3a-Tr	7.9×10^{-3}	0.42	7.7	0.14	20.0	0.19	0.91	0.32	1.05×10^{-8}	0.975
2b-Tr	8.0×10^{-3}	0.43	6.2	0.13	19.7	0.19	1.10	0.28	1.10×10^{-8}	0.974
2a-Tr	7.9×10^{-3}	0.42	5.7	0.11	20.0	0.19	0.96	0.25	1.14×10^{-8}	0.973

Fig. 7.11 shows a comparison of the final ensemble mean of log-hydraulic conductivity along a ZX profile crossing at the center of the numerical model ($Y = 62\,\mathrm{m}$), with four direct-push injection logs (DPIL) located along the same profile (Schneidewind, 2008) and two log-hydraulic conductivity profiles from a thermal flowmeter survey performed in pumping wells B6 and B7 (data not published).

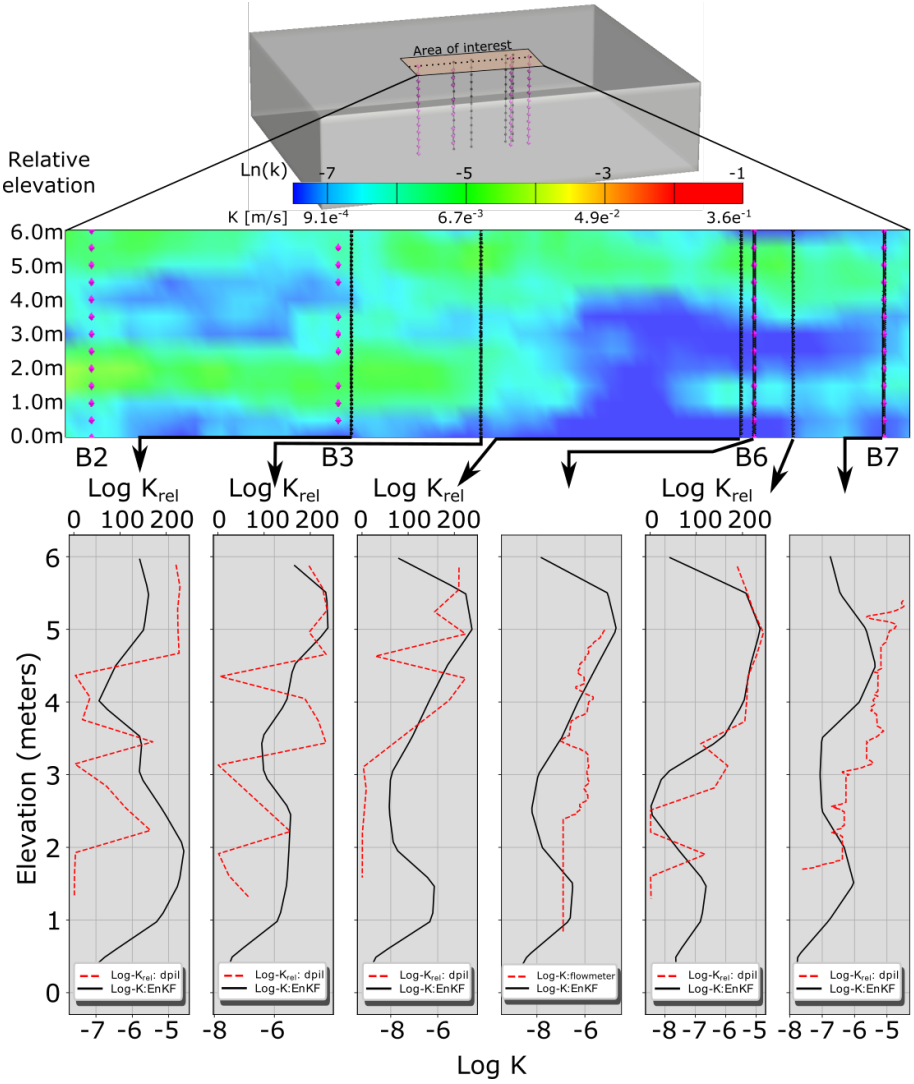


Figure 7.11: Ensemble mean of log-hydraulic conductivity (colored) compared with four DPIL profiles along the cross section represented with a black dotted line on the 3-D view of the model. Black dots indicate the location of the DPIL and flowmeter measurements; Black lines: Log-k values from the ensemble mean; Red lines: Relative log-K values and log-k values of DPIL and flowmeter profiles, respectively.

The exact location of the cross section is represented with a black dotted line on the 3-D view of the model. A visual inspection reveals that the variability of the DPIL and flowmeter profiles are in agreement with the spatially distributed log-hydraulic conductivity field, specially at the surroundings of the extraction wells B6 and B7. The information provided by the breakthrough curves measured in the extraction wells is well utilized by the EnKF, improving the estimation (i.e., lower estimation variance) at the surroundings of the extraction wells. Conversely, the lack of measurements from the injection wells B2 and B3 is reflected in poor estimates in the nearby regions.

The left columns of Figs. 7.12 - 7.15 present the cumulative concentration predictions, as a function of the update step, at selected observation points and for each test. The plots in the right columns of Figs. 7.12 - 7.15 compare the measured breakthrough curves with a full forward model simulation of each tracer test with the updated parameters obtained after all concentration data from the corresponding test were assimilated. Field measurements are represented with black lines, whereas the ensemble median and $90\,\%$ confidence interval of the ensemble predictions are shown with blue lines and a gray zone, respectively. The diagram at the top of each figure shows the spatial distribution of the injection, extraction and observation wells.

In general, the predictions of concentration are gradually improved in the update steps, specially for extraction wells B6 and B7 (Figs. 7.12 - 7.15 left column). Table 7.9 lists the mean relative errors of cumulative concentrations for each test.

Table 7.9: Mean relative error (MRE) of cumulative concentrations, averaged over all assimilation steps of each corresponding test.

Test ID	MRE ——
3b-Tr	5.29×10^{-1}
3a-Tr	6.20×10^{-1}
2b-Tr	6.81×10^{-1}
2a-Tr	4.91×10^{-1}

The wide range covered by the $90\,\%$ confidence interval at some obervation points (e.g., wells w07, w16 and B6 for test 3b-tr and wells w13 and B7 for test 3a-Tr) is attributed to the large increase in S_o variability after assimilation of data from test 3b-Tr. Although the variance of log-aquifer storativity increases thorughout data assimilation, the area covered by the $90\,\%$ is gradually reduced mainly because the variance of the additional transport parameters is also reduced.

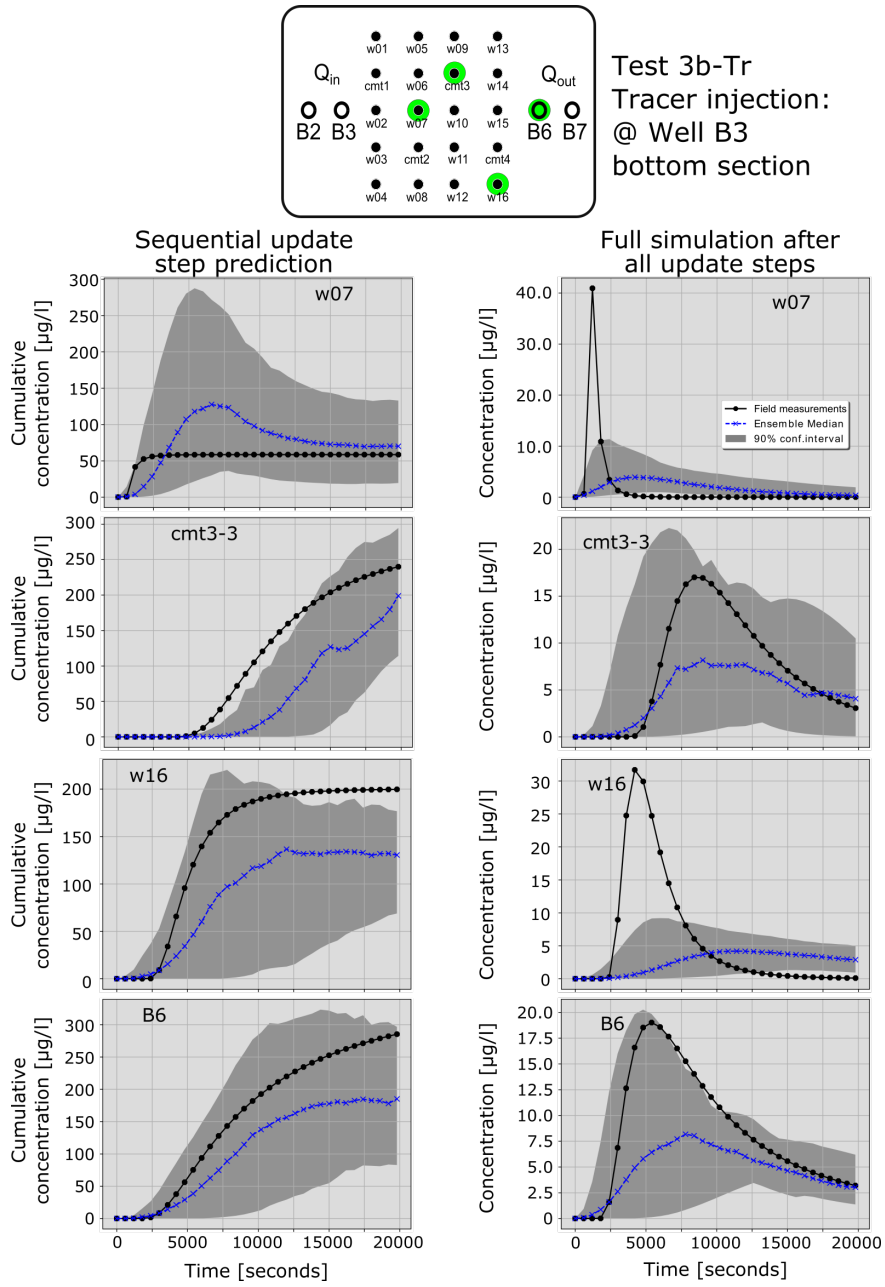


Figure 7.12: Fluorescein breakthrough predictions for four different observation points, as a function of the updating step (left) and full simulation after all 34 assimilation steps were performed (right) for tracer test 3b-Tr. Top diagram shows the spatial distribution of the injection, extraction and observation wells at the field site. Black dots: field measurements; Gray area: $90\,\%$ confidence interval of ensemble predictions. Blue lines: breakthrough ensemble median.

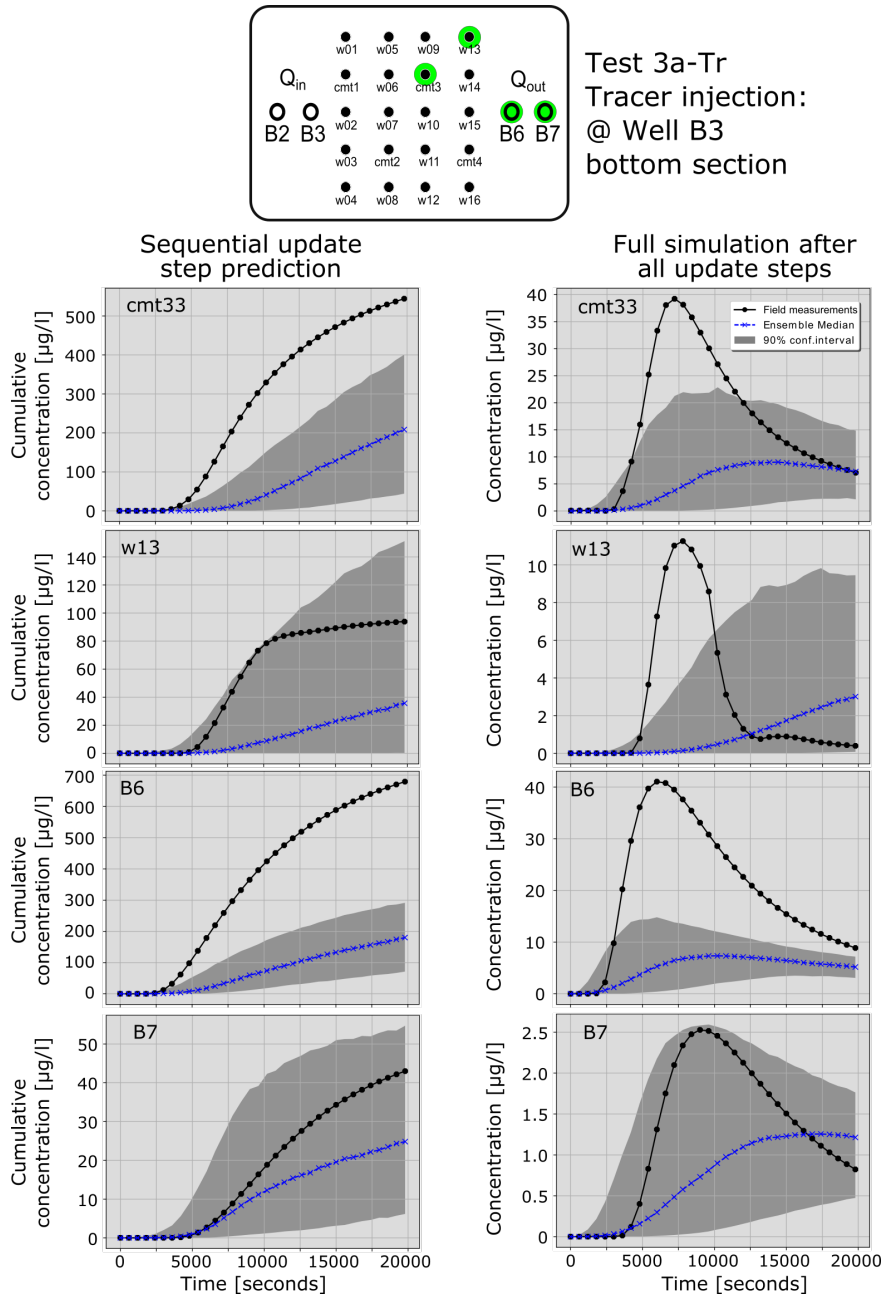


Figure 7.13: Tracer breakthrough predictions for four different observation points, as a function of the updating step (left) and full simulation after all 34 assimilation steps were performed (right) for tracer test 3a-Tr. Top diagram shows the spatial distribution of the injection, extraction and observation wells at the field site. Black dots: field measurements; Gray area: $90\,\%$ confidence interval of ensemble predictions. Blue lines: breakthrough ensemble median.

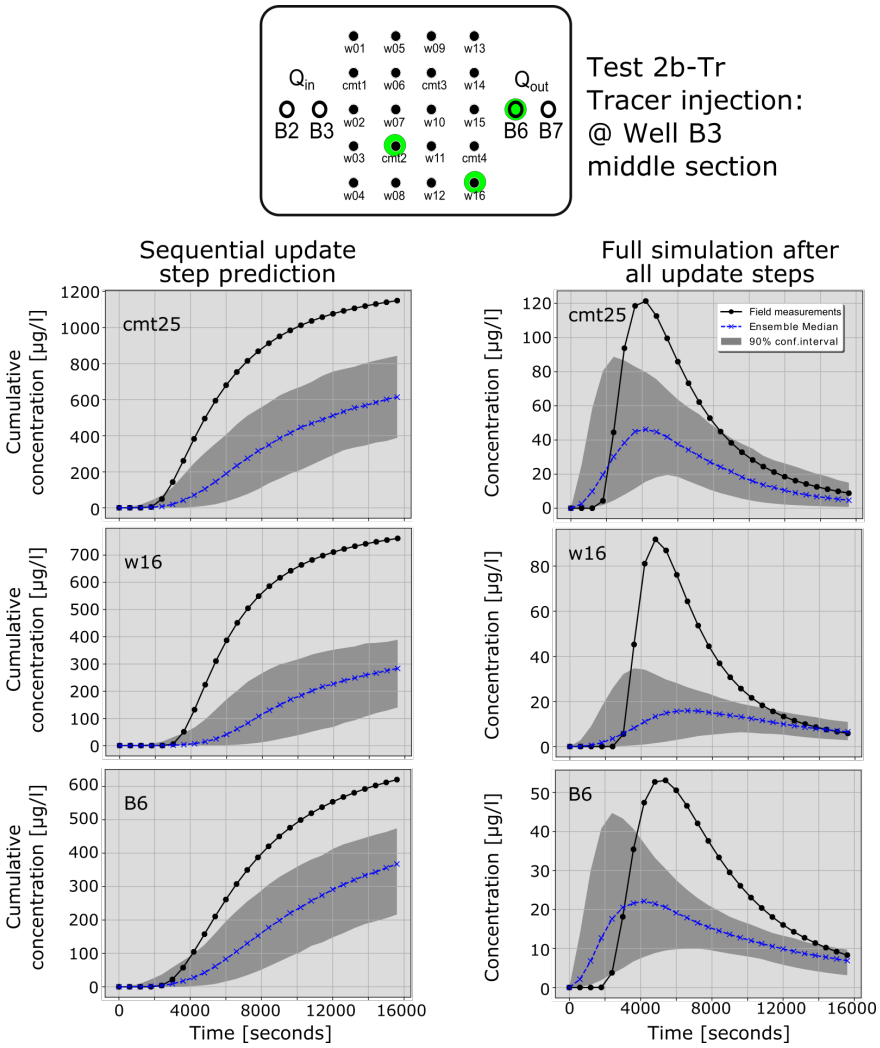


Figure 7.14: Tracer breakthrough predictions for four different observation points, as a function of the updating step (left) and full simulation after all 34 assimilation steps were performed (right) for tracer test 2b-Tr. Top diagram shows the spatial distribution of the injection, extraction and observation wells at the field site. Black dots: field measurements; Gray area: $90\,\%$ confidence interval of ensemble predictions. Blue lines: breakthrough ensemble median.

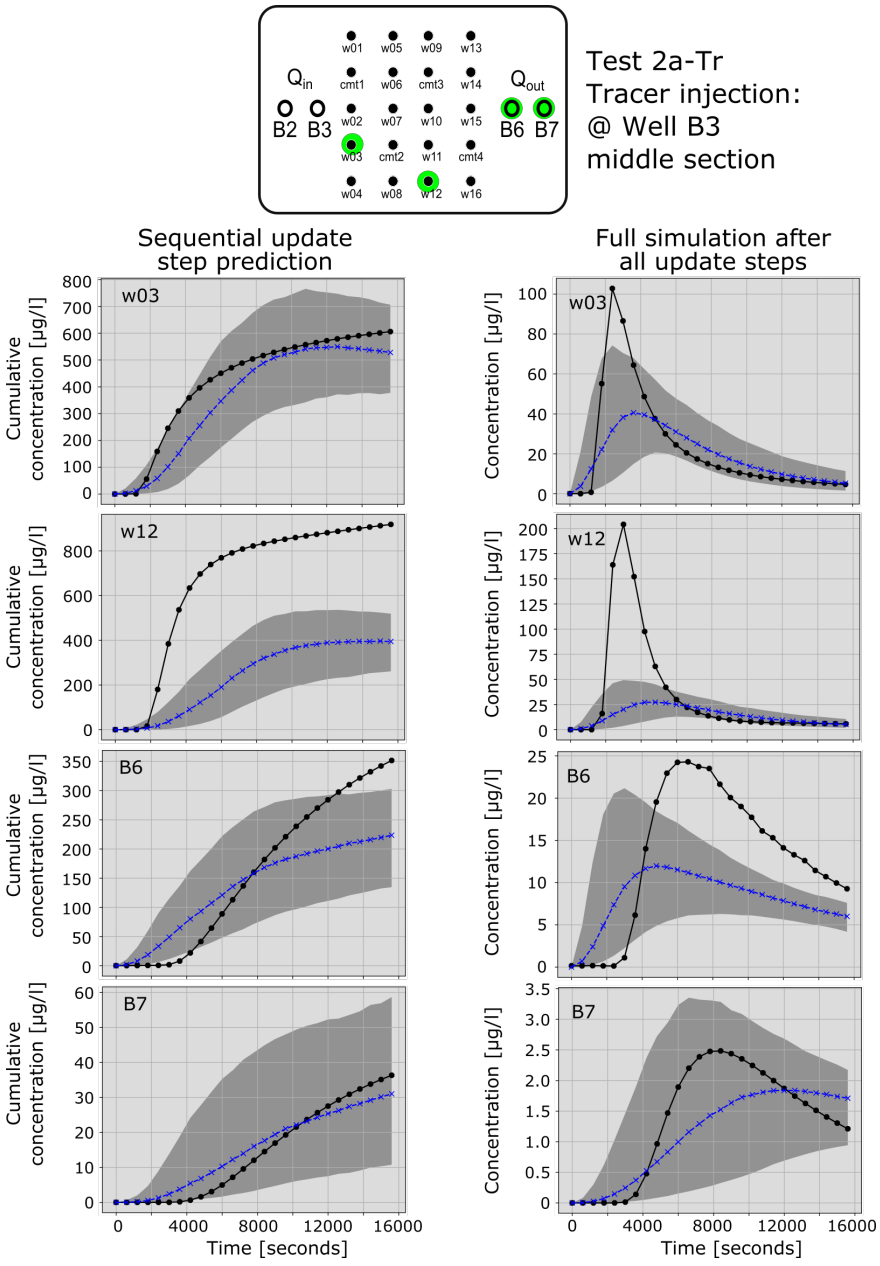


Figure 7.15: Tracer breakthrough predictions for four different observation points, as a function of the updating step (left) and full simulation after all 34 assimilation steps were performed (right) for tracer test 2a-Tr. Top diagram shows the spatial distribution of the injection, extraction and observation wells at the field site. Black dots: field measurements; Gray area: 90 % confidence interval of ensemble predictions. Blue lines: breakthrough ensemble median.

The general tendency to underestimate tracer concentrations (Figs. 7.12 - 7.15 right column) is indicative of excessive numerical dispersion, caused by the relatively coarse grid and large time steps adopted for transport simulations (see e.g., Woods et al. 2003). Numerical dispersion could be reduced while keeping affordable computational times by using an unstructured grid, with refined elements at the area of interest and coarsened elements towards the domain boundary. However, this requires the implementation of a different geostatistical inverse method for the generation of spatially distributed fields (e.g., Li and Cirpka 2006). Further implementations to the version of the EnKF presented in this work are required to optimize the grid resolution with reduced computational costs.

The mismatch in first arrivals and peak concentration between observed and simulated breakthrough curves can be attributed to errors in the estimation of effective transport parameters. Better simulations are generally observed for extraction wells B6 and B7, suggesting that updated effective transport parameters are representative of an average behavior of the aquifer. Effective parameters fail to resolve local variations in the aquifer material, leading to larger discrepancies at observations points constrained to a specific portion of the aquifer (e.g., at the cmt-wells). To improve model simulations, transport parameters (such as n_m and α_l) would most likely have to be treated as spatially distributed during updating, increasing the computational requirements.

Fig. 7.16 shows the model prediction of four tracer breakthrough curves for test 1a-Tr. Data from this test was not included in the assimilation of concentration data. A visual inspection indicates poor prediction capabilities of the ensemble mean of estimated parameters, which is confirmed by a root mean square error for all predicted tracer concentrations of $\sim 5 \times 10^{-2}\,\mathrm{mg\,l^{-1}}$, which is a value higher than the mean observed concentrations.

Observation well w14 is screened over the entire aquifer thickness, therefore the observed breakthrough curve correspond to a depth average record. Better concentration predictions at well w14 strengthen the hypothesis that the estimated effective transport parameters might be representative of average aquifer properties. In contrast, effective transport parameters fail to resolve local material variations, leading to large disagreements between simulated and observed breakthrough curves at the additional observation wells. The mismatch of mean arrival times is observed at multilevel observation points (well cmt3-1 in Fig. 7.16), a wrong recovery is exemplified with the breakthrough curves simulated at wells cmt4-1 and w12, were very low tracer concentrations were expected.

Reasonable model predictions were observed in the simulation of drawdown data from test 2a-Fl (not included in the assimilation), suggesting that parameters conditioned to transport data can still produce accurate groundwater flow predictions. Fig. 7.17 compares the drawdown measured in four different wells (black lines), with two different model predictions. The

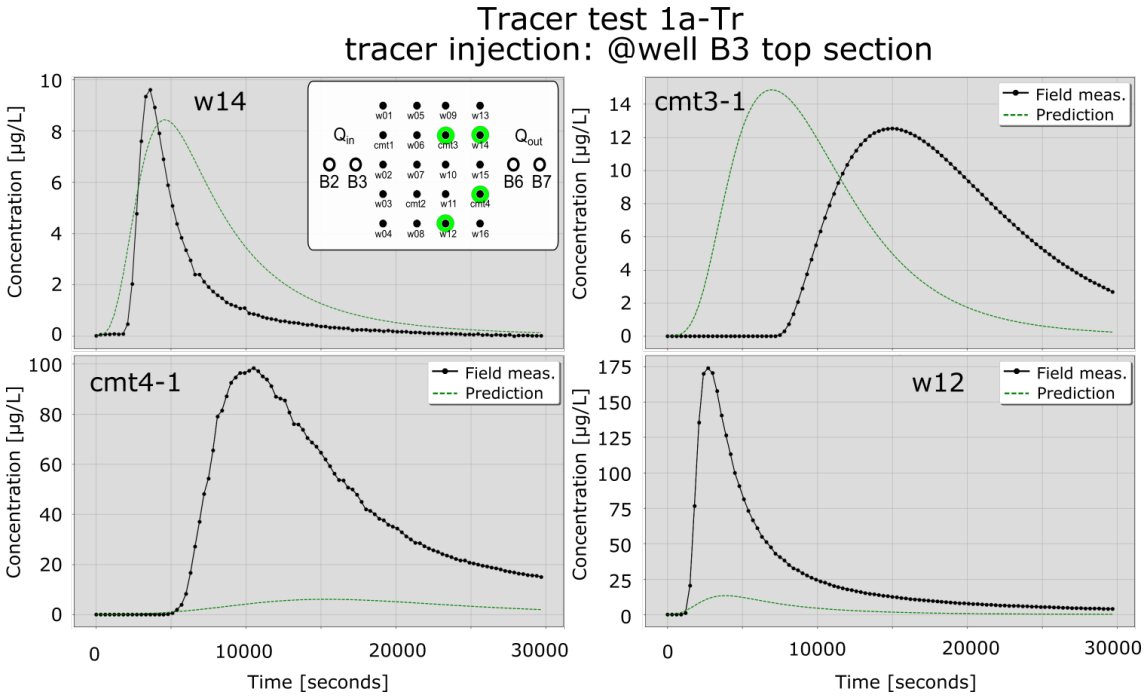


Figure 7.16: Model prediction of four tracer breakthrough curves for test 1a-Tr. The simulation was performed with the ensemble mean of parameters estimated after all concentration data was assimilated. Black lines: measured breakthrough curves; Green dotted lines: model prediction.

same ensemble mean of hydraulic conductivity and transport parameters were used in both simulations, but two different values of aquifer storativity were applied. The first S_o value corresponds to the results obtained after drawdown data were assimilated (red dotted lines in Fig. 7.17), and the second corresponds to the smaller value obtained after assimilation of concentration data (blue lines in Fig. 7.17). Although both values yield the same steady-flow levels, it is noticeable that the transient behavior of the pressure signal is better represented when the (larger) aquifer storativity estimated after assimilation of drawdown data was used. This is confirmed by a RMSE between all measured and predicted drawdowns of $0.02\,\mathrm{m}$ and $0.01\,\mathrm{m}$ for the large and small aquifer storativity values, respectively.

Finally, from the results presented in this chapter the following conclusions can be drawn:

(i) The EnKF is capable of adjusting the hydraulic parameters in order to produce reasonable groundwater flow predictions.

(ii) An improvement in the estimation of parameters is achieved after assimilating concentration data, expressed in a considerable reduction of the estimation variance. This demonstrates the value of the data collected with a tracer tomography experiment.

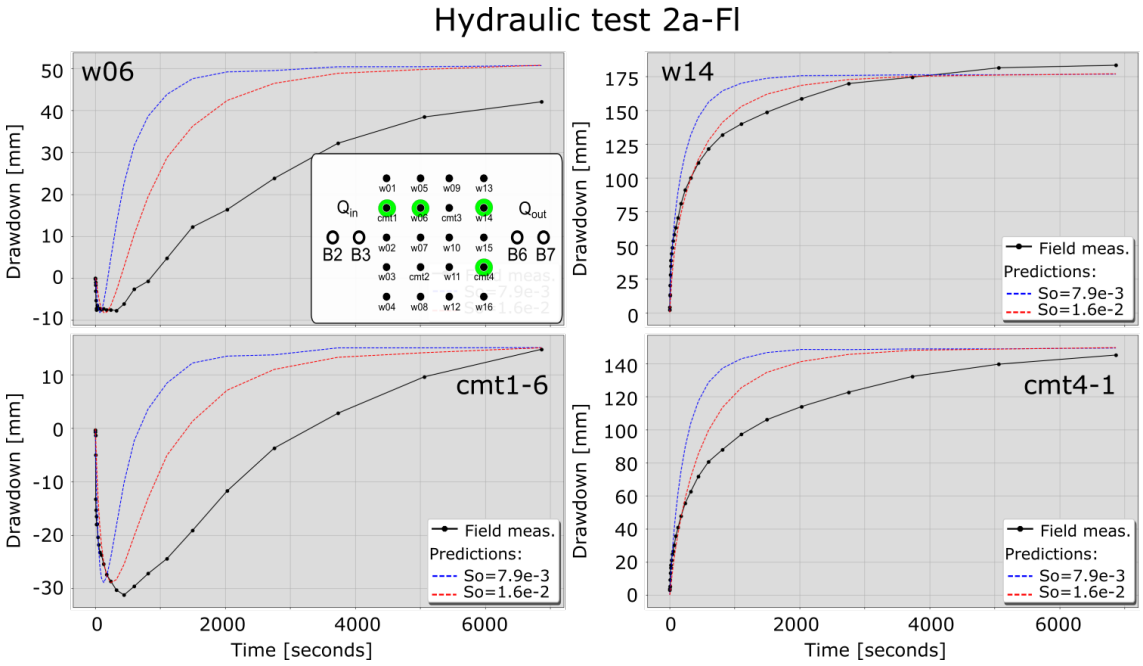


Figure 7.17: Model prediction of four drawdown curves for test 2a-FI. The simulation was performed with the ensemble mean of parameters estimated after all concentration data was assimilated. Black lines: measured drawdown; Blue dotted lines: model prediction with the mean aquifer storativity obtained from the assimilation of transport data $S_o = 7.9 \times 10^{-3}\,\mathrm{m}^{-1}$; Red dotted lines: model prediction with the mean aquifer storativity obtained from the assimilation of drawdown data ($S_o = 1.6 \times 10^{-2}\,\mathrm{m}^{-1}$).

This is in agreement with the results of a synthetic parameter estimation based on tracer tomography presented by Schwede et al. (2014).

(iii) The spatial structure of the ensemble mean of log-hydraulic conductivity estimated with drawdown suggests the distribution of more conductive material in the lower portions of the aquifer. Similar results based on data from a hydraulic tomography experiment were found by Sanchez-León et al. (2016), but are in contradiction with previous studies based on abundant direct-push injection logs and tracer tests, in which a more conductive upper layer overlaying a less conductive and more heterogeneous lower layer was interpreted (Lessoff et al. 2010; Doro et al. 2015). The distribution of log-hydraulic conductivity however, is rearranged after the updates based on tracer concentration, showing a better agreement with the aforementioned studies.

(iv) Parameter fields conditioned to transport data can still produce accurate groundwater flow predictions. Although transport simulations were independent of aquifer storativity (i.e., performed after steady flow is achieved), it was included in the parameter update of concentration data to assess the stability of the assimilation procedure. Results from the

assimilation of transport data yielded a lower aquifer storativity than the one obtained with the assimilation of drawdown data. This might indicate that correlations defined within the auto- and cross-covariance matrices deteriorate throughout the updating. It was observed that the larger aquifer storativity (from the assimilation of drawdown data) is better in predicting the transient behavior of the pressure signal.

(v) Difficulties in reproducing and predicting solute transport may be attributed to measurement errors contained in the measured breakthrough curves, excessive numerical dispersion due to the relatively coarse grid and large time steps adopted, and the update of effective transport parameters rather than considering them spatially distributed.

(vi) Further research is needed to quantitatively assess the errors produced by numerical dispersion and implement the necessary methods to minimize the negative impact in solute transport simulations, while keeping computational costs at reasonable levels.

Chapter 8

Conclusions and Outlook

The present work was motivated by the pressing need for innovation in hydrogeological methods targeting detailed aquifer characterization. Improved quantification of aquifer heterogeneity is key to a better performance of groundwater flow and solute transport models, leading to meaningful predictions that support the implementation of responsible management policies. Consequently, the main goal of this thesis was to contribute to the development of tracer tomography, making a step forward towards its establishment as a feasible approach for improved aquifer characterization on the field scale, and hence, fostering its real-world application. The studies of this dissertation covered experimental work and the implementation of numerical methods using variants of the ensemble Kalman filter.

For the experimental work, I adapted the field method of Doro et al. (2015) using fluorescein instead of heat as a tracer. A multichannel fluorometer was used to simultaneously monitor fluorescein concentrations at 19 different points. I tested the viability of the method by applying it Hydrogeological Research site Lauswiesen, Germany. The solute-tracer tomography experiments reported in this work consisted in the sequential injection of tracer at one out of three isolated intervals generated at an extraction well. The tomographic layout was achieved by shifting the tracer injection to a different injection location, until the three sections were covered. Two tests per injection section were performed, leading a total of six different tracer tests (with their corresponding pumping test). The applicability of the experimental approach was demonstrated with the acquisition of a dataset of 52 drawdown curves and 46 breakthrough curves of fluorescein.

The analysis of drawdown and breakthrough curves with type-curve and temporal moment methods provided mean effective hydraulic conductivity values that served as a basis for parameter estimation. Limitation of these analytical methods produced in some cases untrustworthy

aquifer parameters, e.g., effective porosity $< 1\,\%$ estimated by fitting the 1D analytical solution of the advection-dispersion equation. Except of test 3b (tracer injection in the lower section of well B3), the estimated mass of tracer recovered at the extraction wells was above $62\,\%$, suggesting that a well focused steady flow field was achieved with the nested-cell setup. Over-estimation of the recovered mass in some tests was attributed to complications during water sampling, leading to errors in the scaling of breakthrough curves to units of concentration.

I inverted the collected data with the EnKF, coupled to a three-dimensional groundwater flow and solute transport model that was used to simulate the tracer tomography experiment reported in this work. To the author's knowledge, the estimation of aquifer parameters with the EnKF applied to data from a tracer tomography experiment has not been reported. The performance of the filter was evaluated within a synthetic study, in which synthetic data was generated with numerical simulations of a tracer tomography experiment, using a two-dimensional model.

I found in this work that the update scheme of the classical EnKF leads to mass balance errors during he assimilation of concentration data. Within the update scheme of the Restart EnKF, the steady-flow field is reinitialized and the plume distribution recalculated after each update step, being a suitable option for the assimilation of transient concentration data.

Results of the sequential assimilation of synthetic drawdown and concentration data demonstrated the benefits of the information provided by tracer tomography for parameter estimation and hence, improved groundwater flow and solute transport simulations. Improvements in the estimation using both data types is related to the difference in sensitivity patterns between hydraulic and concentration measurements (e.g., Schwede et al. 2014).

Similar results to the synthetic study were observed in the updating of parameters with drawdown data recorded during the tracer tomography test reported in this work. In general, parameter uncertainties were gradually reduced and groundwater flow predictions improved. The value of combining data of different types is again demonstrated after updating parameters based on the analysis of concentration data, but difficulties in reproducing and predicting solute transport were observed throughout the update steps. This was attributed to measurement errors contained in the measured breakthrough curves, excessive numerical dispersion due to the relatively coarse grid and large time steps adopted, and the estimation of spatially uniform effective transport parameters rather than considering spatially distributed fields of transport parameters.

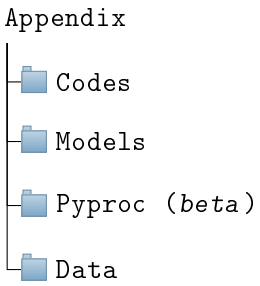
To improve the quality of the field measurements obtained during a field application of tracer tomography, I recommend further development of the sampling system. Measurement errors could be considerably reduced if a better control in the quality of the collected groundwater

samples is guaranteed. I also suggest for future experimental work to extend the monitoring time of tracer concentrations, in order to retrieve breakthrough curves without considerable truncation, as it was often the case for the breakthrough curves obtained in this study.

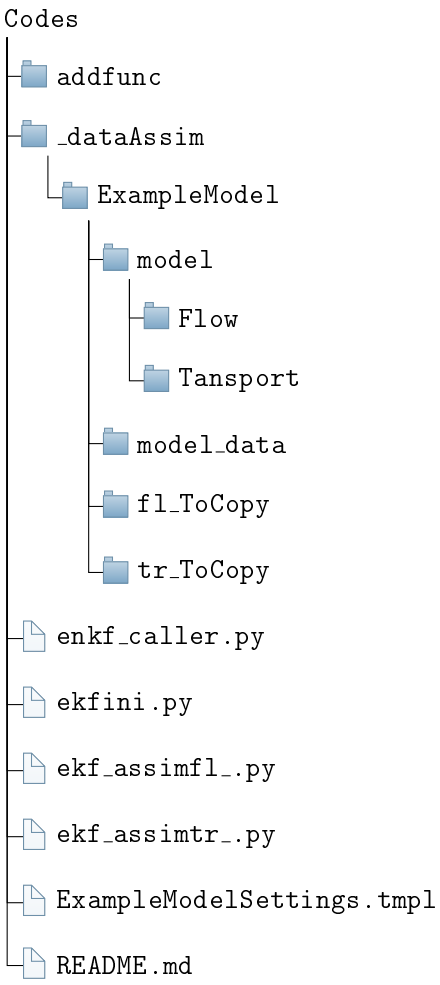
To improve the performance of the EnKF in the assimilation of tracer data, I recommend further research to evaluate the stability of the filter, and the numerical methods used as forward model. Possible research directions include the application of so-called localization methods, in which covariances are computed only for restricted combinations of states and parameters, the implementation of methods to optimize the grid resolution while keeping reasonable model running times and matrices sizes, the assessment of high-performance computing methods in the simulations of flow and transport to reduce numerical artifacts, a systematic assessment of the EnKF methods versus variational data assimilation methods in the context of tracer tomography, and a reassessment whether temporal moments are suitable for condensing transient datasets. In an ongoing study, field-techniques to monitor tracer tests by geophysical methods are tested. Inverting these data will be faced by similar challenges as inverting concentration time series.

Appendix

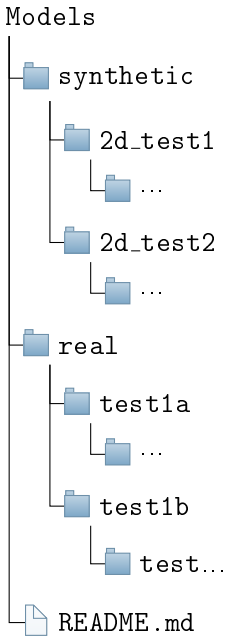
All information contained in the appendix is presented in electronic format. If the folder **Appendix** is provided together with the electronic version of this work, and located at the same directory level, the hyperlink assigned to each node of the following directory trees would automatically open the corresponding directory with the available file explorer. Each subdirectory of **Appendix** contains a *README.md* markdown file with relevant information about the subdirectory content, and can be opened with any text editor. The **Appendix** is organized according to the following directory tree:



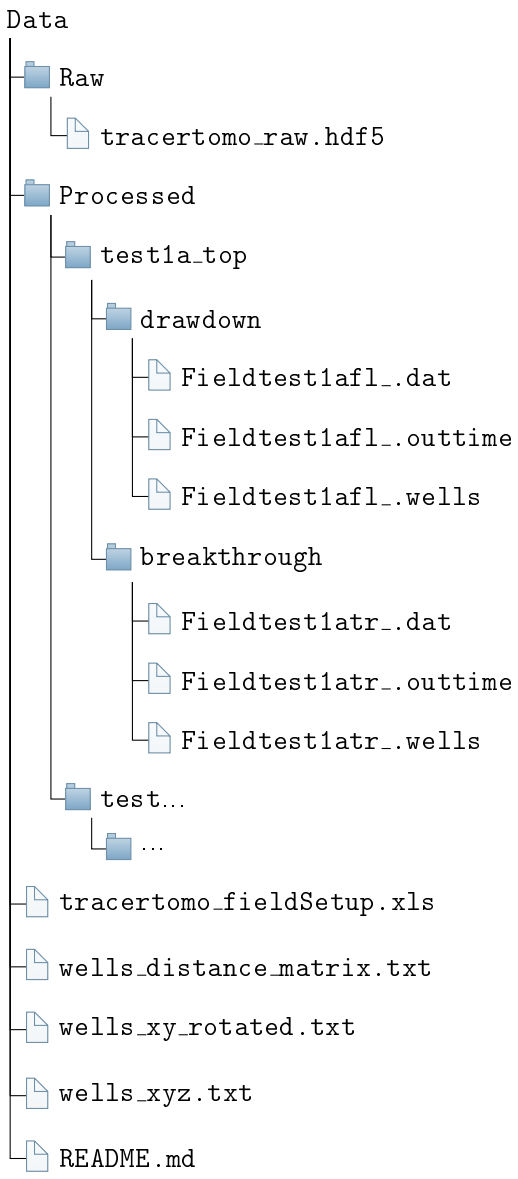
The subdirectory **Codes** contains all necessary scripts to estimate aquifer parameters with the Ensemble Kalman filter (EnKF) and a groundwater flow and/or solute transport model constructed in HydroGeoSphere. Instructions for the installation of necessary software is provided in the corresponding README.md file. A groundwater flow model example (**ExampleModel**), together with a settings file for running the filter are included. The reference model used to generate the synthetic data is included in the aforementioned directory. Note that a valid license for HydroGeoSpehere (revision:1438; build date:20140203; build info: 64bit) would be required. The subdirectory **Codes** is organized as follows:



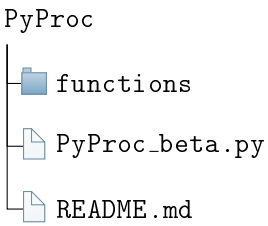
The **Models** subdirectory is separated in two main folders. One folder contains all models used for the synthetic study and the second stores all models used for data assimilation and parameter estimation with data from the field tracer tomography. All subdirectories are organized by the corresponding test ID, and contain the necessary files to run the EnKF algorithm. To minimize storage volume, model output files from HydroGeoSphere have to be generated by performing a forward model run.



Data from the tracer tomography test is provided in the **Data** subdirectory. The raw data is stored in a structured Hierarchical Data Format 5 (HDF5) database with meaningful metadata, and the processed dataset is provided in ASCII files. Instructions on how to download open-source visualization tools are provided in the README.md file, for a quick access to the content of the HDF5 database. The processed data is provided in ASCII files, organized by test ID.



The python-based time-series processing module *PyProc* can be found in the **Pyproc** sub-directory. As mentioned in Chapter 5 Data Processing, *PyProc* is a python module with tools to increase the signal-to-noise ratio of the raw data, detect and remove background values and trends introduced by the sensors, correct signal shifts, fit parametric functions, etc., and a logging feature that records all modifications applied to the original dataset. Instructions on all software needed to run *PyProc* are provided in the README.md file.



References

M. Alexander, S. J. Berg, and W. A. Illman. Field study of hydrogeologic characterization methods in a heterogeneous aquifer. *Ground Water*, 49(3):365–382, 2011.

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

R. T. Bailey and D. Baù. Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother. *Hydrol. Earth Syst. Sci.*, 16(2):287–304, 2012.

D. Béal, P. Brasseur, J. M. Brankart, Y. Ourmières, and J. Verron. Characterization of mixing errors in a coupled physical biogeochemical model of the north atlantic: implications for nonlinear estimation using gaussian anamorphosis. *Ocean Sci.*, 6(1):247–262, 2010.

S. Basiricò, G. B. Crosta, P. Frattini, A. Villa, and A. Godio. Borehole flowmeter logging for the accurate design and analysis of tracer tests. *Groundwater*, 53(S1):3–9, 2015.

J. Bear. *Dynamics of Fluids in Porous Media*. Dover, 1972.

A. A. Behroozmand, K. Keating, and E. Auken. A review of the principles and applications of the nmr technique for near-surface characterization. *Surveys in Geophysics*, 36(1):27–85, 2015.

S. J. Berg and W. A. Illman. Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system. *Water Resources Research*, 47(10): W10507, 2011.

S. J. Berg and W. A. Illman. Field study of subsurface heterogeneity with steady-state hydraulic tomography. *Ground Water*, 51(1):29–40, 2013.

S. J. Berg and W. A. Illman. Comparison of hydraulic tomography with traditional methods at a highly heterogeneous site. *Groundwater*, 53(1):71–89, 2015.

M. Bianchi, C. Zheng, C. Wilson, G. R. Tick, G. Liu, and S. M. Gorelick. Spatial connectivity in a highly heterogeneous aquifer: From cores to preferential flow paths. *Water Resources Research*, 47(5):W05524, 2011.

A. Binley, J. Keery, L. Slater, W. Barrash, and M. Cardiff. The hydrogeologic information in cross-borehole complex conductivity data from an unconsolidated conglomeratic sedimentary aquifer. *GEOPHYSICS*, 81(6):E409–E421, 2016.

G. Bohling and J. J. Butler. Inherent limitations of hydraulic tomography. *Ground Water*, 48 (6):809–824, 2010.

G. C. Bohling. Sensitivity and resolution of tomographic pumping tests in an alluvial aquifer. *Water Resources Research*, 45(2):n/a–n/a, 2009.

G. C. Bohling, X. Zhan, M. D. Knoll, and J. J. Butler. Hydraulic tomography and the impact of a priori information: An alluvial aquifer example. 2003.

G. C. Bohling, J. J. Butler, X. Zhan, and M. D. Knoll. A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities. *Water Resources Research*, 43(5):W05430, 2007.

R. Brauchler, R. Liedl, and P. Dietrich. A travel time based hydraulic tomographic approach. *Water Resources Research*, 39(12):1370, 2003.

R. Brauchler, J. T. Cheng, P. Dietrich, M. Everett, B. Johnson, R. Liedl, and M. Sauter. An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion. *Journal of Hydrology*, 345(3–4):184–198, 2007.

R. Brauchler, G. Böhm, C. Leven, P. Dietrich, and M. Sauter. A laboratory study of tracer tomography. *Hydrogeology Journal*, 21(6):1265–1274, 2013a.

R. Brauchler, R. Hu, L. Hu, S. Jiménez, P. Bayer, P. Dietrich, and T. Ptak. Rapid field application of hydraulic tomography for resolving aquifer heterogeneity in unconsolidated sediments. *Water Resources Research*, 49(4):2013–2024, 2013b.

R. Brown and P. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley and Sons, United States of America, 1997.

G. Burgers, P. Jan van Leeuwen, and G. Evensen. Analysis scheme in the ensemble kalman filter. *Monthly Weather Review*, 126(6):1719–1724, 1998.

J. J. Butler. Pumping tests for aquifer evaluation—time for a change? *Ground Water*, 47(5): 615–617, 2009.

J. J. Butler, P. Dietrich, V. Wittig, and T. Christy. Characterizing hydraulic conductivity with the direct-push permeameter. *Ground Water*, 45(4):409–419, 2007.

J. J. Butler Jr. and W. Liu. Pumping tests in non-uniform aquifers: The radially asymmetric case. *Water Resour. Res.*, 29(2):259–269, 1993.

M. Camporese, G. Cassiani, R. Deiana, P. Salandin, and A. Binley. Coupled and uncoupled hydrogeophysical inversions using ensemble kalman filter assimilation of ert-monitored tracer test data. *Water Resources Research*, 51(5):3277–3291, 2015.

M. Cardiff and W. Barrash. 3-d transient hydraulic tomography in unconfined aquifers with fast drainage response. *Water Resources Research*, 47(12):W12518, 2011.

M. Cardiff, W. Barrash, and P. K. Kitanidis. A field proof-of-concept of aquifer imaging using 3-d transient hydraulic tomography with modular, temporarily-emplaced equipment. *Water Resour. Res.*, 48:1–18, 2012.

M. Cardiff, T. Bakhos, P. K. Kitanidis, and W. Barrash. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resources Research*, 49(9):5395–5410, 2013.

J. Carrera, A. Alcolea, A. Medina, H. Hidalgo, and L. J. Slooten. Inverse problem in hydrogeology. *Hydrogeology Journal*, 13:206–222, 2005.

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

O. A. Cirpka and P. K. Kitanidis. Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data. *Advances in Water Resources*, 24(1):89–103, 2000.

O. A. Cirpka, M. N. Fienen, M. Hofer, E. Hoehn, A. Tessarini, R. Kipfer, and P. K. Kitanidis. Analyzing bank filtration by deconvoluting time series of electric conductivity. *Ground Water*, 45(3):318–328, 2007.

K. H. Coats and B. D. Smith. Dead-end pore volume and dispersion in porous media. 1964.

R. L. Cooley. An analysis of the pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. *Water Resour. Res.*, 36(4):1159–1163, 2000.

H. Cooper and C. Jacob. A generalized graphical method for evaluating formation constants and summarizing well field history. *Am. Geophys. Union Trans.*, 27:526–534, 1946.

E. Crestani, M. Camporese, D. Baú, and P. Salandin. Ensemble kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. *Hydrol. Earth Syst. Sci.*, 17:1517–1531, 2013.

S. A. Crisman, F. J. Molz, D. L. Dunn, and F. C. Sappington. Application procedures for the electromagnetic borehole flowmeter in shallow unconfined aquifers. *Ground Water Monitoring and Remediation*, 21(4):96–100, 2001.

G. de Marsily, G. Lavedan, M. Boucher, and G. Fasanino. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. *In Geostatistics for Natural Resources Characterization, edited by G Verly, M David, André G Journel, and A Marechal*, Part 2:831–849, 1984.

C. R. Dietrich and G. N. Newsam. A fast and exact method for multidimensional gaussian stochastic simulations. *Water Resources Research*, 29(8):2861–2869, 1993.

P. Dietrich and C. Leven. Direct push-technologies. In R. Kirsch, editor, *Groundwater Geophysics*, pages 347–366. Springer Berlin Heidelberg, 2009.

P. Dietrich, R. Helmig, M. Sauter, H. Hötzl, J. Köngeter, and G. Teutsch. *Flow and Transport in Fractured Porous Media*. Springer-Verlag Berlin Heidelberg, 1 edition, 2005.

P. Dietrich, J. J. Butler, and K. Faiss. A rapid method for hydraulic profiling in unconsolidated formations. *Ground Water*, 46(2):323–328, 2008.

H. D.L., B. W., L. C., C. M., C. F., and K. P.K. Hydraulic tomography: Continuity and discontinuity of high-k and low-k zones. *Ground Water*, In Press, 2015.

J. E. Doherty. Ground water model calibration using pilot points and regularization. *Ground Water*, 41(2):170–177, 2003.

J. E. Doherty, M. N. Fienen, and R. J. Hunt. Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. Technical Report Scientific Investigations Report 2010–5168U, United States Geological Survey, 2010.

K. O. Doro, C. Leven, and O. A. Cirpka. Delineating subsurface heterogeneity at a loop of river steinlach using geophysical and hydrogeological methods. *Environmental Earth Sciences*, 69 (2):335–348, 2013.

K. O. Doro, O. A. Cirpka, and C. Leven. Tracer tomography: Design concepts and field experiments using heat as a tracer. *Groundwater*, 53(S1):139–148, 2015.

A. Efoevi Koudou and C. Ley. Characterizations of GIG laws: a survey complemented with two new results. *ArXiv e-prints*, Dec. 2013.

M. D. Einarson and J. A. Cherry. A new multilevel ground water monitoring system using multichannel tubing. *Ground Water Monitoring and Remediation*, 22(4):52–65, 2002.

D. Erdal and O. A. Cirpka. Joint inference of groundwater-recharge and hydraulic-conductivity fields from head data using the ensemble kalman filter. *Hydrology and Earth System Sciences*, 20(1):555–569, 2016.

D. Erdal, I. Neuweiler, and U. Wollschläger. Using a bias aware enkf to account for unresolved structure in an unsaturated zone model. *Water Resources Research*, 50(1):132–147, 2014.

G. Evensen. Using the extended kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research: Oceans*, 97(C11):17905–17924, 1992.

G. Evensen. 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, 1994.

G. Evensen. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.

G. Evensen. *Data Assimilation*. Springer-Verlag Berlin Heidelberg, 2009.

C. E. Feehley, C. Zheng, and F. J. Molz. A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the macrodispersion experiment (made) site. *Water Resources Research*, 36(9):2501–2515, 2000.

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

J. Gómez-Hernánez, A. Sahuquillo, and J. Capilla. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—i. theory. *Journal of Hydrology*, 203(1):162–174, 1997.

J. Gottlieb and P. Dietrich. Identification of the permeability distribution in soil by hydraulic tomography. *Inverse problems II*, pages 353–360, 1995.

P. Hansen. Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms*, 6(1):1–35, 1994.

J. Heinz. *Sedimentary Geology of Glacial and Periglacial Gravel Bodies (SW-Germany)*. PhD thesis, Eberhard-Karls-Universität Tübingen, 2001.

H. J. Hendricks Franssen and W. Kinzelbach. 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):W09408, 2008.

H. J. Hendricks Franssen, A. Alcolea, M. Riva, M. Bakr, N. van der Wiel, F. Stauffer, and A. Guadagnini. 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, 2009.

H. J. Hendricks Franssen, H. P. Kaiser, U. Kuhlmann, G. Bauser, F. Stauffer, R. Müller, and W. Kinzelbach. 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):n/a–n/a, 2011.

F. Händel and P. Dietrich. Relevance of deterministic structures for modeling of transport: The lauswiesen case study. *Ground Water*, pages 1–8, 2012.

R. Hoffmann and P. Dietrich. An approach to determine equivalent solutions to the geoelectrical 2d inversion problem. *Journal of Applied Geophysics*, 56(2):79–91, 2004.

R. Hu, R. Brauchler, M. Herold, and P. Bayer. Hydraulic tomography analog outcrop study: Combining travel time and steady shape inversion. *Journal of Hydrology*, 409(1–2):350–362, 2011.

R. J. Hunt, J. Doherty, and M. J. Tonkin. Are models too simple? arguments for increased parameterization. *Ground Water*, 45(3):254–262, 2007.

B. Iglewicz and D. Hoaglin. *How to Detect and Handle Outliers*. American Society for Quality Control, Milwaukee WI, 1993.

W. A. Illman. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater*, 52(5):659–684, 2014.

W. A. Illman, X. Liu, and A. Craig. Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms. *Journal of Hydrology*, 341:222–234, 2007.

W. A. Illman, A. J. Craig, and X. Liu. Practical issues in imaging hydraulic conductivity through hydraulic tomography. *Ground Water*, 46(1):120–132, 2008.

W. A. Illman, X. Liu, S. Takeuchi, T.-C. J. Yeh, K. Ando, and H. Saegusa. Hydraulic tomography in fractured granite: Mizunami underground research site, japan. *Water Resources Research*, 45(1):W01406, 2009.

W. A. Illman, S. J. Berg, X. Liu, and A. Massi. Hydraulic/partitioning tracer tomography for dnapl source zone characterization: Small-scale sandbox experiments. *Environmental Science and Technology*, 44(22):8609–8614, 2010.

D. J. Irvine, C. T. Simmons, A. D. Werner, and T. Graf. Heat and solute tracers: How do they compare in heterogeneous aquifers? *Groundwater*, 53(S1):10–20, 2015.

J. W. Jawitz. Moments of truncated continuous univariate distributions. *Advances in Water Resources*, 27(3):269–281, 2004.

S. Jiménez, R. Brauchler, and P. Bayer. A new sequential procedure for hydraulic tomographic inversion. *Advances in Water Resources*, 62, Part A(0):59–70, 2013.

B. Jørgensen. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Lecture Notes in Statistics. Springer-Verlag New York, 1 edition, 1982.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

R. Kirsch and U. Yaramanci. Geophysical characterisation of aquifers. In S. B. Heidelberg, editor, *Groundwater Geophysics*, pages 491–509. Springer Berlin Heidelberg, 2009.

P. K. Kitanidis. Quasi-linear geostatistical theory for inversing. *Water Resources Research*, 31 (10):2411–2419, 1995.

P. K. Kitanidis. Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport. *Water Resources Research*, 51(8):5888–5904, 2015.

K. Kleinert. *Das Grundwasser im Kiesaquifer des oberen Neckartals zwischen Tübingen und Rottenburg*. PhD thesis, Eberhard-Karls-Universität Tübingen, 1976.

M. Klepikova, S. Wildemeersch, T. Hermans, P. Jamin, P. Orban, F. Nguyen, S. Brouyère, and A. Dassargues. Heat tracer test in an alluvial aquifer: Field experiment and inverse modelling. *Journal of Hydrology*, 540:812–823, 2016.

W. Käss. Die unmittelbare bestimmung von uranin-spuren bei färbversuchen [immediate detection of uranine for tracer tests]. *Hydrogeol*, (64):37–65, 1964.

M. Laveneu and G. de Marsily. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.*, 37(11):2659–2675, 2001.

J. Lee and P. K. Kitanidis. Large-scale hydraulic tomography and joint inversion of head and tracer data using the principal component geostatistical approach (pcga). *Water Resources Research*, 50(7):5410–5427, 2014.

S. C. Lessoff, U. Schneidewind, C. Leven, P. Blum, P. Dietrich, and G. Dagan. Spatial characterization of the hydraulic conductivity using direct-push injection logging. *Water Resources Research*, 46, 2010.

C. Leven and P. Dietrich. What information can we get from pumping tests?-comparing pumping test configurations using sensitivity coefficients. *Journal of Hydrology*, (319):199–215, 2006.

L. Li, H. Zhou, J. J. Gómez-Hernández, and H.-J. Hendricks Franssen. Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble kalman filter. *Journal of Hydrology*, 428:152–169, 2012.

T. Li. *Evaluation of a thermal tracer test in a porous aquifer*. PhD thesis, Eberhard-Karls-Universität Tübingen, 2008.

W. Li and O. A. Cirpka. Efficient geostatistical inverse methods for structured and unstructured grids. *Water Resources Research*, 42(6):n/a–n/a, 2006.

W. Li, W. Nowak, and O. Cirpka. Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown. *Water Resources Research*, 41(W08403), 2005.

W. Li, A. Englert, O. Cirpka, J. Vanderborght, and H. Vereecken. Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. *Water Resour. Res.*, 43, 2007.

N. Linde, S. Finsterle, and S. Hubbard. Inversion of tracer test data using tomographic constraints. *Water Resources Research*, 42(4):W04410, 2006.

G. Liu, C. Zheng, and S. M. Gorelick. Evaluation of the applicability of the dual-domain mass transfer model in porous media containing connected high-conductivity channels. *Water Resources Research*, 43(12):W12407, 2007.

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

S. Liu, T. J. Yeh, and R. Gardiner. Effectiveness of hydraulic tomography: Sandbox experiments. *Water Resour. Res.*, 38(4), 2002.

J. Luo, O. A. Cirpka, and P. K. Kitanidis. Temporal-moment matching for truncated breakthrough curves for step or step-pulse injection. *Advances in Water Resources*, 29(9):1306–1313, 2006a.

J. Luo, W. Wu, M. N. Fienen, P. M. Jardine, T. L. Mehlhorn, D. B. Watson, O. A. Cirpka, C. S. Criddle, and P. K. Kitanidis. A nested-cell approach for in situ remediation. *Ground Water*, 44(2):266–274, 2006b.

H. Madsen and R. Canizares. Comparison of extended and ensemble kalman filters for data assimilation in coastal area modelling. *International Journal for Numerical Methods in Fluids*, 31:961–981, 1999.

P. M. Meier, J. Carrera, and X. Sánchez-Vila. An evaluation of jacob's method for the interpretation of pumping tests in heterogeneous formations. *Water Resources Research*, 34(5):1011–1025, 1998.

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

L. Naurath, C. Weidner, T. R. Rüde, and A. Banning. A new approach to quantify na-fluorescein (uranine) in acid mine waters. *Mine Water and the Environment*, 30(3):231–236, 2011.

J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965. doi: $10.1093/\mathrm{comjnl}/7.4.308$. URL `+http://dx.doi.org/10.1093/comjnl/7.4.308`.

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

S. P. Neuman. Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response. *Water Resources Research*, 11(2):329–342, 1975.

C. J. Neville, M. Ibaraki, and E. A. Sudicky. Solute transport with multiprocess nonequilibrium: a semi-analytical solution approach. *Journal of Contaminant Hydrology*, 44(2):141–159, 2000.

J. R. Nimmo and K. A. Mello. Centrifugal techniques for measuring saturated hydraulic conductivity. *Water Resources Research*, 27(6):1263–1269, 1991.

W. Nowak. Best unbiased ensemble linearization and the quasi-linear kalman ensemble generator. *Water Resources Research*, 45(4):W04431, 2009.

W. Nowak and O. Cirpka. Geostatistical inference of hydraulic conductivity and dispersivities from hydraulic heads and tracer data. *Water Resources Research*, 42(8):W08416, 2006.

G. Osterman, K. Keating, A. Binley, and L. Slater. A laboratory study to estimate pore geometric parameters of sandstones using complex conductivity and nuclear magnetic resonance for permeability prediction. *Water Resources Research*, 52(6):4321–4337, 2016.

M. Panzeri, M. Riva, A. Guadagnini, and S. P. Neuman. Data assimilation and parameter estimation via ensemble kalman filter coupled with stochastic moment equations of transient groundwater flow. *Water Resources Research*, 49(3):1334–1344, 2013.

M. Panzeri, M. Riva, A. Guadagnini, and S. P. Neuman. Comparison of ensemble kalman filter groundwater-data assimilation methods based on stochastic moment equations and monte carlo simulation. *Advances in Water Resources*, 66(0):8–18, 2014.

M. Panzeri, M. Riva, A. Guadagnini, and S. P. Neuman. Enkf coupled with groundwater flow moment equations applied to lauswiesen aquifer, germany. *Journal of Hydrology*, 521: 205–216, 2015.

D. Pollock and O. A. Cirpka. Temporal moments in geoelectrical monitoring of salt tracer experiments. *Water Resources Research*, 44(12):n/a–n/a, 2008.

D. Pollock and O. A. Cirpka. Fully coupled hydrogeophysical inversion of synthetic salt tracer experiments. *Water Resources Research*, 46, 2010.

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

T. Ptak, M. Piepenbrink, and E. Martac. Tracer tests for the investigation of heterogeneous porous media and stochastic modelling of flow and transport-a review of some recent developments. *Journal of Hydrology*, pages 122–163, 2004.

B. S. RamaRao, A. M. LaVenue, G. De Marsily, and M. G. Marietta. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. theory and computational experiments. *Water Resources Research*, 31(3):475–493, 1995.

J. Rasmussen, H. Madsen, K. H. Jensen, and J. C. Refsgaard. Data assimilation in integrated hydrological modelling in the presence of observation bias. *Hydrol. Earth Syst. Sci.*, 20(5): 2103–2118, 2016.

A. Rein, R. Hoffmann, and P. Dietrich. Influence of natural time - dependent variations of electrical conductivity on dc resistivity measurements. *Journal of Hydrology*, 285(1): 215–232, 2004.

M. Riva, L. Guadagnini, A. Guadagnini, T. Ptak, and E. Martac. Probabilistic study of well capture zones distribution at the lauswiesen field site. *Contaminant Hydrology*, pages 92–118, 2006.

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

J. Rosas, O. Lopez, T. M. Missimer, K. M. Coulibaly, A. H. A. Dehwah, K. Sesler, L. R. Lujan, and D. Mantilla. Determination of hydraulic conductivity from grain-size distribution for different depositional environments. *Groundwater*, 52(3):399–413, 2014.

B. T. Sack-Kühner. *Einrichtung des Naturmeßfeldes Lauswiesen Tübingen, Erkundung der hydraulischen Eigenschaften, Charakterisierung der Untergrundheterogenität und Vergleich der Ergebnisse unterschiedlicher Erkundungsverfahren*. PhD thesis, Eberhard-Karls-Universität Tübingen, 1996.

E. Sanchez-León, C. Leven, C. P. Haslauer, and O. A. Cirpka. Combining 3d hydraulic tomography with tracer tests for improved transport characterization. *Groundwater*, 54(4): 498–507, 2016.

M. Sardin, D. Schweich, F. J. Leij, and M. T. van Genuchten. Modeling the nonequilibrium transport of linearly interacting solutes in porous media: A review. *Water Resources Research*, 27(9):2287–2307, 1991.

J. Sauty. An analysis of hydrodispersive transfer in aquifers. *WATER RESOURCES RESEARCH*, 16(1):145, 1980.

A. E. Scheidegger. General theory of dispersion in porous media. *Journal of Geophysical Research*, 66(10):3273–3278, 1961.

U. Schneidewind. *Determination of the Hydraulic Conductivity using Direct-Push Injection Logger*. PhD thesis, Eberhard-Karls-Universität Tübingen, 2008.

A. Schöniger, W. Nowak, and H. J. Hendricks Franssen. Parameter estimation by ensemble kalman filters with transformed data: Approach and application to hydraulic tomography. *Water Resources Research*, 48(4):W04502, 2012.

R. L. Schwede, O. A. Cirpka, W. Nowak, and I. Neuweiler. Impact of sampling volume on the probability density function of steady state concentration. *Water Resources Research*, 44 (12):W12433, 2008.

R. L. Schwede, A. Ngo, P. Bastian, O. Ippisch, W. Li, and O. A. Cirpka. Efficient parallelization of geostatistical inversion using the quasi-linear approach. *Computers and Geosciences*, 44 (0):78–85, 2012.

R. L. Schwede, W. Li, C. Leven, and O. A. Cirpka. Three-dimensional geostatistical inversion of synthetic tomographic pumping and heat-tracer tests in a nested-cell setup. *Advances in Water Resources*, 63(0):77–90, 2014.

X. Sánchez-Vila, P. M. Meier, and J. Carrera. Pumping tests in heterogeneous aquifers: An analytical study of what can be obtained from their interpretation using jacob's method. *Water Resources Research*, 35(4):943–952, 1999.

M. Somogyvári and P. Bayer. Field validation of thermal tracer tomography for reconstruction of aquifer heterogeneity. *Water Resources Research*, pages n/a–n/a, 2017.

H. W. Sorenson. Least-squares estimation: from gaussian to kalman. *IEEE Transactions on Knowledge and Data Engineering*, 7:63–68, 1970.

S. Straface, T. J. Yeh, J. Zhu, S. Troisi, and C. Lee. Sequential aquifer tests at a well field, montalto uffugo scalo, italy. *Water Resour. Res.*, 43:1–13, 2006.

C. Theis. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Am. Geophys. Union Trans.*, 16: 519–524, 1935.

R. Therrien and E. A. Sudicky. Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *Journal of Contaminant Hydrology*, 23:1–44, 1996.

C. R. Tiedeman and P. A. Hsieh. Evaluation of longitudinal dispersivity estimates from simulated forced- and natural-gradient tracer tests in heterogeneous aquifers. *Water Resources Research*, 40(1):n/a–n/a, 2004.

J. Tong, B. X. Hu, and J. Yang. Assimilating transient groundwater flow data via a localized ensemble kalman filter to calibrate a heterogeneous conductivity field. *Stochastic Environmental Research and Risk Assessment*, 26(3):467–478, 2012.

J. Tong, B. X. Hu, and J. Yang. Data assimilation methods for estimating a heterogeneous conductivity field by assimilating transient solute transport data via ensemble kalman filter. *Hydrological Processes*, 27(26):3873–3884, 2013.

D. W. Vasco and A. Datta-Gupta. Asymptotic solutions for solute transport: A formalism for tracer tomography. *Water Resources Research*, 35(1):1–16, 1999.

D. W. Vasco, A. Datta-Gupta, and J. C. S. Long. Resolution and uncertainty in hydrologic characterization. *Water Resources Research*, 33(3):379–397, 1997.

D. W. Vasco, S. Yoon, and A. Datta-Gupta. Integrating dynamic data into high-resolution reservoir models using streamline-based analytic sensitivity coefficients, 1998/1/1/ 1999.

H. Vereecken, A. Kemna, H.-M. Münch, A. Tillmann, and A. Verweerd. Aquifer characterization by geophysical methods. In *Encyclopedia of Hydrological Sciences*. John Wiley and Sons, Ltd, 2006.

V. Wagner, T. Li, P. Bayer, C. Leven, P. Dietrich, and P. Blum. Thermal tracer testing in a sedimentary aquifer: field experiment (lauswiesen, germany) and numerical simulation. *Hydrogeology Journal*, 22(1):175–187, 2014.

C. Weidner, L. Naurath, T. R. Rüde, and A. Banning. Parameters affecting na-fluorescein (uranine) detection in mine water tracer tests, 2011.

G. Welch and G. Bishop. An introduction to the kalman filter. Technical report, University of North Carolina at Chapell Hill, Department of Computer Sciences, 2006.

C. Welty and L. W. Gelhar. Evaluation of longitudinal dispersivity from nonuniform flow tracer tests. *Journal of Hydrology*, 153(1):71–102, 1994.

X.-H. Wen and W. H. Chen. Real-time reservoir model updating using ensemble kalman filter with confirming option. 2006.

S. H. Wolf, M. A. Celia, and K. M. Hess. Evaluation of hydraulic conductivities calculated from multiport-permeameter measurements. *Ground Water*, 29(4):516–525, 1991.

J. A. Woods, M. D. Teubner, C. T. Simmons, and K. A. Narayan. Numerical error in groundwater flow and solute transport simulation. *Water Resources Research*, 39(6):n/a–n/a, 2003.

J. Xiang, T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen. A simultaneous successive linear estimator and a guide for hydraulic tomography analysis. *Water Resources Research*, 45(2): W02432, 2009.

T. Xu, J. Jaime Gómez-Hernández, H. Zhou, and L. Li. The power of transient piezometric head data in inverse modeling: An application of the localized normal-score enkf with covariance inflation in a heterogenous bimodal hydraulic conductivity field. *Advances in Water Resources*, 54:100–118, 2013.

T. Yeh and S. Liu. Hydraulic tomography: Development of a new aquifer test method. *Water Resour. Res.*, 36(8):2095–2105, 2000.

T.-C. J. Yeh and J. Zhu. Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones. *Water Resources Research*, 43(6):W06435, 2007.

T. J. Yeh, D. Mao, Y. Zha, K. Hsu, C. Lee, J. Wen, W. Lu, and J. Yang. Why hydraulic tomography works? *Ground Water*, pages 1–5, 2013.

D. Yin and W. A. Illman. Hydraulic tomography using temporal moments of drawdown recovery data: A laboratory sandbox study. *Water Resources Research*, 45(1):W01502, 2009.

H. Zhou, J. J. Gómez-Hernández, H.-J. Hendricks Franssen, and L. Li. An approach to handling non-gaussianity of parameters and state variables in ensemble kalman filtering. *Advances in Water Resources*, 34(7):844–864, 2011.

H. Zhou, J. J. Gómez-Hernández, and L. Li. Inverse methods in hydrogeology: Evolution and recent trends. *Advances in Water Resources*, 63(0):22–37, 2014.

J. Zhu and T.-C. J. Yeh. Analysis of hydraulic tomography using temporal moments of drawdown recovery data. *Water Resources Research*, 42(2):W02403, 2006.

J. Zhu and T. J. Yeh. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resour. Res.*, 41, 2005.

J. Zhu, X. Cai, and T.-C. Jim Yeh. Analysis of tracer tomography using temporal moments of tracer breakthrough curves. *Advances in Water Resources*, 32(3):391–400, 2009.

D. A. Zimmerman, G. de Marsily, C. A. Gotway, M. G. Marietta, C. L. Axness, R. L. Beauheim, R. L. Bras, J. Carrera, G. Dagan, P. B. Davies, D. P. Gallegos, A. Galli, J. Gómez-Hernández, P. Grindrod, A. L. Gutjahr, P. K. Kitanidis, A. M. Lavenue, D. McLaughlin, S. P. Neuman, B. S. RamaRao, C. Ravenne, and Y. Rubin. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. *Water Resources Research*, 34(6):1373–1413, 1998.