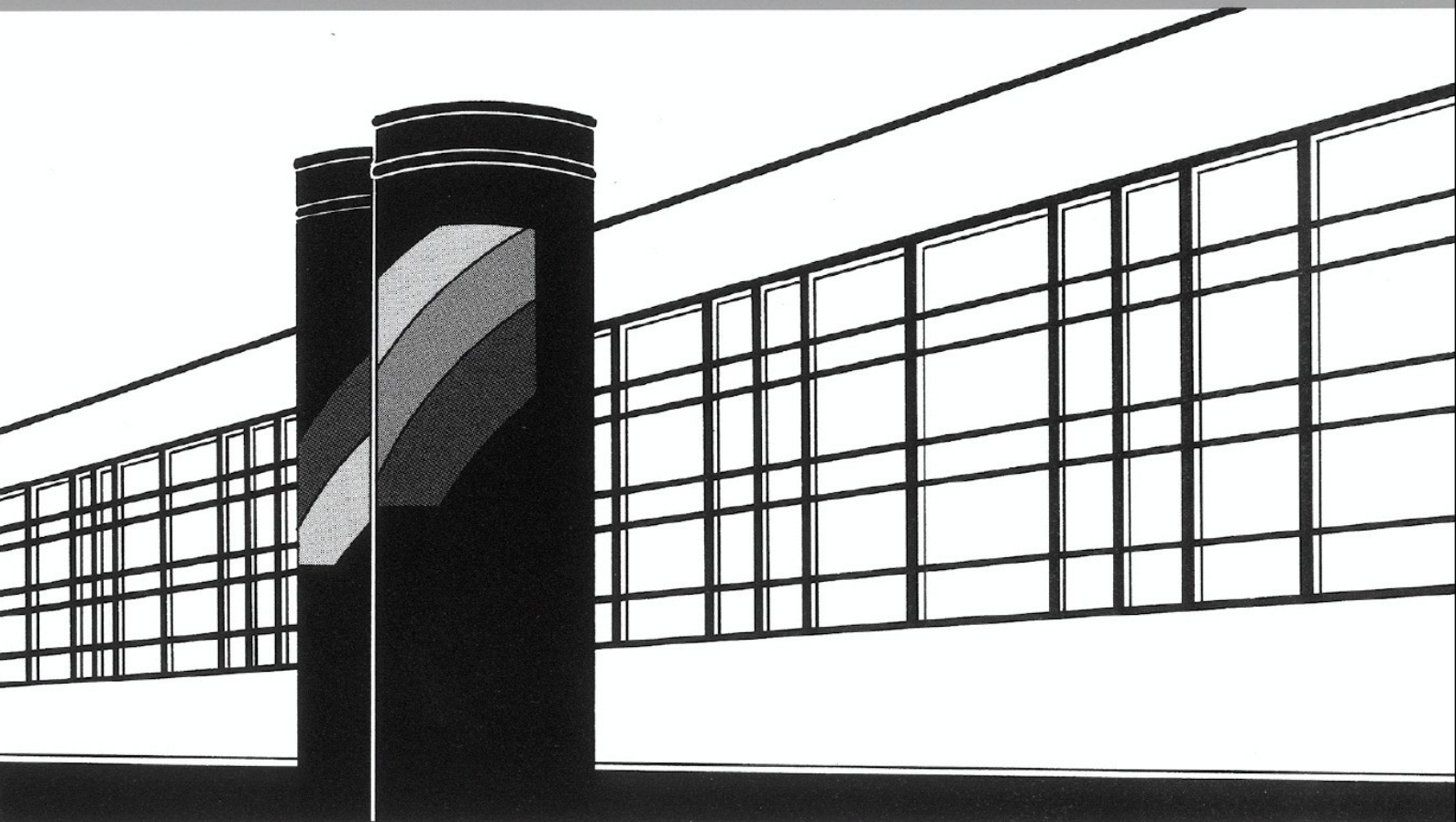


Universität Stuttgart



Institut für Wasser- und Umweltsystemmodellierung

# *Mitteilungen*



Heft 233 Jonas Koch

Simulation, Identification and  
Characterization of Contaminant Source  
Architectures in the Subsurface







# **Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface**

Von der Fakultät Bau- und Umweltingenieurwissenschaften  
der Universität Stuttgart zur Erlangung der Würde eines Doktors der  
Ingenieurwissenschaften (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von

**Jonas Bernhard Maria Koch**

aus Freiburg i. Br.

Hauptberichter:

Jun.-Prof. Dr.-Ing. Wolfgang Nowak

Mitberichter:

Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy

Mitberichter:

Prof. Dr. ir. Timo J. Heimovaara

Tag der mündlichen Prüfung: 05. Juni 2014

Institut für Wasser- und Umweltsystemmodellierung  
der Universität Stuttgart  
2014



Heft 233      Simulation, Identification and  
Characterization of Contaminant  
Source Architectures in the  
Subsurface

von  
Dr.-Ing.  
Jonas Bernhard Maria Koch

**D93 Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface**

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://www.d-nb.de> abrufbar

**Koch, Jonas Bernhard Maria:**

Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2014

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 233)

Zugl.: Stuttgart, Univ., Diss., 2014

ISBN 978-3-942036-37-5

NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen

Gegen Vervielfältigung und Übersetzung bestehen keine Einwände, es wird lediglich um Quellenangabe gebeten.

Herausgegeben 2014 vom Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung

Druck: Document Center S. Kästl, Ostfildern



# Acknowledgments

First of all I would like to gratefully thank my supervisor Wolfgang Nowak for cheerful, inspiring, and relaxed discussions about science, my thesis work, and many other crucial issues. I'm grateful for your guidance into scientific practice and your unshakable faith that everything turns out smoothly in the end. Furthermore, I would like to express my gratitude to my co-supervisors Timo Heimovaara and András Bárdossy. Special thanks are due to Felipe de Barros and Marco Dentz for the insightful and thrilling excursions into flow topology, mixing, Barcelona, and the world of rumors. I had a wonderful time in Barcelona and meeting you guys at conferences and elsewhere was always a pleasure. My last but not least special thanks go to all those who spent time with me - in daily life, Freiburg, the university, summer schools, conferences, and in other exciting and cheerful trips.



# Contents

<b>List of Figures</b>	<b>V</b>
<b>List of Tables</b>	<b>VII</b>
<b>Notation</b>	<b>IX</b>
<b>Abstract</b>	<b>XIII</b>
<b>Kurzfassung</b>	<b>XVII</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 State of the art . . . . .	4
1.3 Goals and approach . . . . .	5
1.4 Structure of the thesis . . . . .	6
<b>2 Fundamentals</b>	<b>9</b>
2.1 Basic concepts to treat uncertainty . . . . .	9
2.2 Important metrics for contaminated site assessments . . . . .	13
2.3 Relevant processes and their interdependencies . . . . .	16
2.4 Crucial issues . . . . .	21
2.5 Geostatistic description of heterogeneous systems . . . . .	24
2.5.1 Descriptive statistics . . . . .	24
2.5.2 Distribution functions . . . . .	25
2.5.3 Geostatistical approach . . . . .	27
2.6 Governing equations . . . . .	32
2.6.1 Fluid and soil parameters . . . . .	32
2.6.2 Multiphase flow equation system of contaminated source zones . . . . .	35
2.6.3 Tailored equation system . . . . .	36
<b>3 Consistent model framework</b>	<b>39</b>
3.1 Goal, approach and contributions . . . . .	42
3.1.1 Goal and approach . . . . .	42
3.1.2 Relation to the state-of-the-art . . . . .	42
3.2 Stochastic DNAPL percolation . . . . .	44
3.3 Groundwater flow, mass transfer and solute transport . . . . .	47
3.4 Quasi-steady-state DNAPL depletion . . . . .	50

---

3.5	Results and discussion . . . . .	51
3.5.1	Test case scenarios . . . . .	52
3.5.2	Discussion of mass discharge results . . . . .	55
3.5.3	Discussion of results for source depletion times . . . . .	57
3.5.4	Local parameter statistics and interdependencies . . . . .	59
3.6	Summary and conclusions . . . . .	59
<b>4</b>	<b>Dirichlet and third-type boundary conditions in PTRW simulations</b>	<b>63</b>
4.1	Goal, approach and contributions . . . . .	65
4.1.1	Goal and approach . . . . .	65
4.1.2	Existing alternative methods . . . . .	66
4.2	Concentration reconstruction . . . . .	67
4.3	Discretizing the boundary condition . . . . .	69
4.4	Controlling the boundary condition . . . . .	72
4.5	Benchmarking . . . . .	75
4.5.1	Validation with analytical solution (test case 1) . . . . .	76
4.5.2	Benchmark against MASST (test case 2) . . . . .	77
4.5.3	Benchmark against FEM (test case 3) . . . . .	83
4.5.4	Third-type boundary condition vs. local chemical equilibrium . . . . .	88
4.6	Summary and conclusion . . . . .	90
<b>5</b>	<b>CSA identification</b>	<b>93</b>
5.1	Goal, approach and contributions . . . . .	94
5.1.1	Goal and approach . . . . .	94
5.1.2	Relation to state of the art . . . . .	96
5.2	Inverse modelling tools . . . . .	99
5.2.1	Bayesian inference . . . . .	100
5.2.2	Adjoint states . . . . .	102
5.3	Reverse-inverse approach . . . . .	104
5.3.1	Assimilating concentration data . . . . .	105
5.3.2	Assimilation of various data types . . . . .	107
5.4	Langrangian reverse formulation . . . . .	108
5.5	Application examples . . . . .	112
5.5.1	Probability of DNAPL presence . . . . .	114
5.5.2	Identification of CSA characteristics . . . . .	117
5.5.3	Conditional probability distribution of mass discharge predictions . . . . .	121
5.5.4	Discussion on inversion performance . . . . .	122
5.6	Summary and conclusions . . . . .	123
<b>6</b>	<b>Plume deformation and mixing of the dissolved contaminant</b>	<b>125</b>
6.1	Goal, approach and contributions . . . . .	126
6.1.1	Goal and approach . . . . .	126
6.1.2	Relation to the state-of-the-art . . . . .	127
6.2	Basics . . . . .	127

---

6.3	Scalar mixing in two-dimensional spatially heterogeneous flow fields . . . . .	128
6.3.1	Local mechanisms . . . . .	129
6.3.2	Global mixing . . . . .	132
6.4	Mechanisms of enhanced mixing in three-dimensional heterogeneous flow fields . . . . .	136
6.4.1	Numerical experiment . . . . .	136
6.4.2	Metrics of scalar mixing enhancement . . . . .	137
6.4.3	Stochastic analysis . . . . .	139
6.4.4	Summary and conclusions . . . . .	141
<b>7</b>	<b>Summary, conclusions and outlook</b>	<b>143</b>
7.1	Summary of working steps . . . . .	144
7.2	Summary of conclusions . . . . .	145
7.3	Overall conclusions . . . . .	146
7.4	Outlook . . . . .	147
	<b>Bibliography</b>	<b>149</b>



# List of Figures

1.1	Illustration of a consistent simulation framework . . . . .	3
1.2	CSA identification from concentration observations . . . . .	3
1.3	Loss of information due to solute mixing and plume deformation . . . . .	3
2.1	Illustration of relevant metrics for contaminated site assessment . . . . .	15
2.2	Stagnant film model . . . . .	19
2.3	Standard-normal probability density function . . . . .	27
2.4	Illustration of a bivariate probability density function . . . . .	28
2.5	Illustration of three covariance functions from the Matérn family . . . . .	30
2.6	Illustration of two random log-permeability fields . . . . .	31
3.1	Proposed consistent model framework . . . . .	40
3.2	Illustration of physically-based random CSAs . . . . .	46
3.3	Mass release rates from a complex-shaped DNAPL source zone . . . . .	49
3.4	Illustration of a dissolved-phase concentration plume . . . . .	49
3.5	DNAPL saturation values at three quasy-steady-states . . . . .	52
3.6	Five scenarios to assess the required model complexity . . . . .	52
3.7	Illustration of base case scenario to assess the required model complexity . . . . .	55
3.8	Probability density functions for log mass discharge for S1-S5 . . . . .	56
3.9	CSA depletion curves and <i>pdf</i> of log-depletion time . . . . .	58
3.10	<i>pdfs</i> of log hydraulic permeability and of log absolute velocity . . . . .	60
4.1	Convergence analysis of the new PTRW method against an analytical solution . . . . .	78
4.2	Benchmarking of the new PTRW method against MASST (test case 1) . . . . .	79
4.3	Comparison of the mass release densities (test case 2) . . . . .	82
4.4	Comparison of the concentration profiles within a control plane (test case 2) . . . . .	83
4.5	Comparison of one-dimensional concentration profiles (test case 2) . . . . .	84
4.6	Benchmarking of the new method against FEM (test case 3, aquifer 1) . . . . .	86
4.7	Comparison of mass release densities (test case 3) . . . . .	87
4.8	Comparison of the concentration profiles within a control plane (test case 3) . . . . .	88
4.9	Comparison of one-dimensional concentration profiles (test case 3) . . . . .	89
4.10	Illustration of mass release densities for third-type boundary conditions . . . . .	90
5.1	Approach of reverse-inverse CSA identification . . . . .	107
5.2	Linking the CSA with downstream concentration values . . . . .	110
5.3	Comparison of concentration profiles obtained with RPTRW and PTRW . . . . .	112
5.4	Illustration of the measurement design and the unknown CSA . . . . .	114
5.5	Apriori probability of DNAPL residence . . . . .	116

---

5.6	Probability of DNAPL presence with respect to $\mathbf{K}_o$ and $\mathbf{S}_{n,o}$ . . . . .	117
5.7	Probability of DNAPL presence with respect to $\mathbf{K}_o$ , $\mathbf{S}_{n,o}$ , and $\mathbf{c}_o$ . . . . .	118
5.8	Comparison of several conditional CSAs to the CSA of the synthetic reality .	119
5.9	Inference of the ganglia-to-pool ratio of the unknown CSA . . . . .	120
5.10	Inference of the area spanned by the unknown CSA in $x_2$ and $x_3$ direction . .	121
5.11	Conditional probability distributions of mass discharge predictions . . . . .	122
6.1	Map of the local values of the Okubo-Weiss parameter . . . . .	133
6.2	Effective Okubo-Weiss parameter and system entropy versus time . . . . .	135
6.3	Experimental setup of the numerical simulations . . . . .	137
6.4	Illustration of the three proposed metrics to quantify mixing enhancement .	138
6.5	Mixing enhancement for various realizations . . . . .	139
6.6	Average dilution curves for different variances of the geostatistical model . .	140
6.7	Asymptotic enhanced mixing versus various mixing mechanisms . . . . .	141



# List of Tables

2.1	Time scales of DNAPL processes in the subsurface . . . . .	23
3.1	Input Parameters of the test case scenarios . . . . .	53
3.2	Scenarios to test the requirement of the proposed model consistency. . . . .	54
4.1	Parameter values used in the test cases (Sec. 4.5). . . . .	77
5.1	Features of the measurement campaign . . . . .	113



# Notation

The following table shows the significant symbols used in this work.

Symbol	Definition	Dimension
<b>Greek Letters:</b>		
$\alpha$	Stretching deformation rate	$[T^{-1}]$
$\alpha_L$	Longitudinal dispersivity	$[L]$
$\alpha_T$	Transversal dispersivity	$[L]$
$\gamma$	Surface tension	$[M T^{-2}]$
$\Gamma$	Boundary domain	
$\varepsilon$	Rescaling vector of unit mass release rates	$[-]$
$\epsilon$	Deformation tensor	$[T^{-1}]$
$\eta$	Shear deformation	$[T^{-1}]$
$\theta_s$	Viscous dissipation rate	$[T^{-1}]$
$\Theta$	Okubo-Weiss parameter	$[T^{-2}]$
$\Theta^e$	Effective Okubo-Weiss parameter	$[T^{-2}]$
$\iota$	Random number uniformly distributed between 0 and 1	
$\kappa_m$	Shape parameter of Matérn family	
$\kappa$	Mass transfer coefficient	$[L T^{-1}]$
$\kappa_{\text{eff}}$	Lumped mass transfer coefficient	$[T^{-1}]$
$\lambda$	Vector of eigenvalues	
$\lambda_b$	Brooks & Corey shape parameter	
$\mu$	Mean value of quantity specified by subscript	
$\mu_\alpha$	Dynamic viscosity of phase $\alpha$	$[M L^{-1} T^{-1}]$
$\xi(t)$	Gaussian white noise vector	
$\rho_\alpha$	Density of phase $\alpha$	$[M L^{-3}]$
$\sigma$	Standard deviation of quantity specified by subscript	
$\Sigma$	Kernel bandwidth matrix	
$\tau$	Transformed time specified in the individual chapters	$[T]$
$v$	Random number, Gaussian distributed	
$\phi$	Porosity	$[-]$
$\phi_e$	Effective porosity	$[-]$
$\varphi$	Contact angle	$[\text{rad}]$
$\chi$	Random vector with zero mean and covariance matrix $\mathbf{R}$	

$\psi$	Adjoint state variable	
$\omega$	Vorticity	$[T^{-1}]$
$\Omega$	Model domain	

**Latin Letters:**

<b>A</b>	Drift vector of particle displacement	$[L T^{-1}]$
$A_i$	Interfacial area	$[L^2]$
$A_{x_2, x_3}$	CSA discharge area to mean flow direction	$[L^2]$
<b>B</b>	PTRW displacement matrix for diffusion/dispersion	$[L T^{-1}]$
$c$	Solute concentration	$[M L^{-3}]$
$c_s$	Solute concentration at chemical equilibrium	$[M L^{-3}]$
$c_o$	Solute concentration observation	$[M L^{-3}]$
$c_{cas}$	Solute concentration within the source zone	$[M L^{-3}]$
$Ca$	Capillary number	[-]
$d$	Dimensionality	[-]
$d_m$	Median grain size	[L]
$D_m$	Molecular diffusion coefficient	$[L^2 T^{-1}]$
<b>D</b>	Dispersion tensor	$[L^2 T^{-1}]$
$E$	Dilution index	$[L^d]$
<b>g</b>	Gravitational vector	$[L T^{-2}]$
<b>H</b>	Sensitivity matrix	
$I_p$	Invasion potential	$[M L^{-1} T^{-2}]$
$k_r$	Relative permeability	[-]
<b>K</b>	Intrinsic permeability	$[L^2]$
<b>K<sub>o</sub></b>	Permeability observation	$[L^2]$
$\ell$	Correlation length as specified by subscript	[L]
<b>M</b>	Boundary related mass matrix	$[T^{-1}]$
$m$	Mass of quantity specified by subscript	[M]
$\dot{m}$	Mass release rate of quantity specified by subscript	$[M T^{-1}]$
$\dot{m}_{csa}$	DNAPL mass release rate emanating from the CSA	$[M T^{-1}]$
$\dot{m}_{tot}$	Total DNAPL mass release rate	$[M T^{-1}]$
<b>n</b>	Unit vector normal to interface	[-]
$p_\alpha$	Phase pressure of phase $\alpha$	$[M L^{-1} T^{-2}]$
$p_c$	Capillary pressure	$[M L^{-1} T^{-2}]$
$p_d$	Entry pressure	$[M L^{-1} T^{-2}]$
$q_\alpha^C$	Source/ sink term of component C in phase $\alpha$	$[M L^{-3} T^{-1}]$
$r$	Pore radius	[L]
<b>R</b>	Covariance matrix of measurement errors	
$S_\alpha$	Saturation of phase $\alpha$	[-]
<b>S</b>	Transport related mass matrix	$[T^{-1}]$
$t$	Time	[T]
<b>T</b>	Modified mass matrix	$[T^{-1}]$
$v_s$	Seepage velocity of the aqueous phase	$[L T^{-1}]$

---

$V$	Support volume as specified by subscript	$[L^3]$
$w$	A posteriori weights	
$\mathbf{x}$	Spatial dimensions	$[L, L, L]$
$X_\alpha^C$	Mass fraction of component $C$ and phase $\alpha$	$[-]$
$\mathbf{X}_p$	Particle position matrix	
$Y$	Log-permeability	

### Operators and Other Symbols:

$\partial ()$	Partial derivative
$\Delta$	Difference operator
$\nabla$	Nabla operator
$\mathcal{L}$	Differential operator of arbitrary order
$() \cdot ()$	Scalar product
$\delta ()$	Dirac function
$\delta_{ij}$	Kronecker symbol
$H ()$	Heaviside function
$E []$	Expected value operator
$Var []$	Variance operator
$Cov []$	Covariance operator
$C()$	Matérn covariance function
$K()$	Kernel function as specified by subscript
$\mathcal{U}(a, b)$	Uniform distribution between $a$ and $b$
$\mathcal{N}(\mu, \sigma)$	Gaussian distribution with mean $\mu$ and standard deviation $\sigma$

**Abbreviations:**

ADE	Advection Dispersion Equation
<i>cdf</i>	Cumulative Density Function
CSA	Contaminant Source Architecture
FEM	Finite Element Method
GTP	Ganglia-to-Pool Ratio
LCE	Local Chemical Equilibrium
MASST	Multiple analytical source superposition technique
<i>pdf</i>	Probability Density Function
PTRW	Particle Tracking Random Walk
RPTRW	Reverse Particle Tracking Random Walk

**Specific super- and subscripts:**

$\alpha$	Phase
$C$	Component
$n$	DNAPL
$w$	Aqueous-phase
$o$	Observation
*	Adjoint state/problem/operator

# Abstract

**Relevance:** Contaminated sites occur abundantly worldwide, posing severe risks for natural resources, the ecosystem and human health. This work focuses on sites contaminated by dense non-aqueous phase liquids (DNAPL). Once a DNAPL is released into the subsurface, it serves as a persistent source of dissolved-phase contamination. The DNAPL migrates through the porous medium and penetrates the aquifer, then it forms a complex pattern of immobile DNAPL saturation, dissolves into the groundwater and forms a contaminant plume within the groundwater, and, finally, it slowly depletes in the long term.

**Open challenges:** Although these issues have been known for decades, several important assessment tasks at contaminated sites (CSs) still pose research challenges. These tasks include a quantitative-risk-based prioritization of CSs for remediation efforts, the choice of a sound remediation strategy under uncertainty, and a robust monitoring of the remediation success. In industrialized countries, the number of such CSs is tremendously high to the point that a ranking from most risky to least risky is advisable. Decisions whether an expensive remediation effort is required or monitored natural attenuation is more adequate, and whether the remediation effort should be focused on the contaminant source, the dissolved plume, or on a few target locations that are sensitive to human activities (such as drinking water wells) are rather difficult [Falta, 2008]. In order to tackle these tasks in an objective and reasonable manner, a proper set of decision metrics that are calculated with a model framework that is consistent in terms of both physical processes and stochastic reasoning is demanded. These metrics include mass flux probabilities indicating the potential impact that a CS may pose on an ecosystem or on human activities, depletion time probabilities indicating the persistence of the potential impact over time and indicating which remediation strategies may be feasible, and characteristics of the contaminant source geometry indicating suitable remediation techniques [e.g., Soga *et al.*, 2004; Lemke and Abriola, 2006]. These metrics are referred to as characteristics of the contaminant source architecture (CSA).

**Focus of work:** A consistent model framework. Two features of a consistent model framework are of particular importance, and thus form the focus of this work. First, a **physically consistent** model framework should treat all crucial physical process sufficiently accurately, including the mutual interplay between the characteristics of a CSA, input parameters (e.g., the permeability field), and intermediate model results, such as the groundwater flow field. Second, a **stochastically consistent** model framework accounts for all relevant uncertainties, including parameter uncertainty (e.g., due to heterogeneous aquifers and only scarce, noisy, and punctual data [Gomez-Hernandez *et al.*, 1997]), observation uncertainty, and model-structural uncertainty. The latter may arise due to unpredictable pore-scale processes

between the percolating contaminant phase and the soil matrix, leading to complex and unpredictable contaminant source geometry [Feenstra and Cherry, 1996]. Although both physical and stochastic requirements are listed here separately, they have to accompany each other in a consistent model framework. Unfortunately, the desired physical and stochastic consistency is often hampered by the fact that the crucial fine-scale process, the desired field-scale answers, and the computational costs counteract against each other in numerical modes. The CS assessment faces strong uncertainties leading to high-dimensional random parameter spaces for both input parameters and predictions. Thus, efficient data assimilation is indispensable in order to specify the model input and the model predictions, and thus to gain prediction confidence. While techniques for the assimilation of permeability and head observations are well explored, there is huge potential of gaining essential insights of CSAs from downstream concentration observations, if they are used in a consistent model framework [e.g. Schwede and Cirpka, 2010; Trolborg *et al.*, 2010]. This, however, requires inverse modeling which is typically non-unique and computationally demanding.

**Goals:** The goal of this work is to develop a physically and stochastically consistent model framework that captures all relevant processes and parameter interdependencies, accounts for all relevant uncertainties, and is still applicable for large Monte Carlo simulations and inverse modelling. Therefore, random space functions of realistic CSAs are generated and incorporated with a transport model that features aquifer heterogeneities and irregular flow fields. This general model framework is accomplished by an efficient reverse-inverse model for assimilating downgradient concentration data in order to infer CSA characteristics, explain mass flux and depletion behaviour, find suitable remediation and monitoring strategies, and assign polluter responsibilities. To this end, a probabilistic simulation framework that estimates probability density functions of mass discharge, source depletion time, and critical concentration values at crucial target locations is conceptualized in this work. Furthermore, it supports the inference of contaminant source architectures from arbitrary site data. To achieve the stated goals, the work is separated into three steps:

1. developing a consistent model that fulfills the desired physical and statistical accuracy, and efficiency for the purpose of explaining, simulating and predicting the fate of DNAPL in the subsurface,
2. coupling this model with an efficient reverse-inverse model within a Bayesian framework to infer CSA characteristics from downgradient concentration observations, and
3. studying mechanisms that enhance the dilution of the dissolved-phase plume since the dilution state of concentration observations strongly influences the inference quality of CSA characteristics.

The three milestones of this thesis build upon each other, and are described in the following three paragraphs.

**Step 1. Consistent model framework:** The question is raised: what is an adequate level of model complexity and how to obtain a physically and statistically consistent model



framework? Therefore, a model framework that avoids as many as reasonably possible simplifications is developed. Then this model is used to assess the impact of specific simplifications on mass discharge and source depletion time predictions, such as block-shaped CSA geometries and homogenized aquifer parameters. The model framework simulates DNAPL infiltration, dissolution and transport in the aqueous phase, and depletion. With this, mass flux and source depletion time probability density functions are estimated in a Monte Carlo simulation. To achieve such a sound model framework three novel methods are developed in this thesis:

1. Random CSAs are generated as functions of non-stationary entry pressure fields by a stochastic invasion percolation (SIP) algorithm based on *Ewing and Berkowitz [1998]*. The well-established pore-scale concept of invasion percolation algorithms was adapted to the Darcy scale and modified to account for the uncertainty of CSAs, which is due to unknown aquifer properties at small scales (below Darcy scale) and instabilities of the drainage process when DNAPL is displacing the aqueous phase (Sec. 3.2)
2. The dissolution of trapped contaminants and the advective-dispersive transport of dissolved contaminant phase are modelled by a particle tracking random walk technique that allows for the implementation of complex-shaped Dirichlet or third-type boundary conditions [*Koch and Nowak, 2014a*]. Once the emission rates from the CSA are obtained, concentration values may be reconstructed within the entire domain by a quick post-processing step.
3. If one is interested in the depletion time of the contaminant phase, one can use the CSA's emission rates as the sink term in the respective contaminant mass balance for the CSA. If the emitter-wise mass balance indicates that the donor phase has been depleted at specific locations within the CSA, the associated terms are simply erased from all equations, and the new PTRW result follows immediately.

The accuracy and performance of the proposed method will be demonstrated in an extensive numerical convergence assessment. In a set of test cases, it will be shown that the proposed model complexity is actually required to account for the key processes and statistical dependencies, and yet the proposed model framework is sufficiently efficient to be suitable for Monte Carlo simulations of CSAs, mass release rates, depletion time, and so forth. This work highlights the relevance of applying physically-based and random CSAs and accounting for their relevant interplays with aquifer parameters and groundwater flow.

**Step 2. Data assimilation via a reverse-inverse approach:** This step derives the theory for inferring CSA via Bayesian updating and presents an efficient reverse-inverse methodology. The prime objectives are to infer CSAs from various field observations such as soil parameters, pure-phase DNAPL presence, and concentration values. Obviously, the previously described parameter dependencies in the forward model (step 1) are also valid and required when inferring one of these parameters. In fact, if a model is calibrated to fit a few observations without maintaining the physical interlinkage between all parameters, the model may lose its ability to interpolate and extrapolate in a physically accurate manner from these measurement locations and times. Therefore, the coupling between CSA and

their associated permeability field always has to be preserved. Within this step, contaminant source zones are conceptualized as random space functions (see work step 1), and are inferred via stochastic inverse modelling. A Bayesian geostatistical approach is chosen in order to cope with the issue of sparse and noisy data, and uncertainty in aquifer structures. This approach solves for adjoint states in a Lagrangian framework leading to a swift transfer function that dramatically reduces the computational burden of common inverse models. With improved knowledge of CSA, better estimates and confidence intervals of contaminant mass flux, total mass, and source depletion times are possible. Step 2 points out the relevance of concentration observations for inferring knowledge on CSA and mass discharge.

**Step 3. Plume deformation and enhanced dilution:** Scalar mixing plays a significant role for transport in geophysical flows because it controls dilution and is a main driver for many chemical reactions. For example, the mixing of dissolved contaminants may diminish the information content of downstream concentration observations dramatically. This step explores the local-scale flow mechanisms that lead to enhanced scalar mixing and studies their impact on the global mixing behavior [*de Barros et al.*, 2012]. A first part of this step analyses the enhanced dilution in a two-dimensional system. Here it is shown that the evolution of the mixing state is directly linked to the flow topology in terms of the Okubo-Weiss parameter  $\theta$  which describes the relative importance between local strain and rotation properties. Sequentially, it is postulated that the same mechanisms are also one of the key drivers in enhanced mixing in three-dimensional porous media. However, they interact differently and additional processes which are unique to three-dimensional systems may arise. Thus, the previously derived link between kinematical features and mixing changes. For instance, three-dimensional trajectories may twist and wrap, which might pull apart initially coherent spatial structures, generating even more contact surface with the surrounding fluid. In general, mechanisms such as shearing deformation, stretching deformation and vorticity gain an additional dimension to work on. In a stochastic framework, kinematical mechanisms and their impact on predefined mixing metrics are analyzed. This is beneficial to the overall understanding of the mixing of a solute cloud and its underlying processes in three-dimensional heterogeneous porous media. Profound knowledge of these mixing mechanisms could be used to conceptualize promising field campaigns whose data is highly informative for the inversion of CSA characteristics and mass discharge estimates.

**Summary:** As an essential novelty, the mutual dependencies of the key parameters and interacting physical processes are taken into account throughout the whole work. In an uncertain and heterogeneous subsurface setting, I will highlight the relevance of three parameter fields: the local velocities, the hydraulic permeabilities and the DNAPL phase saturations. Obviously, these parameters depend on each other during DNAPL infiltration, dissolution, dissolved-phase transport, and depletion. Equivalently, the physical and statistical interplay of these parameters has to be accounted for when CSAs are identified with observed data. In order to solve this non-linear and non-unique inverse problem, an efficient and accurate inverse method is proposed. Since concentration observations are highly valuable data for this purpose, the processes of mixing and plume deformation and their effect on enhanced dilution and thus the diminishing of the data worth is studied.

# Kurzfassung

**Relevanz:** Kontaminierte Standorte (KS) stellen ein hohes Risiko für die menschliche Gesundheit, das Ökosystem und natürliche Ressourcen dar. Weltweit, und besonders häufig in den Industrienationen, wurde in den vergangenen Jahrzehnten der Untergrund kontaminiert. Diese Arbeit befasst sich mit flüssigen Schadstoffen, die schwerer als Wasser sind und deshalb „dense non-aqueous phase liquids“ (DNAPL) genannt werden. Diese wurden und werden zum Beispiel in vielen Produktionsstandorten als Lösungsmittel verwendet. Beispiele für DNAPL sind Trichlorethylen (TCE), Tetrachlorethylen (PCE), oder polycyclische aromatische Kohlenwasserstoffe (PAK). Wenn diese Gefahrstoffe einmal in den Untergrund gelangen, stellen sie eine dauerhafte Kontamination des Grundwasser in gelöster Form dar.

Die schweren Schadstoffe durchwandern als eigenständige Phase das Porengerüst des Grundwasserleiters in komplizierten Wegen und formen individuelle Geometrien. Sie lösen sich im Grundwasser nur sehr langsam auf. Die geringen aber dennoch schädlichen Konzentrationen formen eine Abstromfahne und breiten sich weit im Grundwasser aus. Allmählich löst sich der gesamte Schadensherd schließlich auf. Dieser Prozess kann aber Jahrzehnte oder sogar über ein Jahrhundert andauern.

**Forschungsbedarf:** Obwohl diese Probleme schon seit Jahrzehnten bekannt und Gegenstand intensiver Forschung waren und sind, bleiben wichtige Fragen weiterhin offen. Zum Beispiel ist ungeklärt, wie eine rationale Beurteilung und Einstufung des Risikos verschiedenster kontaminierter Standorte untereinander genau durchzuführen ist. Letzteres setzt, aufgrund der großen Unsicherheiten, die dieses undurchsichtige physikalische System (Grundwasserleiter) in sich birgt, eine quantitative Wahrscheinlichkeits-Analyse voraus. In den Industrienationen ist die Zahl an kontaminierten und möglicherweise kontaminierten Standorten so groß, dass die einzelnen Sanierungen nur nach einer Reihenfolge gemäß des von ihnen ausgehenden Risikos ausgeführt werden sollten.

Falls die Risikoabschätzung ergeben hat, dass ein bestimmter Standort saniert werden sollte, muss die Sanierungstechnik und -strategie unter erheblichen Ungewissheiten gewählt werden. Dies verlangt erneut eine quantitative Wahrscheinlichkeitsbetrachtung aller für die Entscheidung relevanter Größen. Die Fragen, die sich an kontaminierten Standorten stellen, zum Beispiel ob eine Sanierungsverfahren direkt am Schadensherd, an der Konzentrationsfahne, oder nur an besonders wichtigen Orten wie Trinkwasserentnahmestellen ansetzen soll, sind selten eindeutig auflösbar [Falta, 2008].

Demzufolge bedarf es zunächst aussagekräftiger Zustandsgrößen, die die anfallenden Entscheidungen möglichst gut unterstützen. Kenntnisse über diese Zustandsgrößen können

mit Hilfe von computergestützter, stochastischer Simulationsmethoden und im Feld erhobene Daten gewonnen werden. Hierfür bedarf es aber Modelle, welche die wichtigsten physikalischen Zustände und Zusammenhänge sowie die wichtigsten damit einhergehenden Unsicherheiten erfassen und abbilden können. Wahrscheinlichkeitsgrößen sind besonders interessant

1. für den Massenfluss, der die Möglichkeit einer Auswirkung auf Mensch und Umwelt widerspiegelt,
2. für die komplette oder nahezu komplette Auflösungszeit des Schadensherdes, weil es das Andauern der möglichen Auswirkung bemisst, und
3. für geometrische Eigenschaften des Schadensherdes, da mit dem gewonnen Wissen über die Schadensherdgeometrie das Design für ein Sanierungsverfahren erheblich verbessert werden kann.

**Schwerpunkt der Arbeit** ist ein gesamtheitlich konsistenter Modellaufbau. Wie bereits erwähnt, ist eine besonders wichtige Anforderung an die Simulationsmethode, dass physikalische und stochastische Zusammenhänge möglichst genau gewahrt werden. Ein zuverlässiges Modell bildet alle entscheidenden physikalischen Prozesse ausreichend genau ab und berücksichtigt alle relevanten Unsicherheiten. Ersteres beinhaltet die gegenseitige Beeinflussung der oben genannten Charakteristiken des Schadensherdes und der Grundwasserdurchlässigkeiten des Boden bzw. Gesteins. Unglücklicherweise ist die Unsicherheit dieser beiden Größen besonders stark. Zum einen sind beide nur bedingt messbar, da sie räumlich verteilt unterschiedlich Werte annehmen [Gomez-Hernandez *et al.*, 1997], zum anderen sind exakte numerische Simulationen von derartig komplexen Schadensherden besonders kritisch zu sehen [Feenstra and Cherry, 1996]. Dies ist der Fall, weil (1) das hierfür erforderliche detaillierte Wissen über die Bodenstruktur bis hin zur Porenform und Öffnungsweite nicht abrufbar ist und (2) selbst leistungsstarke Computerarchitekturen eine derartig genaue Berechnung nicht ermöglichen. Außerdem sind Messungen im Untergrund grundsätzlich mit einem nicht vernachlässigbaren Fehler behaftet. Der physikalische Zusammenhang und die unvermeidlichen Unsicherheiten sind also untrennbar miteinander verknüpft. Dies begründet die Forderung nach einem gesamtheitlich konsistenten Modell.

Die beschriebenen Unsicherheiten der Eingabeparameter, der Simulation physikalischer Prozesse und der Messungen drücken sich in großen Spektren an möglichen Modellergebnissen aus. Eine effiziente und konsistente Methode, die von unsicheren Messungen auf unsichere Modellergebnisse schließt und dabei die gesamte Unsicherheit reduziert ist daher erforderlich. Problematisch und daher ein zentraler Aspekt dieser Arbeit sind dabei vereinzelte Messungen des angestrebten Modellergebnisses oder der Zwischenergebnisse, wenn diese sich nicht in einen linearen Zusammenhang mit Eingabeparameter stellen lassen. Diese inverse Fragestellung besitzt selten eine eindeutige Lösung. In dieser Arbeit werden gemessene Konzentrationswerte verwendet, um auf mögliche Schadensherdgeometrien rückschließen zu können.

**Ziele:** Ziel dieser Arbeit ist es, wie zuvor gefordert, ein konsistentes Modellgerüst aufzustellen. Dieses soll alle relevanten Prozesse und Parameterabhängigkeiten abbilden und alle relevanten Unsicherheiten berücksichtigen können. Dem zur Folge müssen einzelne Simulationsvorgänge effizient gestaltet werden, um somit eine große Monte Carlo Simulation und das bereits erwähnte inverse Rechnen zu ermöglichen. Zu diesem Zwecke werden Zufallsfunktionen für Schadensherde und Durchlässigkeitswerte aufgestellt, in einem Transportmodell gekoppelt und in ein gesamtheitlich konsistentes Modell eingebaut. Angebunden an dieses Modell wird eine sehr effiziente inverse Methode entwickelt, die adjungierte Zustände nutzt um von gemessenen Konzentrationswerten auf möglich Schadensherde rückzuschließen. Des weiteren soll das Prozessverständnis für Transportvorgänge zwischen dem Schadensherd und den Konzentrationsmessungen ausgebaut werden. Eine wichtige Rolle spielt dabei die auf Grund der Heterogenität des Aquifers verstärkte Durchmischung des im Grundwasser aufgelösten Schadstoffes. Dieser Mechanismus entzieht Konzentrationsmessungen den Informationsgehalt bezüglich feinskaliger Strukturen des Schadensherdes. Es ist daher ratsam, die Konzentrationsmessungen nicht zu weit Unterstrom einzuholen. Auf der anderen Seite verlieren mit zunehmender Nähe zum Schadensherd Konzentrationsmessungen ihren Informationsgehalt bezüglich des großskaligen Musters des Schadensherdes und auf Gesamtmassenflüsse die an dem kontaminierten Standort in das Grundwasser übergehen.

Die Arbeitsschritte dieser Arbeit gliedern sich in drei Hauptteile auf:

**Schritt 1. Konsistenter Modellaufbau:** Ein konsistentes Modell wird entworfen, welches so viele Vereinfachungen vermeidet, wie möglich und sinnvoll ist. Es soll auf Homogenisierung und Hochskalierungsansätze räumlich verteilter Parameterwerte verzichten und somit eine zu ungenaue Beschreibungen des Schadensherdes vermieden werden. Weiterhin werden chemische Gleichgewichte nicht innerhalb unangemessen großskalierten Bereichen angenommen. Der physikalische Raum wird nicht auf ein oder zwei Dimensionen beschränkt. Außerdem muss das Modell Unsicherheiten berücksichtigen können. Im Zusammenhang mit dieser Arbeit kommt es speziell auf die Parameter-, Modellstruktur- und Messunsicherheiten an. Ein hohes Maß an Unsicherheit wird von Aquiferparametern und der Form des Schadensherdes erwartet. Diese Unsicherheiten übertrageb sich auf alle in Beziehung stehenden Modelgrößen.

Ein besonderer Beitrag dieser Arbeit ist der Entwurf einer Methode zur stochastischen Generierung von Schadensherden. Die Schadensherde werden basierend auf physikalischen Zusammenhängen als räumlich verteilte Zufallsfunktionen beschrieben. Weiterhin wird eine neue Methode entworfen, die es erlaubt Massenauflösungsraten von komplizierten Schadensherdgeometrien in einem Lagrangschen Ansatz zu berechnen. Basierend auf dieser Rechnung wird ein Konzept entworfen, das es erlaubt die Auflösungszeiten dieser Schadensherde in adäquater Genauigkeit und Rechenzeit zu berechnen. Diese Simulationmethoden werden dann in einer Monte Carlo Simulation vereint. Somit können Wahrscheinlichkeiten für die wichtigsten Eigenschaften des kontaminierten Standortes berechnet werden.

**Schritt 2. Inverses Problem der Identifizierung von Schadensherden:** Mit einer effizienten inversen Methode, beruhend auf adjungierten Zuständen, werden Konzentrationsbeobachtungen genutzt, um auf mögliche Schadensherde zurückzuschließen. Die Identifizierung von Schadensherden ist nur dann rational nachvollziehbar, wenn sie, entsprechend den zuvor beschriebenen Voraussetzungen, auf einem konsistenten Modell beruht. Eine entscheidende Rolle spielt hierbei das Zusammenspiel von Aquiferparametern und dem Schadensherd. Aufgrund der hohen Unsicherheit in dem zu beschreibenden System ist eine konsistente statistische Beschreibung unabdingbar. Die Verwendung des Bayesschen Wahrscheinlichkeitsbegriffes und -satzes scheint hierfür geeignet. Dem zufolge werden unter der Betrachtung von bedingten Wahrscheinlichkeiten Aussagen über die Plausibilität des Eintretens eines Ereignisses getroffen. Eine besondere Stärke des Baysschen Ansatzes ist es, dass bereits existierendes Wissen über das zu beschreibende System in die Beurteilung mit einfließen und konsistent über bedingte Wahrscheinlichkeiten mit dem Informationsgehalt der Daten verflochten werden kann.

**Schritt 3. Verstärkte Durchmischung des gelösten Schadstoffes:** Der Prozess der Durchmischung von im Grundwasser gelösten Schadstoffen in Heterogenen Aquiferen und ihr Zusammenhang zu kinematischen Größen des Grundwasserflusses wird in diesem Schritt studiert. Diese Durchmischung mindert den Informationsgehalt von Konzentrationsmessungen auf die Geometrie des Schadensherdes mit zunehmender Fließzeit der gelösten Teilchen. Es könnte deshalb ratsam sein die Konzentrationsmessungen in der Nähe des Schadensherdes einzuholen. Dies könnte wiederum aber Rückschlüsse auf den Gesamt-massenfluss, stammend vom Schadensherd, verschlechtern. Deswegen wird diese Arbeit abschließend Geschwindigkeitsfelder aus heterogenen Medien und deren einfluss auf die Durchmischung studieren. Dabei werden bestimmte Eigenschaften des Geschwindigkeitsfeldes mit der Durchmischungsrate in Beziehung gesetzt.

**Zusammenfassend:** Eine Schlüsselerkenntnis dieser Arbeit ist, dass Schadensherde und Durchlässigkeitswerte untrennbar vernetzt sind, was besonders auch bei inversen Methoden berücksichtigt werden muss. Sobald diese Vernetzung, oder auch nur einer dieser beiden Größen nicht den physikalischen Gesetzen folgt, ist auch die Statistik des Modellergebnisses verfälscht. Weiterhin lässt sich ein Viezahl an Charakteristiken des Schadensherdes aus Konzentrationsmessungen ableiten, falls dies im Einklang mit der physikalischen Bedeutung aller Parameter und Modellzwischenergebnissen erfolgt. Die Durchmischung von Schadstofffahnen wird besonders durch lokale Geschwindigkeitsunterschiede des Grundwasserflusses angetrieben. Diese Geschwindigkeitsunterschiede erzeugen Zug- und Schubverzerrungen aber auch vereinzelt lokale Drehungen in der Konzentrationsfahne. Woraufhin diese neue und höhere Konzentrationsgradienten aufweist, welche wiederum durch Diffusions- und Dispersionsprozesse abgebaut werden.

# 1 Introduction

## 1.1 Motivation

**Groundwater contamination with non-aqueous-phase liquids poses severe risks to ecosystems.** Improper storage and disposal of non-aqueous-phase liquids (NAPLs) has resulted in widespread subsurface contamination, threatening the quality of groundwater as freshwater resource. There are various contamination scenarios with non-aqueous phase liquids (NAPL), such as leaking underground storage tanks, improperly sealed hazardous waste sites, landfills, and leaks or spills of industrial chemicals at manufactures. This work focuses on contaminations with NAPLs that are more dense than water (i.e., DNAPLs, such as TCE and PCE) [Pankow and Cherry, 1996]. Due to their immiscibility with and low solubility in the aqueous phase, entrapped DNAPLs remain as a separate phase. They dissolve into the groundwater and spread within the aquifer over long periods of time, before the entrapped DNAPL is fully depleted [Mercer and Cohen, 1990]. Due to their typically high toxicity [US Environmental Protection Agency (EPA), 2007], even low DNAPL concentrations in groundwater may pose high risks on ecosystems and human health.

**Describing the behavior of DNAPLs in the subsurface is far from trivial.** When DNAPLs migrate downward through the porous media of aquifers, they remain in the center of large pore spaces or accumulate on top of capillary barriers with high entry pressure [Page et al., 2007]. The resulting spatial distribution of pure-phase DNAPL in the subsurface is often called DNAPL source zone architecture or contaminant source architecture (CSA for short) [e.g. Lemke et al., 2004; Soga et al., 2004]. The actual CSA geometry in heterogeneous aquifers is highly irregular and not predictable due to the strongly non-linear dependencies of DNAPL migration on material heterogeneity at the pore scale and at the Darcy scale [e.g., Feenstra and Cherry, 1996]. The CSA differs from spill location to spill location (1) due to different spill or leakage scenarios (e.g., point-like or distributed spill patterns, instantaneous or continuous release history), (2) due to different hydrogeological boundary conditions, and (3) due to material heterogeneity [Page et al., 2007]. Yet, the resulting complex and uncertain morphology of CSAs and its interactions with uncertain aquifer parameters and groundwater flow have to be accounted for and need to be resolved at the relevant scale, since the morphology of CSAs is a key control on DNAPL dissolution, solute transport and DNAPL depletion [Marinovich, 2006; Fure et al., 2006; Page et al., 2007].

**Rational risk assessment is in demand.** The abundance of contaminated sites and difficulties of remediation efforts demand rational decisions based on a sound risk assessment.

To this end, screening or investigation methods are applied [e.g., *Trolldborg et al.*, 2008]. These methods assess which sites pose large risks, which ones can be left to natural attenuation, which ones need expensive remediation, and what remediation approach would be most promising. For this type of assessments, it is important to determine relevant characteristics or impact metrics, such as geometric characteristics of the unknown CSA [*Christ et al.*, 2010], total DNAPL mass [*Broholm et al.*, 2005], potential mass removal by remediation [*Freeze and McWhorter*, 1997], emanating dissolved mass fluxes and total mass discharge [*Trolldborg et al.*, 2010] in past and future [*Soga et al.*, 2004], predicted source depletion times [*Kavanaugh et al.*, 2003; *Fure et al.*, 2006; *Basu et al.*, 2008], and the possible impact on drinking water wells [*Einarson and Mackay*, 2001], and thus on human health [*US Environmental Protection Agency (EPA)*, 2007]. The same characteristics are also important for designing monitoring or remediation schemes [*Soga et al.*, 2004; *Kavanaugh et al.*, 2003], with the additional complication that concentration levels and mass discharge often show a rebound effect when remediation ends [*Cohen et al.*, 1994; *de Barros et al.*, 2013].

**Risk assessment in this context requires reasonable and well-designed predictive simulation tools.** Due to sparse data and natural heterogeneity, this risk assessment needs to be supported by adequate predictive models with quantified uncertainty. Assessing the fate and transport of contaminants is commonly supported by simulation techniques that rely on physically-based models. These models require an accurate source zone description, i.e., the distribution of mass of all partitioning phases (DNAPL, water, and soil) in all possible states ((im)mobile, dissolved, and sorbed), mass-transfer algorithms, and the simulation of transport processes in the groundwater (advection, dispersion, and diffusion) [*Freeze and McWhorter*, 1997]. Due to limited knowledge and computer resources, a selective choice of the relevant processes for the relevant states and decisions on the relevant scale is both sensitive and indispensable. Thus, it is an important research question what is a meaningful level of model complexity and how to obtain a physically and statistically consistent model framework.

**Uncertainty and information are two opposing key players in risk assessments.** Almost every estimate of the desired impact metrics will be uncertain due to the typical uncertainty that is inherent in any process description in a heterogeneous subsurface environment, and due to the complex and non-linear interdependencies between aquifer parameters, CSA, groundwater velocities, mass transfer, and so forth. Thus, stochastic methods are indispensable because they can provide reasonable error bars and allow the involved stakeholders to take decisions in proportion to the posed risks of contaminated sites [e.g., *Trolldborg et al.*, 2010; *Bolster et al.*, 2009]. In order to restrict the huge uncertainty stemming from the aquifer parameters and the CSA, field data need to be assimilated by inverse models. To this end, concentration observations possess promising information on CSA geometries, transport processes, and aquifer parameters. Revealing these valuable information, however, requires an efficient inverse model that is again physically and stochastically consistent. In particular, the identification of CSAs has to cope with non-unique problems, non-linear interdependencies, and enhanced mixing and plume deformation in a heterogeneous en-



vironment. Mostly due to mixing and plume deformation, the information content of concentration observations may decrease rapidly with increasing distance between source zone and measurement location.

**This leads to the three main challenges addressed in this work:**

- (I) What is an adequate level of model complexity? And how to obtain a physically and statistically consistent model framework for reliable mass discharge and source depletion time predictions that are adapted to the presence of uncertainty? Fig. 1.1 illustrates a consistent model framework accounting for: (1.) uncertainty of the DNAPL processes: (2.) infiltration, (3.) dissolution, (4.) transport, and (5.) depletion.

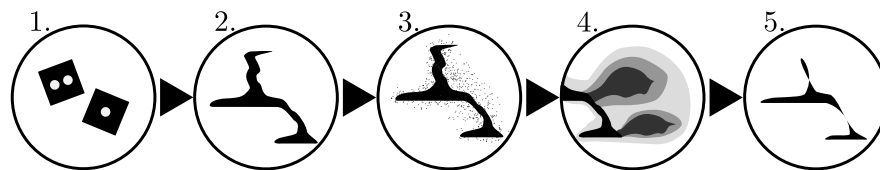


Figure 1.1: Illustration of a consistent simulation framework .

- (II) How to identify CSAs and how to reduce the uncertainty of CSA characteristics, mass discharge, and source depletion times with available field data? A key question for this purpose is: how to assimilate concentration observations in an efficient inverse model, as illustrated in Fig. 1.2?

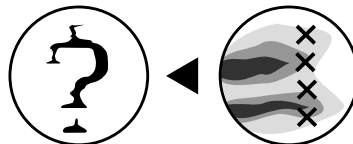


Figure 1.2: CSA identification from concentration observations

- (III) What are the driving mechanisms of enhanced mixing and plume deformation, and how to cope with the related loss of information content of concentration observations in a heterogeneous subsurface environment? This is illustrated in Fig. 1.3.

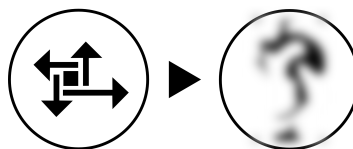


Figure 1.3: Loss of information due to solute mixing and plume deformation

The Figs. (1.1)-(1.3) will be used to guide through this thesis.

## 1.2 State of the art

A holistic analysis of the current practice, state of the art, and needs for scientific research will be given in the individual chapters of this thesis. However, to anticipate the key research questions of this thesis, the outstanding shortcomings of the state of the art are briefly mentioned here. Overall, four major shortcomings have been identified within the scope of this thesis, closely related to the observed challenges.

**1. The scale issue** Although there are many advanced numerical models for simulating complex physics at the pore scale [e.g., *Reeves and Celia*, 1996; *Blunt et al.*, 2002; *DiCarlo*, 2013], the prediction and uncertainty quantification of field-scale quantities is computationally not feasible with such complex models. The state of the art solution to cope with this issue are upscaled models [e.g., *Parker and Park*, 2004; *Jawitz et al.*, 2005; *Christ et al.*, 2006; *Saenton and Illangasekare*, 2007; *Basu et al.*, 2008]. However, field-scale quantities such as mass discharge and source depletion times strongly depend on fine-scale processes (e.g., at the Darcy-scale and pore-scale). Thus, there is a lack of methods that can account for relevant fine-scale processes, such as infiltration patterns and dissolution rates of DNAPL contaminants, and which can at the same time predict the desired field-scale quantities.

**2. The issue of uncertainty** The unpredictability of CSAs due to limited computational resources and limited observable data is well recognized [e.g., *Feenstra and Cherry*, 1996; *Parker and Park*, 2004; *Saenton and Illangasekare*, 2013]. Although random parameters of simplified contaminant sources such as location, time, and source strength have been used to represent the uncertainty of CSAs [e.g., *Wang and Zheng*, 2005; *Chaudhuri and Sekhar*, 2008], there is no model concept that uses physically-based principles to randomize CSAs in order to better quantify the uncertainty of mass discharge and source depletion time predictions. The failure to account for the physical formation processes of CSAs when modeling their uncertainty means to neglect a substantial amount of structural information. This can lead to bias and misspecified uncertainty intervals in predicting impact metrics. Thus, there is to date no model concept that can reasonably represent the uncertainty in predicting CSA geometries and impact metrics.

**3. Overly simplified model concepts to reduce computational costs** In general, stochastic models that treat and quantify uncertainties, and inverse models that assimilate field observations to estimate parameter values and reduce the overall uncertainty, are computationally demanding. Just for that very reason, the physics of these models are dramatically simplified in standard practice by linearization techniques and unrealistic assumptions of the contaminated source zone [e.g., *Snodgrass and Kitanidis*, 1997; *Michalak and Kitanidis*, 2004; *Wang and Zheng*, 2005; *Chaudhuri and Sekhar*, 2008], the model domain is typically reduced to one or two dimensions [e.g., *Lemke and Abriola*, 2006; *Saenton and Illangasekare*, 2013] and the numerical calculations are based on coarse discretizations [e.g., *Parker and Park*, 2004; *Saenton and Illangasekare*, 2007]. It has been recently shown by *Zhang et al.* [2008] and by *Kokkinaki et al.* [2013] that these simplified and upscaled models can not appropriately account for

non-linear relationships between spatially and temporally variable parameter fields. Thus, there is to date no model concept that avoids as much as reasonably possible these simplifications and assesses the minimum requirements on model complexity.

**4. Limited understanding of mixing and plume deformation mechanisms** Overall, the impact of flow variability on scalar spreading and mixing has been topic of intensive research [e.g., *Kapoor and Kitanidis, 1998; Weeks and Sposito, 1998; Fiori and Dagan, 2000; Dentz et al., 2011; Jha et al., 2011*]. Yet, it is still an open question which topological features of groundwater flow directly link to an increase of mixing and dissipation of dissolved-phase concentration. In this context, *Kapoor [1997]* studied the influence of rotating fluid parcels on the destruction of concentration fluctuations, while the influence of pure velocity shear on scalar mixing in porous media has been investigated by *Bolster et al. [2011]*. However, the mechanisms, shear and strain deformation, vorticity, and their interplay are all specific characters of the flow topology affecting all together solute mixing. This mixing and deformation of plumes emanating from CSAs affects the information content of concentration observations, and thus hampers the identification of CSAs. Hence, better understanding of the role of these mechanisms for solute mixing is desired. This helps to design remediation strategies, CO<sub>2</sub> sequestration plants, and, in the context of this thesis, it helps to design measurement campaigns for the CSA identification.

To summarize, the literature does not provide a model framework that reasonably simulates and infers physically-based random CSAs, predicts probability density functions of mass discharge and source depletion times, and aspires and assesses an overall physical and statistical consistency. The main challenge when trying to overcome this gap is indeed to obtain a sound model framework that treats the physically and statistically relevant aspects with adequate complexity in adequate spatial and temporal scales, an efficient inverse model that accounts for the relevant non-linearities between model parameters, field observations, and model predictions, and to obtain a better understanding of mixing and plume deformation mechanisms.

### 1.3 Goals and approach

The overall goal of this thesis is to provide a sound basis for rational decisions that arise in the assessment of contaminated sites. The required knowledge on CSA morphology, mass discharge, and source depletion time probabilities shall be predicted with simulation techniques. In order to answer the questions and challenges posed in Sec. 1.1, a physically and statistically consistent model framework is required. Therefore, existing methods are adopted and, whenever necessary, new methods or method improvements are developed. This is in line with three theses I postulate in the following, and for which I will demonstrate their significance and validity throughout this work.

- (I) The model framework must at least account for the heterogeneity of aquifers, the irregularity of flow fields, realistic and thus complex-shaped CSAs, the three-dimensionality of natural systems, adequate physical interlinkages of the key param-

eters at the adequate spatial and temporal scales, and it must at least treat the uncertainty of aquifer parameters and of the CSA.

- (II) Joint identification of CSAs and aquifer parameters based on concentration observations can be achieved via non-linear and non-unique Bayesian inversion. An accurate and efficient inverse method for this task can be by obtained by applying the method of adjoint states and utilizing the linearity of the transport equation.
- (III) The enhanced mixing of dissolved DNAPL and the solute plume deformation in heterogeneous aquifers significantly influences the inference quality of CSAs from downstream concentration observations. Knowledge on the driving processes of enhanced mixing allows to chose adequate measurement designs.

As will be shown in the individual chapters, improvements or new methods are in particular necessary for (I) generating complex CSAs as random-space functions (see Sec. 3.2), (II) simulating the dissolution and transport of DNAPL in the aqueous phase (see Sec. 3.3), (III) simulating the depletion of complex DNAPL source zones (see Sec. 3.4), and (IV) identifying CSAs from concentration observations (see Chap. 5). Furthermore, the driving mechanisms of enhanced mixing and plume deformation in heterogeneous aquifers need to be identified (see Chap. 6).

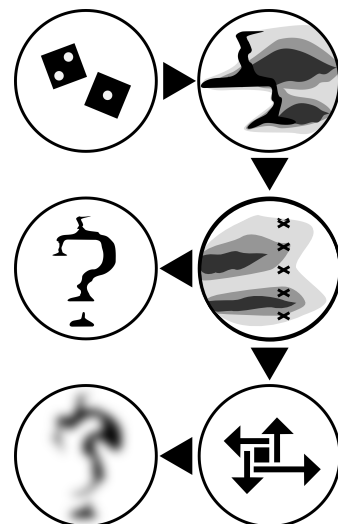
## 1.4 Structure of the thesis

According to the three challenges defined in Sec. 1.1 and the three theses postulated in Sec. 1.3, this work is structured in three steps:

*Step (I)* develops a sound model framework to simulate the fate and transport of DNAPL contaminants and assesses the physical and statistical consistency of this model framework in Chap. 3 and Chap. 4.

*Step (II)* develops a reverse-inverse method to identify CSAs from observed concentration values (Chap. 5), coping with the computational burden of inverse problems, but yet honoring the non-uniqueness and non-linear dependencies of the underlying problem.

*Step (III)* assesses the mechanism of enhanced mixing and plume deformation in heterogeneous subsurface systems (Chap. 6), since these processes are crucial obstacles for the CSA identification from downstream concentration observations.



In more detail, the remainder of this work is organized as follows. Chap. 2 reflects the key parameters and processes at contaminated sites, and repeats the governing physical equations and the geostatistical and Bayesian concept employed throughout the later analysis.

Chap. 3 assembles and introduces adequate physical and stochastic simulation tools to predict mass discharge and source depletion times and to quantify their uncertainty. This includes an adopted stochastic invasion percolation approach to generate random but physically consistent CSAs, the novel implementation of Dirichlet and third-type boundary conditions in PTRW simulations and a novel quasy-steady-state approach for simulating the depletion of CSAs (Step Ia). Chap. 4 introduces the new PTRW method in detail and applies a benchmark against an analytical, a semi-analytical, and a Eulerian numerical method (Step Ib). Chap. 5 derives an efficient reverse-inverse model approach for identifying CSAs using available field data such as concentration observations (Step II). Chap. 6 studies the flow topology and scalar mixing in spatially heterogeneous flow fields on a locally analytical two-dimensional scenario, whose findings are confirmed to be valid also in the global behavior via numerical tests. These findings are then transferred to and analyzed statistically in more realistic three-dimensional flow fields (Step III).

The structure of the present work is partly reflected in a series of papers either published or submitted by the author in cooperation with colleagues. Chap. 3 is based on *Koch and Nowak* [2014b] and applies the method published in *Koch and Nowak* [2014a], which is introduced and benchmarked in Chap. 4. Chap. 6 is based on the study by *de Barros et al.* [2013].



## 2 Fundamentals

Risk assessments of contaminated sites have to cope with many challenges that arise, among others, from difficulties in describing the fate and transport of DNAPL contaminants. Knowing this situation demands a consistent and reliable model framework that accounts for all relevant processes at the relevant scale, uses an adequate physical dimensionality, and quantifies and treats uncertainty in a reasonable manner [Trolborg *et al.*, 2010]. The following chapter discusses these demands and provides the fundamental equations, tools, and methodology which will be used by the studies in the remainder of this thesis. First, I provide a brief summary of the applied concepts to treat uncertainty used within this thesis. Then, the most relevant impact metrics are briefly reviewed, which is followed by a discussion of the relevant processes and their interdependencies at contaminated sites. Subsequently, the most crucial issues when simulating these processes with numerical methods are discussed. Before the governing equations of a multi-phase system at contaminated sites are introduced, I review the principles of geostatistical approaches. Finally, the system of equations for multi-phase flow is tailored for the specific needs of this thesis. These are predicting mass discharge and source depletion times, identifying CSAs and assessing the driving forces of solute mixing.

### 2.1 Basic concepts to treat uncertainty

**Uncertainty:** This thesis tackles three types of uncertainty, namely parameter uncertainties, structural uncertainties, and observation uncertainties.

Parameter uncertainty is typically propagated forward through models from uncertain input parameters to model predictions. A typical source for this type of uncertainty is that hydraulic conductivity varies in space but only point-wise information are available.

Structural uncertainty arises from the lack of knowledge of the underlying physics of the system and the related physical inaccuracy of applied models. For example, the formation of CSAs depends on pore-scale processes which are impossible to simulate within the entire model domain. Although there are models that tackle some of these pore-scale processes of DNAPL, groundwater, and soil interactions [Blunt *et al.*, 2002; DiCarlo, 2013], these models require knowledge of the entire soil and pore structure (e.g., pore geometries and (non-)wetting properties) within the model domain which is unaccessible [Feenstra and Cherry, 1996; Parker and Park, 2004]. Hence, models that can mitigate structural uncertainty via incorporating pore-scale processes fall back to huge parameter uncertainties.

Observation uncertainty may arise from measurement errors. For example, if one could measure a concentration value several times at the same location and at the same time, one would still obtain every time a slightly different value due to imperfect sampling and analysis methods. The measurement error can also be seen as the discrepancy between the observations and the output of the perfect model.

Facing these three types of uncertainties, adequate stochastic methods are indispensable in order to provide reasonable error bars and allow the involved stakeholders to take decisions in proportion to the posed risks of contaminated sites, e.g. *Trolborg* [2010]; *Tartakovsky* [2007].

**Decision theory:** The principal goal of decision theory [e.g., *Luce and Raiffa*, 1957; *Berger*, 1985] is to provide methods for stakeholders to maximize profits and to minimize negative impacts, such as toxic exposure in the context of this thesis. However, the theory about how decisions should be made in order to be rational is not a very unified one. The main issue is how to act if there is a lack of information and decisions have to be made under non-certainty. The literature distinguishes between four categories of decisions according to the knowledge of available alternatives and possible consequences. These decision categories are preferentially ordered, such that achievable rationality in decisions decreases from the first to the last category:

1. decision under certainty, where deterministic knowledge is given,
2. decision under risk, where complete probabilistic knowledge is given,
3. decision under uncertainty, where partial probabilistic knowledge is given, and
4. decision under ignorance where no probabilistic knowledge is given.

The leading paradigm in decision making science is to make decisions under risk by evaluating expected utilities (EU) [e.g., *Fishburn*, 1970], also referred to as probability-weighted utility theory. Any possible alternative has an uncertain utility value that varies under different states of nature. Thus, the utility values are averaged across the probabilities of these states of nature which yields an expected utility value for each alternative. The foundation of maximizing the expected utility is the law of large numbers in probability theory. There is no compelling reason to maximize the *expected* utility (or equivalently to minimize the expected damage) in case-by-case decisions on very rare events that can not average out random effects. However, the expected utility theory provides still a basis for reasonable and transparent decisions. Yet, objective expected utility measures are most powerful in cases when a large number of similar decisions are to be made according to the same decision rule. This is the case, for example, when one has to decide at various contaminated sites whether or not to go for an expensive remediation strategy.

While the first decision category (decision under certainty) is utopian in environmental sciences, the scope of stochastic modeling in environmental sciences is to get as close as possible towards the second category, i.e., allowing for decisions to be taken under knowing the risk of all available alternatives. In contrast to this scope, the practice of environmental consulting typically falls back to the state of making decisions under ignorance. The goal of



this thesis is to install the second category. Thus, it will provide a framework of methods that allow consultants to obtain full probability density functions of the most crucial impact metrics to assess contaminated sites.

**Bayesian probabilities:** Probabilities can be viewed as potential frequencies in the physical world, or alternatively as a degree of confidence of the decision maker. While the first point of view is purely objective, the latter makes use of subjective probabilities that may be different with different persons. These two different approaches of treating probabilities are called frequentist and Bayesian approach, respectively [e.g., Berger, 1985; Press, 2012]. The Bayesian approach enables reasoning with propositions whose truth or falsity is uncertain. The initial degree of confidence (probability in the Bayesian sense) of a proposition is called *a priori* probability, the probability of observing the data when the proposition is true is called *likelihood*, and the degree of confidence after considering additional data is called *a posteriori* probability. The core of this Bayesian framework is the Bayesian theorem that relates a priori probabilities of specified propositions and the likelihood of the data with the a posteriori probability of the propositions conditioned on the observed data. In other words, Bayes theorem is a rule how to update our belief when new data become available. This updating step can be repeated sequentially every time when new relevant data become available, obtaining every time a new posteriori probability of the predefined proposition. If hypothesized states are tested against observed data, the described procedure is called Bayesian inference [e.g. Press, 2012; Kitanidis, 1986]. The relation to standard hypothesis testing in the field of frequency statistics is apparent. However, the Bayesian approach is not concentrating on a single null hypothesis which is either rejected or true, but it is rather balancing the consequences of incorrect decisions across all possibilities. It is then up to the modeler (or consultant) to judge how to use this information.

This thesis shall use the Bayesian framework for both the forward modelling, where parameter uncertainty is modelled and sometimes also reduced by incorporating field-parameter observations (such as permeability values), and the inverse modelling, where parameter and structural uncertainty is reduced by identifying realistic CSAs. Applying the Bayesian framework, one obtains a joint ensemble of input parameters, intermediate model states, and model outcomes always linked to a joint probability for each set. If this ensemble is large enough, one may approximate a full joint probability density function (*pdf*) which is a powerful basis to support rational decisions. As discussed above, it is indispensable to model and infer all model states strictly jointly, since all model states are physically coupled within the overall model framework.

**Stochastic simulation and model complexity:** The tasks of contaminated site assessments are always subject to a high degree of uncertainty in a heterogeneous subsurface environment and thus they demand statistical treatment [Oreskes *et al.*, 1994; Dagan and Neuman, 1997; Rubin, 2003]. Reliable subsurface flow and transport models lean on spatially variable parameter fields. These include variable hydraulic conductivity, groundwater pressure, local velocities, DNAPL saturations, and concentration values of dissolved DNAPL. Information on these parameter fields are typically only accessible by sparse point-like or

integrative observations. This information is insufficient to infer a highly resolved spatial distribution of all relevant material parameters, to be certain about the hydrogeological boundary conditions, or to uniquely infer a single model structure. Yet, heterogeneity of the aquifer material and its type of spatial structure heavily influences the entire system: the highly irregular shapes of CSAs (Sec. 3.2), the local velocities of groundwater flow into which the DNAPL dissolves and downgradient dilution and mixing processes within the dissolved-phase plume (Chap. 6). Overall, heterogeneity controls to a large extent the total dissolution rate, depletion times, and the fate and transport of the emitted contaminant plume (Chap. 3). This leads to the need of quantifying the uncertainty inherent in the model predictions and also to the need of assessing the risk posed by a contaminated site in a probabilistic fashion [Christakos, 1992; Oreskes et al., 1994; Rubin, 2003; Tartakovsky, 2007]. To cope with uncertain model predictions, random or conditional simulation in a geostatistical Bayesian framework has been strongly recommended in many recent studies [Schwede and Cirpka, 2010; Trolldborg et al., 2010; *geostatistics for evaluation of mass discharge uncertainty at contaminated sites* Trolldborg et al., 2012].

The physics of the DNAPL/groundwater system imply non-linear dependencies between the geostatistical model or the actual spatial distribution of porous media parameters, the initial/boundary conditions of the transport equation given by the irregular geometry of the CSA, local groundwater velocities (also influenced by hydraulic boundary conditions), and DNAPL dissolution and degradation kinetics. Ignoring any of these dependencies unavoidably leads to non-realistic, biased, or even unphysical estimates of mass fluxes, depletion times, CSA characteristics, and so forth. Although the model response may match the observed data quite well after calibration, the model predictions outside the calibration range would be unreliable. The underlying problem is that many forms of errors in model structures can be compensated by adjusting parameters to physically or statistically meaningless values, leading to models that seemingly comply with all available data, yet have no predictive power outside the calibrated range. Thus, in order to ensure reliable predictions, the applied model framework has to be physically and stochastically consistent, and this fact demands a minimum degree of model complexity.

The key message is that using wrong physical assumptions (e.g., simplified CSAs) lead to wrong estimates and uncertainty bounds for all dependent quantities. Vice versa, using wrong stochastic assumptions, e.g., ignoring the uncertain nature of complex CSA formation, lead again to valueless estimates and wrong uncertainty bounds for all dependent quantities. This suggests that one has to randomly co-simulate the porous medium, the shape of the pure-phase (DNAPL) contamination (i.e., the CSA), the hydraulic boundary conditions, and the dissolution and transport processes in groundwater.

**Inverse problems:** Incorporating field data into models helps to reduce the uncertainty of predictions [e.g., Rubin and Dagan, 1987; Kitanidis, 1995; Gomez-Hernandez et al., 1997; Franssen et al., 2003], but inevitably leads to so-called inverse problems [Sun, 1994] which imposes yet another computational burden. While soil properties are directly linked to input parameters of the model, observed hydraulic heads and concentration values are (at least in the context of this work) intermediate model results or already part of the aspired model

prediction, respectively, and thus both pose an inverse problem. There are well-established and commonly applied techniques to calibrate a model on soil and hydraulic head observations. Typical assimilation techniques for soil properties, such as permeability and porosity, and hydraulic heads are (co-)kriging [Matheron, 1963], the ensemble Kalman filter [Hendricks Franssen and Kinzelbach, 2008; Nowak, 2009], and Bayesian geostatistical inference [Kitanidis, 1986]. The first two techniques (cokriging and the ensemble Kalman filter) apply a first-order transformation of the first two statistical moments between model outcome (observed data) and input parameters as a swift approximation of the inverse problem. This approach has been successfully applied to link hydraulic head observations with hydraulic conductivities [Kitanidis, 1995, 1997]. However, the non-linear behavior between hydraulic conductivity fields and concentration observations prohibits the application of such simplified methods to assimilate concentration data. Brute-force Monte Carlo methods, such as the Bootstrap filter which uses Bayes theorem to sequentially estimate the posteriori probability distribution, account for non-linear interdependencies. Yet, these methods are often computationally cumbersome for complex systems. Efficient inverse methods that honour the relevant non-linearities are lacking in the literature up to date. This thesis tackles this issue via utilizing actual linearities of the physical system, reverting the causality between boundary condition and concentration observations using adjoint states, and assimilating the observed data types sequentially in an efficient ordered that is proportional to the computational cost for each data type.

Chap. 1 mentioned the endeavor of this work to calibrate a physically consistent model on field-observed concentration values. Although concentration values are only point-like observations, they are highly informative to crucial model states such as the spatial pattern of CSAs and hydraulic permeability fields. This requires, however, appropriate assimilation techniques that honour the previously described and demanded physical and stochastic coherence. The model should reflect the non-uniqueness of the inverse problem and should jointly infer CSAs and permeability fields. Otherwise, if concentration values are either assimilated with linearized inverse models or assimilated to inadequate forward models, they are highly sensitive to support delusive decisions.

If there is a significant difference between the number of observed states (in this thesis typically  $10^1 - 10^2$ ) and the number of states that have to be jointly inferred (in this thesis typically  $10^3 - 10^6$ ), adjoint states [e.g., Sun, 1994; Neupauer and Wilson, 2001] can be used to increase the model efficiency, yet maintaining physical and stochastic consistency. To this end, the causal direction between boundary condition and downstream concentration observation is inverted by solving the advection-diffusion equation (ADE) in reversed direction (i.e., inverting the temporal axis and reverting the groundwater velocity along the streamlines).

## 2.2 Important metrics for contaminated site assessments

There is a wide variety of metrics that could be used to assess the risk at contaminated sites and the success of remediation strategies. Some examples are mass discharge (total emanating mass flux), the spatial mass flux distribution, source longevity, total remaining DNAPL

mass of an older CSA and mass removal achieved by natural dissolution or remediation, dissolved concentration values within and downgradient from the source zone, DNAPL mobility, DNAPL composition and its physical, microbial, and geochemical properties [Kavanaugh *et al.*, 2003]. These assessment metrics are typically prohibitively expensive to measure and thus require computational methods. The most popular risk metrics that support rational decisions are mass discharge  $\dot{m}_{tot}$  and mass dissolution rates (or as illustrated in Fig. 2.1 the sink term  $q_n^n$  of DNAPL component in the DNAPL phase) [Kavanaugh *et al.*, 2003; Soga *et al.*, 2004], source depletion time ( $S_n \approx 0$ ) [Kavanaugh *et al.*, 2003; Basu *et al.*, 2008], and geometric characteristics of the unknown CSA [Christ *et al.*, 2010]. The impact metrics that are important within this thesis are illustrated in Fig. 2.1 and will be discussed in more detail below.

**Dissolved-phase concentration:** The dissolved-phase concentration is the most intuitive metric to assess the impact of contaminants on human and environmental receptors. For example, the concentration of cancerogenic constituents in drinking water is used to evaluate the health risk (e.g., Freeze and McWhorter [1997]). Being a scalar quantity and directly measurable in multilevel wells, the dissolved-phase concentration gained much popularity as basis for decision making purposes. However, often observed rebounds in dissolved-phase concentration after a remediation effort [e.g., Cohen *et al.*, 1994] induce uncertainty on the success of remediation efforts when measuring them through concentrations.

Thus, a probabilistic framework for late-time steady state concentration estimates to design the remediation strategies is advisable [e.g., de Barros *et al.*, 2013]. Moreover, the scalar concentration observations in multilevel wells typically show high variations between zero and high concentration values in all three dimensions [Marinovich, 2006]. Thus, a direct interpretation of the overall impact from contaminated sites from a few concentration observations is not feasible. Still, concentration observations possess high explanatory power on contaminated site properties such as CSA characteristics, transport processes, and aquifer parameters, since downstream concentration values are a result of specific states of these contaminated site properties. Yet, in order to reveal these states from a few concentration observations, a physically and statistically consistent model framework (see Chap. 3), and an efficient and accurate inverse model are indispensable (see Chap. 5).

**Mass discharge and mass flux:** Mass discharge, commonly referred to as total mass flux emanating from the source zone, is an objective measure of the overall impact on ecosystems and potential impact on natural resources and human activities. The value of mass discharge estimates to assess contaminated sites has been increasingly recognized in the last decade. Einarson and Mackay [2001] suggested mass discharge estimates as a measure of potential impact on water supply wells. Soga *et al.* [2004] suggest the use of mass discharge estimates for assessing the efficiency of remediation techniques. Basu *et al.* [2006] proposed a flux-based assessment and management of contaminated sites with the objective to reduce contaminant discharge. They used mass discharge estimates obtained from a few transects of multilevel-measurement wells for the estimation of source strength, receptor loading, and degradation rates in the plume. Trolborg *et al.* [2008] proposed mass discharge estimates to

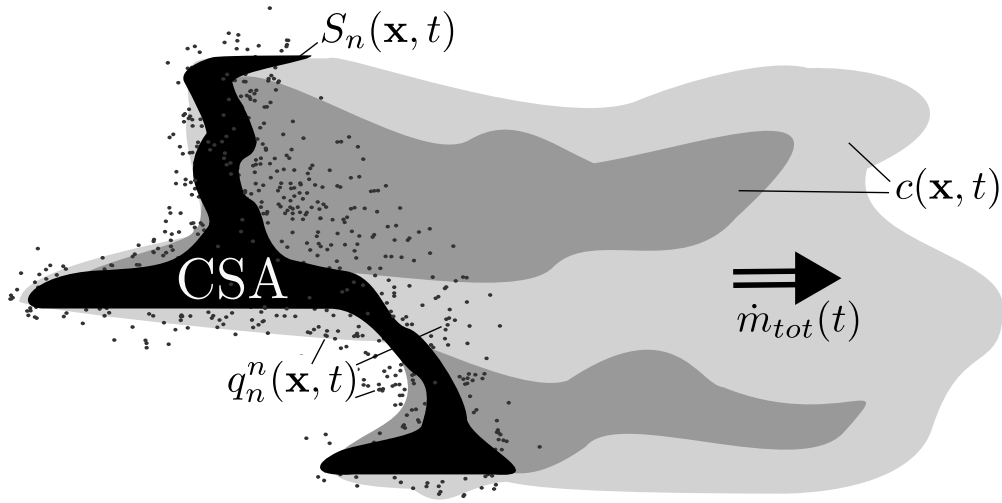


Figure 2.1: Illustration of the relevant metrics for contaminated site assessments: CSA geometry (black area), mass discharge  $\dot{m}_{tot}(t)$ , dissolution rates (sink term  $q_n^n(\mathbf{x}, t)$ ), DNAPL depletion (changes in  $S_n(\mathbf{x}, t)$  over time), and concentration values  $c(\mathbf{x}, t)$

prioritize contaminated sites at the catchment scale. *Newell et al.* [2011] suggest a magnitude classification system based on mass discharge measures in a similar style as the Richter Scale for earthquakes, the Beaufort Scale for wind, and the Saffir-Simpson Scale for hurricanes.

Although the term mass flux has been used to denote mass per time associated with a plume, the traditional chemical engineering nomenclature defines mass discharge as mass per time and mass flux as mass per time per area. Mass flux [ $\text{M T}^{-1} \text{L}^{-2}$ ] is a rate measurement specific to a defined area, which is usually a subset of a plume cross section. Mass discharge [ $\text{M T}^{-1}$ ] is an integrated mass flux estimate, i.e. the sum of all mass flux measures within the cross section if this cross section covers the entire plume. In this case, the mass discharge of the plume also represents the total emanating mass flux from a CSA.

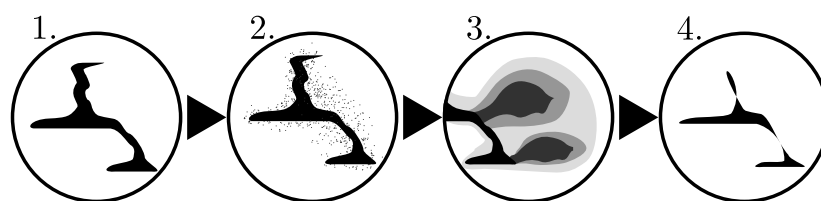
**Source longevity** Accompanying to mass discharge, source longevity is an objective metric for the persistence of the potential impact posed by a contaminated site. Exposure time of environmental receptors is directly correlated to health risk under chronic uptake conditions [e.g., *Byers et al.*, 1988; *Maxwell et al.*, 1998; *Environmental Protection Agency (EPA)*, 2005] and to economic impacts such as downtimes of drinking water wells [*Frind et al.*, 2006; *Enzenhöfer et al.*, 2012, 2014]. Reducing the longevity of the DNAPL source leads to an overall risk reduction [*Kavanaugh et al.*, 2003]. Knowing the natural depletion time and the potential of artificially enhanced source depletion indicates whether remediation is required and beneficial. So, large longevitys of contaminated sites limits the feasibility of monitored natural attenuation and plume management approaches as alternatives to expensive remediation efforts [*Fure et al.*, 2006].

Unfortunately, the source longevity and the depletion behavior are hard to obtain from field

data. A time series of mass discharge estimates (e.g., from a transect) provides insights of the depletion behavior. If the initial total amount of DNAPL is known, balancing the mass with these mass discharge estimates provides an approximation of characteristic depletion times. However, besides the uncertainty in total DNAPL mass, source depletion time predictions also inherit uncertainty from the error-prone process of estimating mass discharge from field data [e.g., *Li et al.*, 2007]. Hence, in order to obtain reliable impact metrics, adequate simulation tools for mass discharge and source depletion time predictions are required. Overall, the necessity for a physically and statistically coherent model framework seems to be ubiquitous for any targeted impact metric.

## 2.3 Relevant processes and their interdependencies

This thesis, features four different sets of processes:



Spilled DNAPL migrates through the aquifer with the trend to follow the easiest way through the aquifer matrix. On its way downward, it leaves behind trapped DNAPL phase or pools above layers it cannot penetrate (1. CSA formation). Once being in contact with the groundwater, DNAPL mass transfers into the aqueous phase (2. dissolution). The dissolved DNAPL is carried away by advective, dispersive, and diffusive transport in the groundwater (3. solute transport). Finally, typically over long time scales, the DNAPL mass is significantly reduced at the source zone (4. source depletion).

For more comprehensive descriptions, the reader is referred to *Abriola* [1989]; *Mercer and Cohen* [1990] and to *Pankow and Cherry* [1996]. As the governing processes at DNAPL contaminated sites are not observable in detail, there will always be uncertainty whether the considered set of governing processes is sufficiently close to reality and complete or not. It is not the scope of this work to discuss each individual physical, chemical, and biological process in detail. For example, I omit all aspects of chemical or microbiological transformation. Instead, the four main processes mentioned above and their interactions are discussed in the following.

**1. CSA formation:** When DNAPL migrates through saturated subsurface environments, it forms complex contaminant source architectures (CSAs). This phenomenon has been intensively studied throughout the last decades [e.g., *Illangasekare et al.*, 1995; *Ewing and Berkowitz*, 2001; *Gerhard and Kueper*, 2003; *Fagerlund et al.*, 2006]. The DNAPL phase pressure typically exceeds the required pressure to penetrate the aquifer, due to the high specific gravity  $g_n$

of DNAPLs compared to that of water  $g_w$  (e.g.,  $g_n/g_w = 1.1 - 1.7$  for most chlorinated solvents). The gravity ratios  $g_n/g_w > 1$ , typical viscosity ratios  $\mu_n/\mu_w < 1$  (where subscript  $n$  marks DNAPL and  $w$  is for water), and spatially variable entry pressures lead to unstable and hardly predictable drainage processes, causing relative rapid migration of DNAPLs along fingers and preferential flow paths [Chambers *et al.*, 2004].

Aquifer heterogeneity is the major control on the resulting shape of CSAs [Illangasekare *et al.*, 1995; Glass *et al.*, 2001; Page *et al.*, 2007]. The DNAPL migration leaves irregularly-shaped lines of residual saturation along its fingers in preferential pathways that are called ganglias. This is a microscopic DNAPL trapping mechanism, where DNAPL blobs are detached from the moving DNAPL phase by a thin water film and then remain captured within large pore volumes. Further, DNAPL accumulates on top of capillary barriers and spreads laterally until it reaches a pathway with sufficiently low entry pressure. It may also pool up until it reaches a sufficiently high DNAPL phase pressure in order to overcome the capillary barrier and advance further downwards. This is a macroscopic DNAPL trapping mechanism and leads to the formation of DNAPL pools with high DNAPL saturations. Both macroscopic and microscopic trapping are controlled by pore-scale quantities such as the pore and throat geometries of the porous medium, the fluids' interfacial tension, and wetting or non-wetting properties between the pore matrix and fluids. The two main CSA features, i.e., pools and ganglias, behave differently in all processes described below. Therefore, it is important to resolve these kind of CSA structures [Soga *et al.*, 2004; Parker and Park, 2004; Lemke and Abriola, 2006].

Studying the DNAPL migration at various groundwater velocities, Luciano *et al.* [2010] and Erning *et al.* [2009] observed that only extremely high groundwater flow velocities can significantly influence the distribution of DNAPL. Likewise, Dekker and Abriola [2000] observed that the hydraulic gradient has only a minor influence on DNAPL migration in heterogeneous aquifers. Thus, the downward migration of DNAPL is mostly governed by capillary, viscosity, and gravity effects. It continues until the macro- and micro-trapping mechanisms install a mechanical equilibrium against gravity and a stable CSA is formed.

**2. Dissolution:** No significant dissolution takes place during the infiltration process, due to the typically low solubility of DNAPLs in combination with the relatively fast process of downward migration [e.g., Grant and Gerhard, 2007]. Once the stable CSA is formed, dissolution and depletion acts on a much longer time scale. Dissolution is the actual trigger of groundwater contamination that could impact human health and the environment far away from the source zone. The overall dissolution rate is limited either by the solubility concentration of DNAPL in water and by the transport of DNAPL-saturated water away from the CSA, or by the pore-scale kinematics of mass transfer (i.e., by the diffusive transport and phase transition of the DNAPL species across the DNAPL/water interfacial area into the aqueous phase) [e.g., Mercer and Cohen, 1990; Feenstra and Guiger, 1996]. In both cases, the ganglia part of a CSA tends to have higher mass release rates than pools, due to the more effective delivery of still undersaturated groundwater through ganglias [e.g., Soga *et al.*, 2004; Parker and Park, 2004; Zhang *et al.*, 2007]. Thus, there is an interdependency between the CSA morphology, dissolution, and advective-dispersive solute transport, where all three are gov-

erned by aquifer heterogeneity.

Results of DNAPL dissolution experiments by, e.g., *Anderson [1988]*, *Schwille and Pankow [1988]*, and *Imhoff et al. [1994]* show that concentrations of dissolved-phase DNAPL close to the aqueous solubility concentration may be obtained in water that passes entrapped DNAPL with velocities from 0.1 to 1 m/d under laboratory conditions. This suggests that the transport in the dissolved phase and the solubility concentration are the limiting factors. In this case, the rate of mass transfer at the scale of representative elementary volumes (REV) mainly depends on (1) the solubility of the donor phase components in the acceptor phase, (2) the current concentration of donor-phase components in the acceptor phase installed by the surrounding mass transfer rates, and (3) the local advective and diffusive fluxes [*Mercer and Cohen, 1990*]. Yet, local hot spots of high groundwater flow and very small interfacial areas may lead locally to pore-scale kinematic limitations of dissolution [*Kokkinaki et al., 2013*]. If the dissolution is limited by mass exchange kinetics, additional factors come into play such as the interfacial area and the diffusion coefficient of DNAPL species across the interface. Hence, two options to model mass transfer rates at the REV scale are: (1) assuming local chemical equilibrium between all phases and components within any REV, or (2) assigning an effective rate coefficient for the mass transfer as a function of interfacial area and pore-scale mass exchange kinetics. These two options lead to Dirichlet or third-type boundary conditions for the dissolved-phase transport, respectively. A method to simulate dissolution into ambient flow for both options is presented in Chap. 4.

To model the non-equilibrium case, the most common approach is a first-order mass transfer model as it arises from the diffusive stagnant-film model [*Levich, 1962; Sherwood et al., 1975*]. This approach conceptualizes a tiny stagnant film near the interface within the aqueous phase, assuming that the two phase velocities are in equilibrium within the film (Fig. 2.2). Hence, mass transport within the stagnant film is essentially governed by molecular diffusion as described by Fick's First Law:

$$\frac{\dot{m}}{A_i} = -\mathbf{n} \cdot (D_m \nabla c), \quad (2.1)$$

where  $\mathbf{n}$  is a unit vector that is normal to the interface pointing into the aqueous phase, and  $A_i$  [ $L^2$ ] is the interfacial area. The stagnant-film model assumes a linear concentration gradient over the boundary thickness  $L_f$  and chemical equilibrium at the interface. This is a simplification of the diffusive double-layer model [*Levich, 1962*] which models such a boundary layer in both fluids. The two boundary layers have two different concentration values at the sharp interface, one arriving from the aqueous phase and one arriving from the DNAPL phase. In the stagnant-film model, the concentration gradient can be approximated by  $(c_w - c_s)/L_f$ , with  $c_w$  and  $c_s$  being the concentration values at the stagnant film boundaries with the aqueous phase and with the interface, respectively. Then, the diffusive flux across the interface  $\dot{m}$  with the units [ $M L^{-2} T^{-1}$ ] is

$$\frac{\dot{m}}{A_i} = \kappa(c_w - c_s), \quad (2.2)$$

where  $\kappa = D_m/L_f$  is the mass transfer coefficient with units [ $L T^{-1}$ ]. Effects from local aqueous velocity and the fluids viscosity are lumped into the assumption of the film thickness.



Modeling the mass transfer in porous media, it is useful to define a mass transfer rate per volume  $\dot{m}/V$  [ $\text{M L}^{-3} \text{T}^{-1}$ ]

$$\frac{\dot{m}}{V} = \kappa_{\text{eff}}(c_w - \hat{c}_s), \quad (2.3)$$

using a lumped mass transfer coefficient  $\kappa_{\text{eff}} = \kappa A_i/V$  with dimensions of  $[\text{T}^{-1}]$ , where  $A_i$  [ $\text{L}^2$ ] is the interfacial area within the volume of porous medium  $V$  [ $\text{L}^3$ ] [Feenstra and Guiger, 1996; Miller *et al.*, 1990]. If  $\kappa_{\text{eff}}$  is significantly higher than a characteristic advective time scale (e.g.,  $t_a = L_a/v_w$ ), for instance due to very small  $L_f$ , the result of the stagnant film model is again equal to assuming local chemical equilibrium in the entire REV  $V$ .

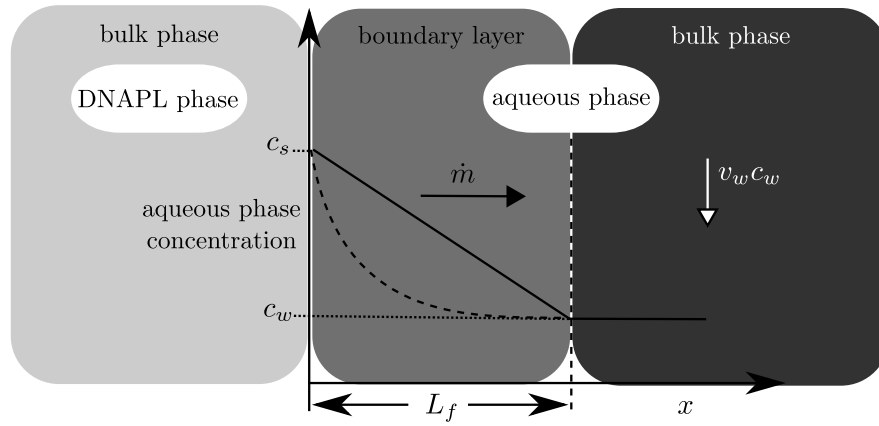


Figure 2.2: Stagnant film model. This sketch is adapted from Miller *et al.* [1990]

Figure Fig. 2.2 illustrates the diffusive mass transfer of a stagnant film model, due to the linear concentration profile within the boundary layer (solid line). However, the actual concentration profile may be in fact non-linear as indicated by the dashed line.

**3. Solute transport:** Upon dissolution, the dissolved DNAPL compounds are transported within the aqueous phase by advection, hydromechanic dispersion, and molecular diffusion. Other processes such as retardation, sorption and chemical reactions may influence the solute transport as well, depending on the site conditions. Dispersion arises from fine-scale irregular advection on the pore scale and on the REV scale in heterogeneous aquifers. This process leads to spreading, dilution, and mixing of the dissolved-phase plume [e.g., Cirpka and Kitanidis, 2000]. The spatial velocity distribution that triggers dispersion can not be resolved in fine-scale detail by numerical methods that are designed for predictions at the field scale. In order to account for the erased small-scale velocity fluctuations, their large-scale effects are simulated as an additional process. Most commonly, this is parametrized through Fick's First Law, lumping molecular diffusion and the effects of dispersion into one scale-dependent term. The larger the grid scale of the model, the larger the required values for dispersion coefficients, and the less the model is able to represent the aqueous phase concentrations that pose the relevant rate limitations for dissolution at the REV scale. Thus,

the degree to which heterogeneity is resolved and how the model equations are discretized are key points for simulating DNAPL dissolution and depletion.

The mentioned solute transport mechanisms install interactions between the dissolution rates at different locations within the CSA. Entrapped DNAPL may be located within the dissolved plumes arising from other parts of the CSA located further upstream, which will reduce the mass release rates at the downstream locations. This effect is triggered mainly by advective solute transport and will be called the advective shadowing effect in the following. While advection causes shadowing effects only from upstream to downstream locations within the CSA, diffusive shadowing acts in all directions, but only across much smaller distances.

**4. Depletion:** The depletion of CSAs is governed by spatially variable DNAPL saturations, local dissolution kinetics, and global solute transport. As all three aspects are highly heterogeneous and thus subject to uncertainty, the change of CSA over time and the final depletion is hard to predict. Experimental rates for the dissolution of entrapped DNAPL suggest that source-zone architectures change very slowly, particularly in sub-zones with high saturation [Miller *et al.*, 1990; Imhoff *et al.*, 1994; Johnson and Pankow, 1992]. It has also been found that ganglias deplete faster than pools, such that ganglia-dominated CSAs deplete faster than pool-dominated ones [e.g., Anderson *et al.*, 1992b; Parker and Park, 2004]. At the field scale, the longevity of DNAPL contaminated sites may vary from several decades up to a few centuries [Kavanaugh *et al.*, 2011; Seyedabbasi *et al.*, 2012]. Due to the different mass release rates and depletion times of ganglias and pools, a minor reduction of CSA DNAPL mass (mainly from ganglias) can already significantly reduce the total mass discharge [Soga *et al.*, 2004]. On the other hand, much more mass removal might be necessary to reach regulatory limits of specific target concentrations [Sale and McWhorter, 2001], and pool removal will not significantly reduce the total mass discharge.

Yet, initial knowledge on macroscopic CSA structures such as ratios between pools and ganglias may not be sufficient to explain the overall depletion process and thus accurately approximate the source longevity for two reasons. First, local depletion of trapped DNAPL reduces the advective and diffusive shadowing effects for other parts of the CSA. Second, reduced DNAPL saturation increases local groundwater velocities due to higher relative permeability values for the aqueous phase. If entrapped DNAPL is locally depleted but still present in the surrounding, the depletion behavior is globally changed in an intricate manner depending on both aquifer and CSA heterogeneities. In more detail, the shadowing effect is partly switched off or reduced due to the depleted DNAPL, which may increase the dissolution at locations that have been shadowed before. Moreover, the depletion of DNAPL increases relative permeabilities, thus changing the velocity field and the related advective-dispersive solute transport.

Overall, these effects lead to a non-linear mutual dependency of aquifer heterogeneity, groundwater flow, local dissolution rates, local depletion, and CSA morphology which frustrates straightforward upscaling or simplifications of the overall depletion process [Soga *et al.*, 2004; Fure *et al.*, 2006]. Once again (as it was also found for the impact metrics), the need for an adequate model framework is ubiquitous.

## 2.4 Crucial issues

The discussion in the previous sections (Secs. 2.2-2.3) revealed several issues that have to be tackled by a sound model framework. When taking severe assumptions such as uniform groundwater flow and simplified CSAs, the discussed processes cannot be reproduced, leading to inaccurate conclusions and misleading statements. For instance, such simplifications caused the debate between *Rao and Jawitz* [2003] and *McWhorter and Sale* [2003], on the statement of *Sale and McWhorter* [2001] that 99% mass removal might be necessary to reach regulatory limits at downstream concentration observations, which was calculated on the basis of a uniform groundwater flow. For reasonable accuracy in predicting mass discharges and depletion times, a list of interdependent physical processes must be accounted for (1) within a suitable physical dimensionality, (2) at a suitable spatial scale, (3) on adequate time scales, and (4) honoring their uncertainty. The challenge is to design a model that is sufficiently complex to cope with these challenges, but is still sufficiently simple to be operable, and to be fast enough to allow for treatment of uncertainties, e.g., by Monte-Carlo simulation. These issues will be discussed next.

**1. Physical dimensionality:** It is a common strategy to approach DNAPL fate and transport by one- or two-dimensional models for the sake of conceptual simplicity and computational ease. Many experiments [e.g., *Glass et al.*, 2000; *Page et al.*, 2007; *Luciano et al.*, 2010] and simulations [e.g., *Fure et al.*, 2006; *Grant et al.*, 2007; *Saenton and Illangasekare*, 2013] have been conducted in that way and brought important insights and relevant contributions to detect the relevant physical interdependencies and their appropriate spatial and temporal scale. However, when the goal is to assess the site-specific fate and transport of DNAPL contaminants, two-dimensional simulations are not appropriate for the following **four reasons**:

**First**, the additional third dimension systematically affects the structural formation of CSAs [*Christ et al.*, 2005], because DNAPL pooling is less likely in three than in two dimensions. **Second**, the ease of groundwater flow bypassing the CSA increases with increasing dimensionality [*Saba and Illangasekare*, 2000; *Christ et al.*, 2009]. Two-dimensional flows can in fact not bypass the downward DNAPL migration path at all. In three-dimensional flows, the effects of line-like ganglia structures on groundwater flow are negligible, and even pools may be bypassed easily. The effect on the groundwater flow is crucial, since groundwater velocities are the key driver for DNAPL dissolution, solute transport and depletion times. The **third** crucial difference is the shadowing effect between upstream and downstream dissolution. Due to the lower dimensionality, two-dimensional analyses observe many more shadowing effects between DNAPL locations, while flow bypassing reduces shadowing effects quite strongly in three dimensions [*Christ et al.*, 2009]. **Fourth**, analytical solutions to the advection-dispersion equation suggest that the dilution rate of peak concentration along streamlines by transverse diffusion/dispersion is roughly proportional to  $t^{-d/2}$ , where  $t$  is time and  $d$  is the dimensionality. Thus, the ratio between advective and diffusive shadowing effects changes with dimensionality. Overall, analyses in two or three dimensions have led to contradicting conclusions in past studies.

**2. Spatial Scales:** The geometric outcome of CSA formation dissolution, and dispersion mechanisms during dissolved-phase transport are heavily influenced by pore-scale processes and by heterogeneities at many scales. Capillary effects often govern the DNAPL infiltration process and hence dictate the final geometry of CSAs. Capillary effects depend on pore-scale features, such as pore geometry, apertures and throats, and wettability of the matrix material. Thus, individual large pore apertures or varying wettability may change the infiltration pathways significantly. These effects are usually not recognized at model scales that are useful to seek answers on the field scale. In such models, the capillary pressure is a function of DNAPL and aqueous phase saturation that are upscaled REV quantities. Since local velocities of the aqueous phase are among the key drivers of DNAPL dissolution, discharge and depletion, the model must resolve the scale of groundwater velocity fluctuations. Thus, the numerical grid scale for groundwater flow and dissolved transport must be at least as fine as the integral scale of aquifer heterogeneity. Therefore, field-scale simulation efforts are commonly limited by difficulties in characterizing the CSAs and aquifer structures in sufficient detail to obtain meaningful predictions [Anderson *et al.*, 1992a; Feenstra and Cherry, 1996; Parker and Park, 2004].

Laboratory experiments commonly observe saturation concentrations in water passing through the trapped DNAPL at typical groundwater flow rates (0.1 - 1 m/d) [Anderson, 1988; Schwille and Pankow, 1988]. Field surveys typically find concentrations of less than 10% of their solubility limits, even when DNAPL is supposed to be present close to the observation location [Mackay *et al.*, 1985; Anderson *et al.*, 1992b]. Due to thin residual zones and dispersed pool zones embedded in heterogeneous aquifers with non-uniform groundwater flow, initially saturated concentration values resulting from local dissolution may become significantly mixed and diluted with pure water [Anderson *et al.*, 1992a; Cirpka and Kitanidis, 2000; Dentz *et al.*, 2011]. A consequence of the scale issue is that field-scale models cannot reproduce this difference, unless invoking artificial rate limitations for mass transfer. Only with decreasing model scale (typically REVs with a few cm<sup>3</sup>), models may take the local equilibrium assumption [Saenton *et al.*, 2002; Grant and Gerhard, 2007; Saenton and Illangasekare, 2007]. In a recent model validation study, Kokkinaki *et al.* [2013] showed that the local chemical equilibrium assumption on very small spatial scales yields a reasonable approximation of effluent concentration values compared to those observed in a laboratory experiment for DNAPL dissolution in a heterogeneous sand packing. Yet, they observed improved accuracy when using a first-order mass transfer model that accounts for local velocities and interfacial areas.

**3. Temporal scale** The processes described above act on various time scales. These time scales span from several seconds or a few minutes for kinetic mass transfer [Miller *et al.*, 1990; Imhoff *et al.*, 1994] up to several decades or a few centuries for complete CSA depletion [Seyedabbasi *et al.*, 2012]. Accounting for all time scales by a transient model that resolves all dependencies from the lowest time scale (dissolution kinetics) up to the largest time scale (complete CSA depletion) is simply not feasible. In between these two extreme time scales, the initial DNAPL downward migration typically takes a few days, dissolved-phase transport through the CSA takes several days up to a few weeks (for typical groundwater

velocities between 0.1 and 1 m/d) and local depletion of individual CSA subregions may take several months up to a few years.

This leads to four clearly distinct time scales on which these interdependent processes act (see Table 2.1), and can be exploited to obtain reasonable model simplifications. First, the time scale for DNAPL infiltration is sufficiently short so that the net effect of dissolution can be neglected during the infiltration process. Thus, one may assume that the infiltration is complete before looking at the depletion process. Second, mass transfer and dissolution is fast enough compared to dissolved-phase transport, so that local chemical equilibrium can be assumed, at least at sufficiently small spatial scales [Saenton and Illangasekare, 2007]. Third, the dissolution and dissolved-phase transport can be considered to be at a quasi-steady state compared to the relatively slow process of local or global CSA depletion. Only when the CSA structure changes by local depletion, the dissolution rates and transport are transient for a few days, to return again to a steady state for several months or a few years until the next local part of the CSA has been depleted. Furthermore, it was shown in a Master project supervised within the framework of this thesis [Rivera, 2013], that the required frequency for updating the relative permeabilities, the groundwater velocity field and solute transport after CSA changes can be chosen to be very low.

**4. Uncertainty:** The previous discussions revealed several sources of uncertainty. To summarize, the heterogeneity of subsurface environments and insufficiently characterized sites demand that decisions about groundwater exploitation and remediation have to be made under uncertainty [e.g., Batchelor *et al.*, 1998; Bolster *et al.*, 2009]. The pore-scale geometry is neither resolvable in numerical models at the field scale, nor is it accessible from field campaigns [Feenstra and Cherry, 1996; Parker and Park, 2004]. Additionally, the instability of DNAPL migration processes induces an apparently stochastic behavior. For example, running two field studies for PCE migration at the same aquifer (Borden Air Force) and under the same boundary conditions, Brewster *et al.* [1995] observed significant differences in the migration behavior. Accordingly, they concluded that even subtle changes in the distribution of hydraulic conductivity and porous media structure can lead to unique DNAPL behavior. As these subtle variations of aquifer parameters are difficult to characterize, site-specific prediction of DNAPL migration will always be subject to a high degree of uncertainty and desire stochastic treatment.

Overall, when modelling the fate and transport of DNAPLs in the subsurface, one faces parameter, model, and observation uncertainty [Kennedy and O'Hagan, 2001]. Models that

Table 2.1: Time scales of DNAPL and its interactions with the groundwater and the porous medium

Process	downgradient migration	dissolution	dissolved transport	CSA changes	final depletion
Time-scale	days	seconds	days	years	centuries

can mitigate structural model uncertainty via incorporating pore-scale processes fall back to huge parameter uncertainties because these models require knowledge of the entire soil structure, e.g., pore geometries and (non-)wetting properties. Facing these uncertainties, adequate statistical and stochastic methods are indispensable in order to provide reasonable error bars, and thus to support the involved stakeholders in taking decisions in proportion to the posed risks of contaminated sites [e.g., *Tartakovsky, 2007; Troldborg et al., 2008*].

## 2.5 Geostatistic description of heterogeneous systems

In this section, concepts to handle the uncertainty of the aquifer model are briefly discussed without raising the claim of completeness. Instead, the interested reader is referred to the standard works of *Kitanidis [1997]* and *Rubin [2003]*. The basic concept of modeling uncertain physical systems, such as contaminated sites, is assigning random variables to all states that possess significant uncertainty. Through these random variables, uncertainty enters the partial differential equation system, posing yet another challenge to the computational methods. Due to the complexity of physical systems, proper models typically demand a huge set of mutually dependent random variables calling for multi-variate statistics.

Descriptive statistics describe the main features of observed data, parameters, and model outcomes. Two special cases of descriptive statistics, namely univariate and bivariate statistics, shall be outlined here in more detail. If only a single variable is considered and analyzed, e.g., all (hydro)geological data collected in a specific area with no respect to space, we speak of univariate statistics. If two variables are considered, e.g., comparing two subsets of collected data towards their dependency, we speak of bivariate statistics. The most prominent metrics in data analysis for both univariate and bivariate statistics shall be briefly described in the following.

### 2.5.1 Descriptive statistics

**Univariate statistics:** Given a set of data  $x_i$ , with  $i = 1 \dots n$ , the most frequently employed univariate statistics are the (arithmetic) mean, variance (or standard deviation), and skewness. The sample mean or, in more popular words, the “average”  $m$  is defined as the normalized sum over  $n$  sample values  $x_i$ :

$$m = \frac{1}{n} \sum_{i=1}^n x_i. \quad (2.4)$$

The sample variance  $s^2$  or, in other words, the data variation around their mean  $m$ , is given by

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2. \quad (2.5)$$

The square root of the variance  $s^2$  leads to the standard deviation  $s = \sqrt{s^2}$ . The asymmetry of variation around the mean is described by the dimensionless skewness  $k_s$ :

$$k_s = \left( \frac{1}{n-1} \sum_{i=1}^n (x_i - m)^3 \right) / s^3. \quad (2.6)$$

Symmetrically distributed values are defined by zero skewness  $k_s = 0$ , whereas asymmetric distributions assume positive or negative values of skewness depending on whether they are right- or left skewed, respectively.

**Bivariate statistics** The most common statistics to measure the combined behavior of two data sets is the covariance. Given two sets of collected data  $x_i$  and  $x'_i$  with  $i = 1 \dots n$ , the similarity of their deviations from the mean can be described by the sample covariance  $q$

$$q = \frac{1}{n-1} \sum_{i=1}^n (x_i - m) (x'_i - m'). \quad (2.7)$$

From that the correlation coefficient  $\rho$  between the two sets  $x_i$  and  $x'_i$  is found by removing the individual standard deviations  $s$  and  $s'$  from  $q$ :

$$\rho = \frac{q}{ss'}. \quad (2.8)$$

### 2.5.2 Distribution functions

A much more complete measure to describe or visualize data sets is their entire distribution of values such as by plotting histograms [Weiss, 2006]. When applying statistical rules of inference, one can obtain information on the distribution of all possible values (called population). Here, distribution functions serve to describe the variability of (hydro)geological parameters and of resulting system states. For example, the possible values that parameters such as hydraulic conductivity may assume are modeled as probability distributions. In the most general way, a random variable  $U$  can be described by its cumulative density function (*cdf*):

$$F(u) = P(U \leq u), \quad (2.9)$$

representing the probability of a random variable  $U$  being lower or equal to a given value  $u$ , with  $F$  taking values in the interval  $[0, 1]$ .

If  $F(u)$  is differentiable for all  $u$ , then  $U$  is a continuous random variable, and its probability density function (*pdf*)  $p(u)$  can be defined as

$$p(u) = \frac{dF(u)}{du} \quad (2.10)$$

where  $p(u) du$  is the probability that  $U$  lies within the infinitesimally small interval between  $u$  and  $u + du$ .

For two jointly distributed variables  $u$  and  $u'$ , the joint (bivariate) *pdf* is defined by  $p(u, u')$ . If both variables are independent, their joint distribution simply factorizes to the product of their marginal distributions  $p(u, u') = p(u)p(u')$ . Integration over  $u$  or  $u'$  by  $\int p(u, u') d u' = p(u)$  and  $\int p(u, u') d u = p(u')$  yields again the marginal distributions  $p(u)$  and  $p(u')$ , respectively. From  $p(u, u')$ , the conditional *pdfs*  $p(u|u')$  and  $p(u'|u)$  can be defined. They represent the distribution of  $u$  given knowledge on  $u'$  and vice versa. Similar to the sample mean and variance described above, distribution functions can be characterized by their means and variances. The mean  $\mu$  of a random variable  $u$  is defined as expectation over the spectrum of possible values by an integral of the form:

$$E[u] = \mu = \int_{-\infty}^{+\infty} u p(u) d u. \quad (2.11)$$

The second central moment, or variance  $\sigma^2$ , is defined as:

$$Var[u] = \sigma^2 = E[(u - \mu)^2] = \int_{-\infty}^{+\infty} (u - \mu)^2 p(u) d u. \quad (2.12)$$

For two variables  $u$  and  $u'$ , the covariance can be calculated as:

$$Cov[u, u'] = E[(u - \mu)(u' - \mu')] \quad (2.13)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (u - \mu)(u' - \mu') p(u, u') d u d u', \quad (2.14)$$

with joint *pdf*  $p(u, u')$ . If  $u$  and  $u'$  are independent variables, their covariance  $Cov[u, u']$  is equal to zero.

In the following, exemplary theoretical distribution functions are described in more detail:

- Gaussian Distribution

A variable  $u$  influenced by a multitude of factors (natural and technical processes), such as measurement errors or marks on a test are often said to have a symmetric, and unimodal distribution. Besides meeting these prerequisites, the bell-shaped, normal or Gaussian distribution is conceptually and statistically attractive because it depends only on two parameters, its mean  $\mu$  and variance  $\sigma^2$ . The Gaussian distribution is defined as:

$$p(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(u - \mu)^2}{2\sigma^2}\right], \quad (2.15)$$

on the interval  $[-\infty, +\infty]$ . For  $\mu = 0$  and  $\sigma^2 = 1$  it is referred to as standard-Gaussian (or normal) distribution. The standard-normal *pdf* is illustrated in Fig. 2.3.

- Log-normal Distribution

If the logarithm of a variable follows a Gaussian distribution, this variable can be considered log-normally distributed, which is defined as

$$p(u) = \frac{1}{u\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(u) - \mu)^2}{2\sigma^2}\right]. \quad (2.16)$$



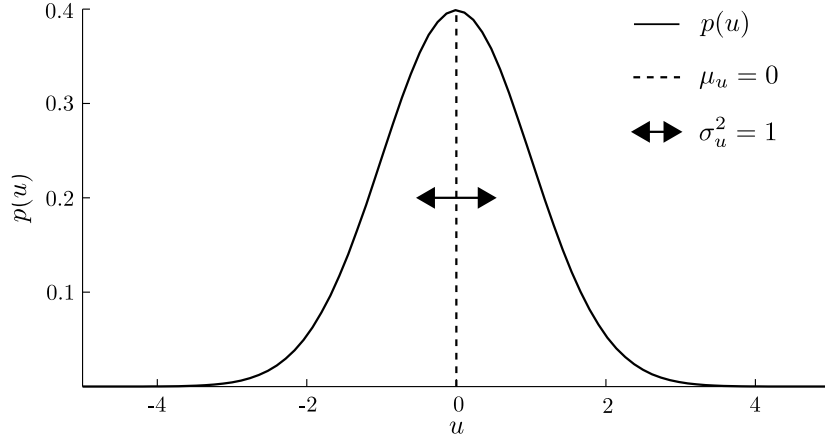


Figure 2.3: Standard-normal probability density function (*pdf*)  $p(u)$  with  $\mu_u = 0$  and  $\sigma_u^2 = 1$

The log-normal distribution is left-bounded on the interval  $[0, +\infty]$ . It is, hence, often applied to parameters or variables that are physically constrained to take positive values. Examples for physically restricted variables are concentrations, permeabilities or precipitation. In geosciences, the most prominent parameter is the hydraulic conductivity, and it is typically assumed to be log-normally distributed [e.g., Freeze, 1975].

- Multivariate Gaussian Distribution

In most natural processes, a multitude of distributed parameters is involved. Typically, the values of these parameters depend on each other and need to be described including their mutual correlation by so-called multivariate distributions. For two Gaussian distributed random variables  $u, u'$  acting jointly, their bivariate distribution is given by:

$$p(\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^2 |\boldsymbol{\Sigma}|}} \exp \left[ -\frac{1}{2} (\mathbf{u} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{u} - \boldsymbol{\mu}) \right], \quad (2.17)$$

with random variables  $\mathbf{u} = \begin{pmatrix} u \\ u' \end{pmatrix}$ , their mean  $\boldsymbol{\mu} = \begin{pmatrix} \mu \\ \mu' \end{pmatrix}$  and covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} Cov[u, u] & Cov[u, u'] \\ Cov[u', u] & Cov[u', u'] \end{pmatrix},$$

where  $Cov[u, u] = \sigma_u^2$  and  $Cov[u', u'] = \sigma_{u'}^2$ , and  $Cov[u', u] = Cov[u, u']$  (see Eq. 2.13). A bivariate distribution is illustrated in Fig. 2.4. In this example the marginal distributions of  $u$  and  $u'$  are Gaussian  $p(u)$  and log-normal  $p(u')$  distributions.

### 2.5.3 Geostatistical approach

In order to characterize and predict the spatial patterns of distributed variables in (hydro)geosystems, the spatial consideration of parameter variations and their mutual depen-

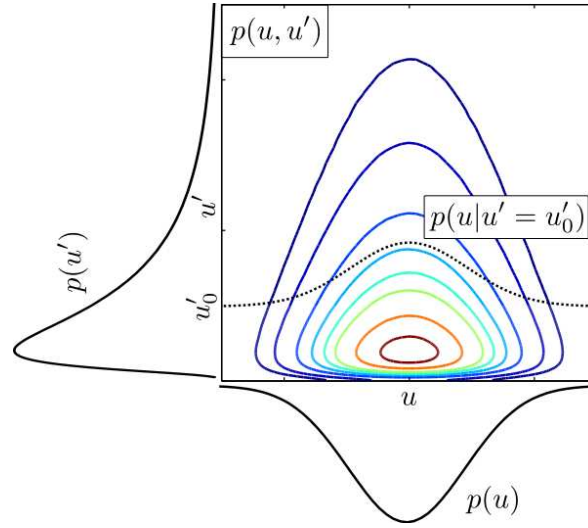


Figure 2.4: Illustration of marginal pdfs  $p(u)$  and  $p(u')$ , joint pdf  $p(u, u')$  and exemplary conditional pdf  $p(u|u' = u'_0)$

dence becomes indispensable. This is where geostatistics drift from univariate statistics. Following Tobler's first law of geography, which states that “*everything is related to everything else, but near things are more related than distant things*” [Tobler, 1970], the statistics of spatial patterns are conceptualized. To this end, experimental *variograms* or more theoretical *covariance functions* can be employed. In this section, I briefly repeat how geostatistics can be employed in order to model and simulate spatial features. The statistical theories, however, consolidating the geostatistical hypothesis are not repeated here. For more information, I refer to *Journal and Huijbregts* [1978], *Matern* [1986], and *Kitanidis* [1997].

Spatial patterns of parameters  $u(x)$  in locations  $x$  can be interpreted as random space functions (RSF). The most common descriptions of RSFs typically rely on the mean  $\mu(x) = E[u(x)]$ , the variance  $\sigma^2(x) = E[(u(x) - \mu(x))^2]$  and the covariance function  $R(x, x') = E[(u(x) - \mu(x))(u(x') - \mu(x'))]$ . With constant mean  $\mu(x) = \mu$  and variance  $\sigma^2(x) = \sigma^2$ , and  $R(x, x')$  depending only on the separation distance  $h = \|x - x'\|$ , a RSF is classified as second-order stationary. Eqs. (2.18)-(2.19) give two exemplary theoretical covariance functions commonly employed in hydro(geo)logical models, the Gaussian and the exponential model. The Gaussian covariance model is defined as

$$R(h) = \sigma^2 \exp\left(-\frac{h^2}{L^2}\right), \quad (2.18)$$

with the integral length parameter  $L$ . The exponential covariance model is defined as

$$R(h) = \sigma^2 \exp\left(-\frac{h}{L}\right). \quad (2.19)$$

$L$  is directly related to the correlation length  $\ell$  by  $\ell \approx 7L/4$  and  $\ell = 3L$  for the Gaussian and exponential covariance model, respectively. Their major difference lies in the fact that

the Gaussian model leads to very smooth patterns whereas the exponential model reveals sharper contours. This is due to the distinctly different gradient  $R(h)'$  for zero distance  $h = 0$ .

To allow for more flexibility with the assumption on the covariance model, the use of the Matérn covariance function [Matern, 1986] is suggested [e.g., Feyen et al., 2003; Nowak et al., 2010]. The Matérn covariance function is:

$$C(l) = \frac{\sigma}{2^{\kappa_m-1}\Gamma(\kappa_m)} (2\sqrt{\kappa_m}l)^{\kappa_m} B_{\kappa_m}(2\sqrt{\kappa_m}l),$$

$$\text{with } l = \sqrt{\left(\frac{\Delta x_1}{\ell_1}\right)^2 + \left(\frac{\Delta x_2}{\ell_2}\right)^2 + \left(\frac{\Delta x_3}{\ell_3}\right)^2}, \quad (2.20)$$

the Gamma function  $\Gamma(\cdot)$ , and the modified Bessel function of the third kind  $B_{\kappa_m}(\cdot)$ . Three examples of Matérn covariance functions are illustrated in Fig. 2.5, each with a different value for  $\kappa_m$ . The shape parameter  $\kappa_m$  controls the shape of the covariance function, e.g.,  $\kappa_m = 0.5$  leads to the exponential (dotted line) and  $\kappa_m = \infty$  to the Gaussian model (dashed line). Covariance functions with  $\kappa_m$  values between 0.5 and  $\infty$  fall in between these two specific cases, e.g.,  $\kappa_m = 2$  (solid line) is shown in Fig. 2.5. Treating spatially distributed parameters as RSFs also allows to include uncertainty in their statistical descriptions, leading to Bayesian geostatistics [Kitanidis, 1986]. As presented by Nowak et al. [2010], the Matérn family of covariance functions can be used in a Bayesian model averaging approach where model structural uncertainty is mapped onto parameter uncertainty.

Given only the mean  $\mu$ , variance  $\sigma^2$  and covariance function  $R(h)$ , spatially distributed parameter fields can be represented by multivariate Gaussian distributions (see Eq. 2.17) because these distribution parameter are sufficient to define the multivariate Gaussian distribution, and because as little as possible additional information is included in the so chosen distribution shape. This assumption has received some criticism because the range of possible pattern types to describe is limited [e.g., Zinn and Harveoy, 2003; Bárdossy and Li, 2008]. For example, it prevents spatially asymmetric arrangements of permeability values, such that low permeability values possess different spatial dependencies as high permeability values, which has been frequently observed in the field [e.g., Haslauer et al., 2012]. Alternative representations resolving more complex spatial features include, among others, copulas [Bárdossy and Li, 2008; Haslauer et al., 2012]. These are not in the scope of this thesis. However, all used methods will be flexible enough such that copulas could be used as well.

**Bayesian geostatistical description of permeability:** Most if not all aquifer parameters are typically subject to uncertainty, e.g., permeabilities  $K$ , porosities  $\phi$ , and pore radii  $r_p$ . In this thesis, the influence of uncertain aquifer parameters on the uncertainty of single phase flow (i.e., groundwater flow in a fully saturated aquifer) is represented by spatially random permeability fields. However, random variables could be also assigned to other parameters. The handling of uncertainty in multiphase-systems that also partially stems

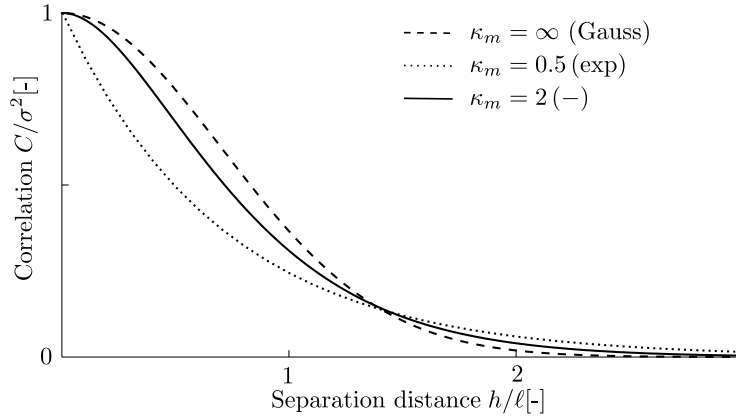


Figure 2.5: Illustration of three specific examples of covariance functions from the Matérn family for different values of the shape parameter  $\kappa_m$

from aquifer parameter uncertainty (e.g., from the uncertainty of pore geometries) and its relation to random permeability fields is discussed in Sec. 3.2.

The log-permeability  $Y = \ln K(\mathbf{x})$  is defined as a discretized multi-Gaussian random space function represented by element-wise constant values on a fine numerical mesh. Here, we assume local isotropy (scalar  $K$ ) for simplicity, but again without loss of generality. Following a stationary geostatistical concept with constant mean  $\langle Y \rangle$  and fluctuations  $Y' = Y - \langle Y \rangle$ , we describe the structure of random heterogeneity by a covariance function  $C(\mathbf{h})$  that only depends on the separation vector  $\mathbf{h}$ . In the Bayesian geostatistical approach [Kitanidis, 1986], the covariance function  $C(\mathbf{h})$  is again uncertain and thus becomes a random variable. Hence, covariance parameters such as the field variance  $\sigma_Y^2$  and the correlation lengths  $\ell_i$  per space direction  $x_i$  (scaling the dependency in the spatial directions  $i = 1, 2, 3$ ) are modeled as random. Applying the Matérn covariance function (Eq. 2.20), the parametric shape of the covariance function can also be modeled as uncertain through its shape parameter  $\kappa_m$  [Feyen et al., 2003; Nowak et al., 2010]. With this set of random covariance parameters, corresponding random permeability fields are generated using the method of Dietrich and Newsam [1993]. Although the simulations in this thesis involve solely multivariate Gaussian permeability fields, more complex geostatistical descriptions could also be used. Two examples of random log-permeability fields are shown in Fig. 2.6. Although they have the same anisotropic correlation lengths ( $\ell_Y = [7, 7, 2]$ ) and same variances ( $\sigma_Y^2 = 2$ ), due to different values of  $\kappa_m$  (left:  $\kappa_m = 2$ ; right:  $\kappa_m = 1$ ), they possess a completely different shape. For  $\kappa_m = 1$  the values of  $Y$  change rapid at already small distances which is similar to the exponential model, while for  $\kappa_m = 2$  the shape of the log-permeability field is already similar to the Gaussian model and thus very smooth.

**Kriging** In the previous paragraph, random permeability fields were generate for various Matérn covariance functions. The selection of a covariance function is equivalent to selecting an ensemble of solutions of the desired random space function. Obviously, if this ensemble

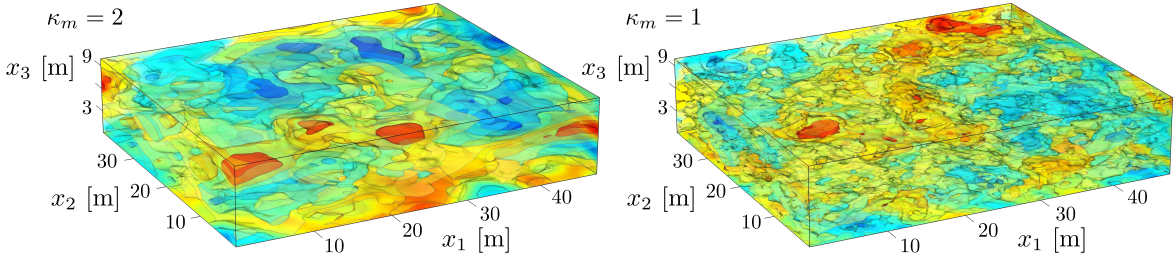


Figure 2.6: Illustration of two random log-permeability fields for different values of the Matérn shape parameter  $\kappa_m$  but same correlation length vector  $\ell_Y = [7, 7, 2]$ , and variance  $\sigma_Y^2 = 2$ . High permeability values are plotted in red and low permeability values in blue.

contains functions that do not honor the observed data, these functions should be eliminated from further considerations. This is achieved in this thesis via kriging for direct permeability measurements. Kriging is a technique to interpolate point-like observations within a spatial random parameter field that utilizes knowledge on the spatial structure of this field. It applies the following linear combination of the measurements to estimate the unknown quantity:

$$s_j = \sum_i \xi_{j,i} y_i, \quad (2.21)$$

where  $y_i$  and  $\xi_{j,i}$  with  $i = 1, \dots, n$  are a set of  $n$  measurement values and the corresponding  $j^{\text{th}}$  set of  $n$  coefficients and  $s_j$  is the  $j^{\text{th}}$  desired quantity. In this thesis only log-permeability measurements  $y$  are kriged within the log-permeability fields.

The basic approach of kriging is selecting the vector coefficients  $\xi$  such that Eq. (2.21) represents a best linear unbiased estimation (BLUE) of  $s$ . That means, on the average over all possible realizations the estimation error of  $s$  must be zero and the mean square estimation error is minimized. To this end, the following linear equation system (so called kriging system) is arranged and solved for  $\xi$ :

$$\mathbf{C}_{yy} \xi = \mathbf{C}_{ys}, \quad (2.22)$$

where  $\mathbf{C}_{yy}$  and  $\mathbf{C}_{ys}$  are matrices whose entries are evaluated by the covariance function  $C(\mathbf{h})$  at the separation distances  $\mathbf{h}$  of the measurement locations and the separation distances between the location of the desired quantity  $s$  and the measurement locations, respectively.

Eq. (2.22) represents the kriging system in the case of a known mean of  $s$ , which is called simple kriging (SK). This SK method is a member of a group of kriging estimators. A comprehensive introduction to the group of kriging estimators is provided by *Kitanidis* [1997] and *Rubin* [2003].

In this thesis, SK is used to generate random permeability fields conditioned directly on permeability observations. Therefore, an unconditional random log-permeability function is generated according to a selected covariance function  $C(\mathbf{h})$  as described above. Then, the

differences of this function to the measurement values at the measurement location is evaluated. These residuals are kriged by SK using  $C_{yy}$  and  $C_{ys}$  for each node of the grid within which the permeability field is discretized. Adding the kriged residual to the values of the unconditional realization shifts this function such that it passes through the measurements. Hence, a conditional realization is obtained which is still characterized by the desired pattern of the spatial correlation, due to  $C(\mathbf{h})$ . If observation uncertainty is considered, a random measurement error is added to the measurement values and the diagonal entries of  $C_{yy}$  are increased by the standard deviation of this measurement errors, assuming independent measurement errors.

In this thesis, the above described procedure of SK is used in a Bayesian framework, where the structural parameters of the covariance function are treated as random variables with probability distributions that reflect the current knowledge (i.e., beliefs in Bayesian terminology) about them [Kitanidis, 1986]. Bayes' theorem is then used to revise the probabilities of these random variables (structural parameters) when new information becomes available, e.g., in the form of permeability data. This is in contrast to the conventional non-Bayesian approach where some point estimates of these parameters are obtained and then treated as perfectly known.

## 2.6 Governing equations

This section briefly introduces the mathematical description of multiphase systems in porous media. The section closely follows the descriptions of *Class* [2001] and *Flemisch* [2013]. For more comprehensive descriptions, the reader is referred to *Scheidegger* [1974], *Abriola* [1989] and *White* [2003]. First, the key parameters involved in the description of contaminated sites are briefly reviewed. After that, I repeat the general mathematical description of two-phase two-component problems in porous media systems. Finally, the universal multiphase flow equation system is tailored to the specific needs of describing the key processes and predicting the most important impact metrics of contaminated sites.

### 2.6.1 Fluid and soil parameters

**Phase and component:** If a medium is filled by several fluids which are immiscible and separated by sharp interfaces, these fluids are referred to as phases. Each phase  $\alpha$  possesses characteristics that are valid throughout the entire phase, such as the phase density  $\rho_\alpha$  and whether the phase behaves affine (wetting-phase) or not (non-wetting-phase) with respect to different soil materials. The phases may be composited by several components. These components can transfer from one phase to another phase, for example when DNAPL components from within the CSA dissolve into the aqueous phase.

**Mass fraction:** Therefore, the composition of a phase  $\alpha$  is described by the mass fractions  $X_\alpha^C$  of the different components  $C$  contained in that phase:

$$X_{\alpha}^C = \frac{m^C}{\sum_C m^C}. \quad (2.23)$$

Obviously, these mass fractions sum up to one for each phases leading to  $\sum_C X_{\alpha}^C = 1$ .

**Porosity:** A porous medium consists of a solid matrix and the pores. The dimensionless ratio of the pore space within the REV to the total volume of the REV is defined as porosity  $\phi$  [-].

$$\phi = \frac{\text{volume of pore space within the REV}}{\text{total volume of the REV}}. \quad (2.24)$$

**Saturation:** The pore space is filled by the different phases. In the REV approach, this is expressed by the saturation  $S_{\alpha}$  [-] of each phase  $\alpha$ , which is defined as the ratio of the volume filled by phase  $\alpha$  within the REV to the volume of the pore space within the REV.

$$S_{\alpha} = \frac{\text{volume of phase } \alpha \text{ in the REV}}{\text{volume of the pore space in the REV}} \quad (2.25)$$

Since the REV is completely filled by the phases  $\alpha = 1, \dots, n_{\alpha}$ , the sum of the phase saturations must be equal one.

$$\sum_{\alpha} S_{\alpha} = 1. \quad (2.26)$$

**Mass and concentration:** According to these definitions, the mass  $m_{\alpha}$  [M] of phase  $\alpha$  occupying the volume  $V$  [L<sup>3</sup>] is

$$m_{\alpha} = \int_V \phi S_{\alpha} \rho_{\alpha} dV, \quad (2.27)$$

and the concentration  $c$  [M L<sup>-3</sup>] of the component  $C$  is

$$c^C = \sum_{\alpha} \phi \rho_{\alpha} X_{\alpha}^C S_{\alpha}. \quad (2.28)$$

**Capillary pressure:** If two phases share one interface within a medium, the phase that is more affine to the medium's material is called the wetting phase, while the other phase is called the non-wetting phase. Although this interface is sharp, different phase pressures  $p_{\alpha}$  are found for the two phases sharing that interface at local mechanical equilibrium with either  $\alpha = w$  or  $\alpha = n$ , referring to wetting and non-wetting phase, respectively. The difference between these two pressures is called capillary pressure  $p_c$  [M L<sup>-1</sup> T<sup>-2</sup>]:

$$p_c = p_n - p_w. \quad (2.29)$$

In the REV approach this capillary pressure is commonly related to saturation values. According to the material affinity of the wetting phase ( $w$ ), it is always willing to occupy the

smallest pores. In a two-phase system, if the saturation of phase  $w$  in the REV is reduced, the saturation of phase  $n$  has increased, and the interfacial area between the two phases has shifted towards smaller pores. This lets the affinity of the wetting phase become even more effective and leads to an increase in capillary pressure, and many multi-phase model concepts describe capillary pressure solely as a function of saturation. Hence, the capillary pressure apparently changes due to changes in saturation. However, this is an extremely complex non-linear relationship posing many open research questions up to date. One example is the observed hysteresis of multiphase flow: the capillary pressure-saturation relationship behaves different whether the wetting phase saturation is currently increasing (which is called inhibition) or whether the non-wetting phase saturation is currently increasing (which is called drainage). The latter is typically the case when a DNAPL infiltrates the subsurface and forms a CSA. The two most common capillary pressure-saturation relationships are Brooks & Corey [Brooks and Corey, 1964] and Van Genuchten [Van Genuchten, 1980].

$$\begin{aligned}
 \text{Brooks \& Corey :} \quad & p_c = p_d S_e^{-1/\lambda_b}, \\
 \text{Van Genuchten :} \quad & p_c = \frac{1}{\beta} (S_e^{-1/m} - 1)^{1/n}, \\
 \text{with} \quad & S_e = \frac{S_w - S_{w,r}}{1 - S_{w,r}}, \\
 & m = 1 - \frac{1}{n},
 \end{aligned} \tag{2.30}$$

and with the shape parameter  $\lambda_b$  [-], the initial entry pressure  $p_d$  [ $\text{M L}^{-1} \text{T}^{-2}$ ] and the effective saturation  $S_e$  [-].

**Velocity and permeability** The seepage-phase velocity  $\mathbf{v}_{s,\alpha}$  with which all containing components are transported is related to the Darcy velocity by

$$\mathbf{v}_{s,\alpha} = \frac{\mathbf{v}_{d,\alpha}}{\phi_e S_\alpha}. \tag{2.31}$$

The Darcy velocity  $\mathbf{v}_{d,\alpha}$  [ $\text{L T}^{-1}$ ] of a phase  $\alpha$  can be described by Darcy's law which was originally obtained experimentally for single phase flow:

$$\mathbf{v}_{d,\alpha} = -\mathbf{K}_f \nabla \left( \frac{p_\alpha}{\rho_\alpha \mathbf{g}} + z \right), \tag{2.32}$$

with  $\mathbf{K}_f$  [ $\text{L T}^{-1}$ ] being the phase-specific conductivity obtained from the ratio of the intrinsic permeability  $\mathbf{K}$  [ $\text{L}^2$ ] and the dynamic viscosity  $\mu_\alpha$  [ $\text{M L}^{-1} \text{T}^{-1}$ ] of phase  $\alpha$ :

$$\mathbf{K}_f = \mathbf{K} \frac{\rho_\alpha \mathbf{g}}{\mu_\alpha} \tag{2.33}$$

and  $p_\alpha/(\rho_\alpha \mathbf{g})$  [L] is the pressure head and  $z$  [L] is the geodetic elevation. This is extended for multiphase systems considering that the phases mutually hamper each other by a relative permeability  $k_{r,\alpha}$  [-]

$$\mathbf{v}_{d,\alpha} = -\frac{k_{r,\alpha}}{\mu_\alpha} \mathbf{K} (\nabla p_\alpha + \rho_\alpha \mathbf{g}), \tag{2.34}$$



with the phase pressure  $p_\alpha$  [ $M L^{-1} T^{-2}$ ] and the gravity vector  $\mathbf{g}$  [ $L T^{-2}$ ]. Following the capillary pressure saturation relationship it is assumed that the relative permeability scales with the saturation values. The most common relative permeability models are again Brooks & Corey [Brooks and Corey, 1964] and Van Genuchten [Van Genuchten, 1980]:

$$\begin{aligned} \text{Brooks \& Corey : } k_{r,w} &= S_e^{\frac{2+3\lambda_b}{\lambda_b}}, \\ k_{r,n} &= (1 - S_e)^2 \left( 1 - S_e^{\frac{2+\lambda_b}{\lambda_b}} \right), \\ \text{Van Genuchten : } k_{r,w} &= \sqrt{S_e} \left( 1 - (1 - S_e^{1/m})^m \right)^2, \\ k_{r,n} &= (1 - S_e)^{1/3} \left( 1 - S_e^{1/m} \right)^{2m} \end{aligned} \quad (2.35)$$

### 2.6.2 Multiphase flow equation system of contaminated source zones

The contaminant source zone is a two-phase two-component system (groundwater and DNAPL as phases and components). Hence, the saturation values of these two phases have to sum up to one

$$\sum S_n + S_w = 1, \quad (2.36)$$

with  $S_n$  and  $S_w$  being the DNAPL and the aqueous-phase saturation, respectively. At local mechanical equilibrium, the corresponding phase pressures  $p_n$  and  $p_w$  are balanced by the capillary pressure  $p_c$

$$p_c = p_n - p_w. \quad (2.37)$$

A very common approach on the REV scale is to define a capillary pressure saturation relationship, such as the Brooks and Corey relationship Eq. (2.30): To this end the entry pressure  $p_d$  is approximated according to the Young-Laplace equation:

$$p_c = \frac{2\gamma \cos \varphi}{r}, \quad (2.38)$$

with  $\gamma$  [ $M T^{-2}$ ] being the surface tension,  $\varphi$  [rad] the contact angle to the aquifer material, and  $r$  [L] being a representative pore radius. To estimate the representative pore apertures  $r$ , empirical correlation rules between the median grain size  $d_m$  [L] and the intrinsic permeability  $K_i$  are used in analogy to Poiseuille's law [Shepherd, 1989].

$$d_m = \left( \frac{K_i}{a} \right)^{\frac{1}{b}}. \quad (2.39)$$

Here, I use the regression parameters  $a = 1923$  and  $b = 1.94$  for units of gallons per day and square foot as given by Bedinger [1961], who analyzed the permeability and grain size of the Arkansas River alluvium. The pore apertures  $r_p$  and pore throat  $r_t$  radii are assumed to be directly proportional to the median grain size with  $r_p = 0.212d_m$  and  $r_t = 0.077d_m$ ,

respectively, and the representative pore radius in Eq. (2.38) is  $r = r_p r_t / (r_p - r_t)$  [e.g., Fetter, 1999].

The resulting transport equation for each component  $C$  in each phase  $\alpha$  can be written in a Eulerian perspective:

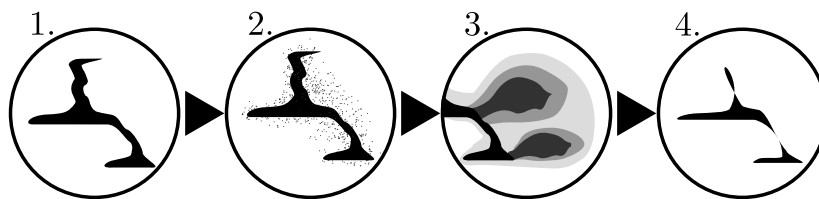
$$\int_V \frac{\partial \phi \rho_\alpha S_\alpha X_\alpha^C}{\partial t} + \nabla \cdot (\phi \rho_\alpha S_\alpha X_\alpha^C \mathbf{v}_{s,\alpha}) - \nabla \cdot (\mathbf{D}^C \rho_\alpha \nabla(X_\alpha^C)) - q^C dV = 0, \quad (2.40)$$

with the four terms: storage, advective flux, diffusive flux, and source or sink terms from left to right, respectively. The source or sink term  $q^C$  represents the exchange of mass of component  $C$  due to interphase diffusion and/or phase change [Mercer and Cohen, 1990].  $V$  represents a Eulerian support volume,  $\rho_\alpha$  is the density and  $S_\alpha$  is the saturation of phase  $\alpha$ .  $X_\alpha^C$  is the mass fraction of component  $C$  in phase  $\alpha$ , and  $\phi$  is the porosity of the medium. With the extended Darcy law Eq. (2.34), the relative permeability saturation relationship 2.35, the capillary pressure saturation relationship Eq. (2.30) and the closure relations (Eqs. 2.25-2.29), the system of equations (Eq. 2.40) is closed.

### 2.6.3 Tailored equation system

The entire system of partial differential equations (PDEs) in Eq. (2.40) could be approximately solved without any modification by standard numerical methods such as finite difference (FDM), finite volume (FVM) and finite element (FEM) methods. However, as briefly addressed in Sec. 2.4, this is not straightforward and demands tremendously huge computational power, if reliable field-scale answers are desired while the relevant scales are resolved. Therefore, it is always advisable to think carefully how to tailor the PDE system in Eq. (2.40) to effectively solve the problem of consideration.

When assessing contaminated sites, one is typically interested in



1. the stable state of DNAPL phase after infiltration, where DNAPL is entrapped in the pore matrix as pools and/or as ganglia (CSA),
2. the dissolution of DNAPL into the aqueous phase after DNAPL infiltration is finished,
3. the subsequent advective, dispersive, and diffusive transport of the DNAPL component within the aqueous phase,
4. and the DNAPL source depletion in the long term,

as discussed in Section 2.3. Thus, the multiphase PDE system Eq. (2.40) can be simplified to the specific needs in three steps.

**Step 1: stable state of DNAPL phase:** Neglecting diffusion and excluding external sources or sinks of the DNAPL component within the DNAPL phase, only the storage and advective term of Eq. (2.40) remains for  $\alpha = n$ . Furthermore, assuming quasi-static conditions (low flow rates of the DNAPL phase) and negligible viscous forces when DNAPL is invading the aquifer, the pattern of the DNAPL percolation depends only on the entry pressure and the hydrostatic phase pressure relationship in Eqs. (2.29-2.30). Assuming stable conditions (i.e., mechanical equilibrium), the entry pressure can be approximated by the Young-Laplace equation (2.38).

The DNAPL CSA stabilizes if the mechanical equilibrium of Eq. (2.29) is fulfilled in the entire source zone. This is a problem with a non-unique solution due to the uncertainty of pore-scale geometries in REV-scale models [e.g., *Glass et al.*, 1995] and due to the instable process of DNAPL infiltration. Thus, I randomly simulate many physically possible CSAs for any given aquifer structure. Each random CSA satisfies the condition of Eq. (2.29) within the uncertainty bounds due to unknown pore-scale features of the aquifer medium. For this task, I adopt a stochastic invasion percolation algorithm that is introduced in Section 3.2.

**Step 2: DNAPL dissolution and dissolved-phase transport** For the dissolution of immobile DNAPL and the dissolved-phase transport, the multiphase mass balance system (Eq. 2.40) simplifies to the advection-dispersion transport equation (ADE) for dissolved DNAPL in the aqueous phase with a dissolved DNAPL concentration  $c$ , an aqueous seepage velocity  $\mathbf{v}_s$ , and a diffusion and dispersion tensor  $\mathbf{D}$ . The ADE can be written as

$$\frac{\partial \phi_e c}{\partial t} + \nabla \cdot (\phi_e \mathbf{v}_s c) - \nabla \cdot (\phi_e \mathbf{D} \nabla c) = 0, \quad (2.41)$$

subject to an appropriate set of boundary conditions.  $\mathbf{D}$  [ $\text{L}^2 \text{T}^{-1}$ ] is the dispersion and diffusion tensor according to *Scheidegger* [1974]:

$$\mathbf{D} = (\alpha_T \|\mathbf{v}_s\| + D_m) \mathbf{I} + (\alpha_L - \alpha_T) \frac{\mathbf{v}_s \mathbf{v}_s^T}{\|\mathbf{v}_s\|}, \quad (2.42)$$

with  $\alpha_L$  [L] and  $\alpha_T$  [L] as the longitudinal and transversal dispersivities, respectively,  $D_m$  [ $\text{L}^2 \text{T}^{-1}$ ] as the molecular diffusion coefficient and  $\mathbf{I}$  as the identity matrix. The aqueous-phase seepage velocity is calculated with the extended Darcy's law (Eq. 2.34), using the relative permeability saturation relationship by Brooks & Corey (Eq. 2.35). The porosity  $\phi_e$  is the effective porosity for the aqueous phase, considering only the pore volume that is not occupied by trapped DNAPL. Please note, that volumetric source terms of the underlying flow, sorption, chemical reactions, and biodegradation are neglected here for the sake of convenience, but without loss of generality. If local chemical equilibrium is an appropriate assumption, Eq. (2.41) is subject to the following Dirichlet boundary conditions:

$$\begin{aligned} \Gamma &= \{(\mathbf{x}, t) \in \Omega \mid S_n(\mathbf{x}, t) > 0\}, \\ \hat{c}(\mathbf{x}, t) &= c_s \quad \forall \mathbf{x} \in \Gamma, \end{aligned} \quad (2.43)$$

where  $c_s$  [ $\text{M L}^{-3}$ ] is the solubility concentration,  $\hat{c}(\mathbf{x}, t)$  are prescribed concentration values,  $\Omega$  denotes the entire model domain, and the boundary domain  $\Gamma$  includes all parts of the domain where there are interfacial areas between water as acceptor phase and pure-phase DNAPL as donor phase, i.e., the CSA. If the mass transfer is not at local equilibrium but subject to a rate-limited mass transfer model, one obtains a third-type boundary condition:

$$\mathbf{n} \cdot [(\phi \mathbf{v}_s c) - (\phi D \nabla c)] = \kappa_{\text{eff}}(\hat{c} - c), \quad \forall \mathbf{x} \in \Gamma \quad (2.44)$$

where  $\mathbf{n}$  [-] is a unit vector orthogonal to the boundary, and  $\kappa_{\text{eff}}$  [ $\text{T}^{-1}$ ] is a lumped mass transfer rate coefficient.

**Step 3: DNAPL depletion:** Obviously, DNAPL mass transfer occurs only within regions that have a non-zero DNAPL saturation. These regions are called emitters in the following, and the entire CSA is conceptualized as a spatially scattered set of emitters. Since the DNAPL phase is assumed immobile after the infiltration has finished, only a storage term and a sink term  $q^C$  [ $\text{M L}^{-3} \text{T}^{-1}$ ] remains in the mass balance equation of the DNAPL phase:

$$\int_{V_k} \frac{\partial S_n(\mathbf{x}) \phi \rho_n}{\partial t} - q^C dV_k = 0, \quad (2.45)$$

where  $V_k$  [ $\text{L}^3$ ] is the support volume of emitter  $k$ . The storage term expresses the mass reduction rate of the immobile DNAPL phase at each emitter. This is equal to the sink term  $q^C$  which results from the mass transfer rate  $\dot{m}_k^C$  [ $\text{M T}^{-1}$ ] at the boundary conditions of the dissolved-phase transport (Eqs. 2.43 or 2.44):

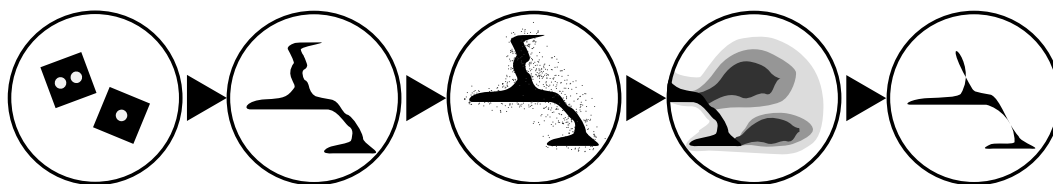
$$\dot{m}_k^C = \int_{V_k} -q^C dV_k. \quad (2.46)$$

For simplicity of notation, I omit the indication of component  $C$  in the following. The total emanating mass flux  $\dot{m}_{CSA}$  from a CSA may directly be calculated from the sum of all  $\dot{m}_k$ :

$$\dot{m}_{CSA} = \sum_{k=1}^{n_k} \dot{m}_k, \quad (2.47)$$

where  $n_k$  is the number of emitters. The emitter-wise mass release rates  $\dot{m}_k$  [ $\text{M T}^{-1}$ ] obviously reduce to zero if the DNAPL is locally depleted in any emitter volume  $V_k$ . Local depletion may change the mass release rates of all other remaining emitters. Hence, all  $\dot{m}_k$  have to be updated whenever an emitter is depleted.

### 3 Consistent model framework



The previous Chap. 2.3 identified a set of relevant processes and their key parameters that significantly influence the assessment of contaminated sites. These key parameters mutually interact through the described processes. Neglected or overly simplified interdependencies may corrupt the entire model framework. Yet, if the relevant processes and the key parameters are coupled in a physically and stochastically consistent manner within the model framework, not only the model predictions gain accuracy and plausibility, they also improve the confidence for decisions within contaminated site assessments. Concluding the discussion of physical processes above, it is clear that misleading impact metrics and remediation efficiencies are predicted in one- or two-dimensional, homogeneous, deterministic models. Hence, a sound model framework shall be conceptualized that couples the process and parameter interplay in a physically and statistically/stochastically consistent manner in reasonable scales, and with sufficient consideration of uncertainty.

Fig. 3.1 shows the entire model concept. It features a list of six physically interdependent aspects (parameter fields or processes). Each of these is highly irregular in space and so demands for a three-dimensional and highly resolved model concept. The most important process and parameter interdependencies are highlighted by arrows and are discussed in the following.

The most crucial uncertainties stem from heterogeneous material parameters such as the permeability field (1.) and from the stochastic aspect of CSA formation (3.). Hence, permeability  $K$  and DNAPL saturation  $S_n$  are modeled as random space functions (this is denoted by dices in Fig. 3.1) and this randomness propagates onto all other involved quantities. The permeability field directly influences the groundwater flow (2.), DNAPL percolation (3.), and solute transport (5.), while relative permeability (1.) is influenced by the DNAPL saturation distribution (3.). Here,  $K$  represents the material heterogeneity and is used to obtain entry pressure by regression. Porosity and other material parameters could be randomized as well, but will be treated as spatially constant and known values for simplicity. Such assumptions should be subject of closer investigation in future studies.

Groundwater velocities (2.) are the key driver of DNAPL dissolution (4.) and advective-dispersive solute transport (5.). Knowing the DNAPL saturation distribution (3.) and the

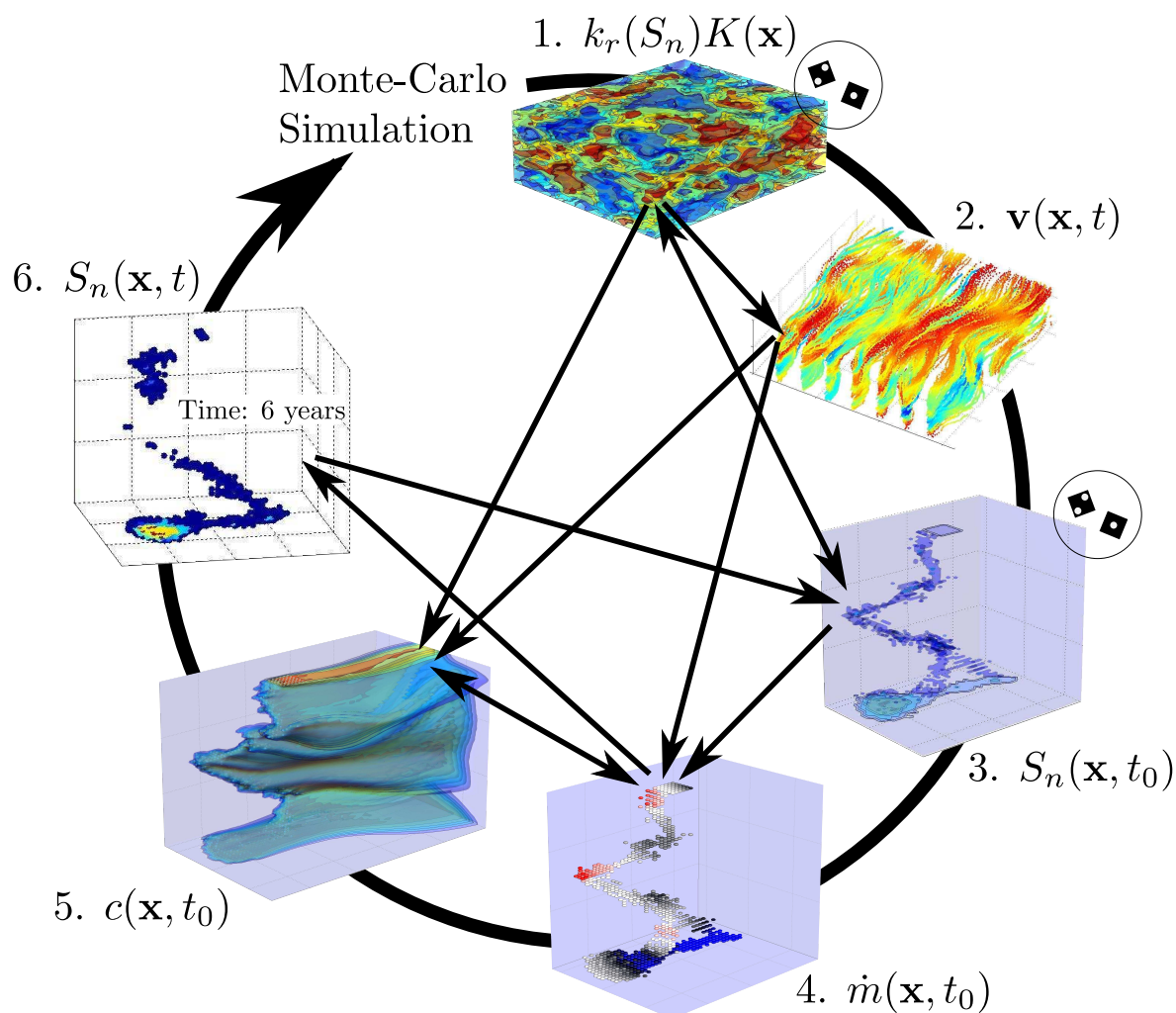


Figure 3.1: Model framework: 1. Intrinsic and relative permeability field, resulting from a geostatistical model  $K(\mathbf{x})$  and the DNAPL saturation distribution  $k_r(S_n)$ , 2. Groundwater flow field, resulting from a numerical solution of the groundwater flow equation using Darcy's law, 3. DNAPL saturation distribution, resulting from a stochastic invasion percolation algorithm, 4. Mass release rates and 5. dissolved-phase concentration distribution resulting from a mass transfer and solute transport model, and 6. DNAPL depletion, assuming stepwise quasi-steady states.

groundwater flow field (2.), the local mass release rates (4.) and solute transport (5.) can be calculated. Solving for the mass release rates requires knowledge on local concentration values, whereas solving for solute transport requires knowledge on mass release rates. Thus, both have to be solved simultaneously. Although the same model framework could be used to simulate rate-limited mass transfer as described in *Koch and Nowak* [2014a], I will mostly assume local chemical equilibrium within this study for simplicity. For the performance of refined rate-limited mass transfer models, I refer to the recent validation study of *Kokkinaki et al.* [2013].

The dissolution of DNAPL mass (4.) into the aqueous phase leads to a depleting DNAPL phase (6.) over long time scales, which then changes the DNAPL saturation distribution (3.) and relative permeability (1.). The temporal evolution of DNAPL dissolution and the CSA depletion (6.) is modeled by a time sequence of steady states, which I call the *quasi-steady-state* approach (QSS). This means that I am not modeling the time-dependent dissolved-phase transport and DNAPL dissolution immediately after a change of CSA structure, but instead directly look at the next temporary phase of quasi-steady-state DNAPL dissolution and transport. This procedure is continued until the CSA is fully depleted or until a characteristic depletion time is reached. The assumption behind this procedure is that groundwater flow through the CSA, which defines the time to reach steady conditions after a change of CSA structure, acts on a different scale than the depletion process (see Sec. 2.4).

The herein presented model framework uses and adopts simulation tools and concepts that were also used in earlier studies:

- Fast generation of random hydraulic permeability fields [*Dietrich and Newsam*, 1993] within a Bayesian framework [*Kitanidis*, 1986];
- Simulating groundwater flow with a Petrov-Galerkin FEM [*Brooks and Hughes*, 1982] as described in *Nowak et al.* [2008];
- Generating random CSAs with a stochastic invasion percolation algorithm (SIP) adapted from [*Glass et al.*, 1995; *Ewing and Berkowitz*, 1998; *Glass et al.*, 2001] (see Sec. 3.2);
- Simulating DNAPL dissolution and dissolved-phase transport via particle tracking random walk (PTRW) [*Koch and Nowak*, 2014a] (see Sec. 3.3);
- And simulating DNAPL depletion with an quasi-steady-state approach (see Sec. 3.4).

The remainder of this chapter will therefore focus only on the new methods for stochastic DNAPL percolation Sec. 3.2, coupled groundwater flow, mass transfer, and solute transport via particle tracking random walk Sec. 3.3, and for quasi-steady-state DNAPL depletion Sec. 3.4.

## 3.1 Goal, approach and contributions

### 3.1.1 Goal and approach

The goal of this chapter is to propose and test a model framework that can predict relevant impact quantities (local dissolution rates, mass discharge, and depletion time) with quantified uncertainty bounds at DNAPL contaminated sites in order to support risk-based decisions. Based on the discussions in Chap. 2, I design a physically and stochastically consistent model framework in order to ensure reasonable and reliable predictions. This demands a certain minimum degree of model complexity, which has yet to be defined. In this chapter, I point out the key role of aquifer heterogeneities, flow irregularity, and complex CSA morphology, and that reverting to overly simplified surrogates of fine-scale physically-based models is not advisable. To account for the involved heterogeneities and uncertainties, aquifer parameters and contaminant source architectures are conceptualized as physically reasonable random space functions (RSF). This results in probability density functions (*pdfs*) for all impact metrics of interest.

The aquifer heterogeneity, the groundwater flow, the CSA, and the resulting impact metrics are strongly interrelated through the joint physical processes of DNAPL infiltration, dissolution, dissolved-phase transport, and depletion. Therefore, these interactions together with their respective inherent uncertainty are considered within a highly-resolved and three-dimensional model framework resting on Bayesian (geo)statistics and stochastic simulation. Based on this novel model framework, I can investigate the importance of these mutual dependencies and uncertainties in a physically and stochastically consistent manner. This is done by corrupting the existing physical and statistical dependencies, replacing them by simplifications found more or less often in the literature. Then, the resulting bias or misspecification of uncertainties are analyzed for the predicted impact metrics compared to the undisturbed model. Finally, I can conclude whether the claimed model complexity is in fact necessary, or whether the model framework could be simplified in a plausible and acceptable manner. If I find the full complexity to be necessary, this does not imply yet that the maximum complexity level considered here is sufficient.

### 3.1.2 Relation to the state-of-the-art

This study is not the first one that seeks to model DNAPL infiltration, dissolution, and depletion in the subsurface. For example, *Parker and Park* [2004] analyzed the mass fluxes and source depletion based on a high-resolution three-dimensional simulation of DNAPL infiltration and dissolved transport for a single realization of a heterogeneous aquifer. However, given the impracticality of such high-resolution simulations when predicting field-scale dissolution kinetics, they proposed to use effective field-scale mass transfer laws that work only with averaged, homogeneous quantities. They calibrated and tested a proposed upscaled mass transfer function with the high-resolution simulation mentioned above. Similar simplified and upscaled models to predict DNAPL dissolution and depletion have been presented within the last decade [e.g., *Jawitz et al.*, 2005; *Christ et al.*, 2006; *Saenton and Illangasekare*, 2007;



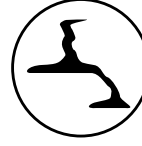
*Basu et al.*, 2008]. The common findings are that prediction accuracy highly depends on accurately capturing the morphology of the DNAPL distribution, and that upscaled predictions of source zone longevity should be handled with caution [e.g., *Christ et al.*, 2006]. In fact, *Zhang et al.* [2008] evaluated these approaches and suggested to revert to highly-resolved three-dimensional flow and transport simulators.

Following this rationale, *Kokkinaki et al.* [2013] investigated the validity of a coupled DNAPL infiltration and dissolution model for simulating three-dimensional laboratory flow cells that contain heterogeneous permeability fields. They could show that a coupled, three-dimensional, and finely discretized numerical model is able to closely predict the experimentally observed DNAPL infiltration pattern and dissolved mass discharge values. Due to extremely controlled experimental conditions and a well-characterized sand packing, their work did not study the issue of uncertainties and their quantification at contaminated sites. *Saenton and Illangasekare* [2013] ran Monte-Carlo simulations (MCS) of the distribution, dissolution, and depletion of DNAPL in two-dimensional intermediate-scale heterogeneous porous media, again in the size of laboratory bench-scale tests. Although they attributed crucial uncertainties to both the geologic heterogeneity and the inability to accurately predict complex CSAs, only hydraulic permeabilities were modeled as random variables. The applied simulation tools would, in principle, allow also for such MCS in three-dimensional and heterogeneous aquifers at the field scale. Yet, to the best of the authors knowledge, this has never been performed in the past. In specific, randomized CSAs that could be used to represent the uncertainty inherent in DNAPL infiltration at the relevant scales have never been used to predict probability distributions of impact metrics. A possible explanation is that the commonly applied simulation tools become prohibitively expensive for MCS in three-dimensional, finely-discretized, heterogeneous aquifers at the field scale required for contaminated sites.

Only a few studies have modeled random contaminant sources and depicted these as a major source of uncertainty in contaminant transport modeling [*Wang and Zheng*, 2005; *Chaudhuri and Sekhar*, 2006, 2008]. They modeled location, time, and strength of simplified contaminant sources randomly through simple geostatistical approaches. Thus, they neglected the physical processes that form DNAPL CSAs, and so these methods can not yet adequately represent the complex morphology of CSAs and their embedding in the heterogeneous aquifer. Such simplifications have substantial consequences, especially because the local hydraulic conductivities and flow velocities in contaminant source volumes have a significant impact on emanating mass fluxes and far-field transport [*de Barros and Nowak*, 2010]. Hence, these approaches exhibit severe bias of the predicted impact metrics, as will be analyzed in a series of test cases in Sec. 3.5.

**Contributions** In summary the work presented here is the first one to jointly consider relevant interdependencies between parameters of heterogeneous aquifers, complex shaped CSAs, and groundwater flow, the three-dimensionality of the physical space, appropriate spatial and temporal scales, the uncertainty in both aquifer parameters and CSA, and the first one that assess the physical and statistical consistency of such a model framework.

### 3.2 Stochastic DNAPL percolation



For generating random CSAs that correspond to the random conductivity fields from Sec. 2.5, I use a quick stochastic invasion percolation algorithm (SIP) adapted from *Ewing and Berkowitz [1998]* and *Glass et al. [2001]*. This is a reasonable choice (1) due to the dominant role of entry pressure for the DNAPL infiltration pattern and (2) since dynamic multiphase processes are not of direct interest here. The algorithm applies the Young-Laplace equation (Eq. 2.38) to calculate a representative capillary pressure  $p_c$  for each element on a fine numerical mesh, using characteristic pore apertures  $r$  and wetting angles  $\varphi$ . The obtained capillary pressure  $p_c$  defines a characteristic entry pressure  $p_d$ , which poses the threshold for initial invasion of the DNAPL phase into any given cell. Upon initial invasion, the entry pressure  $p_d$  is updated with the capillary pressure saturation relationship (Eq. 2.30) to quantify the capillary resistance against further displacement of the aqueous phase and thus against the increase of DNAPL saturation values. If individual elements possess significantly larger pore apertures than adjacent elements, they may be re-invaded several times up to the maximum DNAPL saturation imposed by the residual saturation of water. Assuming hydrostatic conditions, the differences of the phase pressures are calculated by  $\Delta\rho gh$ , and the entry pressures  $p_c$  are calculated with Eq. (2.38) and Eq. (2.30). If the percolating DNAPL is fed from above the aquifer, the hydrostatic DNAPL-phase pressure is increased by  $\rho_n g D_p$ , with  $D_p$  being the elevation of the DNAPL above to the groundwater table.

Theoretically, each element that shares an interface with an already invaded element could be invaded, if its current entry pressure is exceeded. However, these considerations are made at the REV scale, and are thus subject to uncertainty that originates from the pore scale (see Sec. 2.4). Hence, the invasion process is described in a stochastic fashion, where the invasion probabilities of all candidate elements are proportional to the respective difference between threshold pressure and DNAPL phase pressure. This element-wise invasion potential is called  $I_p$  and is given by the inequality of Eq. (2.29) at the current state of the DNAPL percolation model:

$$I_p = (\Delta\rho gh + \rho_n g D_p) - p_c(S_e). \quad (3.1)$$

The actual invaded elements are drawn by a stochastic selection algorithm that has been suggested by *Ewing and Berkowitz [1998]* and *Glass et al. [2001]*. Therefore, the  $I_p$  at all element-wise interfaces (i.e., neighboring elements) are calculated with Eq. (3.1) and listed in descending order. Here, 26 neighboring elements are modeled for each element that is occupied with DNAPL (8 within the same horizontal layer, and 9 above and below). In order to select individual elements for invasion, the sum of all invasion potentials  $\sum I_p$  is multiplied by a random variable  $\iota$  with uniform distribution  $\iota \sim \mathcal{U}(0, 1)$ . Then, from this product the listed  $I_p$  entries are cumulatively subtracted in descending order until further reduction would lead to negative results. The last entry in  $I_p$  that yields to a positive result corresponds to a specific element which is selected for invasion. Results of *Glass et al. [2001]*

suggest to selected a set of elements which are all invaded at once within one percolation step. The number of elements that are invaded at each percolation step, here denoted by  $N_e$ , is a modeler choice (could be also subject to uncertainty). In this case, the selected  $I_p$  entry is a characteristic value (e.g., median, maximum, or minimum) of the  $N_e I_p$  entries that correspond to the set of elements which are chosen for invasion. Then, the DNAPL saturation values of this set of elements are increased by  $\Delta S_n$  and a new list of  $I_p$  is generated according to updated capillary pressure values and/or to a new interface arrangement between DNAPL and aqueous phase. This procedure is continued until the sum of all element-wise DNAPL mass reaches the available DNAPL mass for invasion (e.g., total spilled DNAPL or content of a leaking underground storage tank).

It is well-known that the instability of the displacement front increases inversely proportional to the DNAPL infiltration velocity and the viscosity of the displacing fluid [Lenormand, 1990]. These effects are included in the selection algorithm, by also considering the Capillary number  $Ca = \mathbf{v}_n \mu_n / \gamma$ , where  $\mathbf{v}_n$  [ $\text{L T}^{-1}$ ] is the velocity of the displacing DNAPL,  $\mu_n$  [ $\text{M L}^{-1} \text{T}^{-1}$ ] is the DNAPLs viscosity, and  $\gamma$  [ $\text{M T}^{-2}$ ] is the interfacial tensions [ $\text{M T}^{-2}$ ]. To this end, the random variable  $\iota$  is weighted reciprocal proportional to  $Ca$  by  $\iota' = \iota^{|\nu|}$ , where  $\nu$  is a Gaussian random variable with  $\nu \sim \mathcal{N}(-\log(Ca), 1)$ . If a negative  $\nu$  is drawn it is reflected at  $\nu = 0$ , in order to restricted  $\nu$  to the range  $[0, +\infty]$ . It is  $\iota'$  that is actually used in the stochastic selection algorithm as described above, such that if  $\nu$  is large (e.g.,  $\nu \geq 20$ )  $\iota'$  approaches zero for most values of  $\iota$ . On the other hand, if  $\nu$  is small (e.g.,  $\nu \leq 1$ ), a high degree of randomness is added to the selection procedure and thus a majority of the selected elements is likely to correspond to lower invasion potentials  $I_p$  which stabilizes the DNAPL migration front. Hence, with increasing  $\nu$ , the CSA formation is increasingly influenced by the soil structure and its heterogeneity. Also, if the heterogeneity of aquifer parameters increases, the degree of their influence on CSA formation increases. This is important in the physical context, because stronger contrasts in permeability and pore apertures lead to more pronounced capillary barriers on which DNAPL is forced to accumulate and spread laterally. As a result, the stochastic selection rule proposed by Ewing and Berkowitz [1998] comprises the behavior of stable displacement (large  $Ca$ ) such as the Eden model [e.g., Eden, 1961] and capillary fingering (low  $Ca$ ) such as the invasion percolation (IP) model [e.g., Wilkinson and Willemsen, 1983].

It is experimentally well observed that DNAPL forms various fingers when it migrates through porous media [e.g., Held and Illangasekare, 1995; Glass et al., 2000]. These fingers grow independently and only interact because they are fed by the same source of DNAPL. To account for the independent behavior of individual fingers, an additional finger-wise stochastic selection is suggested by Ewing and Berkowitz [1998] which precedes the element-wise selection rule described above. Which finger may grow in the current percolation step is selected with the same stochastic procedure outlined above, but using a characteristic invasion potential for each finger, e.g., using the maximum invasion potential of each finger. Two independent uniform distributed random variables  $\iota_f$  and  $\iota_e$  and two independent weights  $\nu_f$  and  $\nu_e$  are used for finger- and element-wise selection, respectively. As a consequence of the finger selection, only the neighboring elements of the selected finger are allowed to be invaded at the current percolation step. Hence, the degree of randomness when selecting

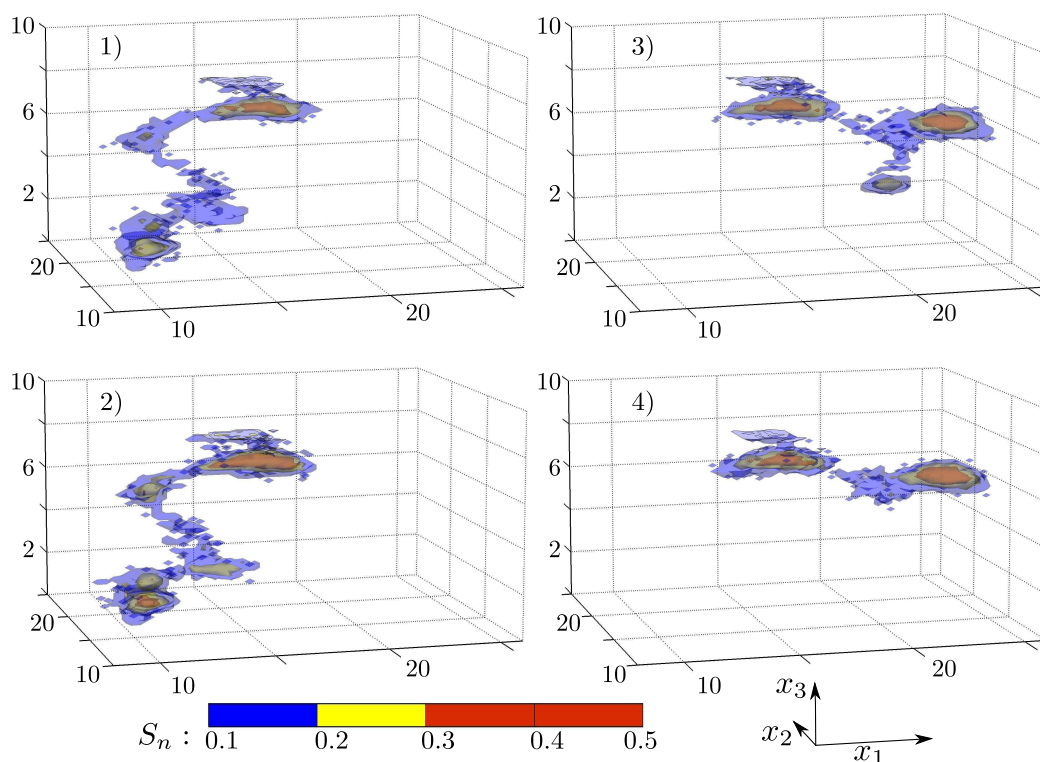


Figure 3.2: DNAPL saturation  $S_n(\mathbf{x})$  in four equally probable realizations of a contaminant source architecture resulting from the same realization of a heterogeneous aquifer (taken from the Scenario S1 in Sec. 3.5)

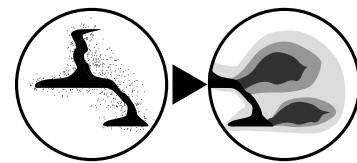
elements for invasion is restricted by the degree of randomness when selecting fingers.

This results in complex CSAs with ganglia and pool structures that are in accordance with aquifer structure and with respective regimes of DNAPL infiltration. The CSAs are defined as saturation values on the same fine numerical mesh that has already been used for the random conductivity field in Sec. 2.5, and that will be used for simulating groundwater flow. Applying the SIP algorithm described here repeatedly leads to an ensemble of various CSAs which are all statistically plausible and physically consistent with the state of knowledge at the chosen REV model scale. For illustration, Fig. 3.2 shows four equally likely realizations of a CSA obtained by the SIP algorithm for a single realization of a heterogeneous aquifer taken from Scenario S1 in Sec. 3.5. CSA 1 and 2 (left) and the CSA 3 and 4 (right) have similar pool arrangements, respectively. It seems to be highly uncertain where the DNAPL leaves the first pool and migrates further downwards. Depending on this decision, completely different CSAs emerge. The only significant difference between CSA 3 and 4 is whether the CSAs stabilize after the second pool, or whether a new ganglia is formed which ends in forming a small third pool below. This type and degree of uncertainty corresponds to field and laboratory experience [Brewster *et al.*, 1995; Glass *et al.*, 2000]. Although the proposed SIP algorithm is obviously a simplification of actual physical DNAPL migration processes and by far not the only possible means of generating physically plausible random CSAs, it

suffices for first analysis of a consistent model framework with random CSAs. Also, it is a significant improvement (both physically and statistically) over simpler CSA models such as upscaled CSAs or geostatistically motivated CSAs.

The stochastic selection rules of selecting fingers and invaded elements has been developed by *Ewing and Berkowitz* [1998], and is only slightly modified here for the purposes of this thesis. The macro modified invasion percolation method (MMIP) developed by *Glass et al.* [2001] used volume averaged saturation-pressure curves to characterize the porous medium and model the percolation in partially saturated porous media. Extending these approaches to the Darcy scale via relating representative pore radii to random permeability fields and conceptualizing the CSAs as random space functions is an essential novelty of this thesis.

### 3.3 Groundwater flow, mass transfer and solute transport



Groundwater flow is solved via a standard Galerkin finite element code as described in [*Nowak et al.*, 2008], using the same mesh as the random conductivity field (see Sec. 2.5). The mass transfer and solute transport are simulated with a recently developed particle tracking random walk method (PTRW) that allows to include Dirichlet or third-type boundary conditions [*Koch and Nowak*, 2014a] (see Chap. 4). Each location with a non-zero DNAPL saturation in the CSA (as generated by the SIP algorithm) forms an individual emitter volume  $V_k$  that imposes a Dirichlet or third-type boundary condition. At first, unit-release transport from each discrete emitter is simulated with standard PTRW techniques. These techniques are described in *Kinzelbach* [1988]; *LaBolle et al.* [1996]; *Wen and Gómez-Hernández* [1996]; *Salamon et al.* [2006]; and *Fernández-García and Sánchez-Vila* [2010], but the technical details are irrelevant here. To assess the boundary condition, the method performs a Galerkin projection of the PTRW-based unit-release particle densities onto control volumes  $V_l$  that discretize the boundary. Emitter-wise mass release rates  $\dot{m}_k$  and domain-wide concentration maps are then obtained in a swift post-processing step that accounts for the mutual dependencies of the mass release rates due to shadowing effects (Sec. 2.4) by solving a linear system of equations. In Chap. 4, this method is described in full detail. Also, it is successfully tested and benchmarked against the conceptually related semi-analytical method MASST (multiple analytical source superposition technique) by *Sale and McWhorter* [2001], and against a streamline-upwinding Petrov-Galerkin finite element method (FEM) [e.g., *Brooks and Hughes*, 1982]. Contrary to FEM and MASST, the PTRW method ensures robust behavior throughout various parameter settings of heterogeneous aquifers and CSAs, and allows for huge Monte-Carlo simulations as it will be demonstrated in Sec. 3.5. In order to provide a sufficient basis and to explain how the PTRW concept is acting within the current context, a brief description will be provided already here.

In general, the goal of this method is to solve the advection-diffusion equation (ADE) with a PTRW technique for cases where Dirichlet or third-type boundary conditions are given within an arbitrarily shaped subvolume (e.g., source zones)  $\Gamma$  of the domain  $\Omega$ . The ADE Eq. (2.41) subject to Dirichlet Eq. (2.43) or third-type Eq. (2.44) boundary conditions is given in Sec. 2.6.3. Since both Dirichlet and third-type boundary conditions refer to concentrations,

they introduce the function  $\hat{c}$  which describes the boundary concentration, e.g., for Dirichlet boundary conditions  $c(\mathbf{x}, t) = \hat{c}(\mathbf{x}, t)$  if  $(\mathbf{x}, t) \in \Gamma$ .  $\Gamma$  may include entire sub-domains as internal boundaries. In the herein presented application of DNAPL dissolution into ground-water flow, this includes all parts of the domain where there are interfacial areas between water as acceptor phase and pure-phase DNAPL as donor phase, forming a source zone. The method allows any geometry for  $\Gamma$ , which could be a fragmented object that is formed by donor-phase ganglia and pools [e.g., *Christ et al.*, 2005; *Page et al.*, 2007]. In this chapter I proceed on the assumption of local chemical equilibrium which leads to Dirichlet boundary conditions. Yet, this is not a restriction of the presented method, but rather a choice merely made for illustrative purposes. If the mass transfer is not at local equilibrium but subject to a rate-limited mass transfer model, one obtains a third-type boundary condition Eq. (2.44). This can be also treated by the PTRW method as addressed in Sec. 4.4.

The underlying principle of the presented method is the following: First, I discretize the volume  $\Gamma$  that participates in the boundary condition into small emitter volumes  $dV$ . Instead of releasing particles in these volumes  $dV$  according to some (yet unknown) mass release rate  $dm$  that would satisfy the Dirichlet (or third-type) boundary condition  $\hat{c}(\mathbf{x})$ , at first a unit mass release rate  $\dot{m}_0$  is released in each  $dV$  and I solve the corresponding set of ADEs by PTRW. If the featured transport equation is a linear PDE (this is the case for the ADE Eq. (2.41)), it holds that

$$\frac{dm}{\dot{m}_0} = \frac{\hat{c}}{c_0} \equiv \varepsilon, \quad (3.2)$$

where  $c_0$  is the resident concentration caused by  $\dot{m}_0$  in the volume  $dV$ , and  $\varepsilon$  is a (yet unknown) correction factor for the emission rate  $\dot{m}_0$  within  $dV$ , required to obtain the desired boundary values  $\hat{c}(\mathbf{x})$ . That means, the desired boundary concentration values  $\hat{c}(\mathbf{x})$  are installed through superpositioning and linear rescaling of the unit-release PTRW results by some factors  $\varepsilon$ . All concentrations in the following will be resident concentrations,  $c = dm/dV$ , which is relevant for defining Dirichlet (or third-type) boundary conditions in the presence of diffusive and dispersive fluxes [*Parker and van Genuchten*, 1984]. Because  $\dot{m}_0$  is a unit mass release,  $\varepsilon$  is in fact the final mass emission rate required to install the respective boundary condition. The remaining task is to find the values of  $\varepsilon$  for all elementary source volumes  $dV$  such that the desired values  $\hat{c}(\mathbf{x})$  are met. To that end, a linear equation system is arranged and solved for the values of  $\varepsilon$  (for more detail see Secs. 4.3 and 4.4):

$$\varepsilon \mathbf{\dot{m}}_0 = \mathbf{T}^{-1} \hat{\mathbf{c}}^*. \quad (3.3)$$

Here,  $\mathbf{T}$  reflects the mutual dependencies through advective-dispersive transport between all emitters that discretize the Dirichlet boundary condition. Its entries  $T_{k,l}$  are calculated via Galerkin projection of the  $k^{th}$  unit-release PTRW simulation with the  $l^{th}$  test function, e.g., using a Gaussian kernel as test function.  $\hat{\mathbf{c}}^*$  is a right-hand side vector that represents the boundary conditions mapped onto the emitter volumes. If the boundary condition shall install solubility concentration  $c_s$  within the entire source zone, the entries of  $\hat{\mathbf{c}}^*$  are

$$\hat{c}_l^* = c_s V_l \quad (3.4)$$

with  $V_l$  being the  $l^{th}$  emitter volume.

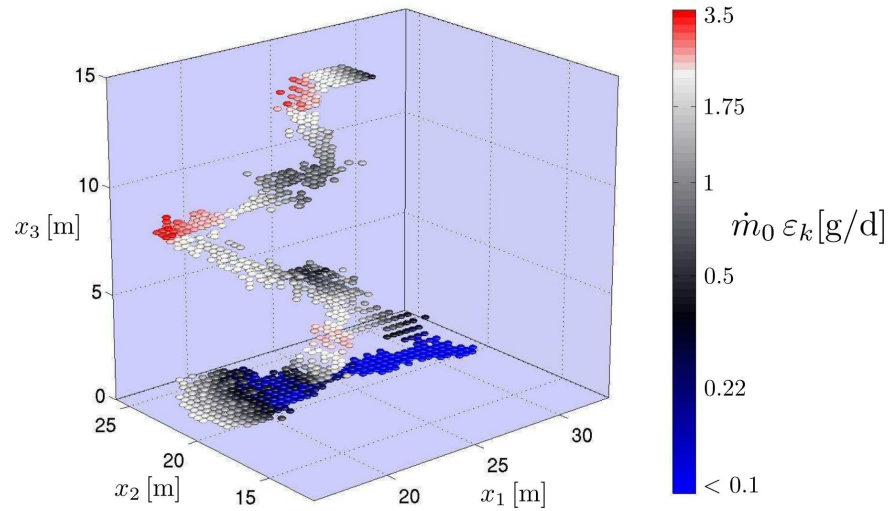


Figure 3.3: Mass release rates from a complex-shaped DNAPL source zone in a heterogeneous aquifer, discretized by Gauss-shaped emitter functions. Please observe the log-scale of mass release rates and the highly non-uniform color scaling.

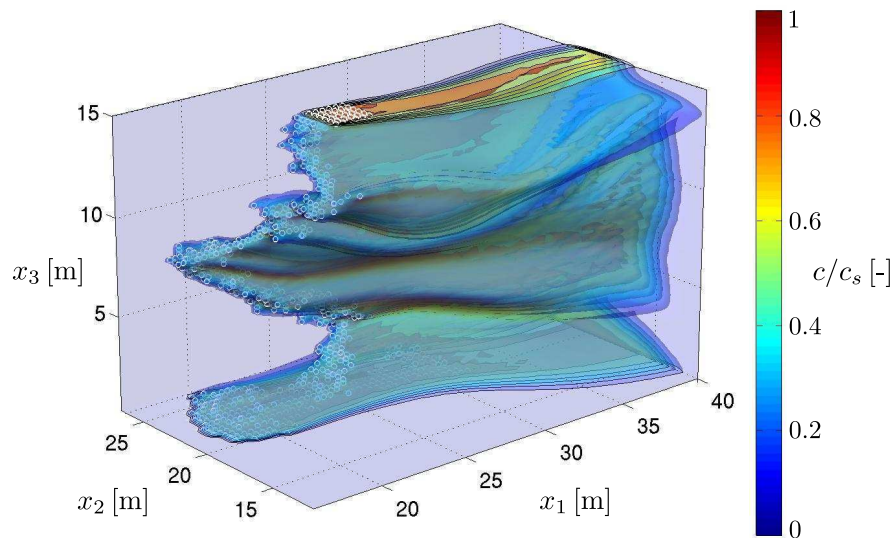


Figure 3.4: Resulting dissolved-phase concentration plume (the values are normalized by the saturation concentration  $c_s$ ) corresponding to the mass release rates shown on the left.

For illustration, the resulting pattern for the mass release rates (Fig. 3.3) and the resulting

dissolved-phase plume (Fig. 3.4) for one specific example of a DNAPL CSA embedded in a heterogeneous aquifer is shown and discussed in the following. The release rate pattern shows large dissolution rates on the upstream faces of the source zone. The DNAPL dissolved in upstream emitters is transported towards emitters located downstream. Thus, reduced emission rates are observed further downstream. Different from this advective shadowing effect, hydrodynamic dispersion and molecular diffusion causes a shadowing effect also between transverse neighboring emitters, allowing emitters with fewer neighbors to release more mass. Hence, slightly increased dilution rates can be observed at ganglias, and at edges and faces of DNAPL pools. Peak dissolution rates can be observed in regions of higher velocity zones due to the irregular nature of velocities in heterogeneous formations. The latter effect, in combination with the irregular geometry of the source, dominates the resulting PTRW transport simulation (Fig. 3.4). The big DNAPL pool at the bottom of the aquifer is only allowed to release mass at its upstream face and is exposed to relatively small groundwater velocities due to the reduced relative permeability at large DNAPL saturations, indicating that this pool will last for a very long time.

### 3.4 Quasi-steady-state DNAPL depletion



Once the emission rates  $\dot{m}_k$  are obtained from the PTRW algorithm described above, the depletion time of each emitter is swiftly obtained by assuming that the mass release conditions are at quasi-steady state on an intermediate time scale between that of DNAPL depletion and that of advective-dispersive transport through the CSA (see Sec. 2.4). If the emitter-wise mass balance (Eq. 2.46 in Sec. 2.6.3) indicates that the donor phase has been depleted at the  $k^{th}$  location, the  $k^{th}$  row and column are simply erased in Eq. (3.3), and the new mass release rates  $\dot{m}_k$  follow immediately. Only if the depletion has non-negligible effects on the flow field (i.e., through changes in relative permeabilities, see Eq. 2.35), the groundwater flow field and the unit-release PTRW simulations have to be updated. In order to manage the entire depletion simulation with only a few updates, the continuous changes of relative permeabilities are considered by an intermediate post-processing procedure after each update.

Thus, the quasi-steady-state (QSS) depletion approach introduces two distinct model simplifications, since two processes influence the dissolution rates at two different time scales. First, the emitter-wise dissolution rates  $\dot{m}_k$  change every time the CSA structure changes (due to depleting emitters), which is accounted for via a simple rearrangement of the matrix system that results from the PTRW simulation as described in Sec. 3.3 and in more detail in Chap. 4. Second, the dissolution rates change continuously with changing groundwater velocities, which are influenced in turn by increasing relative permeabilities. Hence, incorporating this effect requires a recalculation of groundwater flow and solute transport.

These calculations become computationally cumbersome if they are repeated at high frequencies. Motivated by the large time scale of the second process compared to the first one and by the computational costs for the second update, it is desirable to perform such updates as seldom as possible. Still, keeping the relative permeability  $k_r$  constant within any



time interval underestimates the continuous increase of local velocities at the emitter locations. This may lead to a sudden increase of local velocities and dissolution rates when updating the relative permeability from  $k_r(S_n^*(t_{i-1}))$  to  $k_r(S_n(t_i))$ , with  $S_n^*(t_{i-1})$  being the result of the preceding update and correction as described below. This indicates that too little mass has been depleted during the preceding time interval  $\Delta t_i$  of constant  $k_r(S_n^*(t_{i-1}))$ . To correct for this error, an exponential decay model is used to interpolate the emitter-wise dissolution rates between two consecutive time steps  $t_i$  and  $t_{i-1}$ :

$$\dot{m}_k^*(t_{i-1} + \tau) = \dot{m}_k(t_{i-1}) \exp(b_{k,i} \tau), \quad (3.5)$$

with

$$b_{k,i} = \frac{1}{\Delta t_i} \log \left[ \frac{\dot{m}_k(t_i)}{\dot{m}_k(t_{i-1})} \right], \quad (3.6)$$

where  $\tau$  is the time coordinate between the updates. Over the time interval  $\Delta t_i$ , this leads per emitter to a total released mass  $\Delta m_{k,i}$ :

$$\Delta m_{k,i} = \int_0^{\Delta t_i} \dot{m}_k(t_{i-1}) \exp(b_{k,i} \tau) d\tau = \frac{\dot{m}_k(t_i) - \dot{m}_k(t_{i-1})}{b_{k,i}}. \quad (3.7)$$

Thus, the DNAPL saturation value of emitter  $k$  is updated by  $S_{n,k}^*(t_i) = S_{n,k}^*(t_{i-1}) - (\Delta m_{k,i} / \rho_n V_k)$ , with  $V_k$  being the volume of emitter  $k$  and  $\rho_n$  being the DNAPL density. Emitters that run empty within  $\Delta t_i$  are not considered by this post-processing step. Also, the depletion times of emitters that run empty due to the correction step are not dated back but simply erased at the current time. Fig. 3.9a, in Sec. 3.5.1 will show that the twofold QSS allows for large time intervals between updates of the groundwater velocity, yet keeping a high accuracy of the depletion time and pattern.

For illustration, the shape and the DNAPL saturation values of the CSA already presented in Fig. 3.3 are shown in Fig. 3.5 at three different times: after 5, 10, and 35 years of depletion from left to right, respectively. A few components of the CSA are already depleted after 5 years. These substructures are also highlighted in Fig. 3.3 as regions with extremely high dissolution rates (red spheres), relative to the dissolution of the remaining components of the CSA. The long ganglia ( $x_1 \approx 20 - 30\text{m}$  and  $x_3 \approx 7 - 0\text{m}$ ) pointing slightly in mean flow direction depletes slower than typical ganglia, due to shadowing effects. Thus, parts of this ganglia still remain after 10 years of depletion. Finally, after 35 years, all ganglia parts of the CSA are depleted and only the pool on top of the aquitard (and the ganglia shadowed by it) remains and depletes further on for about another 30 years (not shown here).

### 3.5 Results and discussion

Finally, the resulting simulation framework is analyzed. Most of all, I investigate the importance of the individual model components and of their interdependence. To do so, I introduce several physical and statistical simplifications in five different scenarios (see Fig. 3.6, Tab. 3.2 and Sec. 3.5.1) that cut one of the interdependencies at a time, and I discuss their

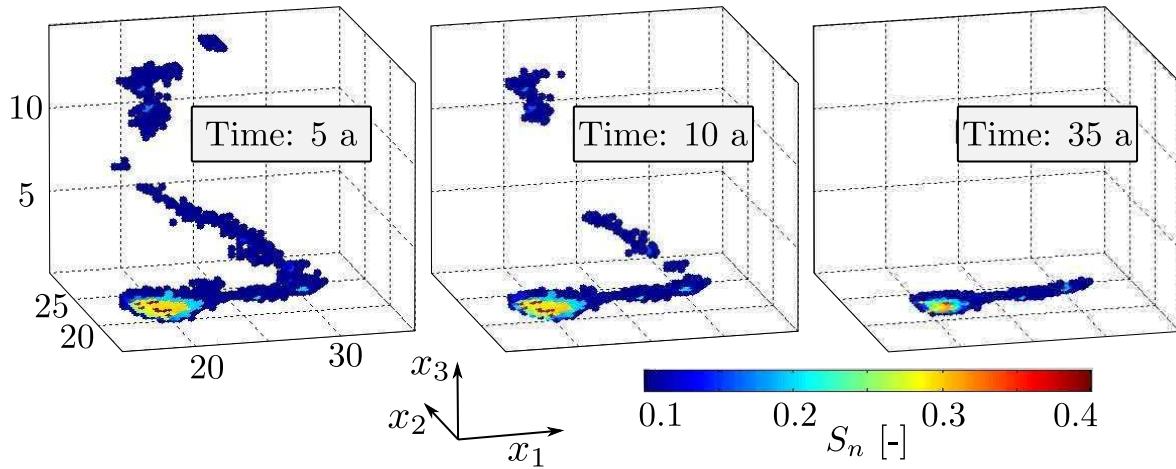


Figure 3.5: DNAPL saturation values at three quasy-steady-states after 5, 10, and 35 years of depletion from left to right, respectively.

impact on the obtained *pdfs* for mass discharge and source depletion time. It is not unambiguously possible to determine whether the presented model framework is sufficiently complex and detailed or not. However, I can test whether commonly applied simplifications are too strong, or whether they still lead to similar results.

### 3.5.1 Test case scenarios



Figure 3.6: Five scenarios to test whether the proposed complexity (physical and stochastic consistence) is required.

**Scenario S1** is the base case scenario, where I apply the presented model framework without modifications. The interdependent random parameter fields for permeability, the CSA (DNAPL saturation), and groundwater velocity are generated and solved within a model domain of  $48 \times 40 \times 10$  m. The domain is discretized with rectangular elements sized  $0.4 \times 0.4 \times 0.2$  m in mean flow, lateral horizontal, and vertical direction, respectively. The uncertainty of these parameter fields and of the model predictions (i.e., mass discharge and source depletion times) are quantified by a Monte-Carlo simulation with 1000 realizations. A summary of the used parameter values is provided in Tab. 3.1. Please note, that any parameter modeled deterministically here might just as well be considered to be uncertain, and the uncertainty choices made here are merely a scenario assumption. The results of Scenario S1

Table 3.1: Input Parameters, with  $i = 1, 2, 3$ ,  $\iota \sim \mathcal{U}(0, 1)$  and  $\log(v) \sim \mathcal{N}(0, 1)$ 

<b>Model Domain</b>			
Domain size	$L_{1,2,3}$	[m]	[40, 48, 10]
Grid spacing	$\Delta x_{1,2,3}$	[m]	[0.4, 0.4, 0.2]
<b>Geostatistic</b>			
Mean exp $Y$	$\mu_Y$	[m/s]	$10^{-4}$
Variance $Y$	$\sigma_Y^2$	[(m/s) <sup>2</sup> ]	$(\iota + 1)^2$
Correlation length	$\lambda_i$	[m]	$2\iota + [6, 6, 2]$
Matérn shape	$\kappa_m$	[-]	$v + 0.5$
<b>Flow and Transport</b>			
Hydraulic gradient	$\nabla h$	[-]	$10^{-2}$
Diffusion	$D_m$	[m <sup>2</sup> /s]	$1 \times 10^{-9}$
Dispersion	$\alpha_l, \alpha_t$	[m]	[ $10^{-2}, 10^{-3}$ ]
Effective porosity	$\phi_e$	[-]	0.3
<b>DNAPL Source</b>			
Spill center	$\mathbf{x}_s$	[m]	[20, 20, 15]
Total DNAPL mass	$m_n$	[t]	[2]
Solubility concentration	$c_s$	[kg/m <sup>3</sup> ]	1
DNAPL density	$\rho_n$	[kg/m <sup>3</sup> ]	1465
DNAPL viscosity	$\mu_n$	[Pa s]	$5.7 \times 10^{-4}$
Brooks-Corey shape	$\lambda_b$	[-]	2

are a *pdf* of dissolved mass discharge immediately after CSA formation (Fig. 3.8), and a *pdf* of source depletion time (Fig. 3.9). Also, I can analyze the probability distributions of absolute velocity and conductivity at the locations where the CSA resides (Fig. 3.10). These *pdfs* serve as reference to compare and discuss the corresponding results from Scenarios S2-S5.

**Scenario S2** investigates how important it is to use complex CSA shapes that are conditional to any given aquifer structure. For this analysis, I decouple the CSA and  $K(\mathbf{x})$  results of Scenario S1: The CSA obtained by the SIP algorithm for hydraulic conductivity realization  $i$  is placed in the conductivity field of realization  $j$ , where always  $i \neq j$ . After this exchange, I proceed to follow the framework of Scenario S1 without further modifications. This destroyed link introduces additional uncertainty, as the physical consistency between stochastically generated CSAs and aquifer realizations (as imposed by the SIP algorithm) is replaced by pure unconditional randomness.

**Scenario S3** investigates the importance of using (or knowing about) CSAs with adequate spatial structure. For this analysis, I simply ignore the physical formation process for CSAs and emplace a simple cubic CSA in the domain center. The cube is designed to have the same cross-sectional area orthogonal to the direction of mean flow as the average cross-sectional

Table 3.2: Scenarios to test the requirement of the proposed model consistency.

	CSA		Aquifer		Coupling	CV*
	physically	random	physically	random		
S1	✓	✓	✓	✓	✓	0.64
S2	✓	✓	✓	✓	–	0.86
S3	–	–	✓	✓	–	1.15
S4	✓	✓	–	–	–	0.29
S5	✓	✓	✓	✓	–	0.63

\* CV: coefficient of variation

area of the CSAs generated in Scenario S1, the same total mass, average thickness and average saturation. Under uniform flow and advective-dominated dissolved-phase transport, this would lead to identical initial mass fluxes, but already to a different depletion behavior.

**Scenario S4** investigates the impact of aquifer heterogeneity during the dissolution and depletion phase. Here, I replace the heterogeneous aquifers from Scenario S1 by an equivalent homogeneous medium that leads to the same average groundwater velocity as the ensemble average in S1, but I use again the complex-shaped CSAs from Scenario S1. Since the random CSAs are the only remaining source of uncertainty, I can analyze how this uncertainty is propagated to the resulting mass discharge *pdf*, isolated from the uncertainty that originates from randomly heterogeneous conductivity fields.

**Scenario S5** investigates the importance of considering relative permeability effects. Here, I simply set all relative permeability values to one, regardless of DNAPL saturations. The groundwater flow through the CSAs will be directly influenced by this simplification, leading to larger dissolution rates, larger mass discharges, and faster depletion.

To illustrate the base case scenario, Fig. 3.7 shows DNAPL saturation  $S_n(\mathbf{x})$ , emitter-wise dissolution rates  $\dot{m}_k$ , DNAPL dissolved-phase concentration  $c/c_s(\mathbf{x})$ , and log-conductivities  $\ln K(\mathbf{x})$  for an individual realization taken from Scenario S1. Here, one can see how the DNAPL migrated down along irregularly shaped ganglia, forming four distinct pools. Each pool resides in a high-conductivity zone on top of a capillary barrier at  $x_1 \approx 9, 17, 18,$  and  $22\text{m}$ , denoted in the following by P1, P2, P3, and P4, respectively. Each pool can only dissolve (if at all) on its upstream face, while the core and back side are fully shadowed by the pool itself. The upstream located pool P1 has relatively large dissolution rates at its front and top face (green emitters), which casts a shadow on the ganglia that is connecting this pool with P2. Thus, the dissolution rates from this ganglia are restricted (white emitters). The two pools P2 and P3 release only little mass, due to lower local groundwater velocities compared to those at P1. Because the dissolved DNAPL compounds emanating from P3 and from the ganglia between P3 and P4 are transported towards P4, the dissolution rates from P4 are restricted, although P4 experiences high local groundwater velocities (large blue

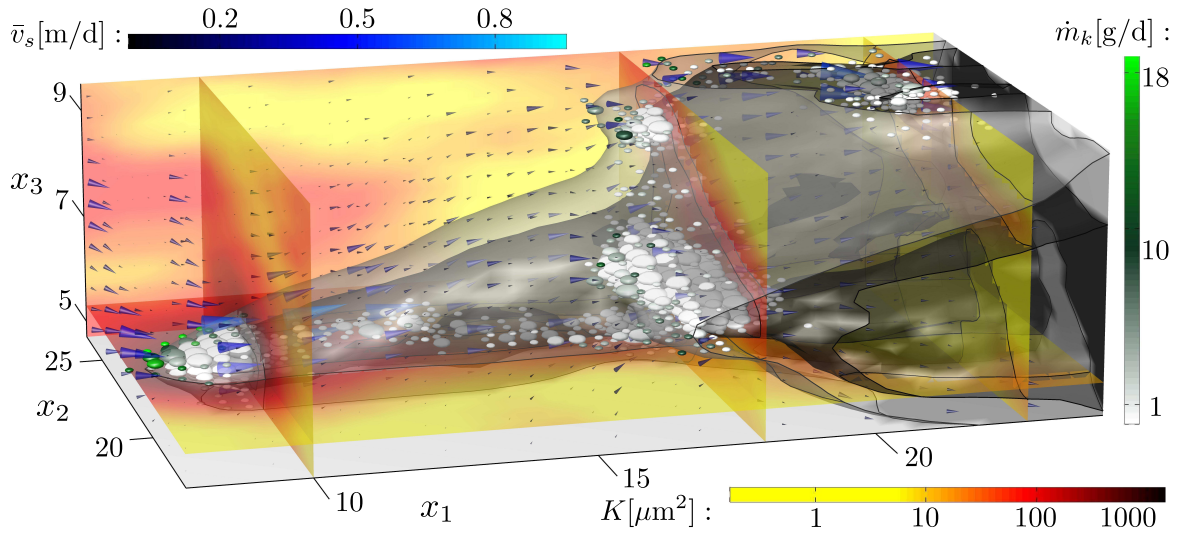


Figure 3.7: Illustration of base case scenario (S1) by one realization of the CSA ( $S_n(\mathbf{x})$ : size of spheres), dissolution rates ( $\dot{m}_k$ : color of spheres), DNAPL dissolved-phase concentration ( $c/c_s(\mathbf{x})$  [-]: transparent isosurfaces at 0.5 and 0.9 in light and dark gray, respectively), and log-conductivity ( $\log K(\mathbf{x})$ : background color).

cones). Overall, the shown CSA has a relatively small cross-sectional area orthogonal to the direction of mean groundwater flow, because the DNAPL migration path is mostly oriented along the main flow direction in this specific realization. This causes massive shadowing effects and a small mass discharge, and thus leads to an extremely low depletion rate of this CSA.

### 3.5.2 Discussion of mass discharge results

As shown in Fig. 3.8, the most consistent framework (S1) has a clear trend towards much higher mass discharges compared to the Scenarios S2-S4. Hence, the applied simplifications in the 2nd, 3rd, and 4th Scenarios lead to severe underestimation of mass discharge as relevant impact metric for assessing contaminated sites. In contrast to this, ignoring the reduced relative permeability values  $k_r$  as in Scenario S5 obviously leads to higher mass release rates.

Although the DNAPL source zones in Scenario S2 and S3 experience the same local conductivity and velocity statistics (see Fig. 3.10 and corresponding discussion in Sec. 3.5.4), and have the same average discharge area, source thickness, and DNAPL saturation by design, their estimated mass discharge *pdfs* reveal clear differences. To be more specific, the mass discharge *pdfs* of both S2 and S3 spread over more orders of magnitudes than all other scenarios, with a clear shift towards smaller values compared to S1. As a result, S2 and S3 yield the highest coefficients of variation  $CV = \sigma_{\dot{m}}/\mu_{\dot{m}}$  (see Tab. 3.2). Comparing S1 to S2, these differences originate only from the structural difference between realistic CSAs with and without correspondence to the aquifer structure. In S3, they originate from the simplistic assumption of cube-shaped CSAs. For S3, the reason for the smaller mean value lies in the

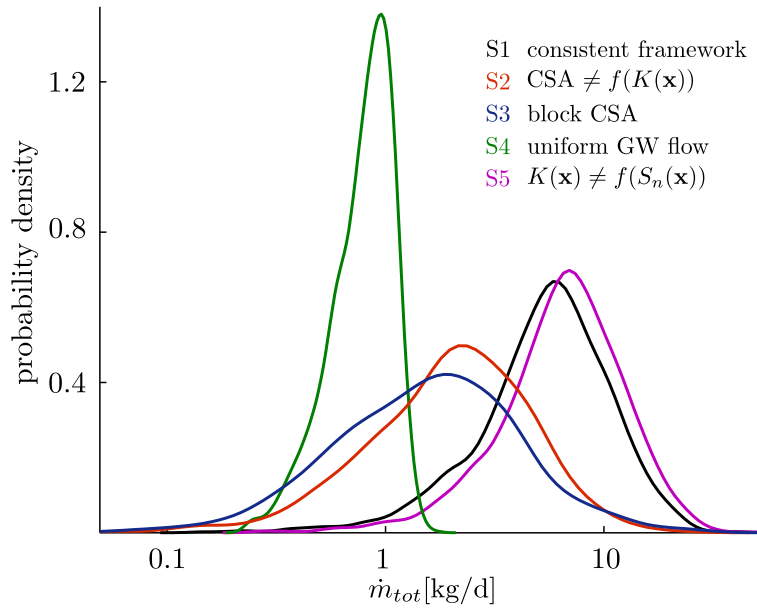


Figure 3.8: Probability density functions for log mass discharge resulting from the five Scenarios S1-S5.

extreme degree of shadowing between downstream and upstream emitters within the cube and in the reduced relative permeabilities on a relatively large spatial scale. In combination, I can conclude that ignoring the physics of CSA formation leads to severe bias.

The increase in uncertainty for Scenario S2 and S3 stems from the fact that the CSAs in S1 are conditional on the respective conductivity field, while they are unconditionally random in S2 and unconditionally cubic in S3, respectively. Thus, one can interpret the physical process description for the CSA formation process within the SIP algorithm as information, which reduces the uncertainty in predicting mass fluxes when used properly as in S1. Overall, these insights highlight the importance of utilizing physical knowledge on the CSA formation processes and on the resulting parameter interdependencies when simulating the fate of and transport from DNAPL contaminated sites.

The worst result is obtained when the groundwater flow irregularity is ignored (S4). Due to the drastic assumption of deterministic and uniform flow, this scenario shows a highly overrated confidence. Also, it shows the strongest underestimated mean value of mass discharge, caused by the destroyed link between DNAPL residence location and groundwater velocity. The mechanism behind the downward bias will become clear in Sec. 3.5.4. The combination of overrated confidence and underestimated impact is a serious and dangerous situation for decision making purposes. The remaining uncertainty in mass discharge ( $\dot{m}$  still varies over one magnitude) is solely caused by the random character of CSAs generated by the SIP algorithm and points out that CSA uncertainty cannot be ignored. Together with the results from S2 and S3, this shows that CSA structures must be simulated with adequate methods that cover the following aspects: (1) they must produce adequate CSA structures in correspondence to aquifer heterogeneity and groundwater velocity fields, and

(2) they must account for the uncertainty in the DNAPL infiltration process which remains even if the aquifer structure was perfectly known.

Scenario S5 results in a slight overestimation of mass discharge. Of course, the effect of the simplifications in S5 depends on the choice of the relative permeability saturation model and its parameter values (here I use the Brooks & Corey model in equation (2.35) with  $\lambda_b = 2$ ). At least in the given example, the total effect is surprisingly small. The main reason is that the lowest values of  $k_r$  usually appear in the core of pools with high DNAPL saturation values, where no dissolution takes place anyway due to shadowing effects. Thus, setting  $k_r = 1$  within pools causes only a small increase in mass fluxes. Also, the low saturation values typically found in ganglias cause only changes in hydraulic permeability within at most one order of magnitude (i.e.,  $k_r$  varies between 0.1 and 1). As discussed in Sec. 2.4, ganglia structures of the CSA may affect the local velocities to some degree but can not significantly change the global groundwater flow. Thus, setting  $k_r = 1$  within ganglias has, again, only a small effect towards larger mass fluxes. Contrary, groundwater velocities may increase in the direct surrounding of DNAPL pools because the groundwater is forced to bypass the zone of extremely low relative permeability values inside the pool. This accelerates the transport of dissolved DNAPL away from the pool's surface to a small degree. Thus, when setting  $k_r = 1$  within pools, there is a secondary effect that may lead to lower mass discharges. In total, however, the first two effects are stronger, leading to a slight net increase of mass discharges when artificially setting  $k_r = 1$ . To conclude, changes in relative permeability values seem to influence the mass discharge (at least at early times and in the investigated scenario) only to a small extent.

### 3.5.3 Discussion of results for source depletion times

The previous section revealed that any of the applied simplifications in Scenarios S2-S4 has severe effects on mass discharge estimates at early times, i.e., just after the CSA has been formed. It is intuitively clear that these biases and misspecifications of uncertainty will not disappear for any good reason during later stages of the depletion process. Therefore, at least the same degree of model complexity is required for predicting source depletion time as the one I already observed for mass discharge estimates. For that reason, I discontinue the discussion of scenario S2-S4.

For predicting the depletion process, the interdependence between all processes requires an adequate updating of several relevant quantities as the depletion process proceeds. In particular, changes in the CSA due to dissolution affects relative permeabilities (via equation 2.35) and hence the groundwater flow field (via equation 2.34). Also, the CSA depletion process changes shadowing effects within the CSA. For example, whenever an upstream emitter disappears, it will lift its shadowing effect from a downstream emitter. Altogether, this modifies dissolution rates, and hence feeds back into the dissolution process.

Continuing the comparison between S1 and S5, I analyze the importance of updating the groundwater velocities due to changes in DNAPL saturation and related changes in relative permeabilities when simulating source depletion. All other simplifications (Scenarios S2-S4) do not introduce an additional bias (compared to the error mechanisms already observed for

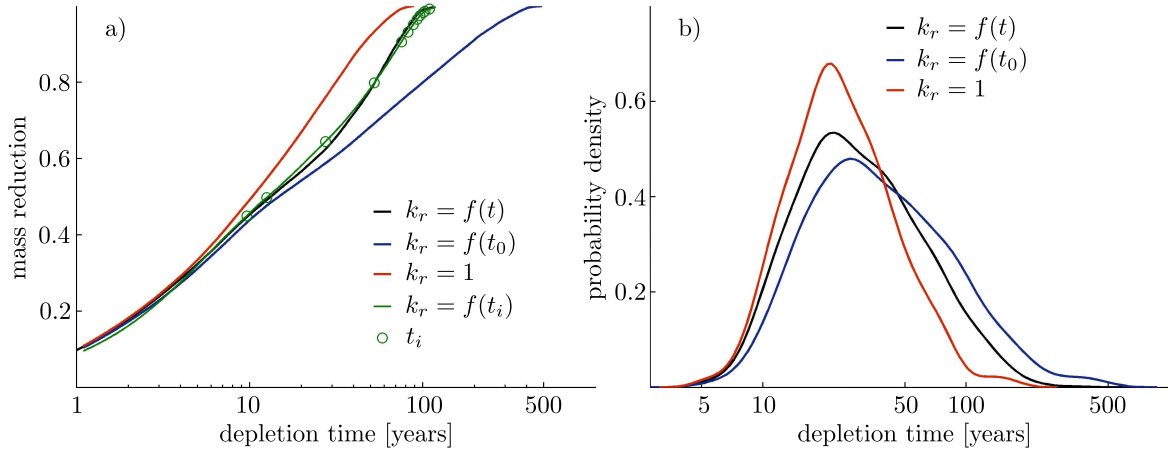


Figure 3.9: a) CSA depletion curves for various settings of the relative permeability, b) *pdf* of log-depletion time (90% mass reduction).

the initial mass discharge) with progressing time, since the aquifer heterogeneity (neglected in S4) does not change over time and because the formation of CSAs (neglected in S2-S3) is settled before depletion starts.

The depletion curve (removed mass over time) of a composite CSA (i.e., a CSA composed of both pools and ganglias) is shown in Fig. 3.9a. The different curves correspond to different simplifications for the temporal updating of relative permeabilities  $k_r(t)$ . For the sake of comparison, a reference case with a yearly updating procedure is simulated (shown by the black line in Fig. 3.9a). The longest depletion times are obtained if the relative permeability is not updated during depletion, but remains at the initial values that correspond to the initial DNAPL saturation  $S_n(\mathbf{x}, t_0)$  (blue line). The other extreme is to entirely neglect relative permeability effects and assign  $k_r = 1$  throughout the entire depletion process (red line), leading to an underestimation of depletion time. The proposed quasi-steady-state depletion simulation with intermediate post-processing steps (green line) follows the reference depletion pattern (black line) quite accurately.

This approach yields a computational speed-up of about 10, and so enables the estimation of *pdfs* for characteristic depletion times with feasible computational effort. The results are shown for 90% mass reduction as the black *pdf* in Fig. 3.9b. Again, the two extreme cases (blue and red *pdf*, as described above) lead to severe bias in the *pdf* (please observe the log-scale of the time axis). Also, they result in misspecifications of depletion time uncertainty: All three *pdfs* start similar in the low-duration range, but extend up to different high-duration extreme values. Freezing  $k_r$  at initial values overestimates uncertainty, whereas entirely neglecting relative permeability underestimates uncertainty. In combination with the biases towards higher and lower values, respectively, the  $k_r = 1$  simplification should never be used, while freezing  $k_r$  at initial values is at least a conservative simplification. However, an adequate updating scheme is advised.

The results confirm the statement that upscaled models (here: for DNAPL dissolution, depletion, and mass fluxes) go bad if they do not account for higher-order dependencies



between all individual influential variables, such as discharge area, groundwater velocity, DNAPL saturation, and CSA geometrical parameters [Zhang *et al.*, 2008; Kokkinaki *et al.*, 2013]. Simply taking the average of these parameters leads to a bias in the model output and does not reliably support the decision making process.

### 3.5.4 Local parameter statistics and interdependencies

The effects of all simplifications discussed above can be rooted back to the involved physical processes. The involved error mechanisms can be revealed and explained quite well when examining the mutual parameter dependencies of the involved spatially variable parameter fields such as DNAPL saturation  $S_n(\mathbf{x})$ , permeability  $K(\mathbf{x})$ , and groundwater velocity  $\mathbf{v}(\mathbf{x})$ .

Both Figs. 3.10a and b show a drastic interlinkage between the parameters for aquifer heterogeneity, groundwater flow, and CSA residence for Scenario S1. Fig. 3.10a shows how the *pdf* of permeability ( $K$ ) and of permeability times relative permeability ( $k_r K$ ) differ between locations where the CSA resides ( $S_n > 0$ ), and the corresponding global, unconditional *pdfs*. This is caused by the physics of CSA formation, where DNAPL is prevented from entering regions with high entry pressure, i.e., with low permeability (via equation 3.1). Fig. 3.10b shows how the systematics of DNAPL infiltration propagate onto the groundwater velocities experienced by the CSA: Although relative permeability effects reduce velocities (compare black *pdf* to red *pdf*), the effect that the CSA resides primarily in high permeability regions prevails (compare red *pdf* to blue *pdf*).

In Scenarios S2 and S3, the link between CSA and heterogeneous aquifer has been corrupted by design. Thus, the CSAs in S2 and S3 may experience any permeability and velocity value that is randomly drawn from the unconditional *pdfs*  $p(K)$  and  $p(v_s)$ , respectively. Replacing the conditional relation between  $K$  and CSA by unconditional randomness is the source of bias and misspecified uncertainty found in sections 3.5.2 and 3.5.3. Again, the introduced physical knowledge on the CSA formation process (see Sec. 3.2) systematically shifts the model output, and significantly reduces the model output uncertainty (see Fig 3.8). The resulting necessity to account for physically plausible CSA structures poses a major challenge for field investigations and for upscaling the CSA depletion process, because such non-deterministic, non-linear and complex processes on CSA formation (and the resulting CSAs) can hardly be observed in required detail or represented on larger scales.

## 3.6 Summary and conclusions

This chapter conceptualized and assessed a sound model framework to predict mass discharge and source depletion time, equipped with *pdfs* for uncertainty quantification. The adequate model formulation and the achieved uncertainty quantification are important because both quantities are relevant impact metrics for contaminated site assessments, yet subject to a high degree of prediction uncertainty. To this end, I presented a methodology that avoids commonly accepted simplifications of DNAPL dissolution and solute transport models in three steps:

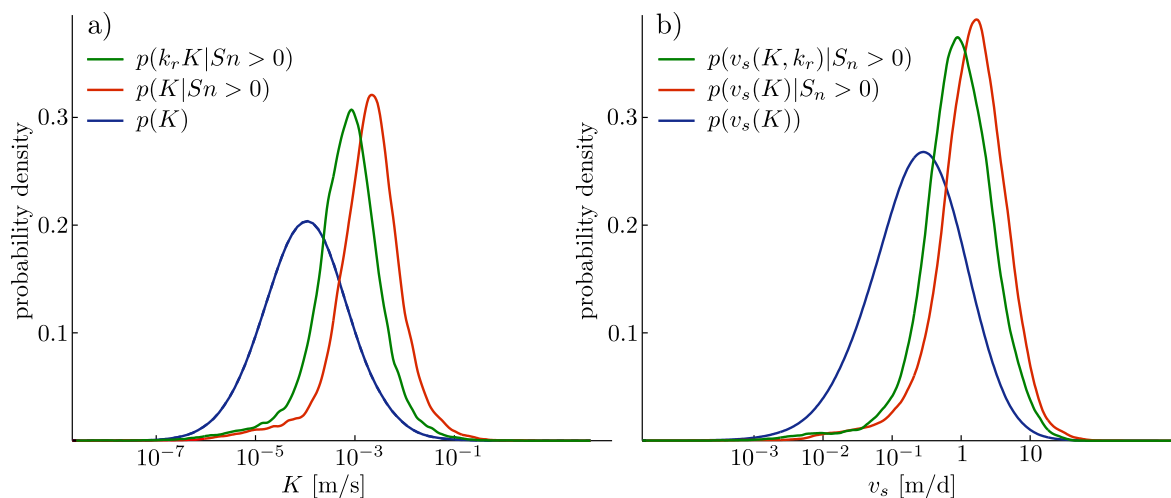


Figure 3.10: *pdfs* of log hydraulic permeability (left) and of log absolute velocity (right) conditional and unconditional to DNAPL presence at any given location.

1. aquifer parameters (here represented by permeability values) and CSAs are conceptualized as random space functions.
2. the relevant processes that control mass discharges and depletion times are conceptualized and resolved at appropriately small spatial and temporal scales and at the actual physical dimensionality.
3. these processes and parameters are coupled through their relevant interdependencies.

In other words, I avoid as much as reasonably possible simplifications like homogenization, upscaling, simplified contaminant source zones (e.g., block-shaped), equilibrium assumptions on too large scales, and one- or two-dimensional simplified conceptualizations.

With this model framework, I assessed the validity of common simplifications and the necessity for more complex models. As one crucial result, I could show that it is important to generate and implement physically-based and random contaminant source architectures (CSAs) into solute transport models. Overall, I observed substantial bias and misspecified confidence intervals in mass discharge and depletion time *pdfs* for all scenarios that introduced simplifications compared to my consistent model framework. Based on the discussion in Chap. 2 and the analyses in the current chapter, I identified a set of relevant processes and their key parameters that significantly influence the assessment of contaminated sites. In particular, the heterogeneity and uncertainty of aquifer parameters, the groundwater flow irregularity, the complex fine-scale structures and uncertainty of CSAs, and their interlinkages need to be considered. Moreover, if physical process understanding (e.g., on the CSA formation processes) is used in the solute transport model and if the relevant processes and the key parameters are coupled in a physically and statistically consistent fashion, model predictions gain accuracy and plausibility. This, in turn, leads to improved confidence for model-based decisions within contaminated site assessments. Contrary, if crucial parameter or process uncertainties are ignored or not considered appropriately, the predicted risk from

such models tends to be underestimated. Thus, decisions on the basis of simplified models for DNAPL contaminated sites should be taken with caution.

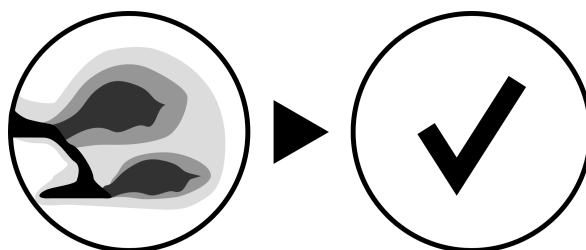
The strongest simplification I still use is a quasi-steady-state (QSS) concept, justified by the discussion of time scales and their separation in Sec. 2.4. The QSS assumes piecewise (in time) constant DNAPL dissolution rates and piecewise steady-state solute transport. Every time a substructure of the CSA has been depleted, new dissolution rates follow immediately without re-computing the groundwater velocity field. The velocity field is updated with new relative permeabilities only when significant changes in the CSA structure have been observed. The fact that relative permeabilities and groundwater velocities change continuously between these updates is accounted for by a time interpolation scheme. The QSS approach yields a computational speed up by a factor of roughly ten. The two most interesting benefits of the QSS depletion approach (due to the achieved speed-up) are the ability to obtain (1) a complete probabilistic description of characteristic source depletion times, and (2) the temporal evolution of the total emanating mass discharge. Both can be used to address a wide field of typical questions in practical applications, such as the risk assessment and prioritization of contaminated sites or the identification of contaminant source architectures, which in turn can be useful to optimize remediation strategies.

#### Key conclusions

- For reliable mass flux and source depletion time probability estimates, one needs at least to account for adequate aquifer heterogeneity and randomness which is associated with non-uniform groundwater flow fields, complexity and uncertainty of CSAs, and for adequate physical interlinkages for the governing parameters.
- CSA geometry plays a key role for characterizing mass discharge and source depletion behavior. Therefore, decisions on the basis of simplified models for DNAPL contaminated sites should be taken with caution.
- Relative permeability exerts a secondary influence on these processes compared to the influence of complex CSA geometry. However, its effect on depletion times is non-neglectable. Thus, the only acceptable simplification (out of all tested simplifications in this thesis) seems to be a quasi-steady-state approach that approximates reasonably well the time evolution of the CSA and of the related relative permeabilities.



## 4 Dirichlet and third-type boundary conditions in PTRW simulations



In Chap. 3, I introduced and assessed a novel simulation framework that couples the relevant physics and treats the relevant uncertainties in a consistent manner. The central numerical part of this framework is a novel PTRW method that allows for implementing complex-shaped Dirichlet and third-type boundary conditions within non-uniform flow fields. The novel PTRW framework is a new method for simulating dissolution and dissolved-phase transport, and it is suitable for the quasi-steady-state source depletion concept as introduced and applied in Chap. 3. Due to its central role within this thesis, it is introduced in more detail and benchmarked against well established methods (MASST and FEM) in the current chapter.

**A quick overview:** Efficient and accurate implementation of Dirichlet or third-type boundary conditions in particle tracking random walk (PTRW) simulations of advective-dispersive transport in heterogeneous flow fields is a challenge. The challenge lies in the fact that defining concentrations for these boundary conditions requires to invoke control volumes of some kind, which are not natural to the Lagrangian-based PTRW concept. The method proposed here performs a Galerkin projection of PTRW-based particle densities onto control volumes that discretize the boundary. Thus, the method obtains concentration values at the boundary condition and can control the particle release rates such that the prescribed boundary values are met. This allows for complex-shaped internal and external boundaries, where concentration values are fixed to prescribed values. Third-type boundary conditions can be addressed as well. I test and illustrate the properties and behavior of the proposed method in a series of test cases. The results are benchmarked against the conceptually related semi-analytical method MASST (multiple analytical source superposition technique) [Sale and McWhorter, 2001] and to those of a well established finite element method (FEM). While MASST is restricted to uniform velocity fields due to the underlying analytical solutions, FEM is limited in heterogeneous velocity fields at large Péclet numbers by numerical

dispersion in the feasible discretization range. The results of the benchmarks demonstrate that the proposed method performs better than the other methods in both regimes.

**Advantages of PTRW:** The advection-dispersion transport equation (ADE) is often solved via Lagrangian techniques such as particle tracking random walk (PTRW), especially in the advection-dominated regime with large Péclet numbers. The main benefits of PTRW are its computational efficiency and the absence of numerical dispersion and unphysical oscillations [Kinzelbach, 1988; LaBolle *et al.*, 1996; Salamon *et al.*, 2006]. Being a Lagrangian technique, PTRW discretizes and tracks the dissolved mass via discrete carriers (called particles), while Eulerian descriptions discretize concentrations in small elementary volumes  $dV$  that are fixed in space. PTRW works straightforwardly, if only Neumann boundary conditions appear. These are easily implemented by releasing a certain number of particles per time and volume.

**The challenge: Dirichlet boundary conditions in PTRW.** However, Neumann boundaries are an exception rather than the rule in many physical systems. In specific, the Secs. 2.3 and 3.3 discussed and demonstrated the need for Dirichlet or third-type boundary conditions when simulating the dissolution and depletion of CSAs. Dirichlet boundary conditions are rather natural in Eulerian frameworks, but are hard to implement with Lagrangian techniques. Dirichlet boundary conditions impose fixed concentration values at specific locations, defined as a mass  $dm$  within a given volume  $dV$ . In PTRW, mass  $dm$  is directly given by a number of particles  $n$  and their individual masses  $m_p$ ,  $dm = \sum_{p=1}^n m_p$ . To define Dirichlet boundary conditions in PTRW, an intermediate jump to Eulerian thinking is required, since a control volume  $dV$  of some kind has to be invoked for defining concentration values. Third-type boundary conditions are a mixture between Dirichlet and Neumann boundary conditions. They fix the flux across the boundary as a function of concentration at the boundary. Because they also rely on concentration values, they share the necessity of jumping to Eulerian thinking.

Dirichlet boundary conditions are often caused due to local chemical equilibrium by mass transfer up to solubility concentration from one phase (donor phase) into another phase (acceptor phase). At the scale of representative elementary volumes (REV), this causes a fixed-concentration condition in the acceptor phase at all locations where the two phases coexist. This region of coexistence is called *source zone* in the context of this thesis. One specific example is the dissolution of non-aqueous-phase-liquids (NAPLs) present in the pore space into the pore water as it is simulated in Chap. 3. The rate of mass transfer at the REV scale mainly depends on:

1. the solubility of the donor phase components in the acceptor phase,
2. on the current concentration of donor-phase components in the acceptor phase installed by the surrounding mass transfer rates,
3. on local advective and diffusive fluxes, and
4. accounting also for the mass transfer rate coefficient in between the phases leads to third-type boundary conditions.

The resulting source terms in the acceptor phase (i.e., the amount of released particles per volume and time in PTRW terms) are function of all these kinematic parameters, and thus vary both in time and in space.

**CSA as irregular shaped Dirichlet boundary conditions:** In realistic physical systems, the donor phase often forms a complex and possibly irregular shape in all three spatial dimensions, called CSAs in this thesis. The resulting source zone volume is unlikely to form points, lines or planes that are orthogonal to the respective local flow direction. In such realistic conditions, the particle release rate per location can not be set as a straightforward fixed-flux (Neumann) boundary condition beforehand. This is caused by mutual dependencies between the inter-phase mass transfer rates at different locations within the source zone, e.g., by diffusive and dispersive fluxes between neighboring locations, and by advective transport from upstream to downstream locations within the source zone. Sec. 2.4 discussed these so-called shadowing effects in detail. Assessing and updating the particle release rates in an iterative loop until all Dirichlet boundary conditions are met is a possible solution, but computationally cumbersome. Contrary to using an iterative loop, the presented method will run the PTRW simulation only once, using unit particle release rates from a number of REV's (called emitters that represent the CSA as an outcome of the SIP algorithm in Sec. 3.2) at all source locations. Mass release rates and concentration profiles are then obtained in a swift post-processing step that accounts for the mutual dependencies of the mass release rates by solving a linear system of equations.

## 4.1 Goal, approach and contributions

### 4.1.1 Goal and approach

The goal of this chapter is to introduce in detail and test the novel PTRW method that has been applied in Chap. 3. In particular, the accuracy of this method when solving for advective-dispersive transport in three-dimensional non-uniform flow fields is tested. This method preserves the advantages of Lagrangian techniques, but still can implement concentration-dependent boundary conditions. The ADE is entirely solved by PTRW, and the Eulerian support volume concept is only applied to control the mass release rates which are necessary to enforce the concentration values imposed by Dirichlet or third-type boundary conditions. This leads to several conceptual advantages:

1. In contrast to standard Eulerian and Lagrangian hybrid methods (ELMs, see Sec. 4.1.2), the proposed method is not alternating between Eulerian and Lagrangian maps, i.e., I avoid the operator splitting and thus prevent the propagation and accumulation of transformation errors that could lead to artifacts such as artificial numerical dissipation of peak concentration values.
2. The method is mesh-free. Thus, any boundary condition geometry can be placed freely within the entire domain.

3. The method works very robustly under a large range of regimes and parameter setups. This holds, in particular, at high Péclet numbers, where existing methods struggle with numerical dispersion, oscillations, or with the necessity of prohibitively fine spatial discretizations.

Conceptually, the method is similar to the (semi-) analytical approach of *Sale and McWhorter* [2001], called multiple analytical source superposition technique (MASST). MASST uses a convolution-based analytical ADE solution for continuous cubic source blocks. Then, it superposes the cubic blocks to approximate source zones, and finally solves a linear system of equations in order to fix the mass release rate per cube element such that the specified Dirichlet boundary condition in each cube is met. Prescribed concentration values are ensured only at a single point in the center of each cube. MASST requires a uniform flow field and constant parameters within an infinite domain in order for the analytical solution to hold. Because the ADE is solved analytically, the overall accuracy only depends on the boundary discretization by the cubic elements and on the control of concentration values at only single points per control volume.

The herein proposed approach is to discretize boundary conditions and source zones by small elementary volumes, obtain the element-wise transport solutions for superpositioning by PTRW (not analytically), measure the concentration values within the entire volume (not only at one point) of each elementary volume, and finally account for their interaction by solving the same linear system that is also used in MASST. This leads to the desired numbers of particles to be released per time in each of the elementary volumes. Contrary to MASST, the transport processes are calculated in a numerical Lagrangian framework by PTRW, and thus do not require the strong assumptions necessary for analytical solutions to hold. Also, the volume-based assessment of concentration values within the elementary volumes is superior to the point-like assessment.

The basic approach has already been touched upon Sec. 3.3. Before presenting the approach in more detail (Sec. 4.3 and Sec. 4.4), I review existing alternative methods (Sec. 4.1.2), and discuss the reconstruction of concentrations from particle distributions (Sec. 4.2). Finally, I test and illustrate the method in the context of subsurface DNAPL dissolution in a test case series. In these test cases, I compare the results of the method to those of standard numerical (FEM) and (semi-)analytical methods (MASST), show the convergence behavior, and assess the computational effort (Sec. 4.5).

#### 4.1.2 Existing alternative methods

Efficient methods to solve transport problems in three-dimensional and highly irregular flow fields with Dirichlet or third-type complex-shaped boundary conditions are lacking up to date. Standard Eulerian methods such as finite differences [*Smith*, 1985], finite volumes [*Ferziger and Perić*, 1996] or finite elements [*Zienkiewicz and Taylor*, 1977] suffer from numerical dispersion, unphysical oscillations, and convergence issues at larger Péclet numbers in the feasible discretization range of field-scale domains. Lagrangian-based approaches such as PTRW techniques are limited to Neumann boundary conditions, while approaches based on the method of characteristics are restricted to purely advective cases. To date, there is no



method that allows to implement Dirichlet or third-type boundary conditions into PTRW simulations. Yet, several hybrid methods have been developed in order to compensate for these issues, often called Eulerian-Lagrangian methods (ELMs). Based on operator splitting ideas, they split the ADE into two parts, treating the advective fluxes with the method of characteristics (Lagrangian type), while other terms such as diffusive and dispersive fluxes and boundary conditions are assessed in a nodal concentration framework (Eulerian type) [e.g., *Russell and Celia*, 2002]. Existing ELM methods include approaches based on finite differences [e.g., *Konikow et al.*, 1996], finite volumes [e.g., *Healy et al.*, 1993] and finite elements [e.g., *Douglas and Russel*, 1982; *Pironneau*, 1982; *Li et al.*, 2000].

The family of Eulerian-Lagrangian localized adjoint methods (ELLAM) [e.g., *Celia et al.*, 1990; *Heberton et al.*, 2000; *Russell and Celia*, 2002; *Binning and Celia*, 2002] is based on a specific weak form of the ADE, using specialized test functions that follow the characteristics of advection in space and time, obtained from the localized adjoint method (LAM) [*Herrera et al.*, 1993]. The developments within the ELLAM family seek to reduce numerical dispersion, guarantee mass conservation, and provide access to all desired types of boundary conditions. However, ELLAMs approximate the integrals of the weak form in terms of nodal concentration values at discrete time levels. Consequently, Dirichlet boundary conditions are naturally implemented, but the Lagrangian advantages in terms of local mass conservation and the absence of numerical dispersion and of unphysical oscillations may fade. The reason is that their Eulerian-Lagrangian operator splitting requires to transform back and forth between Eulerian and Lagrangian maps at each time step. Each transformation introduces a numerical transformation error, and these errors accumulate and propagate through time [*Russell et al.*, 2003].

## 4.2 Concentration reconstruction

Mass in PTRW is given by the particle masses. Eulerian approaches invoke so-called support or elementary volumes to define concentration. In order to reconstruct concentration values from PTRW-based particle distributions, the PTRW methods often borrow the support volume concept from the Eulerian approach. According to *Bagtzoglou et al.* [1992] and *Fernández-García and Sánchez-Vila* [2010], a variety of kernel functions can be used to reconstruct concentrations from PTRW-based particle clouds, such as box, triangular or Gaussian kernel functions. Conceptually, these kernels are assigned to each particle, equipping them with a non-zero volume. Alternatively, the same kernel can be used to define fixed-in-space support volumes to define concentration values in the Eulerian sense. While the box kernel function introduces discontinuities at the box edges, the triangular and Gaussian kernels generate continuous and infinitely smooth concentration distributions, respectively.

**Mathematical formulation of the kernel-based concentration reconstruction:** In general, using a yet unspecified kernel function  $K(\mathbf{x})$ , the concentrations can be reconstructed with:

$$c(\mathbf{x}, t) = \sum_{p=1}^n V_p^{-1} m_p K_p(\mathbf{x} - \mathbf{x}_p(t)), \quad (4.1)$$

where  $\mathbf{x}_p(t)$  are the time-dependent particle locations from PTRW,  $m_p$  is the mass of the  $p^{\text{th}}$  particle and  $V_p$  is the particle's support volume, as given by the kernel function:

$$V_p = \int_{\Omega} \phi(\mathbf{x}) K_p(\mathbf{x}) d\mathbf{x}. \quad (4.2)$$

The most common situation assumes that the underlying flow field (e.g., of the acceptor phase) is at steady state, and that the dissolution is either an instantaneous event, followed by transient transport, or that it is a continuous release, followed by steady-state transport in the long-time limit. A common technique to obtain continuous-release results from instantaneous-release PTRW simulations is to convolve the instantaneous-release results in time, see *Robinson et al.* [2010]. In this context, the particles are viewed as mass release rates and hence denoted by  $\dot{m}_p$  [ $\text{MT}^{-1}$ ] (i.e., particle mass per time step). Thus, Eq. (4.1) modifies to

$$c(\mathbf{x}, \tau) = \int_0^{\tau} \sum_{p=1}^n V_p^{-1} \dot{m}_p K_p(\mathbf{x} - \mathbf{x}_p(t)) dt, \quad (4.3)$$

where the limit for  $\tau \rightarrow \infty$  yields the long-term steady state, and the steady state  $c(\mathbf{x}, \tau \rightarrow \infty)$  is denoted by  $c(\mathbf{x})$  for short. For the sake of brevity but without loss of generality, I will consider only the late-time steady-state case in the following. For transient transport cases, all the following equations would simply inherit the time dependence of Eqs. (4.1) or (4.3), but otherwise remain identical.

**Choosing the kernel:** If a Gaussian kernel is applied, then  $K_p(\mathbf{x}) = \exp(-\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x})$ , and the support volume becomes:

$$V_p = \int_{\Omega} K_p(\mathbf{x}) d\mathbf{x} = \phi \sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}, \quad (4.4)$$

where  $d$  is the physical dimensionality and  $\boldsymbol{\Sigma}$  is a (typically diagonal) covariance matrix with the entries  $\Sigma_{i,i} = H_i^2$ , where  $H_i$  is the kernel bandwidth in direction  $x_i$  that scales the Gaussian kernel. From the broad range of applications for kernel density estimation in mathematical statistics and information theory [e.g., *Scott*, 1992], it is well known that there is no clearly optimal choice of kernel functions and their width. Instead, it is a problem-specific compromise between (1) overestimating the smoothness of the underlying function and (2) obtaining noisy results if the kernel bandwidth is too small. This is discussed for PTRW applications by *Fernández-García and Sánchez-Vila* [2010] and in more general terms by *Silverman* [1986], *Scott* [1992], and *Härdle et al.* [2004].

A common selection rule of an optimal bandwidth  $H^{opt}$  is derived by minimizing the asymptotic error of the resulting density estimate, while assuming that the underlying probability density function belongs to the family of normal distributions. This leads to the simplified reference rule  $H_i = (4/(d+2))^{1/(d+4)} \tilde{\sigma}_i n_p^{-1/(d+4)}$ , where  $d$  is the dimensionality,  $n_p$  is the sample size (here: the number of particles), and  $\tilde{\sigma}_i$  is a data-based estimate of the  $i^{th}$  marginal standard deviation [e.g., Härdle *et al.*, 2004]. However, the relevant literature provides a wide range of alternative bandwidth selection techniques. The suitability of specific kernel functions for the presented method is discussed in Sec. 4.3.

### 4.3 Discretizing the boundary condition

**The twofold role of kernels in the proposed method:** Basically, there are two volumetric objects that need to be considered if Dirichlet boundary conditions are implemented in PTRW: (1) volumes  $V_k$  in which particles are released at individual mass release rates  $dm_k$ , and (2) Eulerian elementary volumes  $V_l$  for assessing and controlling the resulting concentration values at the boundaries. Both  $V_k$  and  $V_l$  are used to discretize the volume  $\Gamma$  that participates in the boundary condition and are implicitly included in Eq. (3.2). For the sake of consistency, the particle release procedure is attached to the same support volumes as the boundary condition control procedure, so that  $V_k \equiv V_l$ . In the following, these volumes are called *emitters*, in accordance with their role as volumes where particles are being released. Due to the symmetry of the applied kernels,  $K_p(\mathbf{x}, \mathbf{x}_p) = K_p(\mathbf{x}_p, \mathbf{x})$ , and because all particles have the same kernel, one may conceptually transfer the kernel in Eq. (4.1) from the particles to the emitter centers and model the particles as points instead, without affecting the result. This suggests to represent the *emitters*  $V_k$  by kernel functions  $K_k(\mathbf{x}, \mathbf{x}_k)$ , representing a Eulerian definition of boundary conditions.

Choosing (weighted) volumes for emission and control bears a similarity to Galerkin methods in finite element approaches (e.g., [Zienkiewicz and Taylor, 1977; Helmig, 1997]). The kernels  $K_k$  and  $K_l$  that define the emitter and control volumes  $V_k$  and  $V_l$  relate to Ansatz and test functions, respectively. Please note that, although Ansatz and test functions are involved, this does not imply that only a weak form of the ADE is solved. The proposed method applies the Galerkin procedure with its Eulerian character only to discretize the boundary condition and the corresponding emission rates  $dm_k$ . All transport-related terms in the ADE are solved in the Lagrangian way by standard PTRW techniques.

**Optimal kernel bandwidth in the current context:** Hence, the key step of this method is a proper discretization of the boundary condition. Various kernel functions such as triangular, box, or Gaussian functions could be used for this purpose [Bagtzoglou *et al.*, 1992; Fernández-García and Sánchez-Vila, 2010]. As discussed in Sec. 4.2, their suitability is problem-dependent. For example, Gaussian functions relate to the homogeneous solution of advective-dispersive transport and suitably approximate smooth geometries of the boundary condition. These benefits can turn into disadvantages if the problem under consideration

suggests a very sharp outline of the source zone and, at the same time, there are neither diffusive nor dispersive fluxes. In this case, using a box function may be more beneficial. Box functions are flush-mounted within the source zone, i.e., they cover the entire source volume without gaps or overlapping, and add up to unity within the entire source volume when interpreted as weighting functions. Gaussian functions are smooth, weighted windows that have to overlap widely in order to properly cover the whole volume and in order to approximately add up to a spatially constant value. In addition to the optimal bandwidth selection for kernel density estimators (see Sec. 4.2), the length-scale variables for the volumes  $V_k$  and  $V_l$  possess a physical foundation because they discretize the boundary condition.

Intuitively, the Gaussian kernel bandwidth should be chosen in a range such that its support volume according to Eq. (4.4) is roughly identical to the support volume of flush-mounted box-shaped kernels. Defining a box function with dimensions  $\Delta x_i$  per dimension  $i = 1, \dots, d$ , this intuitive choice demands that

$$\sqrt{(2\pi)^d |\Sigma|} = \prod_{i=1}^d \Delta x_i. \quad (4.5)$$

Thus, one possible choice of bandwidth is:

$$\Sigma_{i,i} = (2\pi)^{-1} \Delta x_i^2. \quad (4.6)$$

This choice of  $\Sigma_{i,i}$  ensures that  $\int K_G(x) dx = \int K_{\text{box}}(x) dx$ , where  $K_G$  denotes the Gaussian kernel and  $K_{\text{box}}$  denotes the box kernel. However, when introducing the mathematical formulation for this method in the remainder of this section and in Sec. 4.4, the relevant term  $\int K(\mathbf{x})^2 dx$  will also appear in Eqs. (4.11), (4.13), and (4.17), where the squared kernel function has an effective volume of:

$$\int_{\Omega} K_G^2(\mathbf{x}) d\mathbf{x} = \sqrt{(\pi)^d |\Sigma|}. \quad (4.7)$$

Thus, it would be equally plausible to desire  $\int K_G(\mathbf{x})^2 dx = \int K_{\text{box}}(\mathbf{x})^2 dx$ . Hence, I introduce a bandwidth scaling factor  $a^2$  for the Gaussian kernel, demand  $\sqrt{(\pi)^d |a^2 \Sigma|} = \prod_{i=1}^d \Delta x_i$ , and obtain that Eq. (4.5) corresponds to  $a^2 = 1$ , whereas  $a^2 = 2$  ensures  $\int K_G(\mathbf{x})^2 dx = \int K_{\text{box}}(\mathbf{x})^2 dx$ . For the choice of Kernel spacing and Kernel scaling  $a^2$ , three conditions should be considered in accordance with Sec. 4.2:

1. a good compromise between avoiding oversmoothing and bias for defining concentration values ( $H_i \approx H_i^{\text{opt}}$ ),
2. a sufficient overlapping between adjacent emitters to discretize the emitting volume without holes in between the individual emitters, and
3. an appropriate physical representation of the boundary condition outline and geometry.

That is, if the first and third requirement demand a smaller  $a^2$  and hence smaller kernels, the second requirement demands increasing the number of emitters at smaller distances in between. The influence of Kernel bandwidth (scaled by  $a^2$ ) and more effects of the kernel choice on the solution quality are also discussed in Sec. 4.5.

**Boundary discretization approach:** The overall PTRW solution  $c(\mathbf{x})$  will be built from the  $k$  individual emitter-wise PTRW simulations  $c_k(\mathbf{x})$  for unit-rate particle releases in the volumes  $V_k$  (see Eq. (3.2) in Sec. 3.3), using the emitter kernels  $K_k$  as density functions for the particle injection, and the yet unknown weighting factors  $\varepsilon_k$  for each emitter:

$$c(\mathbf{x}) = \sum_{k=1}^{n_k} \varepsilon_k c_k(\mathbf{x}). \quad (4.8)$$

In Eq. (4.8), the functions  $c_k(\mathbf{x})$  can be seen as transport-related Ansatz functions. These are not to be confused with the boundary-related Ansatz functions, i.e., with the Kernel functions  $K_k$ . The yet unknown emission rates  $\varepsilon_k$ , analogous to Eq. (3.2), rescale and weight the transport Ansatz functions  $c_k(\mathbf{x})$  to meet the concentration values  $\hat{c}(\mathbf{x})$  prescribed by the Dirichlet boundary condition. This requires a number of steps, described in the remainder of this section and in Sec. 4.4.

**The boundary-related mass matrix:** The boundary condition values  $\hat{c}(\mathbf{x})$  have to be approximated by  $\tilde{c}(\mathbf{x})$ , using the Kernel functions  $K_k$  as boundary-related Ansatz functions with the coefficients  $\tilde{c}_k$ :

$$\tilde{c}(\mathbf{x}) = \sum_{k=1}^{n_k} \tilde{c}_k K_k(\mathbf{x} - \mathbf{x}_k). \quad (4.9)$$

In order to obtain the boundary condition coefficients  $\tilde{c}_k$  in Eq. (4.9), the residual  $(\tilde{c}(\mathbf{x}) - \hat{c}(\mathbf{x}))$  between discretized and prescribed boundary concentrations is projected onto the test functions  $K_l$  and inserting Eq. (4.9) leads to:

$$\int_{-\infty}^{\infty} \left( \sum_{k=1}^{n_k} \tilde{c}_k K_k(\mathbf{x} - \mathbf{x}_k) - \hat{c}(\mathbf{x}) \right) K_l(\mathbf{x} - \mathbf{x}_l) dx = 0 \quad \forall l, \quad (4.10)$$

Now reformulating Eq. (4.10) to

$$\int_{-\infty}^{\infty} \hat{c}(\mathbf{x}) K_l(\mathbf{x} - \mathbf{x}_l) dx = \sum_{k=1}^{n_k} \tilde{c}_k \int_{-\infty}^{\infty} K_k(\mathbf{x} - \mathbf{x}_k) K_l(\mathbf{x} - \mathbf{x}_l) dx, \quad (4.11)$$

and solving for the vector  $\tilde{c}_k$  of boundary condition coefficients leads to Eq. (4.12). In a more convenient matrix notation, the solution to Eq. (4.11) can be written as:

$$\tilde{\mathbf{c}} = \mathbf{M}^{-1} \hat{\mathbf{c}}^*, \quad (4.12)$$

where  $\mathbf{M}$  is a boundary-related mass matrix with elements

$$M_{k,l} = \int_{-\infty}^{\infty} K_k(\mathbf{x} - \mathbf{x}_k) K_l(\mathbf{x} - \mathbf{x}_l) dx, \quad (4.13)$$

and  $\hat{\mathbf{c}}^*$  is a right-hand side vector with elements

$$\hat{c}_l^* = \int_{-\infty}^{\infty} \hat{c}(\mathbf{x}) K_l(\mathbf{x} - \mathbf{x}_l) dx. \quad (4.14)$$

Thus, the vector  $\hat{\mathbf{c}}^*$  represents the boundary conditions mapped onto the emitters. If  $K_k$  and  $K_l$  are Gaussian kernels, the mass matrix  $M_{k,l}$  according to Eq. (4.13) can be computed analytically as

$$M_{k,l} = \exp \left[ -\frac{1}{4} (\mathbf{x}_k - \mathbf{x}_l)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_k - \mathbf{x}_l) \right] \sqrt{(\pi)^d |\boldsymbol{\Sigma}|}. \quad (4.15)$$

The term of *mass matrix* is borrowed from the finite element method, where the matrix of inner products between all Ansatz and test functions bears the same name and is in fact defined as in Eq. (4.13). In the current context, the mass matrix  $\mathbf{M}$  quantifies the degree of overlapping between all possible pairs of emitter and control kernels  $K_k$  and  $K_l$ . Thus, if  $K_k$  and  $K_l$  are flush-mounted functions such as the box functions,  $M_{k,l}$  is the identity matrix times the emitter volume  $V_k$ .

## 4.4 Controlling the boundary condition

**The transport-related mass matrix:** Now, repeating the Galerkin projection, but this time applying it to the residuum  $(\hat{c}(\mathbf{x}) - c(\mathbf{x}))$  between prescribed boundary conditions  $\hat{c}(\mathbf{x})$  and the desired PTRW solution  $c(\mathbf{x})$ , controls the boundary condition. This enforces that the targeted overall PTRW solution  $c(\mathbf{x})$  according to Eq. (4.8) meets the boundary condition  $\hat{c}(\mathbf{x})$ :

$$\int_{-\infty}^{\infty} (c(\mathbf{x}) - \hat{c}(\mathbf{x})) K_l(\mathbf{x} - \mathbf{x}_l) dx = 0. \quad (4.16)$$

Substituting the prescribed boundary concentration  $\hat{c}(\mathbf{x})$  by its approximation  $\tilde{c}(\mathbf{x})$  from Eq. (4.9) and expressing the overall PTRW solution  $c(\mathbf{x})$  by Eq. (4.8) leads to

$$\sum_{k=1}^{n_k} \varepsilon_k \int_{-\infty}^{\infty} c_k(\mathbf{x}) K_l(\mathbf{x} - \mathbf{x}_l) dx = \sum_{k=1}^{n_k} \tilde{c}_k \int_{-\infty}^{\infty} K_k(\mathbf{x} - \mathbf{x}_k) K_l(\mathbf{x} - \mathbf{x}_l) dx, \quad (4.17)$$

or

$$\sum_{k=1}^{n_k} \varepsilon_k S_{k,l} = \sum_{k=1}^{n_k} \tilde{c}_k M_{k,l}, \quad (4.18)$$

where  $M_{k,l}$  is the boundary-related mass matrix from Eq. (4.13) and  $S_{k,l}$  is a new transport-related mass matrix with elements:

$$S_{k,l} = \int_{-\infty}^{\infty} c_k(\mathbf{x}) K_l(\mathbf{x} - \mathbf{x}_l) dx. \quad (4.19)$$

The matrix  $\mathbf{S}$  in the above equation resembles the inner products between all boundary-control kernels  $K_l$  and all unit-release PTRW-based transport Ansatz functions  $c_k(\mathbf{x})$  for the overall solution to Eq. (2.41) in Sec. 2.6.3. Thus,  $\mathbf{S}$  reflects the mutual dependencies (shadowing effect) through advective-dispersive transport between all emitters that discretize the Dirichlet boundary condition. Its entries  $S_{k,l}$  are calculated via Monte Carlo integration using the particle locations  $\mathbf{x}_{p,k}$  from the  $k^{\text{th}}$  unit-release PTRW simulation to approximate the density  $c_k(\mathbf{x})$ , with  $m_p = n_p^{-1}$  and  $V_p = \phi^{-1}$  analogous to Eq. (4.1):

$$S_{k,l} \approx \phi^{-1} n_p^{-1} \sum_{p=1}^{n_p} K_l(\mathbf{x}_{p,k} - \mathbf{x}_l). \quad (4.20)$$

**Modified mass matrix:** Similar to MASST, the presented method assumes a prescribed spatial pattern of mass release density within each support volume  $V_k$ , which would not necessarily conform to an assumption of constant concentration values within these support volumes. In fact, the concentration value within each emitter builds up gradually from upstream to downstream regions within each emitter. For example, in the case of uniform mass release by box-shaped emitters, a linear increase in concentration can be observed within the box. More generally, if concentration values are assessed in the midpoint of any emitter kernel with symmetric test and Ansatz functions (as in MASST), the midpoint concentration is the arithmetic mean between the concentration upstream and downstream of the emitter volume. The same is true when assessing concentration values as integral average over each support volume  $V_l$  (as in the proposed method). Without further correction, this would lead to twice the necessary release rate at emitters that are solely effected by their own mass release rate. This fact requires a modification of  $\mathbf{S}$  given by:

$$T_{k,l} = S_{k,l} + \delta_{k,l} S_{k,l}. \quad (4.21)$$

Here,  $\delta_{k,l}$  is the Kronecker delta and  $T_{k,l}$  substitutes  $S_{k,l}$  in Eq. (4.18) and in all following calculations. Overall, this correction doubles the values of all elements on the diagonal of  $\mathbf{S}$ .

Finally, I solve for the emission rates  $\varepsilon_k \dot{m}_0$ . To this end, I write Eq. (4.18) in matrix notation and rearrange:

$$\varepsilon \dot{\mathbf{m}}_0 = \mathbf{T}^{-1} \mathbf{M} \tilde{\mathbf{c}} = \mathbf{T}^{-1} \hat{\mathbf{c}}^*, \quad (4.22)$$

where  $\varepsilon$  is the solution vector that rescales the unit mass release rates  $\dot{\mathbf{m}}_0$  to the individual emission rates of each emitter. The boundary-related mass matrix  $\mathbf{M}$  cancels out after

inserting Eq. (4.12). The latter does not imply that mass conservation is canceled out of consideration, but only that the mutual overlap between the emission kernels as quantified by  $\mathbf{M}$  cancels out between the discretization of the boundary condition and the projection of the overall PTRW solution in Eq. (4.17). The transport-related mass matrix  $\mathbf{T}$  is the remaining representative of overlap and interaction (shadowing effect) between all emitter kernels and all their individual unit-release PTRW solutions.

**Overall solution:** If local chemical equilibrium is assumed for the mass transfer such that a given solubility concentration  $c_s$  is reached within each emitter, Eq. (4.14) simplifies to:

$$\hat{c}_l^* = c_s V_l, \quad (4.23)$$

with  $V_l$  according to Eq. (4.2). Then, the emission coefficients  $\varepsilon_k$  become

$$\varepsilon \dot{\mathbf{m}}_0 = \mathbf{T}^{-1} \mathbf{w}; \quad w_k = c_s V_k. \quad (4.24)$$

If local chemical equilibrium is not an appropriate assumption, emitter-specific and rate-limited mass transfer models which are defined at the scale of the emitter kernels  $K_k$  have to be included in Eq. (4.22). I focus here on the first-order mass transfer model as it arises from the diffusive stagnant-film model [Levich, 1962; Sherwood *et al.*, 1975]:

$$\frac{\dot{m}}{dV} = (c - \hat{c}_s) \kappa_{\text{eff}}, \quad (4.25)$$

where  $\kappa_{\text{eff}}$  is the effective first-order mass transfer coefficient. Analogous to the MASST approach, this modifies Eq. (4.22) to:

$$\varepsilon \dot{\mathbf{m}}_0 = (\mathbf{T} + \boldsymbol{\kappa}^{-1})^{-1} \hat{\mathbf{c}}^*, \quad (4.26)$$

where  $\boldsymbol{\kappa}$  denotes a diagonal matrix with the emitter-specific rate coefficients  $\kappa_k$  on its diagonal. Eq. (4.26) expresses how the mass release rates are restricted by  $\boldsymbol{\kappa}$ . In fact, Eq. (4.26) resembles third-type boundary conditions. The rate limitation through Eq. (4.25) vanishes and Eq. (4.26) returns to the Dirichlet boundary condition in Eq. (4.22) when the elements of  $\boldsymbol{\kappa}^{-1}$  become much smaller than the elements on the diagonal of  $\mathbf{T}$ , i.e., if the mass transfer is sufficiently fast to always install local equilibrium.

Once the emission rates are obtained from Eqs. (4.22) or (4.26), the final PTRW solution of the ADE is reconstructed by a quick post-processing step according to Eq. (4.8). Furthermore, if one is interested in the depletion time of the donor phase at any emitter location  $k$ , one can use  $d\dot{m}_k = \dot{m}_0 \varepsilon_k$  as the sink term in the respective donor-phase mass balance of each emitter volume  $V_k$  (see Eq. (2.45)). If the emitter-wise mass balance indicates that the donor phase has been depleted at the  $k^{\text{th}}$  location, the  $k^{\text{th}}$  row and column are simply erased from all equations, and the new PTRW result follows immediately. Only if the depletion has non-negligible effects on the flow field (e.g., through relative permeability changes), the unit-release PTRW simulations have to be updated. The total mass transfer rate  $\dot{m}_{\text{tot}}$  from the



donor to the acceptor phase (i.e. the total dissolved mass flux within the acceptor phase) is given through Eq. (3.2) by summing up all emission rates:

$$\dot{m}_{tot} = \sum_{k=1}^{n_k} \dot{m}_0 \varepsilon_k. \quad (4.27)$$

The accuracy of the emission rates calculated for each emitter increases with finer discretization of the Dirichlet boundary condition. Obviously, the mass release rates obtained by solving Eqs. (4.22) and (4.26) can vary in space and time. To account for variable mass release rates between the emitters in Eq. (4.8), one can either adjust the particle masses for each emitter by the corresponding value of  $\varepsilon_k$ , or change the number of released particles for each Ansatz function. However, the mass for each individual particle during its transport is always constant. This ensures the mass conservation property of PTRW as required for conservative transport, and would be different if non-conservative solutes were tracked.

## 4.5 Benchmarking

Some of the major benefits of Lagrangian techniques compared to Eulerian techniques are the absence of numerical diffusion, the inherent local conservation of mass, the robust convergence and computational efficiency [LaBolle *et al.*, 1996; Salamon *et al.*, 2006]. The herein presented method, however, may forfeit some of these benefits since a translation between Lagrangian and Eulerian technique is required for discretizing the boundary condition (see Secs. 4.3 and 4.4). Testing this is the goal of the current section.

Non-appropriate discretization of the boundary condition may induce inaccuracy due to improper representation of the source-zone geometry and due to the non-local mass conservation properties of the Galerkin projection. Together, this can induce artificial smearing, leading to numerical dilution effects. A second aspect is that the accuracy and computational time of any PTRW technique strongly depends on the mass discretization by particles (i.e., on the number of particles) and on the temporal discretization for tracking the particles through space. For the second aspect, I suggest the following references: Kinzelbach [1988], Bagtzoglou *et al.* [1992], and Hassan and Mohamed [2003]. Here, I assess the accuracy of approximating the source term geometry by looking at total mass fluxes, the dilution state of the simulated plume downstream of the boundary condition and the  $L_2$ -error of reconstructed concentration profiles.

In the following, the accuracy of the presented method is benchmarked via comparing my results to results of established numerical, analytical, and semi-analytical methods. This benchmark includes a streamline-upwinding Petrov Galerkin finite element method (FEM) [Zienkiewicz and Taylor, 1977; Brooks and Hughes, 1982] in an implementation as described in Nowak *et al.* [2008]. It uses tri-linear Ansatz functions to discretize the solution. The considered semi-analytical method is MASST as described in Sale and McWhorter [2001]. The benchmark will investigate the convergence behavior of all three methods in the three aspects mentioned above against discretization level and computational time. For the sake of

discussion, I implemented several versions of this PTRW method, varying in the shape of test and Ansatz functions (Gauss, box, and Dirac). I compare  $\text{PTRW}_G$  (both Ansatz and test functions are Gaussian),  $\text{PTRW}_{\text{box}}$  (both are box functions), and  $\text{PTRW}_{\text{box}}^c$  (one box and one point-like Dirac-function similar to MASST).

It is well known that the results of both FEM and MASST are only reliable in very specific conditions. While MASST requires a uniform flow field, a constant diffusion/dispersion coefficient and an infinite domain, FEMs are strongly dependent on the discretization, which may be cumbersome for realistic three-dimensional simulations in heterogeneous aquifers with large Péclet numbers. Hence, I set up three different test cases: one where a straightforward analytical solution exists, one where MASST can be applied, and one where only FEM can be applied.

All test cases consider a cubic source zone whose faces are either orthogonal (inlet and outlet) or parallel (top, bottom, left and right) to the uniform flow field or to the average flow direction in heterogeneous aquifers. The (mean) flow direction is defined as  $x_1$ . The source-zone volume is  $0.125 \text{ m}^3$  (0.5 m edge length in all three directions) and its center is placed at  $\mathbf{x} = [2.25, 2, 2]$ , within a rectangular domain. The rectangular domain has 12 m edge length in  $x_1$  and 4 m edge length in the two transversal directions  $x_2$  and  $x_3$ .

The benchmarks comprises the following test cases:

**In test case 1** (see Sec. 4.5.1) I consider purely advective transport such that I can analytically calculate the exact mass flux that has to result from the cubic Dirichlet boundary condition as the simple product of the total volumetric flux through the cube and the prescribed concentration value.

**In test case 2** (see Sec. 4.5.2) diffusion and transverse as well as longitudinal hydromechanic dispersion are considered. Here, the advection-based analytical solution is invalid, but the MASST method still applies.

**In test case 3** (see Sec. 4.5.3) I switch from uniform flow fields to the more realistic case of heterogeneous permeability fields that result in irregular streamlines. In this case, the assumptions behind MASST do not hold, and only FEM can serve for benchmarking.

#### 4.5.1 Validation with analytical solution (test case 1)

The obvious solution of this test case is given by:  $\dot{m}_{tot} = A c_s q_x = 1 \text{ kg/d}$ , where  $A = 0.25 \text{ m}^2$  is the area of the source cube perpendicular to the uniform Darcy velocity  $q_x = 4 \text{ m/d}$ , and  $c_s = 1 \text{ kg/m}^3$  is the assumed solubility concentration enforced by the Dirichlet condition. In this case, all possible sources of errors vanish for the semi-analytical method MASST. The same holds for the PTRW solution if box kernel functions are applied for both Ansatz and test functions ( $\text{PTRW}_{\text{box}}$ ). However, the resulting mass flux is consistently overestimated if Gaussian kernels are used as Ansatz and test functions ( $\text{PTRW}_G$ ). Although the mass flux converges towards the correct value for finer discretization by more (and hence also smaller) emitters as shown in Fig. 4.1, the smooth shape of the Gaussian function apparently does not suit well for this example with a sharp source and sharp solutions in the

Table 4.1: Parameter values used in the test cases (Sec. 4.5).

Parameters		Test 1	Test 2	Test 3
<b>Flow</b>				
Permeability $\exp(\bar{Y})$	[m/s]	$10^{-3}$	$10^{-3}$	$10^{-3}$
Hydraulic gradient $\nabla h$	[-]	0.47	$1.7 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
Darcy velocity (average) $\bar{q}$	[m/d]	4	0.075	0.15
<b>Transport</b>				
Effective porosity $\phi$	[-]	0.3	0.3	0.3
Molecular diffusion $D_m$	[m <sup>2</sup> /s]	0	$10^{-9}$	$10^{-9}$
Longitudinal dispersivity $\alpha_L$	[m]	0	$1 \cdot 10^{-1}$	$2.5 \cdot 10^{-2}$
Transversal dispersivity $\alpha_T$	[m]	0	$2 \cdot 10^{-2}$	$5 \cdot 10^{-3}$
Boundary domain $\Gamma$	[m <sup>3</sup> ]	0.125	0.125	0.125
<b>Geostatistics</b>				
Variance $\sigma_Y^2$	[-]	-	-	1 and 4
Correlation length $\ell_Y$	[m]	-	-	[0.4 0.4 0.4]

purely advective case. In other words, applying the Gaussian Ansatz functions is equivalent to assuming rather smooth source zones compared to the sharp edges of the perfect cube. From this point of view, each PTRW<sub>G</sub> result in Fig. 4.1 is correct, but referring to a different smoothness assumption for the source zone. Which level of smoothness is suitable in real applications is a problem-specific question. Here, I set the scale parameter  $a^2$  in Eq. (4.5) to the relatively small value of  $a^2 = 1/2$  in order to get an appropriate approximation of the sharp cubic source geometry. Larger values of  $a^2$  lead to a smoother source zone approximation with even larger resulting fluxes, while smaller values of  $a^2$  would underestimate the mass transfer due to insufficient overlapping between the kernels (when keeping the spacing between the emitters constant).

#### 4.5.2 Benchmark against MASST (test case 2)

In the second test case, the solutes are allowed to disperse by molecular diffusion with  $D_m = 10^{-9} \text{m}^2/\text{s}$  and transverse and longitudinal dispersion denoted by  $\alpha_T$  and  $\alpha_L$ , respectively. As a relatively common assumption, I assume  $\alpha_L = 5\alpha_T$  and set  $\alpha_T = 2 \text{ cm}$ . The second test case is set up such that good performance of the FEM is ensured. Thus, I install dispersion values that can lead to favorable values of the grid Péclet number ( $Pe_g = \Delta x/\alpha_T$ ), which is required to avoid oscillations in FEM-based solutions [Brooks and Hughes, 1982; Daus et al., 1985]. Please note that  $\alpha_T = 2 \text{ cm}$  by far exceeds the typical range of a few millimeters or even less that would be physical for the hydromechanic dispersion in porous media which is the one actually relevant for describing dilution and dissolution rates [Cirpka et al., 2006; Benekos et al., 2006, and references therein]. Just for that very reason, the proposed method (see also Sec. 2.4 and Chap. 3) avoids the use of macroscopic dispersivities. Thus, the performance of FEM would be worse compared to what is shown here, when using FEM

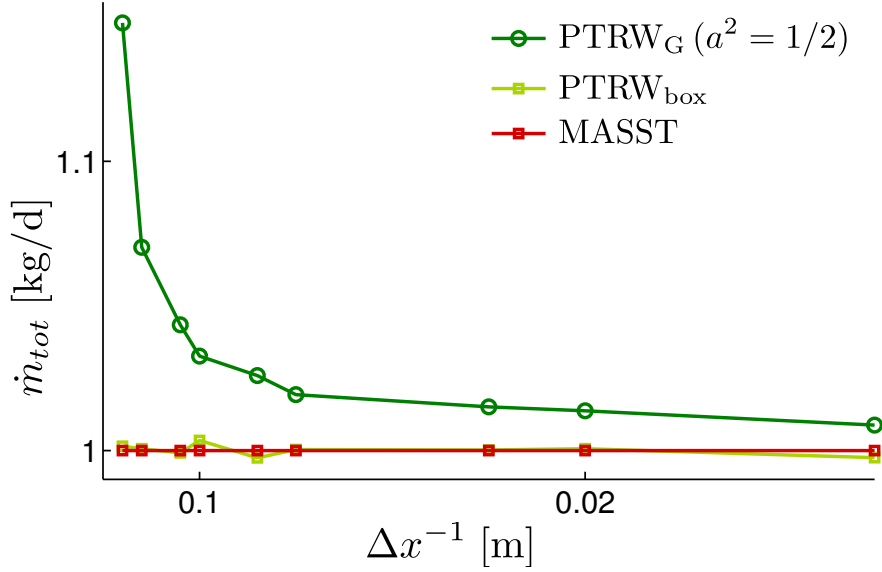


Figure 4.1: Convergence analysis (test case 1). Total mass flux emanating from a cubic source aligned with the uniform flow against various discretization levels for three different methods. The theoretical correct value is  $\dot{m}_{tot} = 1$  kg/d.

under realistic conditions in fine-scale simulations. In the following, all discretization levels will be expressed as inverse grid Péclet numbers ( $Pe_g^{-1} = \alpha_T/\Delta x$ ), where large numbers relate to fine discretization.

**Quality metrics used for benchmarking:** Besides the total mass fluxes, I calculate the dilution index ( $E$ ), assess the computational time, and evaluate the normalized root mean square error ( $NRMSE$ ) for all participating methods at various discretization levels, plotted in Fig. 4.2a, b, c, and d, respectively.

**The dilution index** was introduced by *Kitanidis* [1994] in order to quantify the degree of dilution in advective-dispersive transport. It is defined by:

$$E = - \exp \left( \int_A p(\mathbf{x}) \log(p(\mathbf{x})) dA \right), \quad (4.28)$$

where  $p(\mathbf{x})$  is the density distribution function of the mass (i.e., the concentration field normalized to unit mass) and  $A$  is the cross-sectional area. The  $E$  can be physically interpreted as effective volume (or area) covered by the plume [*Kitanidis*, 1994].

**The NRMSE** is defined as the  $L_2$  norm of the error between any tested result against a reference result, divided by the  $L_2$  norm of the reference:

$$NRMSE = \sqrt{\frac{\sum (c_i - \hat{c}_i)^2}{\sum \hat{c}_i^2}}. \quad (4.29)$$

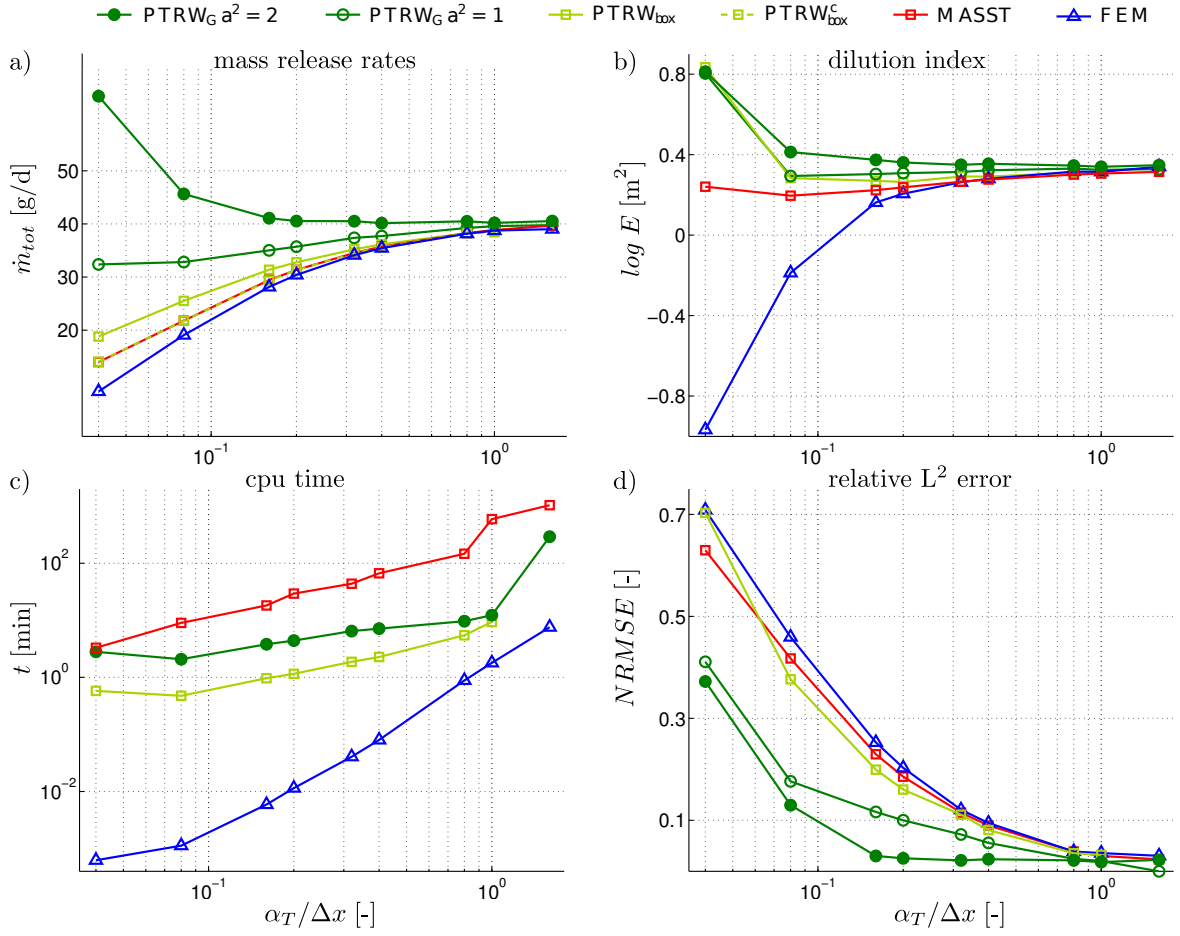


Figure 4.2: Convergence analysis and benchmarking of the new PTRW method by comparison to the semi-analytical MASST method and to FEM (test case 2). Top left (a): total mass release rate; top right (b): dilution index; bottom left (c): CPU time; bottom right (d): normalized root mean square error

The choice of the reference result will be discussed later. The  $E_s$  and  $NRMSE_s$  are obtained at the cross section in the  $x_2, x_3$  - plane 1 m downstream from the source zone inlet, i.e., at  $x_1 = 3m$ . Hence, these two metrics quantify both the initial smoothing or error at the source zone as caused by discretization, and the additional smoothing or error due to numerical dispersion and concentration reconstruction.

**Overall convergence:** The most evident result of this test case is that the PTRW results show the best convergence behavior in all three metrics, as shown in Fig. 4.2a, b, and d. All methods converge towards the same value for both total mass flux and dilution index (see Fig. 4.2a and b). Also, the  $NRMSE$  decreases to an acceptable range at finer discretization for all methods. However, results of FEM, MASST, and PTRW<sub>box</sub> converged surprisingly slow even in this simple scenario, while PTRW<sub>G</sub> converges much faster.

In Fig. 4.2, I show two curves for  $\text{PTRW}_G$  with two different values for the bandwidth scaling factor  $a^2$  ( $a^2 = 1$  and  $a^2 = 2$ ), that relate to the flush-mounted kernels as discussed in Sec. 4.3. For both values of  $a^2$ , I observe good convergence behavior in all three aspects (Fig 4.2a, b, and d). If  $a^2 = 2$ ,  $\text{PTRW}_G$  significantly smooths the cubic source at lower discretization levels but then converges very fast. If  $a^2 = 1$ , good results are already obtained at low discretization levels, followed by slightly slower convergence (Fig. 4.2a and b).

In the current test case with diffusion and dispersion, both versions of  $\text{PTRW}_G$  perform better than  $\text{PTRW}_{\text{box}}$ , see Figs. 4.2a, b and d. Thus, the results of the test cases 1 and 2 suggest to choose Ansatz and test functions whose shapes relate to the ADE solution, i.e., box functions for the purely advective cases and Gaussian functions for advective-dispersive cases. For  $\text{PTRW}_{\text{box}}$ , I consider here the two different subversions  $\text{PTRW}_{\text{box}}$  and  $\text{PTRW}_{\text{box}}^c$ .  $\text{PTRW}_{\text{box}}^c$  is conceptually identical to MASST and thus leads to identical mass release rates as shown by the exactly matching curves in Fig. 4.2a. However, consistently using boxes for both Ansatz and test functions is superior to the concept behind MASST, as can be seen in the faster convergence of  $\text{PTRW}_{\text{box}}$  in Fig. 4.2a.

**Dilution:** Since MASST and PTRW simulations are mesh free, the accuracy of evaluating concentration values at the downstream control plane does not depend on the discretization level of the source term. Instead, only the initial smoothing by improper discretization of the boundary condition determines the artificial numerical dilution. This explains the very quick convergence of all PTRW versions and MASST in terms of the dilution state in the downgradient control plane (Fig. 4.2b). The overall poor convergence of FEM is based on three effects:

1. the mesh-based character of FEM affects the accuracy of evaluating concentration values.
2. FEM is consistently releasing too little mass (Fig. 4.2a) due to problems of discretizing a sharp cubic shape with the trilinear Ansatz functions used in this specific implementation.
3. FEM is consistently under predicting the dilution state in the featured control plane (Fig. 4.2b).

The latter is counterintuitive against the known problems of FEM with numerical dispersion, but can be explained by two effects that artificially decrease the dilution state. First, a plume with less mass leaving the source area will also occupy a smaller effective volume and hence have a smaller  $E$ . Second, FEM at coarse discretization levels is suffering from unphysical oscillations in the direct vicinity of the sharp-edged Dirichlet boundary condition. This leads to negative concentration values next to the lateral and upstream faces of the source zone. These unphysical values maintain steeper concentration gradients down-gradient of the source, also decreasing the simulated dilution state.

**Computational time:** With respect to computational time, FEM performs best in test case 2 for all levels of discretization, as shown in Fig. 4.2c. Yet, this has to be relativized with the

poor convergence of FEM. Besides, the computer time results shown in Fig. 4.2c do not suit for a fair quantitative assessment of computational cost across different semi-analytical and numerical methods, due to different code-implementation levels. While PTRW and MASST are performed with codes solely implemented for this study, the FEM code used here has a long-lasting history of code optimization [e.g., Nowak, 2005; Nowak et al., 2008; Leube et al., 2012]. Moreover, the differences in CPU-time between the methods may vary case-by-case. The CPU time of FEM increases at least proportional to the total number of nodes within the domain, while the CPU time of PTRW simulations is sensitive to the number of particles and emitters and to domain length, but not to domain volume. Also, parallel computing of the presented concept is trivially implemented, since each row of the transport-related mass matrix  $\mathbf{T}$  is independently calculated. The purpose of Fig. 4.2c is solely to show that computational times of  $\text{PTRW}_G$  and  $\text{PTRW}_{\text{box}}$  are in a feasible range, and that they do not scale worse than those of FEM.

**Root mean square error:** Comparing the *NRMSE* (Fig. 4.2d) according to Eq. (4.29) requires an unavoidably problematic choice of reference solution among different numerical approximation methods. Here, I used the finest  $\text{PTRW}_G$  discretization ( $\alpha_T/\Delta x = 1$ ) as reference. The reason for my choice is that  $\text{PTRW}_G$  shows the most robust convergence state in Figs. 4.2a and b already at the third discretization level ( $\alpha_T/\Delta x = 0.16$ ). This choice is confirmed by the fact that I observe a monotone decrease for all *NRMSE* curves, even for FEM at the finest discretization ( $\alpha_T/\Delta x = 1.6$ ). Independent from the, the  $\text{PTRW}_G$  results are very accurate in terms of mass release,  $E$ , and *NRMSE*.

**Mass release patterns:** In Fig. 4.3, the local (emitter-wise) mass release rates of the cubical source zone are plotted for the discretization level of  $\alpha_T/\Delta x = 1$ , resulting in  $25^3 = 15,625$  emitters for each method. All cubes have several common aspects. First, the main mass is released at their inlet planes, while downgradient emitters emit much less or even no mass.. This is caused by the advective shadowing effect (see Sec. 2.3). Second, corner points and edges show higher mass release rates than emitters on faces or in the core of the cube, because corners and edges have the largest lateral exchange with clean water by transverse dispersion, relating to the diffusive shadowing effect (see again Sec. 2.3).

While the MASST and  $\text{PTRW}_{\text{box}}$  cubes produce sharp mass release patterns due to their flush-mounted Ansatz and test functions, both  $\text{PTRW}_G$  and FEM produce slightly smoother mass release patterns, yet with slightly more pronounced corner mass release rates as effect of their overlapping Ansatz and test functions.  $\text{PTRW}_G$ ,  $\text{PTRW}_{\text{box}}$ , and MASST release mass also in the second and third row of emitters. This is due to backward dispersion and diffusion which lets the second and third row of emitters still have an influence onto the installed concentration values in the inlet plane. This is not occurring in the FEM results. As mentioned in Section Sec. 3.3, the applied FEM method uses a streamline-upwinding scheme. This scheme shifts the tri-linear test functions upstream, with the result that downstream emitters do not affect at all the upstream concentration values, explaining the sharp contrast between first and second row of mass release rates from FEM.

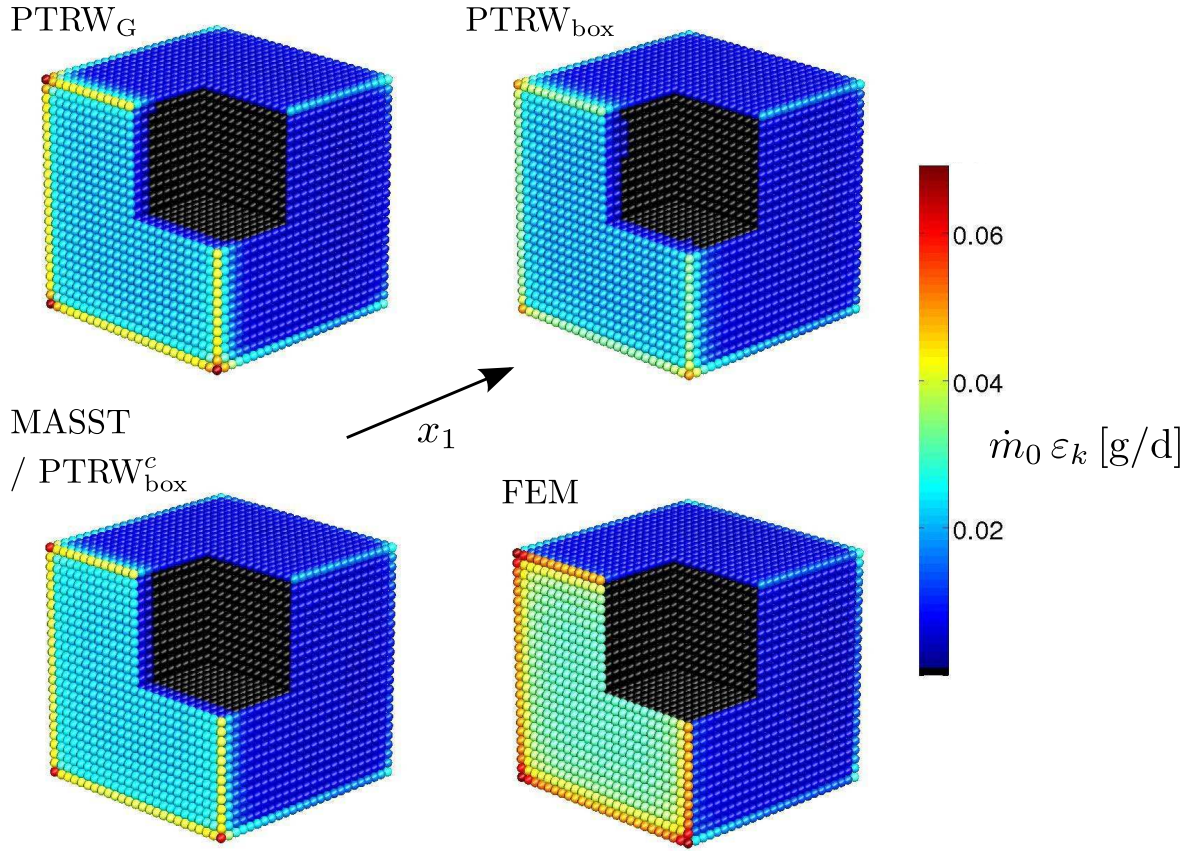


Figure 4.3: Comparison of the mass release densities within the cubic source zone at  $\alpha_T/\Delta x = 1$  (test case 2) as obtained by  $\text{PTRW}_G$  (top left),  $\text{PTRW}_{\text{box}}$  (top right), MASST and  $\text{PTRW}_{\text{box}}^c$  (bottom left), and FEM (bottom right).

**Downgradient plume cross-section:** Both Fig. 4.2b and 4.2d indicate that the downstream concentration profiles from  $\text{PTRW}_G$  are very accurate from moderate (third level) to fine discretization levels. This is analyzed in Fig. 4.4 and Fig. 4.5 in more detail. Fig. 4.4 shows the concentration values within the cross-section that was already used in Fig. 4.2. Fig. 4.5 features the concentration profiles along the horizontal center line through Fig. 4.4. Both figures show three different discretization levels (coarse  $\alpha_T/\Delta x = 0.04$ , moderate  $\alpha_T/\Delta x = 0.16$ , and fine  $\alpha_T/\Delta x = 1$ ) for the three methods,  $\text{PTRW}_G$ , MASST, and FEM. In accordance with the trilinear test functions of FEM, all FEM results are linearly interpolated. For both MASST and  $\text{PTRW}_G$ , the evaluation of concentration values is mesh free, thus arbitrary fine and does not depend on the discretization of the boundary condition. Hence, the resulting cross-sectional values in Fig. 4.4 reflect the physically required smoothness of the transport signal throughout all discretization levels. Fig. 4.5 shows the accurate agreement of the concentration profiles for all three methods at very fine discretization levels (dashed lines,  $\alpha_T/\Delta x = 1$ ). It is evident that  $\text{PTRW}_G$  is quite accurate already at moderate discretization levels (dotted lines,  $\alpha_T/\Delta x = 0.16$ ), and that all three methods strongly



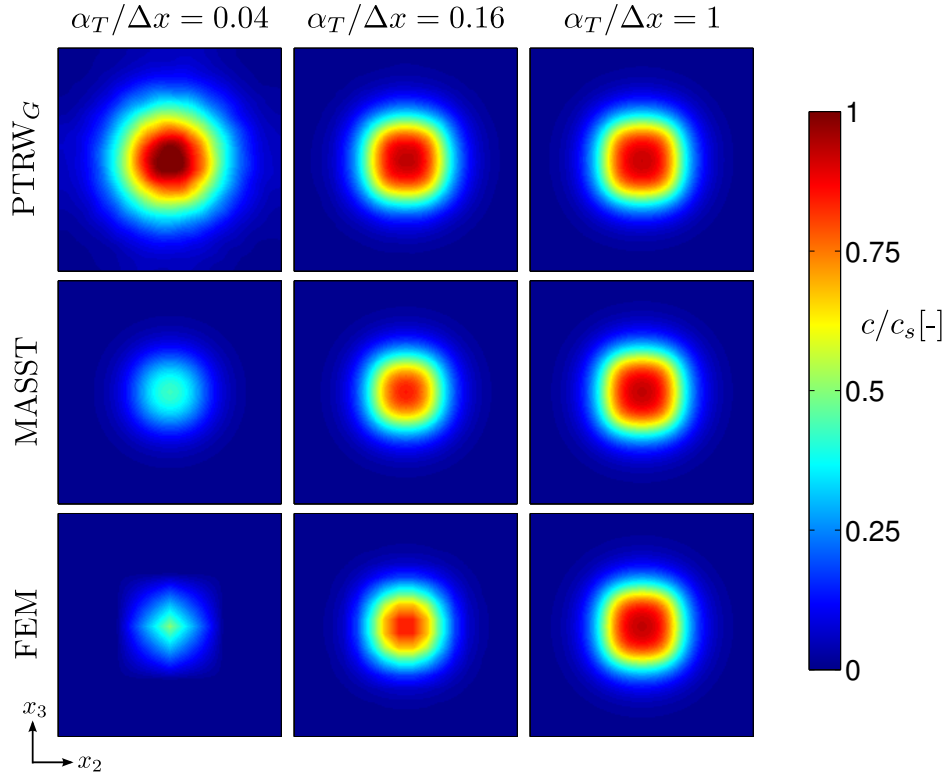


Figure 4.4: Comparison of the concentration profiles (test case 2) within the control plane 1 m downstream from the source cube at various discretization levels (left, center, right) for PTRW<sub>G</sub> (top), MASST (middle), and FEM (bottom)

disagree and poorly match the results from the reference resolution at coarse discretization level (solid line,  $\alpha_T/\Delta x = 0.04$ ). The FEM results at the coarsest shown discretization level show slightly negative values (belong the dashed grey line), resulting from the already mentioned unphysical oscillations around the source cube. Overall, I can conclude clearly that the new PTRW method outperforms FEM and MASST in this test case.

### 4.5.3 Benchmark against FEM (test case 3)

In order to get closer to the realistic application scenario from Chap. 3, I compare results of FEM and PTRW<sub>G</sub> for transport through heterogeneous porous media with non-uniform flow conditions, but still feature the cubic source zone. Since the performance of both methods depends on the properties of the aquifer, I compare them for two different aquifers, where I vary the spatial variance  $\sigma_Y^2$  of the logarithmic hydraulic permeability values ( $Y = \log K$ ) from  $\sigma_Y^2 = 1$  (aquifer 1) to  $\sigma_Y^2 = 4$  (aquifer 2). Aquifer 1 shall be used to conduct the same quantitative benchmarking and convergence analysis (Figs. 4.6 and 4.7) as in test case 2 (compare Figs. 4.2 and 4.3), which is followed by a qualitative comparison between aquifer 1 and aquifer 2 (Figs. 4.8 and 4.9).

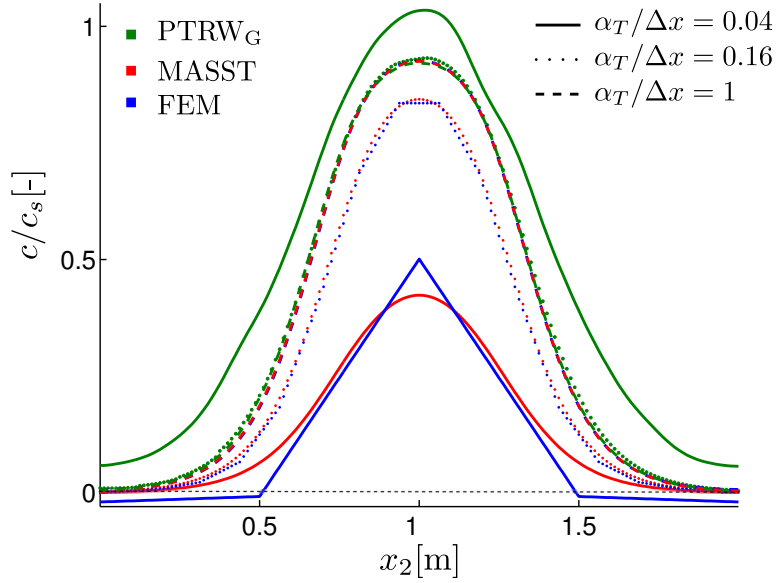


Figure 4.5: Comparison of one-dimensional concentration profiles (test case 2) at various discretization levels for PTRW<sub>G</sub> (red), MASST (green), and FEM (blue)

**Why there is no reference solution available:** Since there is no reference solution available for test case 3, I can only benchmark the results and the properties of the new PTRW method to those of FEM. Unfortunately, one can not even expect robust behavior for FEM, due to well-know numerical issues of unphysical oscillations and numerical dispersion in heterogeneous flow fields and at large Péclet numbers. These issues already occurred in test case 2, and will be even more problematic in test case 3, where local velocities at all angles against the FEM grid will occur. Yet, utilizing all available computational power (i.e., running FEM with very fine discretization levels on a specialized high-performance cluster that provides 250 GB RAM per node), I could come relatively close to (but could not fully reach) grid convergence for FEM. This can be seen in Fig. 4.6. Some of the issues of FEM are highlighted either by magenta or gray. Magenta line segments in Fig. 4.6b indicate that I had to ignore unphysical negative concentration values when calculating the  $E$ , because their logarithm would not be a real number. In Fig. 4.6c, magenta indicates that the memory requirements exceeded the limit of 32 GB. Hence, all FEM results at the 8<sup>th</sup> ( $\alpha_T/\Delta x = 0.25$ ) and 9<sup>th</sup> ( $\alpha_T/\Delta x = 0.4$ ) discretization levels were calculated on a high performance cluster with 250 GB RAM per node. The gray and blue line segments in Fig. 4.6c indicate a switch of the used linear solver. Blue and magenta parts could be solved with a geometric multi grid solver [e.g., *Alcouffe et al.*, 1981], and thus are significantly faster. However, this multi grid solver did not converge for lower discretization levels (gray line segments in Fig. 4.6), and had to be replaced with a much slower Bi-CGstab solver [e.g., *Van der Vorst*, 1992]. Both solvers produced identical results at the 5<sup>th</sup> ( $\alpha_T/\Delta x = 0.08$ ) discretization level and above.

**General convergence:** For all metrics in Figs. 4.6a, b and d, the PTRW<sub>G</sub> with both scaling factors ( $a^2 = 1$  and  $a^2 = 2$ ) converged much earlier than FEM. For that reason, it is not unambiguously clear to use the well-established method (here: FEM) as reference solution towards which the new PTRW method would have to converge. In Fig. 4.6a, both PTRW<sub>G</sub> curves overestimate the mass flux at coarse discretization levels, but they converge very quickly to an accurate mass flux of  $\dot{m}_{tot} = 70$  [g/d].

**Dilution:** The dilution index calculated with PTRW<sub>G</sub> behaves similar as in test case 2. In both test cases, accurate values are observed already for  $\alpha_T/\Delta_x = 10^{-1}$ . FEM, however, behaves quite differently in test case 3. As mentioned above, oscillations and unphysical negative concentration values lead to problems in evaluating the  $E$ , here indicated by magenta. On top of that, FEM overestimates the dilution index at coarse discretization levels due to numerical dispersion. Numerical dispersion is much stronger in test case 3 than in test case 2, as the non-uniform flow field leads to flow directions at all angles against the principal grid direction. Once numerical dispersion is sufficiently suppressed by finer discretization levels, FEM starts to underestimate the dilution index like in test case 2. This is again caused by artificially steepened concentration gradients caused by oscillations. Throughout the entire test case 3, these two error sources within FEM mutually interact and sometimes even compensate each other. This is in particular evident in the non-monotone convergence behavior of FEM in the medium resolution part of the  $E$  curve.

**Computational time:** The computational time for FEM using the Bi-CGstab solver exceeds the computational time of PTRW<sub>G</sub> far before the FEM metrics converge (see Fig. 4.6c). Using a multi grid solver, FEM is again significantly faster, but there is only a small window (blue part in 4.6c) within which this solver can be used for finer discretization levels on standard contemporary desktop computers with typically less than 32 GB RAM.

**Root mean square error:** To calculate the  $NRMSE$ , there is again no real reference solution available. Still, I used the finest FEM result as reference, in order to avoid testing the new method against itself. At very fine discretization levels, both methods produce quite similar results such that the choice of the reference solution is not very sensitive. Obviously, in the heterogeneous case, the convergence behavior in Fig. 4.6 depends on the convergence of the flow solution for both PTRW<sub>G</sub> and FEM. As PTRW<sub>G</sub> is mesh free, it can handle any discretization level of the flow field. To show the effect of flow field convergence on the PTRW<sub>G</sub> result, I plotted in Fig. 4.6d two curves for each PTRW<sub>G</sub> setting ( $a^2 = 1$  and  $a^2 = 2$ ). Results of PTRW<sub>G</sub> using the same discretization level for both the flow field and the boundary condition are plotted in green curves, while all results shown in the red curves are obtained with a fine-scale flow field (corresponding to  $\alpha_T/\Delta_x = 0.25$ ). These results (red curves) are only shown in 4.6d, since the benefits of the accurate flow field were only visible for the  $NRMSE$ . The  $NRMSE$  values of PTRW<sub>G</sub> (all versions) are significantly lower than those observed with FEM throughout all discretization levels, although both FEM and PTRW<sub>G</sub> are tested against the FEM result at the finest discretization level.

At lower discretization levels, I observe similar results for FEM and PTRW<sub>G</sub> when using  $a^2 = 2$  and the same discretization level for flow and the boundary conditions (green lines). This is caused by the commonly used and poorly converged flow field at this discretization level. The stagnating moderate values of  $NRMSE \approx 0.05$  for PTRW<sub>G</sub> at fine discretization levels indicate a remaining discrepancy between both methods after PTRW<sub>G</sub> has converged, and can not be blamed onto PTRW<sub>G</sub> due to known issues and incomplete convergence of FEM.

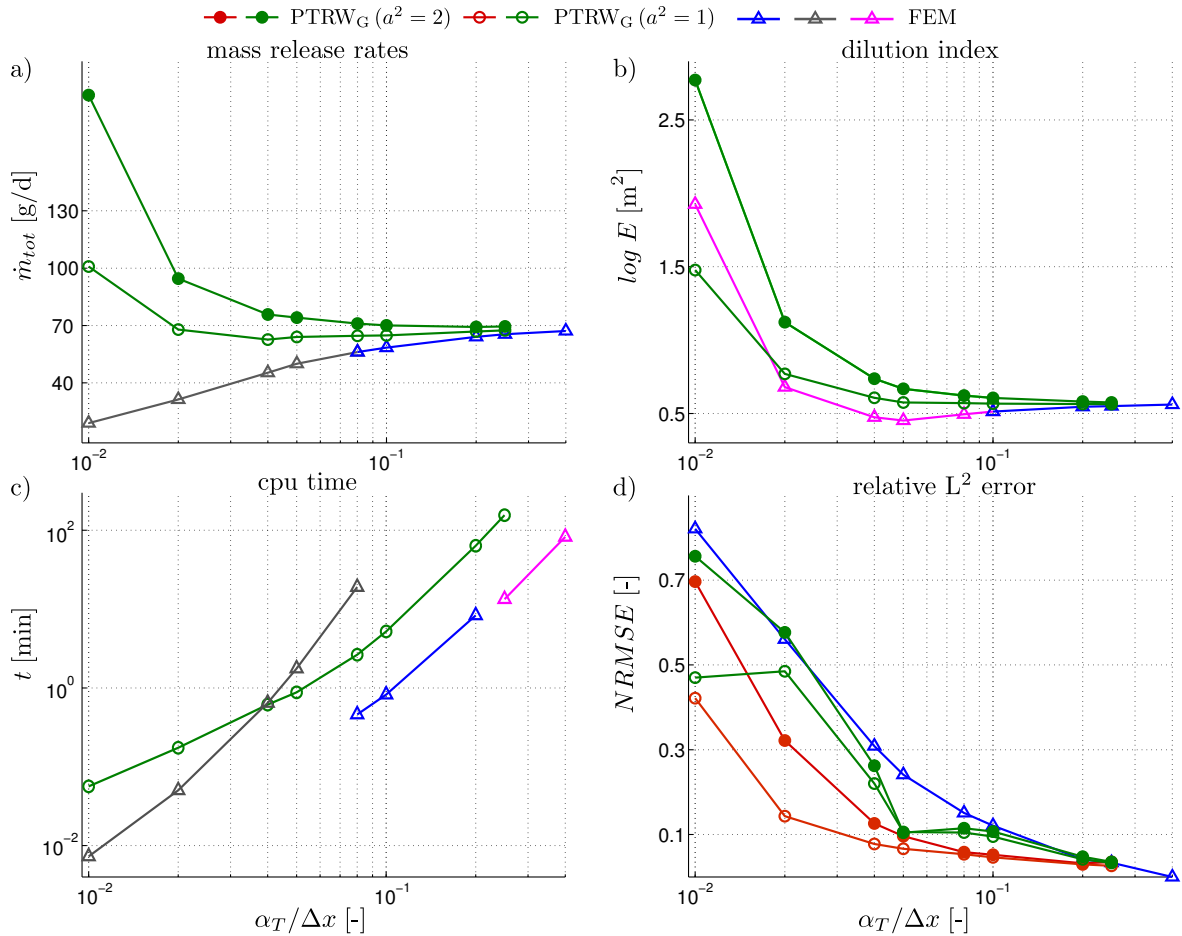


Figure 4.6: Convergence analysis and benchmarking of the new method by comparison to FEM (test case 3, aquifer 1). Top left (a): total mass release rate; top right (b): dilution index; bottom left (c): CPU time; bottom right (d): normalized root mean square error

**Mass release patterns:** In Fig. 4.7 again the mass release densities of the cubic source is plotted for test case 3. Compared to test case 2 (Fig. 4.3), the heterogeneous flow field of test case 3 leads to a pronounced spatial variation of mass release densities. Still, the two cubes show similar features as in test case 2. Due to backward dispersion, the PTRW<sub>G</sub> cube releases

a significant amount of mass in the second row of emitters, while the FEM cube covers almost the entire mass release within the inlet plane (here slightly shifted away from the rest of the cube for better visibility of the cube's core). The great discrepancy between the inlet plane of PTRW<sub>G</sub> and FEM is again an effect of the FEM's streamline-upwinding scheme. However, summing up the first and the second row of the PTRW<sub>G</sub> cube approximately leads to the values found in the inlet plane of FEM. As hydrodynamic dispersion scales with the local velocity in the classical Scheidegger parameterization used here [e.g., *Daus et al.*, 1985], the backward dispersion effects in PTRW<sub>G</sub> are enhanced at the same  $x_2$  and  $x_3$  coordinates where both FEM and PTRW<sub>G</sub> show enhanced mass release rates due to larger local velocities at the inlet plane. As in the second test case, the PTRW<sub>G</sub> result emphasizes corners and edges slightly more than FEM does. The black edges (zero mass release) within both cubes at the top ( $x_3 = \max x_3$ ) and the side face ( $x_2 = \min x_2$ ) are quite identical and indicate the influence of local velocities pointing out of the cube's face, here  $v_2 < 0$  and  $v_3 > 0$ , respectively.

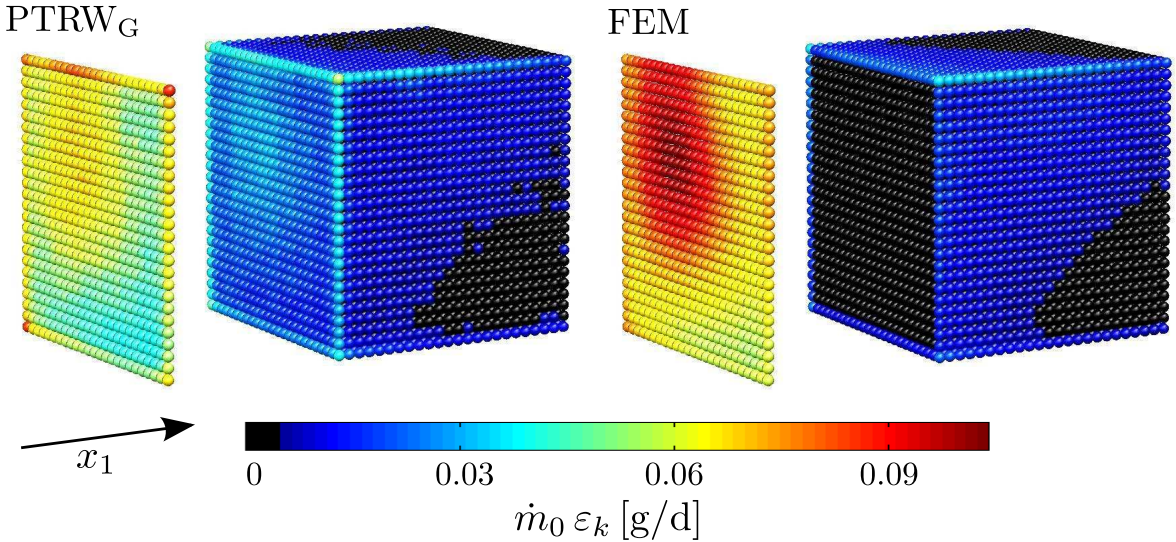


Figure 4.7: Comparison of mass release density between PTRW<sub>G</sub> and FEM at  $\alpha_T/\Delta x = 0.25$  (test case 3, aquifer 1 with  $\sigma_Y^2 = 1$ ).

**Downgradient plume cross sections:** Several concentration profiles at the control plane 1 m downstream from the source inlet for the two different heterogeneous aquifers are compared in Fig. 4.8. The concentration profiles in the top row correspond to aquifer 1 ( $\sigma_Y^2 = 1$ ) and to the results shown in Figs. 4.6 and 4.7, while the concentration profiles in the bottom row correspond to aquifer 2 ( $\sigma_Y^2 = 4$ ). In both aquifers, PTRW<sub>G</sub> produces valuable results already at moderate discretization levels (top:  $\alpha_T/\Delta x = 0.04$ , bottom:  $\alpha_T/\Delta x = 0.1$ ), while FEM is suffering from unphysical oscillations with negative concentration values (highlighted in magenta) and numerical dispersion. The numerical dispersion is visible as underestimation of peak concentration values (absence of dark brown) in the center plot in the bottom row of Fig. 4.8. FEM produced quite plausible results at a very fine discretization level

( $\alpha_T/\Delta x = 0.4$ ) shown in the right column of plots in Fig. 4.8. However, this required far more than 32 GB RAM and still does not serve as an unambiguous reference.

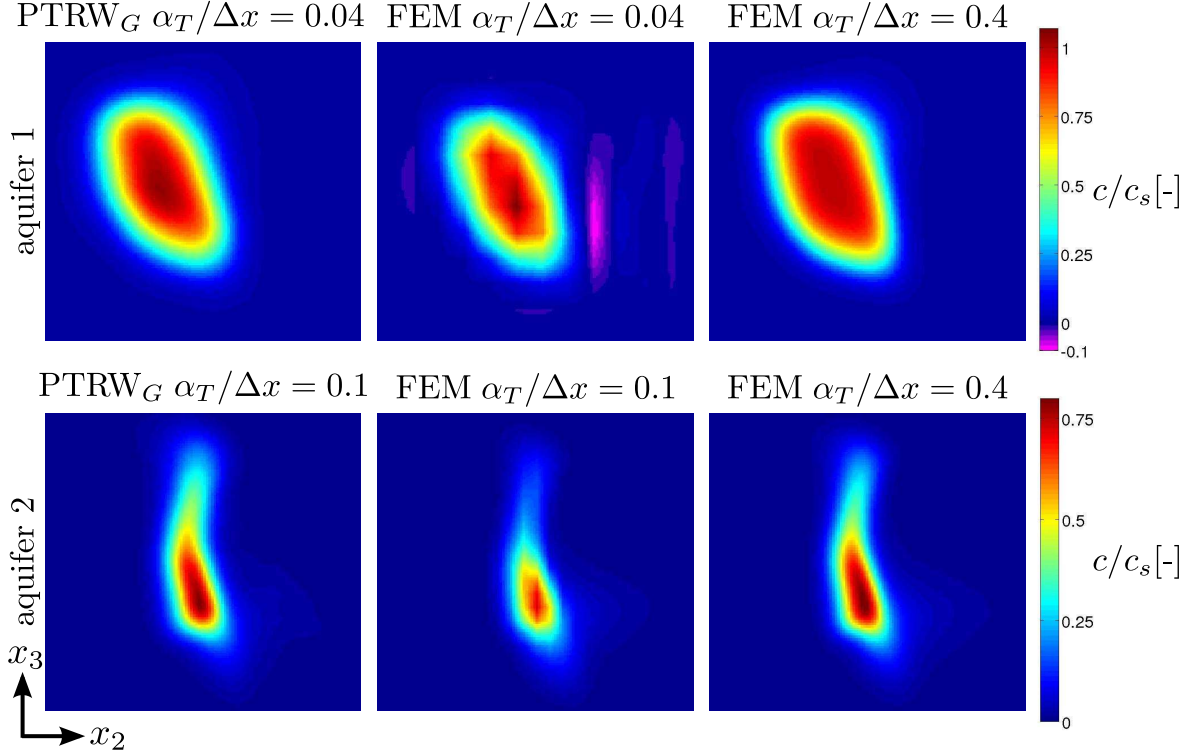


Figure 4.8: Comparison of the concentration profiles (test case 3) within a subpart of the control plane 1 m downstream from the source cube at various discretization levels for two different heterogeneous aquifers  $\sigma_Y^2 = 1$  (top, aquifer 1) and  $\sigma_Y^2 = 4$  (bottom, aquifer 2) for PTRW<sub>G</sub> (left) and FEM (middle and right)

The observations from Fig. 4.8 are supported by Fig. 4.9, where one-dimensional concentration profiles through the cross-section are compared at the same discretization levels as in Fig. 4.8. It shows a quite accurate agreement between PTRW<sub>G</sub> results at moderate ( $\alpha_T/\Delta x = 0.1$ ) and fine discretization levels ( $\alpha_T/\Delta x = 0.25$ ) with the FEM results at a very fine discretization level ( $\alpha_T/\Delta x = 0.4$ ). The only significant difference between the very finely discretized FEM result and these two PTRW<sub>G</sub> results appears in the left plot (aquifer 2), where either FEM underestimates or PTRW<sub>G</sub> overrates the concentration plateau on the left side. At the coarse discretization level ( $\alpha_T/\Delta x = 0.04$ ), the negative concentrations in FEM solutions become visible quite clearly (left plot in Fig. 4.9). Again, I can conclude overall that the new PTRW method clearly outperforms FEM.

#### 4.5.4 Third-type boundary condition vs. local chemical equilibrium

In the setting of the above test cases, I assumed local chemical equilibrium. In order to support this assumption, I will repeat the simulation of test case 2 in this section. However, this

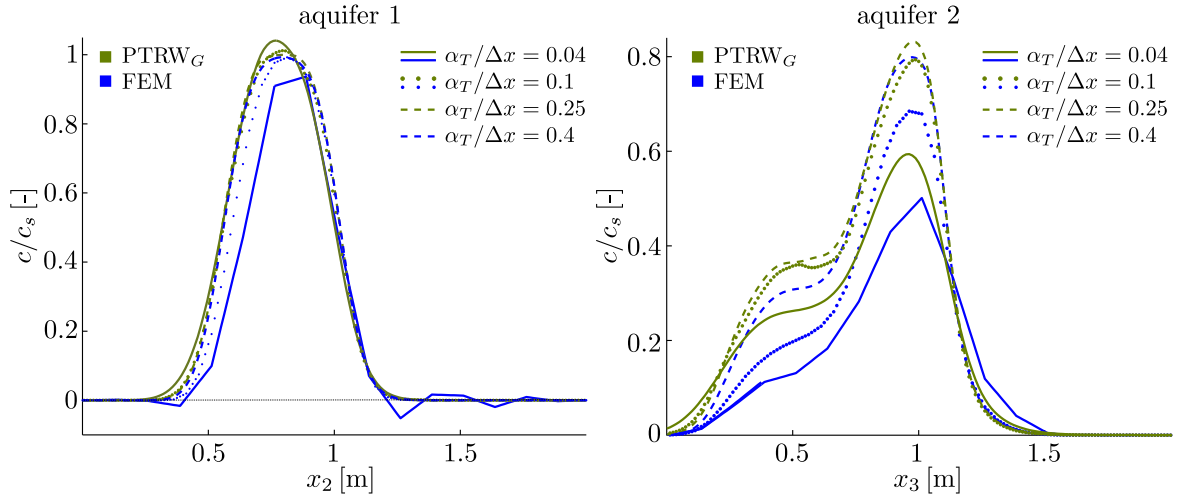


Figure 4.9: Comparison of one-dimensional concentration profiles (test case 3) through the control plane shown in Fig. 4.8 at various discretization levels for two different heterogeneous aquifers with  $\sigma_Y^2 = 1$  (left, aquifer 1) and  $\sigma_Y^2 = 4$  (right, aquifer 2) for PTRW<sub>G</sub> (red), and FEM (blue)

time I consider third-type boundary conditions, and thus solve for Eq. (4.26). As discussed in Sec. 4.4, the parameter  $\kappa_{\text{eff}}$  applied here depends on the scale of the support volumes. Test case 2 possesses very small support volumes (e.g.,  $\Delta x = 0.02\text{m}$ ) that correspond to the scales of experimentally observed mass transfer rate coefficients [e.g., *Miller et al.*, 1990; *Imhoff et al.*, 1994]. In this scale, experimentally observed values for  $\kappa^{-1}$  vary from  $10^{-2}$  d to  $10^{-3}$  d.

Using again the setting of test case 2, but setting  $\kappa_{\text{eff}}^{-1} = 10^{-2}$  and solving for Eq. (4.26), we can quantify the error for the local chemical equilibrium assumption. Comparing the local mass release rates emitter-wise, we observe the normalized root mean square error  $NRMSE \approx 3.8 \cdot 10^{-11}$  and a maximum difference at the edges of 0.09 %. The total mass flux differs only by 0.0018 g/year ( $1.2 \cdot 10^{-5}\%$ ), but this is caused by the fact that downstream emitter can compensate for the reduced release rates at the inlet, which would be different for less compact sources. Repeating this simulation whilst increasing the seepage velocity stepwise from  $v_1 = 0.25$  m/d to  $v_1 = 64$  m/d, we observed  $NRMSE$ s in the range of  $10^{-10} - 10^{-7}$ . This supports the assumption of local chemical equilibrium in test case 2.

Yet, to make the kinematic effect visible, I set  $\kappa_{\text{eff}}^{-1} = 40$  d and increased the Darcy velocity to  $q = 60$  [m/d]. The resulting mass release rates are shown in Fig. 4.10. For the sake of better comparison to Fig. 4.3, the color scale in Fig. 4.10 is the one from Fig. 4.3, linearly scaled by the ratio of velocities.

In the results shown here, emitters in the core of the source cube are still releasing zeros mass. However, if dissolution is limited by kinematic mass transfer the first row of emitters releases less mass compared to the results of test case 2 (Fig. 4.3), and it is not able to install solubility concentration. Therefore, shadowed emitters in a few subsequent rows are

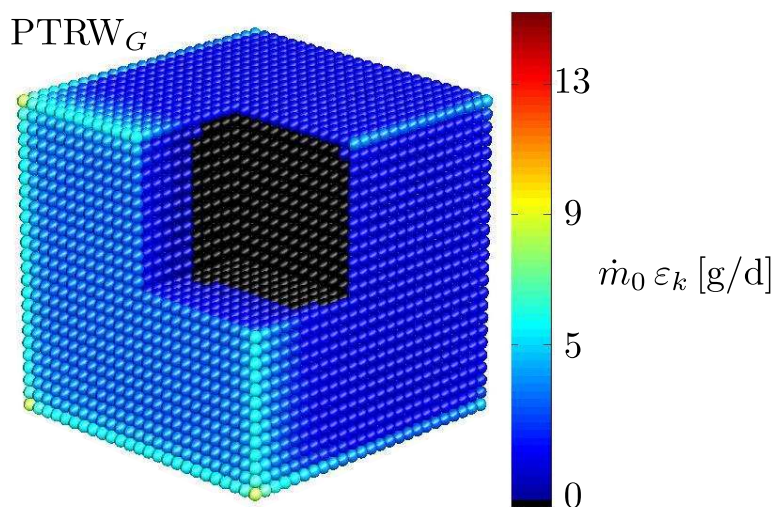


Figure 4.10: Illustration of mass release densities for third-type boundary conditions, if dissolution is clearly limited by mass transfer kinetics with  $\kappa_{\text{eff}}^{-1} = 40$  d and  $q = 60$  m/d

enabled to contribute to the overall mass flux, as can be seen in the first few rows of the cube in Fig. 4.10. Despite these dramatic changes in  $\kappa_{\text{eff}}$ ,  $q$ , and in the mass release pattern, the total mass flux difference with and without local chemical was still only 3.9%. Of course, the effect on total mass release rates would be different for less compact sources, as already mentioned above. The emitters most affected by the kinematic mass transfer limitation are those at the edges and corners (see Fig. 4.10). Without kinematic limitation, they had the lowest level of limitations imposed by shadowing effects. Thus, they show the largest (relative) impact of the new kinematic limitation.

## 4.6 Summary and conclusion

I presented and tested an efficient novel method to implement Dirichlet boundary conditions in PTRW simulations. It allows to directly calculate concentration values at any place in space and time, the source strength resulting from the boundary condition, total mass fluxes and the corresponding depletion time of the donor phase. The key conceptual step is to discretize the geometry of the boundary condition by a set of Ansatz functions, and to project the entire PTRW simulation result onto a set of test functions in order to control the particle release rates at the boundary. The boundary condition can have any arbitrary geometrical shape, including complex-shaped subvolumes of the domain as internal boundaries. Although Ansatz and test functions are involved and I use a Galerkin projection procedure, this does not lead to a weak form of the discretized equation. Instead, the proposed method solves the transport equation entirely by PTRW, inheriting the local and global mass conservation properties from PTRW. The Galerkin procedure only serves to discretize and control the boundary condition. Thus, if the mass release is adequately discretized, the over-



all method is in fact mass conservative. Increasing this discretization controls the accuracy of the method, but also increases the computational demand.

I illustrated the presented method in an application example, featuring a highly irregular geometry of the boundary condition (as it arises in simulation of DNAPL dissolution) and transport through a heterogeneous porous medium. The accuracy of the presented method is demonstrated in three test cases. One test case is designed for comparison to an analytical solution, one is designed for comparison to the semi-analytical method MASST [Sale and McWhorter, 2001], and one is designed such that both analytical and semi-analytical solutions do not apply, and only finite element methods (FEM) or comparable numerical schemes are suitable for benchmarking. I observed very robust convergence behavior of the proposed PTRW method in all test cases, and the new method outperforms both MASST and FEM under realistic application conditions – both in terms of accuracy and computational efficiency. The performance was investigated for different kernel functions to discretize the boundary condition (Gaussian and box function), and different scaling factors  $a^2$  of the kernel bandwidth.

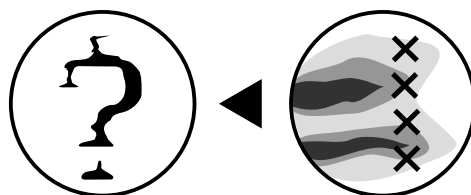
The relevant application range of this method includes simulations of transport in large domains of heterogeneous porous media resolved at fine scales, and at large Péclet numbers. The computationally most demanding steps are trivial to parallelize on computing clusters. The method also allows quick updating of all results, if locations disappear from the boundary condition, e.g., if a donor phase at any location has been depleted, without running any additional simulation. The presented method is feasible even for extremely high-contrasted three-dimensional heterogeneous porous media, where Eulerian-based methods struggle with unphysical oscillations and numerical dispersion unless when reverting to strictly prohibitive fine numerical resolutions.

### Key conclusions

- Dirichlet and third-type boundary conditions in PTRW simulations is challenging, but possible using the developed PTRW method.
- The requirements for a consistent model framework as discussed in Chap. 3 are fully met. According to these requirements, the method can account for complex-shaped and adequately discretized boundary conditions (e.g., by smooth-emitter volumes), is suitable for the quasi-steady-state depletion concept, and avoids artificial numerical dissipation of concentration values at high Péclet numbers.
- The new PTRW method outperforms other well-established methods like, the multiple analytical superposition technique (MASST) and a Petrov-Galerkin FEM.



## 5 CSA identification



**Motivation** Modelling the fate and transport of DNAPL contaminants in the subsurface is subject to a high degree of uncertainty. This was discussed and assessed in Chap. 2 and Chap. 3 for mass discharge and source depletion time predictions. The current chapter addresses the question of how this uncertainty can be reduced with data that is obtained from field campaigns. Due to the unpredictability of the actual CSA and its interdependencies to aquifer parameters, groundwater flow, and dissolution rates, the CSA morphology was revealed as a relevant source of uncertainty for mass discharge predictions in Chap. 3. Hence, the key task of this chapter is the identification of CSAs and their characteristics. This helps to reduce the uncertainty of predicted impact metrics, such as mass discharge and source depletion time. Increased knowledge of CSA characteristics and of these impact metrics helps to determine the need for remediation and supports the decision for a specific remediation strategy.

**CSA identification is a challenge** The direct and simple use of data for the desired predictions is highly limited by the heterogeneity of the sought natural states and the tremendous cost per sampled data point. A detailed program of exhaustive soil sampling would be necessary to estimate the actual and complex shape of the CSA, due to the spatial variability of the DNAPL distribution [Feenstra, 1996]. Inverse models enable the assimilation of indirect data to numerical models as briefly discussed in Sec. 2.1. The main challenges of inverse modeling stem from the typical non-linearity, non-uniqueness, and computational demand of inverse problems. Thus, methods for efficient inverse modeling are indispensable.

**Goal and structure** The key task of inverse models in the context of this thesis is the identification of possible CSAs from observable field data. Since this is typically a non-unique problem, a set of possible CSAs shall be identified instead of one best estimate of the unknown CSA. For this purpose, an efficient and accurate reverse-inverse methodology for CSA identification is derived and presented in the following. First, the principal

goal, approach, and contributions of this chapter are discussed in Sec. 5.1. Sec. 5.2 summarizes state-of-the-art tools for efficient inverse modeling which are applied in the proposed reverse-inverse method. The basic framework of this method is introduced in Sec. 5.3, and the Lagrangian-reverse formulation to link concentration observations directly with the CSA is presented in 5.4. Finally, the performance of the proposed reverse-inverse method is demonstrated in several inversion examples in Sec. 5.5.

## 5.1 Goal, approach and contributions

### 5.1.1 Goal and approach

**Goal** The goal of this chapter is to propose an accurate and efficient inverse model framework to identify CSAs and to reduce the uncertainty of predicted impact metrics. A key challenge for this purpose is the assimilation of concentration values, because they possess strongly non-linear relationships to various model parameters that are subject to uncertainty [Schwede and Cirpka, 2009]. The relevance of a physically and statistically consistent model framework for mass discharge and source depletion times was demonstrated in Chap. 3. This requires that aquifer heterogeneity, groundwater flow irregularity, complex and physically-based CSAs, and their interdependencies are accounted for. The observed key requirements of such a sound model framework are also relevant for the assimilation of field data and CSA identification techniques. Here, the physical and statistical interdependency between CSA morphology and aquifer parameters is of particular importance, since both features are subject to a high degree of uncertainty. Thus, CSA and aquifer parameter need to be inferred jointly.

**Shortcomings of standard methods** Honoring the non-linearity of all relevant interdependencies and facing the non-uniqueness of the source identification problem under limited computer power is the major challenge of accurate and efficient inverse models. The DNAPL source identification problem is inherently non-unique: (1) the unknown parameters (required to describe complex CSA and aquifer heterogeneity) by far exceed the feasible number of observations, (2) redundant information are observed from the field, and (3) certain types of information are lacking in general (e.g., pore-scale information). A unique solution to this inverse problem only exists in the sense of a best estimate, or if prior assumptions significantly constrain the solution space. The latter typically demands overly strong and unphysical assumptions. To obtain a best estimate, optimization algorithms navigate the search for the desired information and suggest a solution that is satisfactory according to a predefined criterion. Gradient-based search methods [e.g., Doherty *et al.*, 1994; Sciortino *et al.*, 2000] are one example out of many. These types of approaches drastically impose unphysical constraints on the parameter values and their interplay when calibrating them with observed data. As a result, the calibrated parameter set (in the context of this thesis to describe both aquifer heterogeneity and complex CSA) does not necessarily obey the underlying physical processes. This and the inability to obtain uncertainty bounds would lead to an unreliable basis for risk assessment at contaminated sites [Schwede and Cirpka, 2010].

**Approach and ingredients** The current chapter of this thesis tackles these issues of inverse problems via joint geostatistical inversion of the heterogeneous permeability and the complex CSA in a Bayesian framework based on Monte Carlo simulation and bootstrap filtering [e.g., *Gordon et al.*, 1993; *Ho and Lee*, 1964]. Soil properties, DNAPL saturation, and solute concentration observations are used for the geostatistical inversion in the present application in Sec. 5.5. However, other types of field observations such as mass flux, head, and groundwater flow could also be used in the same framework that is introduced in Sec. 5.3. Following the Bayesian approach to inversion, probabilities of CSA and aquifer parameters which are a priori constrained to be physically plausible are evaluated. This is an important contrast to the source identification approaches which can be found in the literature (see Sec. 5.1.2). The physical constraints dramatically accelerate the search for possible solutions, which may compensate for the absence of a navigating optimization algorithm in a brute-force Monte Carlo-based search. In order to enable huge Monte Carlo simulations, a bundle of well-chosen fast algorithms are proposed:

1. the use of fast assimilation techniques to calibrate the geostatistic model [e.g., *Fritz et al.*, 2009; *Nowak and Litvinenko*, 2013],
2. fast random generation of physically-based CSAs (see Sec. 3.2),
3. and a swift transfer function from observed concentration values to the joint likelihood of the proposed aquifer model and CSA (see Sec. 5.3 and Sec. 5.4).

To this end, actual linearities of the physical system are utilized, the causality between boundary condition and concentration observations is reverted in the sense of adjoint-state approaches [e.g., *Sun*, 1994; *Neupauer and Lin*, 2006; *Michalak and Kitanidis*, 2004], and observed data is assimilated sequentially in an efficient, and physically reasonable order. The latter implies the following order of assimilating data:

1. fitting the geostatistical model parameters (i.e., variance of  $Y = \log(K)$   $\sigma_Y^2$ , the correlation lengths in three dimensions  $\ell$ , the Matérn shape parameter  $\kappa_M$ ) and interpolating the permeability values  $Y = \log K$  due to direct soil observations is performed. The entry pressure field is related to permeability values by linear regression (see Sec. 2 and *Shepherd* [1989]).
2. the randomly generated CSAs (see Sec. 3.2) are filtered with observed saturation values.
3. the preconditioned aquifer parameter fields and CSAs are jointly inferred with respect to concentration observations.

This last step of conditioning possesses strong non-linearities and faces the non-uniqueness discussed above [*Schwede and Cirpka*, 2009, 2010]. However, utilizing the linearity of the featured transport equation, adjoint-states sensitivities of resident concentration values for measurement locations with respect to the source zone are swiftly calculated without loss of accuracy. As will be discussed in Sec. 5.3, these adjoint states are used to derive the probability of a proposed set of aquifer parameters and CSA. These adjoint states can be calculated explicitly by solving the reversed transport equation with an accurate and efficient Lagrangian method. Applying Bayesian inference, a priori knowledge on CSAs and aquifer

parameters, measurement uncertainty, and the non-uniqueness of CSA and aquifer combinations that lead to the observed concentration values are considered in a probabilistic consistent manner.

### 5.1.2 Relation to state of the art

The unknown locations and source geometries of possible DNAPL contaminations in the subsurface have been subject to intensive research in the last decades, and many valuable contributions have been made. Additionally, valuable contributions have been made to assimilate concentration data to geostatistical models [e.g., *Kitanidis, 1996; Schwede and Cirpka, 2009*]. *Huang et al. [2009]* applied an Ensemble Kalman filter (EnKF) to calibrate a heterogeneous transmissivity field with head and concentration observations with an unknown contamination source. Concentration observations have also been used to evaluate the uncertainty of mass discharge estimates at contaminated sites [*Schwede and Cirpka, 2010; geostatistics for evaluation of mass discharge uncertainty at contaminated sites Trolldborg et al., 2012*]. However, I will focus in the remainder of this section on inverse approaches that aspire to identify contaminant sources. In general, there are deterministic and stochastic inversion approaches to infer information of the unknown source [*Sun, 2007*]. Their main difference is whether or not the unknown source information is modeled as a random process.

**Deterministic inversion** Deterministic approaches select a small set of source zone parameters by assuming a deterministic, simple, and parametric CSA geometry. The values of these parameters are then optimized to find a best estimate. Such a best estimate of a set of parameter values is typically evaluated with respect to their power of explaining the observed data on the basis of a least mean square error [e.g., *Gorelick et al., 1983; Sun et al., 2006; Yeh et al., 2007*]. This forms an objective function whose minimum can be found with various optimization algorithms. Examples along these lines include:

- *Sciortino et al. [2000]* applied a gradient-based optimization to identify the source location and shape of a DNAPL pool based on a three-dimensional analytical transport model.
- *Aral et al. [2001]* used a genetic algorithm to infer the release history and source location.
- *Sun et al. [2006]* formulated a constrained robust least squares estimator that minimizes the worst-case error and applied it to a source identification problem with an uncertain but homogeneous permeability field.
- *Yeh et al. [2007]* incorporated a combined simulated annealing (SA) and tabu search (TS) algorithm with a three-dimensional groundwater flow and transport model (MODFLOW-GWT) to estimate the source location and the release concentration and period in heterogeneous aquifers. They produced a trial solution of candidate source location with TS, and the release period and concentration value at this location with SA, and then they employed MODFLOW to calculate sampling concentrations at the

monitoring points. Finally, when the objective function has converged to the stopping criteria, the last trial source information (location, release period, and concentration) which yields the best objective function value is considered the final solution.

- *Saenton and Illangasekare* [2004] determined the entrapment architecture along a vertical DNAPL source line using downgradient mass flux measurements. Using the optimization-based inverse modeling algorithm PEST [*Doherty et al.*, 1994] in combination with forward flow and transport simulations via MODFLOW [*Harbaugh et al.*, 2000] and RT3D [*Clement et al.*, 1998], they were able to infer a unique solution of the DNAPL distribution perpendicular to the mean flow direction of a two-dimensional transport problem. Their downgradient observations were obtained from a DNAPL infiltration and solute transport experiment applying the same heterogeneous porous medium.

**Reverse transport for CSA identification** The approach of using backward solutions of the transport equation to identify contaminant sources also has a long history. *Skaggs and Kabala* [1995] applied the method of quasi-reversibility of the diffusion equation within a backward-in-time moving coordinate system. Since the coordinate system is moving backward with respect to advection, a reverse solution of the advection-dispersion equation was obtained and the history of a solute contaminant plume from present observations was recovered in a one-dimensional domain. The quasi-reversibility scheme theoretically permits the use of spatially and temporally variable transport parameter, which is the main advantage over previous approaches (e.g., Tikhonov regularization [*Skaggs and Kabala*, 1994]). *Bagtzoglou et al.* [1992] and *Ababou et al.* [2010] transformed downstream contaminant plumes in particle density fields in order to back track these particles via random walk. While *Bagtzoglou et al.* [1992] identified probabilities of point sources by reversing a known two-dimensional non-uniform velocity field, *Ababou et al.* [2010] extended this approach with an anti-diffusive scheme, such that they are able to identify multi-modal point sources in a two-dimensional homogeneous aquifer. The anti-diffusive scheme is a censored random walk, which forces particles to tend towards local mass centers of the particle distribution.

**Stochastic inversion** In contrast to deterministic approaches, stochastic approaches avoid overly strong and restrictive low-parametric assumptions on the unknown source function. Through using stochastic and statistical concepts, they are able to reflect the non-uniqueness of the underlying source identification problem. The most significant contributions in this area are:

- *Woodbury et al.* [1998] recovered the source release history of a three-dimensional plume within a steady and uniform flow via linear inversion with an extended minimum relative entropy method in an analytical framework.
- *Neupauer and Wilson* [2001, 2002] derived location and travel-time probabilities from adjoint states of resident concentrations via solving the adjoint equation of a forward contaminant transport model.

- *Neupauer and Lin* [2006] applied this reverse technique for conditioning the location and travel-time probabilities on measured concentrations.
- *Snodgrass and Kitanidis* [1997] applied the Quasy-Linear Geostatistical Approach [*Kitanidis, 1995*] to find a best estimate of the release history at a known source location, conditioned on concentration observations some time after the release. In this Bayesian approach, they search for an estimate of the source release function that maximizes its a posteriori probability in a gradient-based iteration procedure (e.g., Gauss-Newton iteration). To this end, they define a random release function with known correlation structure but unknown structural parameters (this is comparable to the covariance functions applied in Sec. 2.5.3). The source identification problem is transferred to an optimization problem of these structural parameters. Their work demonstrates that the problem of estimating solute release from downstream concentration observations is actually a linear one, which is also utilized in the reverse-inverse methodology of this work.
- *Michalak and Kitanidis* [2004] extended the work of *Snodgrass and Kitanidis* [1997] to the identification of the historical two-dimensional spatial distribution of a contaminant in two-dimensional deterministic heterogeneous porous media. Using adjoint states (similar to *Neupauer and Wilson* [2001]), they presented a stochastic inverse methodology, which is theoretically applicable for three-dimensional heterogeneous porous media and complex source zones. The historical contaminant distributions at a selected set of points in space and time are inferred from concentration data.
- *Sun* [2007] formulated a robust min-max problem applying the geostatistical inversion of *Michalak and Kitanidis* [2004] within a probabilistic setting for the aquifer model and considering measurement errors of concentration observations.
- Recently, *Troldborg et al.* [2010] applied a Bayesian geostatistical inversion framework to evaluate the uncertainty from mass discharge estimates based on several conceptual models that represented the uncertainty of the site condition. For example, the uncertainty of whether or not the contaminant penetrated the aquifer and accumulated on top of the subjacent aquitard.
- *Hosseini et al.* [2011] obtained probabilities of areal source sizes via Monte Carlo simulation under uncertain heterogeneous transmissivity and uncertain dissolution rate and biodegradation rate. They reconstructed areal source zones via the distance-function approach (DF) and conditioned the geostatistically generated transmissivity fields on hydraulic conductivity and head measurements using the sequential self-calibration approach [*Gomez-Hernandez et al., 1997*].
- *Zeng et al.* [2012] combined a stochastic collocation method (SCM) and Monte Carlo Markov Chain (MCMC) method to infer a selected set of source parameters within a deterministic homogeneous and heterogeneous transmissivity field. With the SCM, they constructed a surrogate system of the forward problem in polynomial form. This surrogate is then used by the MCMC method in order to alleviate the computational burden of standard MCMC methods. Unfortunately, the construction of the surrogate is only feasible for a small set of random source parameters. For the homogeneous



case, the location of a point source, its release concentration, and duration were randomly drawn. For the heterogeneous case, only the release concentration was modeled as uncertain, while source location and release duration were assumed to be known.

**Summary of shortcomings** Solving a deterministic optimization problem to find the best estimate of the contaminant sources necessarily disregards the non-uniqueness of the underlying problem. Several authors studied the identification of probable source locations in homogeneous porous media in stochastic inversion approaches [e.g. *Skaggs and Kabala, 1995; Woodbury et al., 1998; Sun et al., 2006; Ababou et al., 2010*]. Other authors tackled aquifer heterogeneity in this context in two-dimensional flow and transport models [e.g. *Bagtzoglou et al., 1992; Saenton and Illangasekare, 2004*], followed by studies using (geo)statistical methods to treat the uncertainty of the unknown source function and the aquifer model [e.g. *Snodgrass and Kitanidis, 1997; Michalak and Kitanidis, 2004; Sun, 2007; Troldborg et al., 2010*]. Nevertheless, all described inversion techniques apply strong assumptions to the location and geometry of the source zone. They either infer source information from a known source location [e.g., *Snodgrass and Kitanidis, 1997; Sun, 2007*], or infer the location of a few point sources [e.g., *Neupauer and Lin, 2006; Yeh et al., 2007; Ababou et al., 2010*], or the inversion is based on a few scenarios of conceptual source models [*Troldborg et al., 2010*]. Using a random space function, the stochastic inversion approaches of *Michalak and Kitanidis [2004]* and of *Hosseini et al. [2011]* could also infer information on the areal shape or size of the source zone in a two-dimensional aquifer with heterogeneous transmissivities. However, these studies assume statistical independence between the proposed source functions and aquifer properties. In conclusion, there is no inversion approach that utilizes physical knowledge on the DNAPL source formation process and its interdependency to aquifer properties in order to restrict the possible outcomes of contaminant distribution a priori. Hence, efficient and accurate inverse methods to jointly infer complex CSAs and aquifer parameters in a physically and statistically consistent setting are lacking in the literature up to date. Overcoming this gap is the goal and contribution of the current chapter.

## 5.2 Inverse modelling tools

In the previous section, I reviewed a variety of inverse approaches to solve the source identification problem. These approaches are part of the current state-of-the-art, but by far do not represent an exhaustive review of existing inverse tools. However, if the inverse problem faces highly non-linear interdependencies and non-uniqueness, and if one desires to adequately address these issues, the list of feasible and appropriate techniques deflates dramatically. From the literature discussed above, I conclude that a fully Bayesian-geostatistical framework in conjunction with an efficient adjoint method, as it was applied by *Michalak and Kitanidis [2004]*, is the most promising one to cope with the posed challenges and yet to maintain high physical and statistical accuracy. Hence, Bayesian inference within a Bayesian geostatistical framework and adjoint states form the basis of the reverse-inverse approach presented in this thesis. The principles of these two techniques are introduced in the following.

### 5.2.1 Bayesian inference

The general goal is to make a statistical conclusion about a unknown parameter  $u$  due to the observed value of  $y$ . In Bayesian inference, this conclusion is made in terms of probability statements which are conditional on the observed value of  $y$ . This probability statement is commonly written as  $p(u|y)$ , i.e., as conditional probability distribution. The fundamental concept of Bayesian probabilities has been briefly discussed in Sec. 2.1. In this section, I will briefly introduce the basic mathematics and notation of Bayesian inference which is then used in the methodology of the reverse-inverse approach in Sec. 5.3. This is a brief summary of what is written in the standard textbooks by *Gelman et al.* [2000], *Weiss* [2006], and *Gamerman and Lopes* [2006], to which I refer to for more comprehensive descriptions.

**Bayes' theorem** A fundamental ingredient in parameter inference is the so-called Bayes' theorem. It is derived using the basic rules of probability theory [e.g. *Weiss*, 2006]: the rule of total probability,  $p(y) = \int p(y,u)du$  and the rule of conditional probability,  $p(y|u) = p(y,u)/p(u)$ . With these two, one can write

$$p(u|y) = \frac{p(y|u)p(u)}{p(y)}, \quad (5.1)$$

where  $p(y) = \int p(y|u)p(u)du$ . Bayes' theorem describes how additional new information can reduce prior uncertainty. To this end, Eq. (5.1) considers the following *pdfs*.

The prior *pdf*  $p(u)$  expresses the state of knowledge of some quantity  $u$  before viewing the data.

The posterior *pdf*  $p(u|y)$  defines the reduced state of uncertainty about  $u$  after having considered observations  $y$  distributed according to  $p(y)$ .

The likelihood *pdf*  $p(y|u)$  describes how likely it is to observe the data given a state of  $u$  drawn from the prior *pdf*  $p(u)$ .

Hence, both prior and posterior knowledge are related to each other via the so-called likelihood function  $p(y|u)$  in the Bayes' theorem.

Constructing a proper prior distribution is a crucial step in Bayesian inference. The choice of  $p(u)$  often results in flat (least subjective as possible) priors when there is insufficient information to otherwise define the prior [*Kass and Wasserman*, 1994]. On the other hand, ignoring relevant information and thus underestimating initial knowledge on  $u$  may lead to misspecifications of uncertainty in the posterior distribution  $p(u|y)$ . The latter may occur, in particular, if only a few observations of  $y$  are available, so that the choice of purportedly "noninformative" prior specifications makes a difference [*Gelman et al.*, 2000].

**Inference** Bayes' theorem Eq. (5.1) can be applied directly to inference problems. In this context, a  $n \times 1$  vector  $\mathbf{u}$  of model parameter values or model outcomes is the desired quantity for the inference problem and the  $m \times 1$  vector  $\mathbf{y}$  summarizes all available evidences on

$\mathbf{u}$  from observed data ( $m$  being the number of sampling points). In the subsequent application, the data set in  $\mathbf{y}$  includes concentration values and the vector  $\mathbf{u}$  refers to CSAs and heterogeneous conductivity fields. The corresponding model with its governing equations is denoted as  $\mathbf{y} = f(\mathbf{u})$ . The fact that  $\mathbf{u}$  is non-linearly interlinked in the model equation with still uncertain model parameters is a striking problem of data assimilation. The principal approach is to claim that

$$\mathbf{y} = f(\mathbf{u}) + \boldsymbol{\chi}, \quad (5.2)$$

where  $\boldsymbol{\chi}$  is a random variable with zero mean and the covariance matrix  $\mathbf{R}$ . This means that, besides a lumped measurement and model error  $\boldsymbol{\chi}$ , the model outcome  $f(\mathbf{u})$  has to conform with the observed data  $\mathbf{y}$ . Due to the typical non-uniqueness of such problems, the unknown parameters  $\mathbf{u}$  are assessed in a probabilistic framework. Thus, the *pdf*  $p(\mathbf{u}|\mathbf{y})$  of  $\mathbf{u}$  given the data  $\mathbf{y}$  is inferred. To this end, at first a prior *pdf*  $p(\mathbf{u})$  is assigned, from which propositions of  $\mathbf{u}$  are drawn and tested against the data  $\mathbf{y}$ . The latter step requires the likelihood  $p(\mathbf{y}|\mathbf{u})$ , representing the probability of observing the observed data  $\mathbf{y}$  provided that hypothesized values in  $\mathbf{u}$  and Eq. (5.2) are actually true. In order to evaluate this likelihood, however, a likelihood function with a predefined shape (i.e., a probability distribution for  $\boldsymbol{\chi}$ ) has to be assigned. It is common practice to use a Gaussian kernel for this purpose, due to the absence of better knowledge and in the hope that the smoothness and the symmetry of Gaussian kernels minimizes the risk of unjustified assumptions. Therefore, the Gaussian likelihood in inverse problems is given by

$$p(\mathbf{y}|\mathbf{u}) = (2\pi)^{-d/2} \|\mathbf{R}\|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - f(\mathbf{u}))^T \mathbf{R}^{-1}(\mathbf{y} - f(\mathbf{u}))\right). \quad (5.3)$$

For use in Eqs. (5.2) and (5.3), the remaining task is to define a proper model function  $f(\mathbf{u})$ . In the best case, this function is linear such that

$$f(\mathbf{u}) = \mathbf{H}\mathbf{u}, \quad (5.4)$$

where  $\mathbf{H}$  is a  $m \times n$  sensitivity matrix. A linear model has the striking advantage that it allows for a swift transfer from hypothesized  $\mathbf{u}$  to observed data  $\mathbf{y}$ . Linearization of  $f(\mathbf{u})$  is a common practice, if only mildly non-linear interdependencies are considered, for instance when head observations  $\mathbf{h}$  are used to infer a vector of hydraulic conductivities  $\mathbf{K}$ . In this thesis, the source identification problem is formulated such that we have an actually linear transfer function that relates resident concentrations between measurement locations and boundary condition. The sensitivity matrix  $\mathbf{H}$  of  $\mathbf{y}$  with respect to  $\mathbf{u}$  is given by the Jacobian:

$$H_{i,j} = \frac{\partial(y_i - \varepsilon_i)}{\partial u_j}. \quad (5.5)$$

In inference problems and sensitivity analyses, the differentials in  $\mathbf{H}$  are typically defined by small perturbations:  $\partial y / \partial u \equiv y' / u'$ . Hence,  $\mathbf{H}$  expresses how a variation,  $\mathbf{u}'$ , in  $\mathbf{u}$  cause a variation,  $\mathbf{y}'$ , in  $\mathbf{y}$ . Evaluating how  $\mathbf{u}'$  propagates onto  $\mathbf{y}'$  requires the solving of a system of stochastic differential equations, which may be computationally cumbersome. In specific cases, using adjoint states,  $\mathbf{H}$  can be evaluated much faster without loss of accuracy. This is described in the following section.

### 5.2.2 Adjoint states

The following description of adjoint states closely follows the standard textbook by *Sun* [1994], to which I refer to for more comprehensive descriptions. The straightforward calculation of  $\mathbf{H}$  according to Eq. (5.5) typically requires  $n$  model runs, when  $n$  is the number of uncertain model parameters. This may be cumbersome, depending on the size of the inferred parameter vector  $n$ , and on the costs of individual model runs. If the number of measurements  $m$  is smaller than the number of parameter values  $n$  to be inferred, it is advisable to transpose the direction of solving for the transfer function  $\mathbf{H}$ . This is because one solution of the primary problem  $P$  (in the case of solute transport,  $P$  is given by the ADE in Eq. 2.41) defines a full column of the sensitivity matrix:

$$\mathbf{H}_{i=1..n,j} = P(\mathbf{y}, u_j), \quad (5.6)$$

while one solution of the adjoint problem  $P^*$  defines a full row of the sensitivity matrix:

$$\mathbf{H}_{i,j=1..m} = P^*(y_i, \mathbf{u}^*). \quad (5.7)$$

This approach is based on the adjoint state method [e.g., *Sun*, 1994; *Neupauer and Wilson*, 2002]. Adjoint states are a reformulation of the primary problem  $P$  (i.e., the ADE), such that the sensitivity of one output state with respect to the input parameters is obtained with one model run of the adjoint problem. It is obvious that, when using  $P^*$ ,  $m$  model runs are required to obtain  $\mathbf{H}$ , while  $n$  model runs are required when using  $P$ . Since the differential operators in both  $P$  and  $P^*$  are equivalent and often can be solved with the same numerical subroutine, only the difference between  $m$  and  $n$  matters. Thus, if  $m \ll n$ , as will be the case in the application of this thesis, the adjoint state method is a powerful tool for model inversion.

To obtain the adjoint problem  $P^*$  of the adjoint state  $u^*$ , a weak formulation of the primary problem  $P$  is constructed. For example, let  $\mathcal{L}u(\mathbf{y}, \mathbf{x}, t) = 0$  describe the primary problem  $P$  with  $u$  being the primary state,  $\mathbf{x}$  the vector of independent variables, and  $\mathbf{y}$  the vector of unknown parameters. It can then be written that:

$$\int_0^{t_f} \int_{\Omega} \mathcal{L}u(\mathbf{y}, \mathbf{x}, t) \psi(\mathbf{x}) dt d\mathbf{x} = 0, \quad (5.8)$$

with the test function  $\psi(\mathbf{x})$ , and  $t_f$  the final time which is equal to the time of sampling  $u$ . In this weak formulation, the differential operators can be shifted from the primary state to the test function  $\psi(\mathbf{x})$  through integration by parts, i.e., using Greens' theorem.

The key approach of the adjoint state method is to eliminate all differential terms that contain the unknown primary state in Eq. (5.8). So far, the test function  $\psi(\mathbf{x})$  may have any form. In the adjoint state method,  $\psi(\mathbf{x})$  is defined such that it satisfies the adjoint problem

$$\mathcal{L}^* \psi(\mathbf{y}, \mathbf{x}, t) = 0 \quad (5.9)$$

with  $\mathcal{L}^*$  being the adjoint operator associated with the primary operator  $\mathcal{L}$  of Eq. (5.8), and with  $\psi(\mathbf{x})$  being the adjoint state  $\psi(\mathbf{x}) = u^*$ . After eliminating those differential terms that

depend on the primary state, the weak form of the primary problem only depends on the adjoint state. The adjoint state is obtained by solving Eq. (5.9). The obtained adjoint state is then inserted to simplify the remaining terms of the primary problem Eq. (5.8) and to solve for the sensitivity matrix  $\mathbf{H}$  of Eq. (5.5).

Although the adjoint problem has the same form as the primary problem, the source/sink term and the initial and boundary condition may be different. For instance, instead of an initial condition  $u(\mathbf{x}, 0) = f_0$ , a final condition  $\psi(\mathbf{x}, t_f) = 0$  is required. Applying the following time transformation

$$\tau = t_f - t \quad (5.10)$$

to the adjoint problems  $\mathcal{L}^*\psi(\mathbf{p}, \mathbf{x}, \tau) = 0$ , the initial condition of the adjoint problem is  $\psi(\mathbf{x}, 0) = 0$ . Applying the adjoint state method on the ADE (Eq. 2.41) the adjoint state  $\psi(\mathbf{x}, t)$  has to satisfy

$$\begin{aligned} -\frac{\partial \phi_e \psi}{\partial \tau} - \nabla \cdot (\phi_e \mathbf{v} \psi) - \nabla \cdot (\phi_e \mathbf{D} \nabla \psi) &= 0, \\ \psi|_{\tau=0} &= 0, \\ \psi|_{\Gamma_1} &= 0, \quad \phi_e \mathbf{D} \nabla \psi \cdot \mathbf{n}|_{\Gamma_2} = 0 \end{aligned} \quad (5.11)$$

The transformed time led to the change of sign in the storage and advective term. Due to the symmetry of the diffusive term, its adjoint is equal to the original term. Effectively, this means to solve an ADE backward in time and with reversed velocity field. First-type (Dirichlet) boundary conditions of the primary problem remain unchanged in the adjoint problem, while second-type (Neuman) boundary conditions become third-type and vice versa. However, I focus in this chapter on first-type boundary conditions, for reasons that are discussed in Sec. 5.4.

**Analytical adjoint function of the solute transport problem** *Michalak and Kitanidis [2004]* applied the adjoint state method to infer the historical contaminant distribution at some point in time from concentration observations at some later time. As a similar problem will be tackled in Sec. 5.3, I will briefly repeat the adjoint-derived solution to their homogeneous-aquifer example. The unknown distribution of a contaminant was restricted to predefined locations  $\mathbf{x}_h$  within a given area  $\Omega_h$  and at a given time  $t_h$ :  $c(\mathbf{x}_h, t_h)$ ,  $c_h$  for short, and the concentration observations were defined at specific locations  $\mathbf{x}_f$  at a given time  $t_f$  some time after  $t_h$ :  $c(\mathbf{x}_f, t_f)$ ,  $c_f$  for short. The task of the adjoint method in this example is to evaluate the sensitivity matrix of the observation concentrations  $\mathbf{c}_f$  on the historical concentration distribution  $\mathbf{c}_h$ :

$$H_{i,j} = \frac{\partial(c_{f,i} - \varepsilon_i)}{\partial c_{h,j}}, \quad (5.12)$$

for the  $i^{\text{th}}$  observation location and the  $j^{\text{th}}$  source location. Please note, although the term source location is used for  $x_h$  here, it is not necessarily the origin of the contamination, but rather some location that had been occupied by the solute plume at some earlier time in the past. For the homogeneous example, the primary problem has an analytical solution:

$$c_{f,i} = \int_{\Omega_h} c_{h,j} P(\mathbf{x}_f, \mathbf{x}_h, \Delta t) d\Omega_h, \quad (5.13)$$

with the transfer function:

$$P(\mathbf{x}_f, \mathbf{x}_h, \Delta t) = D \exp \left[ -\frac{(x_{1,f} - x_{1,h} - v_1 \Delta t)^2}{4D_1 \Delta t} - \frac{(x_{2,f} - x_{2,h} - v_2 \Delta t)^2}{4D_2 \Delta t} \right]. \quad (5.14)$$

with  $\Delta t = t_f - t_h$  and  $D = \frac{1}{2\sqrt{\pi D_1 \Delta t}} \frac{1}{2\sqrt{\pi D_2 \Delta t}}$ . This solution could be used to obtain  $H_{i,j} = \Delta x_1 \Delta x_2 P(\mathbf{x}_{f,i}, \mathbf{x}_{h,j}, \Delta t)$ . However, in this case the impact of each  $c_{h,j}$  on the observation locations  $x_f$  has to be calculated individually. This may become computationally cumbersome for problems where the analytical solution does not hold. *Michalak and Kitanidis* [2004] showed that the  $i^{th}$  row of  $\mathbf{H}$  can be directly obtained from the  $i^{th}$  adjoint state function  $P_{\psi_i}$  with respect to the  $i^{th}$  observation location. To this end, they followed the adjoint approach that is outlined above. Having inverted the time axis as described in Eq. (5.10), the adjoint problem inherits the form of the primary problem and the initial condition remains unchanged. As they were facing an infinite domain in the analytical solution, there are no boundary conditions in this example. The adjoint transfer function is:

$$P_{\psi}(\mathbf{x}_h, \mathbf{x}_f, \Delta \tau) = D_{\psi} \exp \left[ -\frac{(x_{1,h} - x_{1,f} + v_1 \Delta \tau)^2}{4D_1 \Delta \tau} - \frac{(x_{2,h} - x_{2,f} + v_2 \Delta \tau)^2}{4D_2 \Delta \tau} \right]. \quad (5.15)$$

with  $\Delta \tau = -(t_f - t_h)$  and  $D_{\psi} = \frac{1}{2\sqrt{\pi D_1 \Delta \tau}} \frac{1}{2\sqrt{\pi D_2 \Delta \tau}}$ . Note, the only difference between  $P$  and  $P_{\psi}$  is the time axis, the sign of velocity, and that  $\mathbf{x}_h$  and  $\mathbf{x}_f$  have been swapped. Hence, the sensitivity matrix  $\mathbf{H}$  can be expressed by:

$$\mathbf{H} = \Delta x_1 \Delta x_2 \begin{pmatrix} P_{\psi}(\mathbf{x}_{h,1}, \mathbf{x}_{f,1}, \Delta \tau) & \dots & P_{\psi}(\mathbf{x}_{h,n}, \mathbf{x}_{f,1}, \Delta \tau) \\ P_{\psi}(\mathbf{x}_{h,1}, \mathbf{x}_{f,2}, \Delta \tau) & \dots & P_{\psi}(\mathbf{x}_{h,n}, \mathbf{x}_{f,2}, \Delta \tau) \\ \vdots & \ddots & \vdots \\ P_{\psi}(\mathbf{x}_{h,1}, \mathbf{x}_{f,m}, \Delta \tau) & \dots & P_{\psi}(\mathbf{x}_{h,n}, \mathbf{x}_{f,m}, \Delta \tau) \end{pmatrix}$$

where in total  $m$  vector operations are required, one for each row.

This thesis is interested in the contamination origin of a steady-state solute plume that arises from a three-dimensional heterogeneous solute transport problem. The origin is supposed to be inferred using again some concentration measurements of the solute plume. Different to the example above, the solution of both the primary and the adjoint problem has to be obtained numerically. The adjoint problem, Eq. (5.11), is solved in a Lagrangian fashion by reverse particle tracking random walk (RPTRW). Different to the primary problem solved by PTRW in Chap. 4, the time is inverted as described above and the observed resident concentrations (somewhere downstream from the source) become Dirichlet boundary conditions at the positions of data collection. This is analogous to the adjoint transfer function derived by *Michalak and Kitanidis* [2004] for the analytical solution which is shown above (see Eq. 5.15).

### 5.3 Reverse-inverse approach

Probability density functions of the total mass flux emanating from a hypothetical contaminated site were presented in Chap. 3, with uncertain CSA and uncertain hydraulic conductivities. Since the model is constrained to physically-based CSAs, the spectrum of possible

model responses is already restricted. In order to increase the applicability of this framework to actual sites and to improve prediction confidence, I now investigate how conditioning on all available site data can be achieved.

To this end, the current section proposes an efficient conditional simulation method for uncertainty reduction using concentration observations somewhere downgradient from the source zone, DNAPL-phase saturation observations and hydrogeological data within the source zone. The problem of jointly inferring both the parameters of a flow and transport model and random CSAs is clearly non-unique. Therefore, I use the Bayesian inference approach, implemented via the rejection method or weighted resampling [e.g., *Gamerman and Lopes, 2006*] within a Bayesian geostatistical framework [e.g. *Michalak and Kitanidis, 2004*; *Snodgrass and Kitanidis, 1997*]. Using rejection sampling, I obtain a conditional ensemble of equally likely combinations of CSAs and hydraulic conductivity fields  $\mathbf{K}$ . From this ensemble, one can read conditional probabilities of mass discharge and CSA characteristics, which can be used to assess the risk posed by the contaminated site. Equivalently, this can be used to analyze conditional *pdfs* of peak concentrations, source depletion times, and well contamination levels. Likewise, with more information on the CSA, better remediation strategies could be designed.

I will first outline the approach of assimilating concentration data to the consistent model framework that has been introduced in Chap. 3. Then, I will briefly discuss how this conditional simulation technique is complemented by assimilating also soil and DNAPL observations (in form of  $K$  and  $S_n$ , respectively). Please note that this is not necessarily the order that is applied when actually assimilating the data.

### 5.3.1 Assimilating concentration data

As it was discussed in Chap. 3 and Chap. 4, I assume that the CSA imposes Dirichlet boundary conditions within the source zone that is spanned by the CSA geometry. Hence, the concentration distribution within the source zone, denoted in the following by  $\mathbf{c}_{csa}$ , is solely a function of the CSA geometry:  $\mathbf{c}_{csa} = f(\text{CSA})$ . Of course, the CSA is again a function of aquifer parameters, such that the concentration distribution  $\mathbf{c}_{csa}$  and the permeability  $\mathbf{K}$  are interdependent. As already mentioned above, I use Bayesian inference to jointly identify CSA, represented by  $\mathbf{c}_{csa}$ , and aquifer parameters, represented by  $\mathbf{K}$ , that can in conjunction explain a set of downstream concentration measurements  $\mathbf{c}_o$ . Applying Bayes Theorem (Eq. 5.1) to this context yields:

$$p(\mathbf{c}_{csa}, \mathbf{K} | \mathbf{c}_o) \propto p(\mathbf{c}_o | \mathbf{c}_{csa}, \mathbf{K}) p(\mathbf{c}_{csa}, \mathbf{K}), \quad (5.16)$$

where  $p(\mathbf{c}_{csa}, \mathbf{K} | \mathbf{c}_o)$  is the posterior probability density for combinations of CSAs and conductivity fields, here represented by a vector  $\mathbf{K}$  of element-wise values on a dense Cartesian grid.  $p(\mathbf{c}_{csa}, \mathbf{K})$  is the joint prior *pdf* of  $\mathbf{c}_{csa}$  and  $\mathbf{K}$ . This prior distribution is derived using the geostatistical approach discussed in Sec. 2.5 and using the SIP approach introduced in Sec. 3.2.  $p(\mathbf{c}_o | \mathbf{c}_{csa}, \mathbf{K})$  is the likelihood of observing a data set of concentrations  $\mathbf{c}_o$  which are observed further downstream in a multilevel control plane, for given CSA and  $\mathbf{K}$ .

The dissolved contaminant transport problem is linear in its source terms, if the involved flow and transport equations are linear PDEs [e.g., *Snodgrass and Kitanidis, 1997*]. Thus, the model expression for downstream concentration observations can be written in the form

$$\mathbf{c}_o = \mathbf{H}(\mathbf{K}) \mathbf{c}_{csa} + \boldsymbol{\chi}, \quad (5.17)$$

where  $\boldsymbol{\chi}$  is a random Gaussian vector with zero mean and known covariance matrix  $\mathbf{R}$ . The latter matrix represents both the measurement error and any numerical or conceptual inaccuracies. Apart from these errors, the linear function is an exact calculation of  $\mathbf{c}_o$ .  $\mathbf{H}(\mathbf{k})$  is an  $m \times n$  sensitivity matrix of the  $m \times 1$  concentration data vector  $\mathbf{c}_o$  with respect to the  $n \times 1$  vector  $\mathbf{c}_{csa}$  that describes the CSA geometry. Following the approach of adjoint states (Sec. 5.2.2),  $\mathbf{H}(\mathbf{k})$  is obtained by back-tracking particles using the reverse particle tracking random walk (RPTRW) method [*Bagtzoglou et al., 1992*]. When applying this method, the particles start from the  $m$  measurement locations, and particle densities at the  $n$  cells of the proposed source zone are computed. This is explained in more detail in the subsequent Sec. 5.4. Having evaluated the sensitivity matrix  $\mathbf{H}$ , the likelihood  $p(\mathbf{c}_o | \mathbf{c}_{csa}, \mathbf{K})$  can be calculated via

$$p(\mathbf{c}_o | \mathbf{c}_{csa}, \mathbf{K}) = (2\pi)^{m/2} (|\mathbf{R}|)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{c}_o - \mathbf{H}(\mathbf{K})\mathbf{c}_{csa})^T \mathbf{R}^{-1} (\mathbf{c}_o - \mathbf{H}(\mathbf{K})\mathbf{c}_{csa})\right). \quad (5.18)$$

The Bayesian inference procedure in short is:

1. Realizations of  $\mathbf{K}$  and CSA from the prior distribution are drawn using the methods described in Sec. 2.5 and Sec. 3.2.
2. The sensitivity matrix  $\mathbf{H}$  is evaluated according to a given set of concentration observations  $\mathbf{c}_o$  using RPTRW (see Sec. 5.4).
3. With  $\mathbf{H}$  and Eq. (5.17), the likelihood of observing the data with the current set of  $\mathbf{c}_{csa}$  and  $\mathbf{K}$  is evaluated in Eq. (5.18).
4. Once a huge ensemble of realizations and their corresponding likelihoods has been obtained, the posteriori probability distribution can be obtained either directly via converting the likelihoods to weights and then evaluating weighted ensemble statistics, or via filtering these realization with rejection sampling.

For the latter, each realization is rejected or accepted as a conditional realization using the following rule:

$$\frac{w(\mathbf{c}_{csa}, \mathbf{K} | \mathbf{c}_o)}{\max(w(\mathbf{c}_{csa}, \mathbf{K} | \mathbf{c}_o))} \begin{cases} > \iota & | \text{accept}; \\ < \iota & | \text{reject}; \end{cases} \quad (5.19)$$

where  $w \propto p(\mathbf{c}_{csa}, \mathbf{K} | \mathbf{c}_o)$ , and  $\iota$  is a realization-wise random number which is uniformly distributed between zero and one.

Fig. 5.1 illustrates the reverse-inverse CSA identification approach. A synthetic CSA is sketched by pink spheres. The black markers at a downstream control plane (right-hand side) indicate the positions of concentration measurements. For each measurement position,



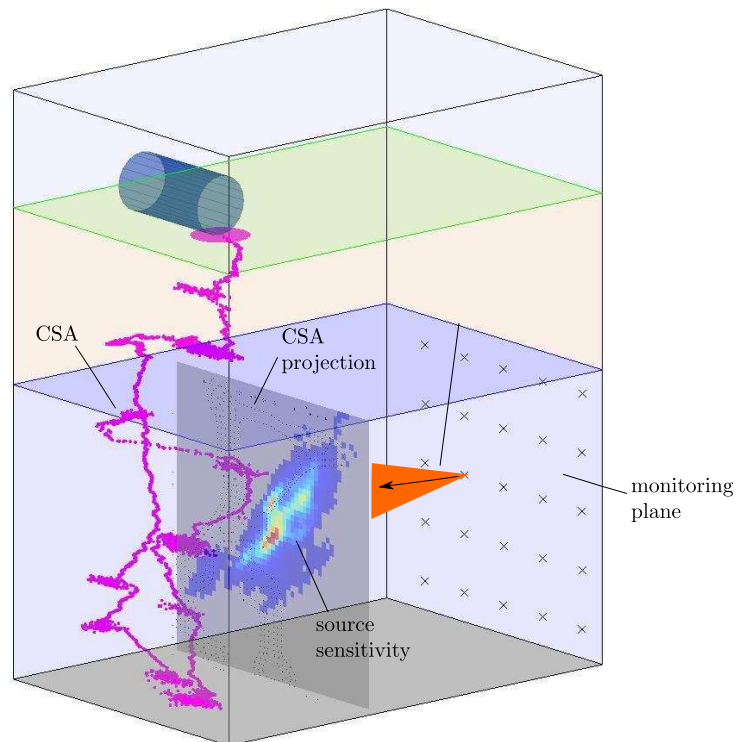


Figure 5.1: Approach of reverse-inverse CSA identification

a sensitivity with respect to the CSA is obtained via reverting the advective-dispersive transport. The latter is indicated by the orange plume. To illustrate the obtained source sensitivity, sensitivities are shown within a control-plane at the downstream face of the source zone (colormap with red for high and blue for low sensitivity). Within the same control plane, contour lines of concentration values as a result of the CSA are mapped (dotted lines).

### 5.3.2 Assimilation of various data types

So far it has not been considered that also other types of data are usually available as well, which may help to further improve the inference of the CSA and  $\mathbf{K}$ . If various data types are assimilated, it is advisable to choose the most appropriate assimilation technique for each data type. This leads to a sequential assimilation of various data types whose order is organized according to the following two principles. The order of assimilation should:

1. honor all interdependencies between the data types and nuisance parameters, i.e., geological data is assimilated before pure-phase and dissolved-phase DNAPL observations;
2. maintain efficiency of the overall inference model, i.e, pure-phase DNAPL observations are inferred before assimilating dissolved-phase concentration data.

This leads to the following approach. Direct data of conductivity can be included by fitting the geostatistical covariance parameters with rejection sampling and by the standard kriging-based correction of unconditioned (only conditioned on structural parameters) conductivity fields with the techniques described in Sec. 2.5. The structural parameters which are fitted here are the variance of log-conductivities  $\sigma_Y^2$ , the correlation lengths  $\ell$  for each principal direction, and the Matérn shape parameter  $\kappa_M$ . Once such a conditional realization of  $\mathbf{K}$  is obtained, a random CSA is generated with the technique described in Sec. 3.2. Both  $\mathbf{K}$  and CSA are then filtered in conjunction against observed DNAPL saturation values. This means that CSA and  $\mathbf{K}$  are jointly rejected if the CSA does not show saturation values ( $S_n > 0$ ) at locations where pure-phase DNAPL was observed in the field, and vice versa if the randomly generated CSA shows saturation values at locations where pure-phase DNAPL has not been observed. Measuring DNAPL saturation is difficult [Mercer and Cohen, 1990; Feenstra, 1996]. Thus, the presence of DNAPL is typically only observed qualitatively and only localized core-wise (i.e., making a checkmark for each drilling core when DNAPL was observed). This is considered by the filter by allowing a vertical discrepancy of  $\pm 0.5\text{m}$  between observed DNAPL locations and the randomly generated ones. Each assimilation technique considers additional uncertainty due to measurement errors and model inaccuracies. This is embedded in the Bayesian framework, where probabilities are updated instead of seeking a best estimate.

The overall proposed procedure can be summarized in 3 conditioning steps:

1. **Conditioning on geological data:** Conditional realizations of  $\mathbf{K}$  due to geological data are generated. First structural parameters are conditioned via rejection sampling. With these parameters  $\mathbf{K}$  fields are generated and corrected via direct Kriging. The applied methods are described in Sec. 2.5.
2. **Conditioning on pure-phase DNAPL observations:** A random CSA for each  $\mathbf{K}$  is generated, with the technique described in Sec. 3.2, and these CSAs are filtered with respect to pure-phase DNAPL observations.
3. **Conditioning on dissolved-phase DNAPL concentration:** With the remaining set of CSAs and  $\mathbf{K}$ , the Bayesian inference using the adjoint state method as described above (Sec. 5.2) is applied.

Likewise, conditioning on hydraulic heads may be achieved efficiently, e.g., via an Ensemble Kalman Filter for parameter estimation [e.g. Nowak, 2009; Huang et al., 2009] or iterative co-Kriging [e.g., Kitanidis, 1997]. However, the latter will not be considered in the presented example in Sec. 5.5.

## 5.4 Lagrangian reverse formulation

This section describes the implementation of a Lagrangian reverse formulation in order to solve the adjoint ADE. This will enable an efficient evaluation of the sensitivity matrix  $\mathbf{H}$  and thus yields a swift transfer function for the inverse problem (Eq. 5.2). As described in Sec. 5.2.2, this approach is superior to the forward simulation if the number of concentration

observations  $m$  is smaller than the number of source terms  $n$  (in the form of discrete Dirichlet boundary conditions, i.e., the emitters that resemble the CSA).

The Langrangian technique for solving the transport problem has several advantages compared to standard Eulerian approaches, such as avoiding numerical dispersion and unphysical oscillations (see Chap. 4). These advantages remain the same whether the forward transport problem or the adjoint problem is solved. Moreover, the mesh free and explicit character of Langrangian techniques allows for the efficient evaluation of the individual states of interest. Its computational effort mainly depends on the number of source terms, as discussed in Chap. 4. Hence, reducing the number of source terms from a huge number of emitter volumes of the CSA in the forward formulation to a few concentration observations in the reverse formulation dramatically reduces the computational effort of the Lagrangian model.

In the standard PTRW method, the movement of each particle obeys the following algorithm:

$$\mathbf{X}_p(t + \Delta t) = \mathbf{X}_p(t) + \mathbf{A}(\mathbf{X}_p, t)\Delta t + \mathbf{B}(\mathbf{X}_p, t)\boldsymbol{\xi}(t)\sqrt{\Delta t}. \quad (5.20)$$

$\mathbf{X}_p(t)$  is the position of a particle at time  $t$ ,  $\Delta t$  is the time step,  $\mathbf{A}$  is a drift vector,  $\mathbf{B}$  is the displacement tensor, and  $\boldsymbol{\xi}(t)$  is a Gaussian white noise characterized by zero mean and unit variance. The displacement matrix  $\mathbf{B}$  defines the strength of diffusion and dispersion.  $\mathbf{B}$  is related to the dispersion tensor  $\mathbf{D}$  (see Sec. 2.6) by  $2\mathbf{D} = \mathbf{B}\mathbf{B}^T$ . The drift vector is  $\mathbf{A} = \mathbf{v} + \nabla\mathbf{D}$ .  $\mathbf{A}$  accounts for advective displacement due to groundwater velocities  $\mathbf{v}$ , and a correction by  $\nabla\mathbf{D}$  such that the PTRW algorithm (Eq. 5.20) solves the ADE. Without this correction, the PTRW algorithm would solve for the Fokker-Planck equation [Kinzelbach, 1988]. Obviously, the PTRW approach requires a large number of particles to accurately approximate the ADE [Bagtzoglou et al., 1992; Hassan and Mohamed, 2003].

RPTRW is the associated numerical method to solve for the adjoint ADE. Accordingly, the advective term is reversed, flipping the sign of time and velocity, and leaving the dispersive term unchanged as discussed in Sec. 5.2.2 and by [Bagtzoglou et al., 1992]. This yields

$$\mathbf{X}_p(\tau - \Delta\tau) = \mathbf{X}_p(\tau) + \mathbf{A}^*(\mathbf{X}_p, \tau)\Delta\tau + \mathbf{B}(\mathbf{X}_p, \tau)\boldsymbol{\xi}(\tau)\sqrt{\Delta\tau}, \quad (5.21)$$

with  $\mathbf{A}^* = \nabla\mathbf{D} - \mathbf{v}$ . Since the concentration data are observed from a steady-state plume in the scenario featured here, it does not matter whether  $t$  or  $\tau$  (defined in Eq. 5.10) is used in Eq. (5.21). Using  $\tau$  might even be misleading, since it suggests that the time of contamination can be dated back simply by applying the RPTRW algorithm. Stationarity of the RPTRW results is ensured via convolving the particle locations over time as described in Chap. 4. Different to the PTRW approach in Chap. 4, the convolution integrates over  $\tau$  from  $\tau = -\text{inf}$  to  $\tau = 0$ . Of course, for practical reasons, the integral is truncated before  $\tau = -\text{inf}$ .

Physically speaking, this means that the solutes which arrive at the measurement locations are tracked backwards in time in order to find their possible origin. The irreversibility of diffusion and dispersion does not allow individual solute particles to be back-tracked in a deterministic fashion. However, if a huge number of particles is released at each measurement location and tracked backward by RPTRW, one can estimate the composition of upstream locations that contribute to the concentration observation at this concentration measurement location.

The standard PTRW method discretizes and tracks the dissolved mass via discrete carriers (particles) according to advective and dispersive transport (see Chap. 4). In the RPTRW concept used here, however, the tracked particles process signals of concentration values. That means, after the reverse particle trajectories are known, one simply has to add up all source terms along these trajectories in order to know the concentration value delivered by each trajectory.

Chap. 4 discussed how to model the CSA as a complex-shaped Dirichlet boundary condition. Thus, the task is now to link the Dirichlet boundary conditions with downstream concentration values. Therefore, the fraction of particle trajectories who have been in contact with a Dirichlet boundary condition (out of all particles that arrive at the measurement location) is evaluated. The assumption of local chemical equilibrium (LCE) dramatically simplifies this approach, since it leads to a homogeneous and predetermined Dirichlet boundary condition  $\hat{c} = c_s$ . Thus, the number or duration of contact does not matter. This is a direct analogy to the shadowing effect, which has been intensively discussed in Chaps. 2-4. Fig. 5.2 illustrates the mapping of arbitrary distributed Dirichlet boundary conditions to a downstream concentration value (e.g., at the measurement location). This two-dimensional illustration highlights again the importance of three-dimensional simulations, because it makes a significant difference whether the shadowing effect is assessed in two or in three dimensions. For example, in Fig. 5.2 one can not see whether the red trajectories really cross through the CSA or bypasses the CSA in the third dimension.

Although the method could be refined to work also for third-type boundary conditions, for the time being, I will continue with the assumption of LCE. Further, I assume conservative solutes here and in the example applications (Sec. 5.5). Nevertheless, linear decay models (e.g., due to biodegradation) can be directly accounted for by keeping track also of the travel times, and reducing the particle masses and concentrations accordingly.

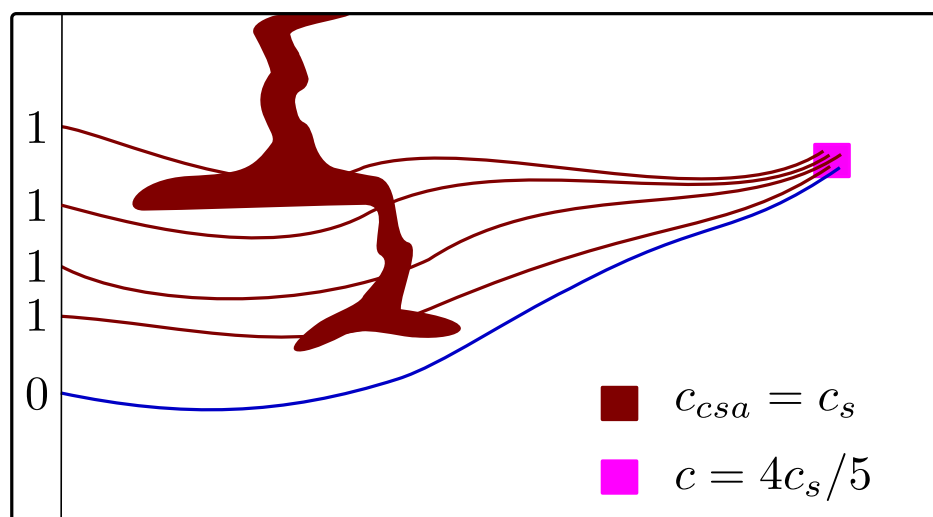


Figure 5.2: Linking the CSA with downstream concentration values, where  $c_s$  is the solubility concentration and  $c_{csa}$  the concentration within the CSA

Finally, one can efficiently approximate the sensitivity matrix  $\mathbf{H}$  and obtain the transfer function in the following simple form:

$$c_{o,j} = \frac{1}{n_p} \sum_i^{n_p} s_i \hat{c} \quad (5.22)$$

with  $c_{o,j}$  being the concentration observation at the  $j^{th}$  measurement location,  $s_i$  being the signal of the  $i^{th}$  particle that states whether the trajectory has ( $s_i = 1$ ) or has not ( $s_i = 0$ ) touched the CSA, and  $n_p$  being the number of released particles at the measurement location  $j$ .

**Test of the proposed reverse formulation** Theoretically, both the forward PTRW method proposed in Chap. 4 and the reverse PTRW mapping should lead to equivalent results of downstream concentration values. However, numerical approximation errors for the trajectories could lead to different results between the backward and forward approach. Moreover, although both the forward and backward models are Lagrangian concepts based on PTRW simulation, their numerical subroutines differ significantly.

1. The forward simulation uses Gaussian kernels to discretize the CSA, since their performance was found to be superior to the performance of box kernels (see Chap. 4). Yet, the reverse simulation discretizes the CSA with flush-mounted box kernels. This is because in the inversion model a different compromise between simulation time and numerical accuracy had to be found, since a tremendously huge ensemble of realizations is required to infer posteriori probabilities of CSA and  $\mathbf{K}$ .
2. Using the forward model, an irregular particle distribution is obtained within the downstream-control plane, which is a consequence of irregular advective and dispersive transport (e.g., due to flow focusing and spreading). In contrast, the reverse model releases particles regularly within this control plane and then tracks them backwards towards the CSA. As such, the reconstruction of concentration values at this control plane poses a different pattern of inaccuracies whether it is obtained from forward (local errors) or backward (global error) tracked particles.
3. Additionally, the information content of individual particles differs whether they are tracked forward or backward. Due to the shadowing effects, the forward particles possess redundant information, which is considered by reducing the particles mass (see Chap. 4), while each particle treats the shadowing effect individually in the reverse simulation.

Nevertheless, Fig. 5.3 shows good agreement between the reverse (left side) and forward (right side) simulation of the concentration distribution in a downstream control plane at 15m distance from the initial spill area of the CSA, applying the model setting as described in Chap. 3 for the forward simulation and in Sec. 5.5 for the reverse simulation.

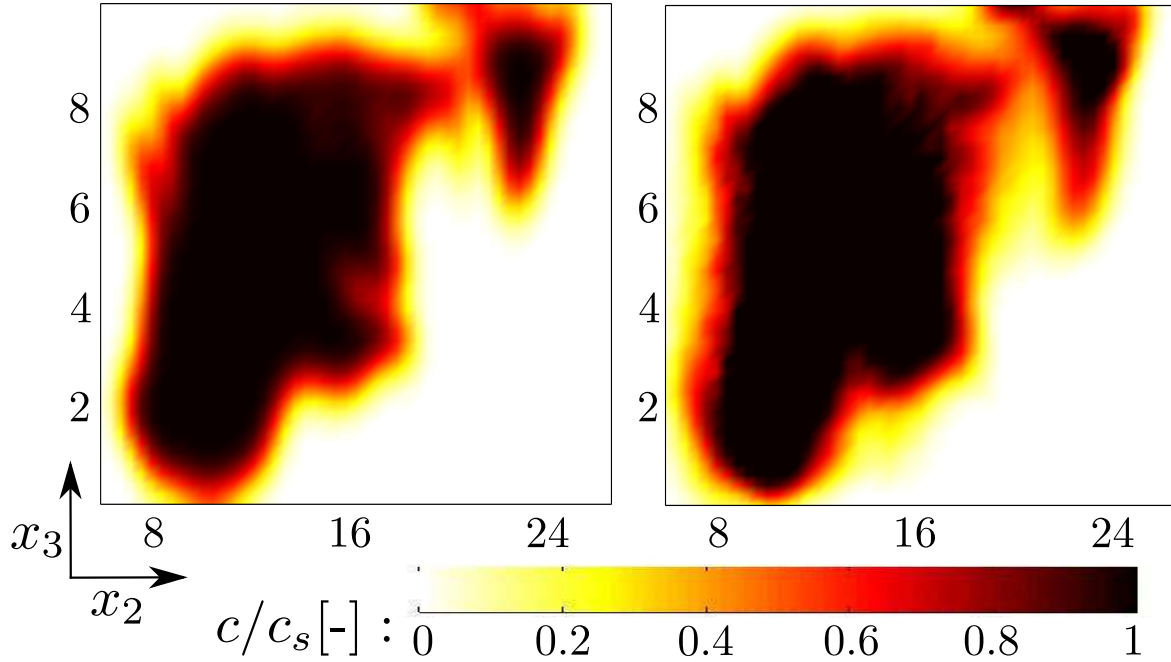


Figure 5.3: Comparison of RPTRW with PTRW: Reverse (left) and forward (right) simulated concentration distribution 15m downstream from the CSA

## 5.5 Application examples

In the following the presented approach of reverse-inverse CSA identification from downstream concentration observations is illustrated and applied to the model framework presented in Chap. 3 (Scenario 1). With the methods presented in Chaps. 2-3, realizations of interdependent random parameter fields for permeability  $\mathbf{K}$  and CSA are generated, within a model domain of  $48 \times 40 \times 10$  m. This domain is discretized with rectangular elements sized  $0.4 \times 0.4 \times 0.2$  m in mean flow, lateral horizontal, and vertical direction, respectively. A summary of the used parameter values is provided in Tab. 3.1 (see Chap. 3).

In order to demonstrate the proposed method, synthetic data is generated with the simulation tools described in Secs. 2.5.3, 3.2, and 5.4. Thus, I generate one realization of permeability field and CSA, both are drawn from the a priori distribution (according to Scenario 1 of Chap. 3), and simulate the corresponding downstream concentration distribution. This is an arbitrary chosen realization that serves in the following as “*synthetic reality*”, providing a synthetic data set. Of course, it would be more interesting to use actual measurements from a contaminated site [e.g., *Brewster et al.*, 1995; *Rivett et al.*, 2001], or taking data from laboratory experiments [e.g., *Zhang et al.*, 2007]. However, applying the proposed reverse-inverse method on actual field (or laboratory) data is not in the scope of this work. The simulated measurement campaign involves 9 drilling cores (DC) for  $\mathbf{K}$  and  $S_n$  observations ( $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$ , and 9 multilevel wells for concentration observations ( $c_o$ ). The design is illustrated in Fig. 5.4, and its specific features are listed in Tab. 5.1. Although it is unknown, the CSA that

serves as synthetic reality is also plotted in Fig. 5.4 for comparison. The presence of DNAPL is only known when this CSA crosses one of the 9 drilling cores, which is obviously the case for DC 5 at the top ( $x_3 = 9.9$ ), since this drilling core crosses the spill center, which is assumed to be known in this specific example. For this specific synthetic reality, DNAPL is also found within DC 7 at  $x_3 = 9.1, 9.3$ , and  $9.5$ . If the CSA touches a DC but does not cross through it (i.e., it occupies an adjacent element), as it occurs for DC 8 at  $x_3 = 9.5$ , viewing the DC and its borehole might reveal some hints that DNAPL is present in the direct vicinity. Hence, one can neither force the conditional CSAs to bypass nor force them to cross the DC at this location. That means, I do not condition the CSAs on these observed values of  $S_n = 0$  at this part of the DC 8 ( $\pm 0.5\text{m}$  in  $x_3$ ).

Table 5.1: Features of the measurement campaign

Type:	permeability	pure-phase DNAPL	dissolved-phase concentration
Symbol:	$\mathbf{K}_o$	$\mathbf{S}_{n,o}$	$\mathbf{c}_o$
Technique:	drilling cores (DC)	drilling cores (DC)	multilevel wells (MW)
Analyse:	quantitative	qualitative	quantitative
Number of observed values:	450	450	45
Locations:	vicinity of spill area	vicinity of spill area	15 m in $x_1$ from spill center
Spatial arrangement:	$3 \times 3 \times 50$ $\Delta x_{1,2} = 3.2\text{m}$ , $\Delta x_3 = 0.2$	$3 \times 3 \times 50$ $\Delta x_{1,2} = 3.2\text{m}$ , $\Delta x_3 = 0.2$	$1 \times 9 \times 5$ $\Delta x_2 = 3\text{m}$ , $\Delta x_3 = 2\text{m}$

Having generated the data set, I proceed as described in Sec. 5.3.2 to generate conditional realization on the basis of the observed data ( $\mathbf{K}_o, \mathbf{S}_{n,o}$ , and  $\mathbf{c}_o$ ), (see Tab. 5.1). To that end, a Monte Carlo simulation with  $10^6$  realizations is run. These realizations are first conditioned on  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$  via rejection sampling on the covariance parameters with respect to  $\mathbf{K}_o$ , direct kriging on the permeability fields, and rejection sampling on the CSA and  $\mathbf{K}$  with respect to  $\mathbf{S}_{n,o}$ . With this, 190 000 pre-conditioned realization are obtained to infer CSA and  $\mathbf{K}$  with respect to  $\mathbf{c}_o$  applying the proposed reverse-inverse inference method Sec. 5.3.

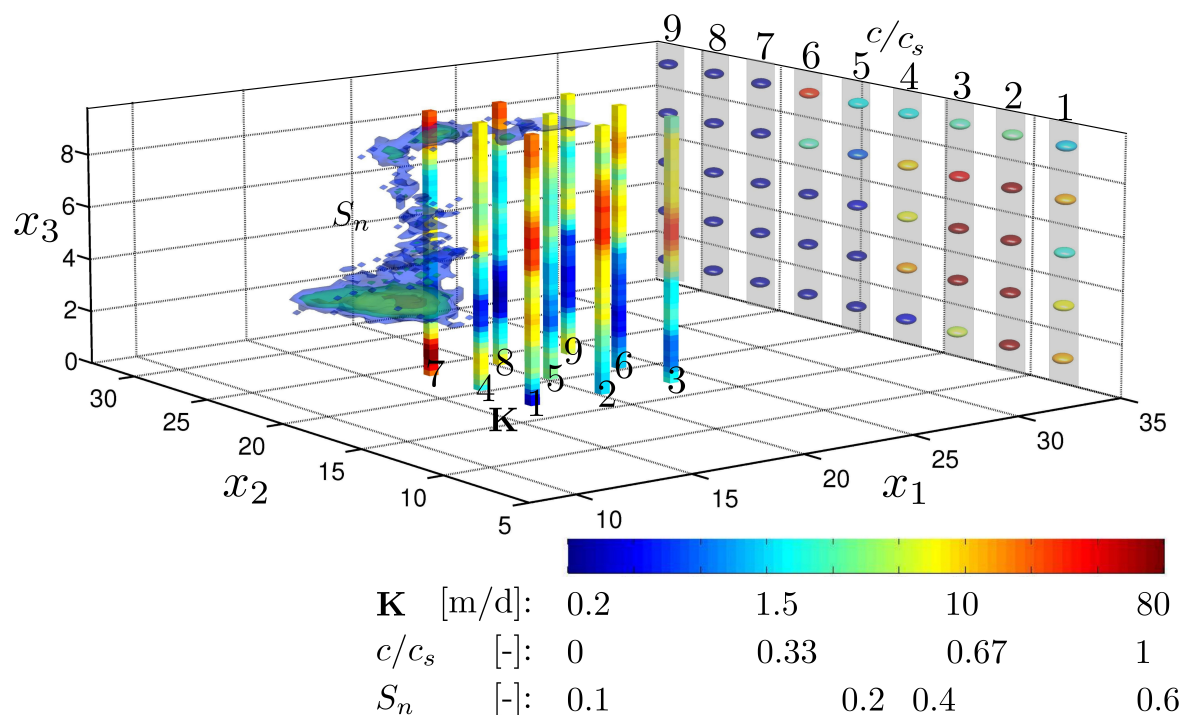


Figure 5.4: Illustration of the measurement design and the unknown CSA of the example application

### 5.5.1 Probability of DNAPL presence

This section will infer knowledge on the spatial patterns and trends of DNAPL residence. Using the entire ensemble of pre-conditioned CSAs (190 000 realizations), one can count the DNAPL occupation frequencies of individual elements in the model domain and then estimate the probability of whether or not this specific element is occupied by DNAPL. In the following, I will analyze the conditional CSAs in two steps: first, only conditioned on  $\mathbf{K}_o$  and  $S_{n,o}$  in Fig. 5.6, and then conditioned also on  $c_o$  in Fig. 5.7. Fig. 5.5 shows the a priori probability of DNAPL residence before assimilating these observations. Obviously, the probability of observing DNAPL saturation within the known spill area equals 1 within all conditioning states (see purple area). From Fig. 5.5 it is evident that the probability of observing DNAPL is rapidly decreasing with increasing distance from this spill area. Hence, the identification of the CSA solely with DNAPL saturation observations is very inefficient.

**Conditioning on  $\mathbf{K}_o$  and  $S_{n,o}$ :** The probability of DNAPL presence within the predefined spill area is obviously one, while the probabilities in the rest of the model domain are influenced by the observations at the 9 DCs. High probability is obtained in the vicinity of the drilling core in which DNAPL presence has been detected in the field. Although the DNAPL data show DNAPL presence at one specific location, the considered measurement error of the qualitative DNAPL observation leads to a reduced probability at this location (less than



one). The probability is increased within the vicinity of the location where DNAPL has been observed for three reasons:

1. Although a perfect match of pure-phase DNAPL observations and simulated DNAPL distribution was not demanded, the conditional CSAs are forced to show DNAPL presence in the direct vicinity of observed DNAPL locations (i.e., within an error bound of  $\pm 0.5\text{m}$ ).
2. The same drilling core that detects DNAPL saturation (DC 8) shows, at the same location, high permeability values and a capillary barrier directly below this DNAPL observation (see Fig. 5.4). This is considered in the conditional  $K$  realizations via Kriging. Hence, pooling of pure-phase DNAPL on top of this capillary barrier is very likely and occurs in many realizations.
3. The spatial distribution of DNAPL saturation values within a physically-realistic CSA possess a correlation that is different to zero. The structure of correlation is clearly anisotropic and behaves completely different at pools than at ganglias. That is, for pools the correlation in the horizontal direction is significantly larger. Thus, detecting the presence of DNAPL at one specific location increases the probability of DNAPL presence also in the direct vicinity.

In summary, the combination of  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$  observations yields a high probability for DNAPL pooling in the direct vicinity of DC 8 at  $x_3 = 9 - 10\text{m}$  (see Fig. 5.6). Therefore, demanding physical and statistical consistency of all involved variables allows valuable inference on the CSA using only the DC observations ( $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$ ).

**Conditioning on  $c_o$ :** In order to increase the knowledge on the unknown CSA, concentration observations  $c_o$  are indispensable. The increased CSA inference performance when using also  $c_o$  can be seen from the clear contrast between the two probability density fields of DNAPL presence in Fig. 5.6 and Fig. 5.7. In the latter, the DNAPL presence of the individual CSA realizations is weighted by the weights obtained by the Bayesian inference with respect to  $c_o$  (see Eq. (5.16) in Sec. 5.3). The spatial pattern of the probability density field in Fig. 5.7 builds upon the spatial pattern observed in Fig. 5.6. Furthermore, the updated spatial pattern reveals significantly increased probabilities for DNAPL presence in the direct vicinity of DC 8, and an expanded region of higher probabilities in the negative direction of  $x_1$ . Also, the probability of DNAPL presence is increased for  $x_2 < 20$  and dramatically reduced for  $x_2 > 20$ .

The obtained probability density field of DNAPL presence indicates several trends on the geometry of the unknown CSA, yet it does not provide evidences for unique conclusions on the unknown CSA. There is a clear trend that, after having passed DC 8, the CSA moves first in the negative direction with respect to  $x_1$  and  $x_3$ , and then bends in the negative direction to  $x_2$ . After this bend the probability drops dramatically, which could be caused by a thin ganglia whose path and shape is highly uncertain. Nevertheless, the lowest iso-surfaces of possible DNAPL presence suggest two principal directions: either the CSA moves several meters in the negative direction of  $x_1$  (10-15m from the spill), or bends furthermore towards lower  $x_2$  ( $x_2 \approx 10$ ). This effect is mainly triggered by the heterogeneous permeability field

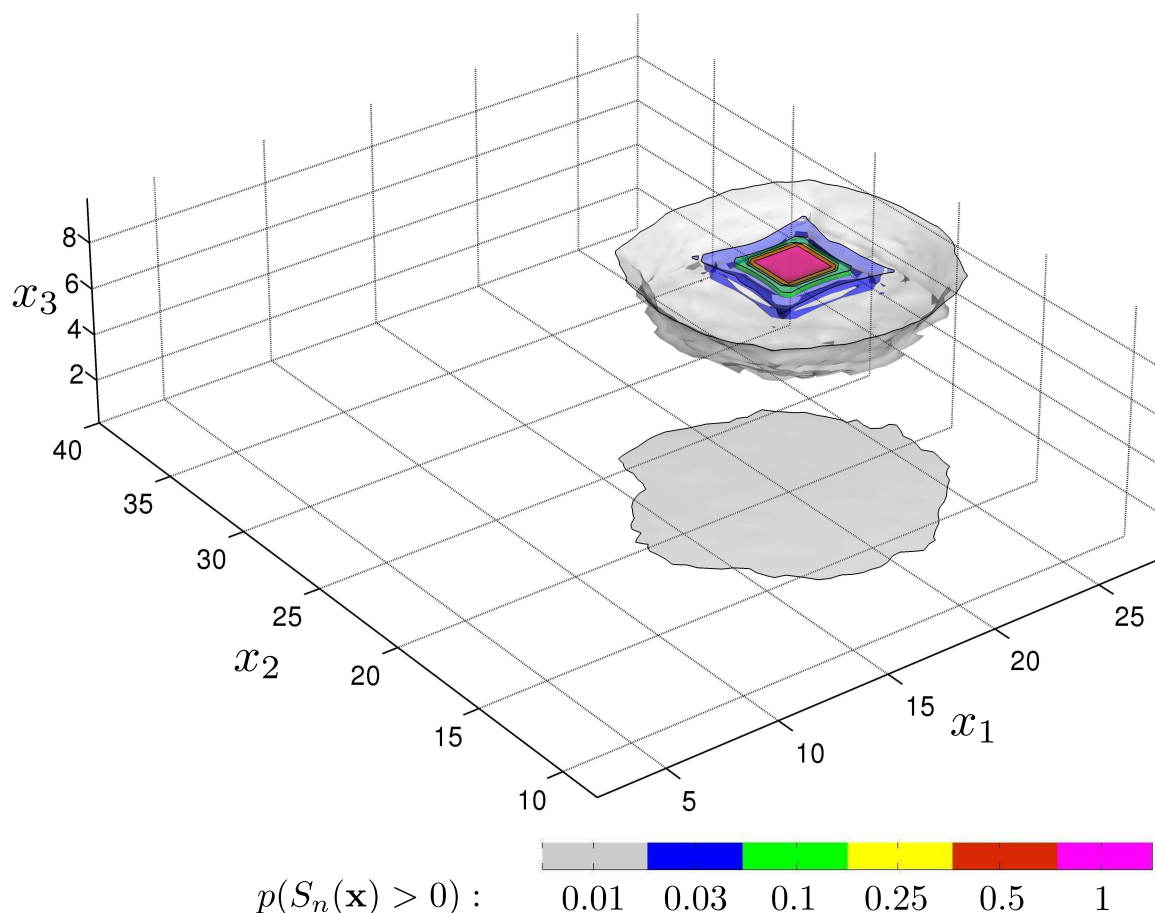


Figure 5.5: Normalized frequency of DNAPL presence for unconditional CSAs

and groundwater flow of the synthetic reality, where solute trajectories moved from the CSA, which has formed a pool around  $\mathbf{x} \approx (10, 15, 5)$ , towards significantly lower values of at least  $x_2 = 8$ . Hence, depending on the permeability field, the conditional realization demands either longer distances between solute release and the multilevel control plane or the release of solutes have to occur far out at low  $x_2$  values, in order to explain the high concentration observations at MW 1. Thus, maintaining the interdependency between CSA and aquifer parameters improves the inversion performance significantly.

For illustration purposes, I plotted a subset of these conditional CSAs and indicated their weights obtained from the Bayesian inference. These weights have also been used for the weighted re-sampling in Fig. 5.7 and in the following discussions in Secs. 5.5.2-5.5.3 for the inference of CSA characteristics and mass discharges, respectively. The CSA of the synthetic reality is plotted at the very top left, for the purpose of comparison. Fig. 5.8 highlights again that various geometries of the CSA are able to reflect the observed data within the measurement error bounds ( $\sigma_\varepsilon^2 = 0.04$ ). This is in line with the statement that the CSA and aquifer parameters have to be inferred jointly. While a few CSA mimic some features of the synthetic reality quite well (e.g., pooling of DNAPL at  $x_3 \approx 5$  and  $x_2 \approx 10 - 15$  in the 4<sup>th</sup>

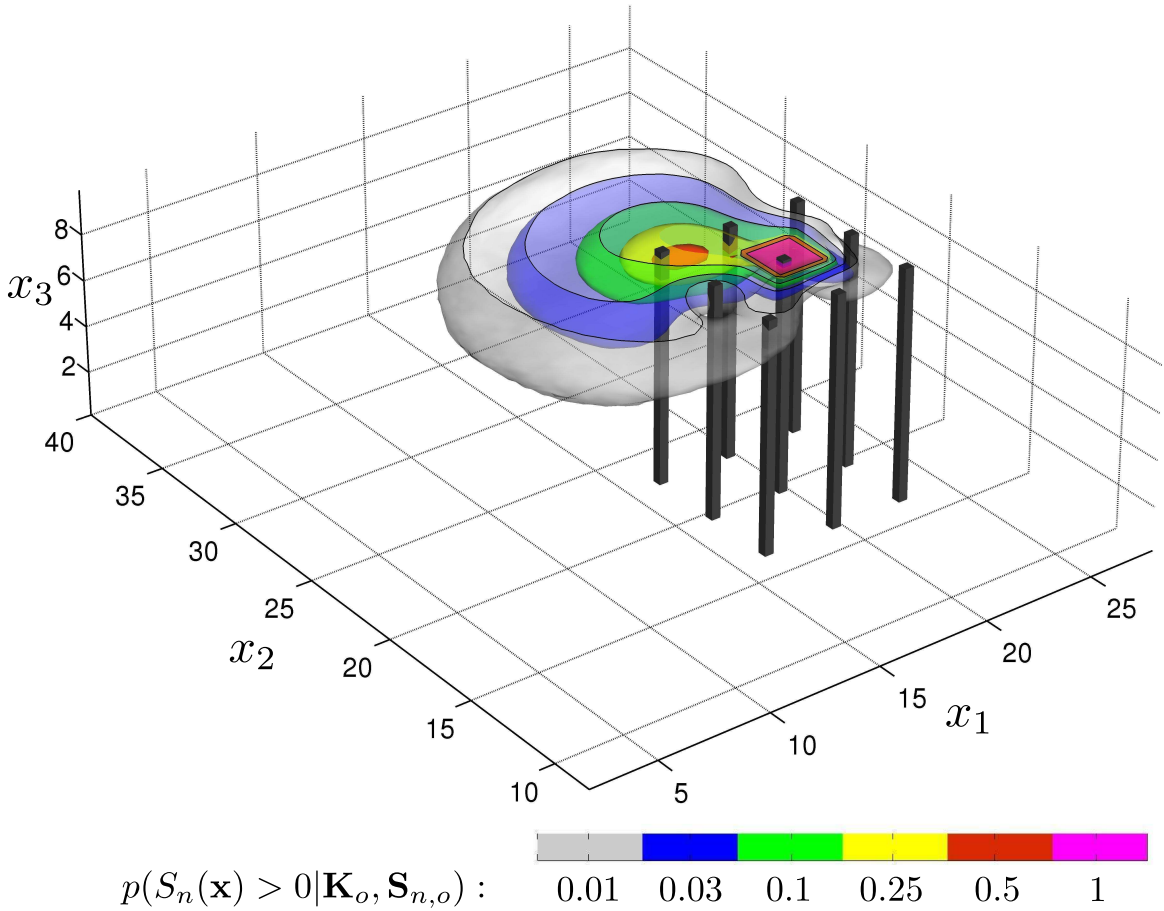


Figure 5.6: Normalized frequency of DNAPL presence for conditional CSAs, after assimilating  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$

row), some CSAs turn out to be significantly different (e.g., the CSAs in the 2<sup>nd</sup> row).

### 5.5.2 Identification of CSA characteristics

Although Fig. 5.7 reveals valuable hints on possible DNAPL locations, the inference of the exact geometry of the unknown CSA is an unfeasible task. However, inferring knowledge on specific characteristics of the CSA could help for designing remediation strategies or assessing the impact of the unknown CSA on natural resources and the drinking water supply. For example, a compact CSA might release less mass per time than a CSA that spans over a wide area perpendicular to the mean flow direction. The current section infers knowledge on CSA characteristics that might help to describe the behavior of the unknown DNAPL/groundwater system. To this end, I selected two common characteristics, which are the ganglia to pool ratio (GTP) and the discharge area of the CSA with respect to flow in  $x_1$  direction ( $A_{x_2, x_3}$ ) [e.g., *Christ et al.*, 2006; *Lenke and Abriola*, 2006; *DiFilippo and Brusseau*, 2011]. These are introduced in the following.

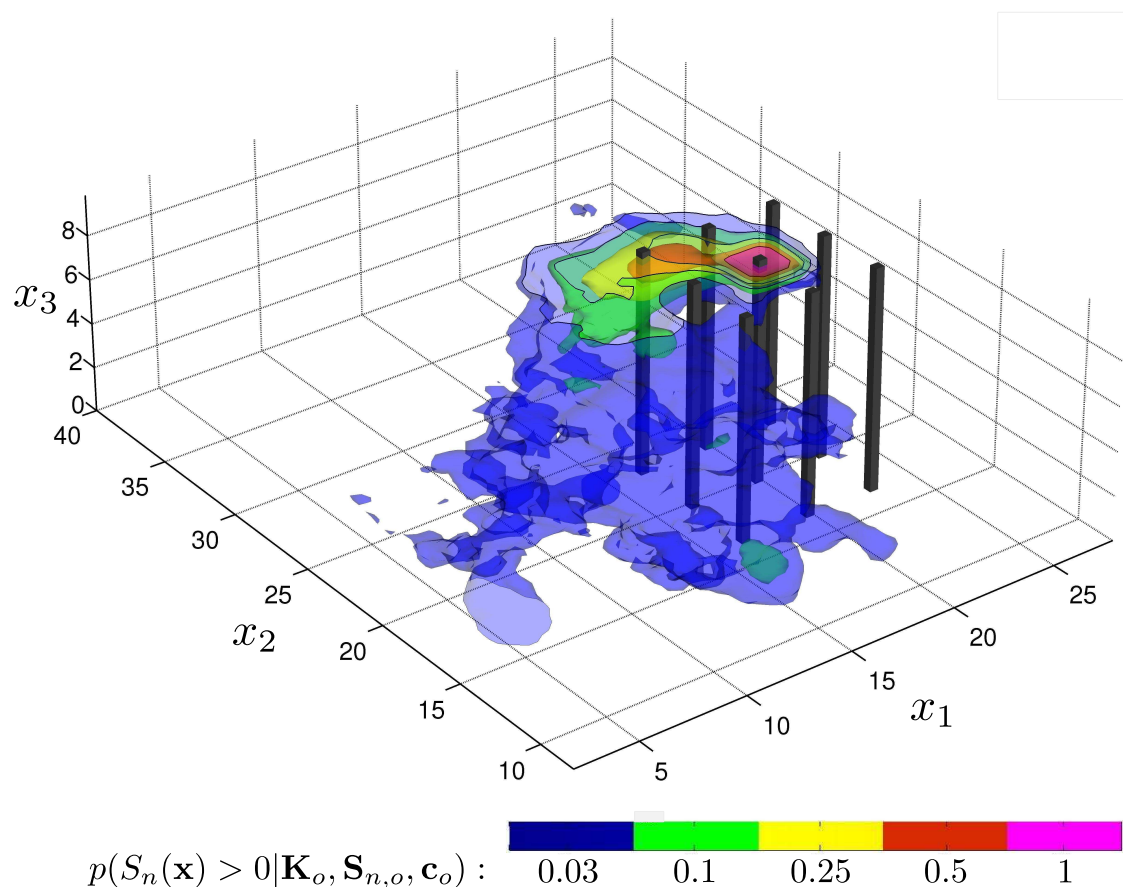


Figure 5.7: Normalized frequency of DNAPL presence for conditional CSAs, after assimilating  $\mathbf{K}_o$ ,  $\mathbf{S}_{n,o}$ , and  $\mathbf{c}_o$

**Ganglia to pool ratio** The GTP is often used to characterize and upscale contaminated sites, because it is said to correlate with the mass discharge emanating from the CSA [e.g., *Christ et al.*, 2006; *Lemke and Abriola*, 2006; *DiFilippo and Brusseau*, 2011]. It is the ratio of DNAPL mass present at less than the residual saturation ( $S_{n,r} = 0.15$ ) to the DNAPL mass present at higher saturation values. Ganglia-dominated CSAs typically lead to higher mass release rates than pool-dominated CSAs. This is because ganglias have a larger specific interfacial area with the groundwater flow available for mass transfer, and because DNAPL pools may act as flow barriers that reduce the effective mass transfer into the groundwater.

**Discharge area** The  $A_{x_2,x_3}$  is the area when the CSA is projected one to one onto a cross section orthogonal to  $x_1$ . Hence it is the CSA discharge area with respect to the mean flow direction  $x_1$ . Accordingly, the label discharge area is actually only valid for homogeneous or axis-symmetric stratified flow fields ( $v_{2,3} = 0$ ). Equivalent to the GTP, it can be expected that for the same total DNAPL mass, larger discharge areas lead to higher mass release rates and to shorter depletion times.

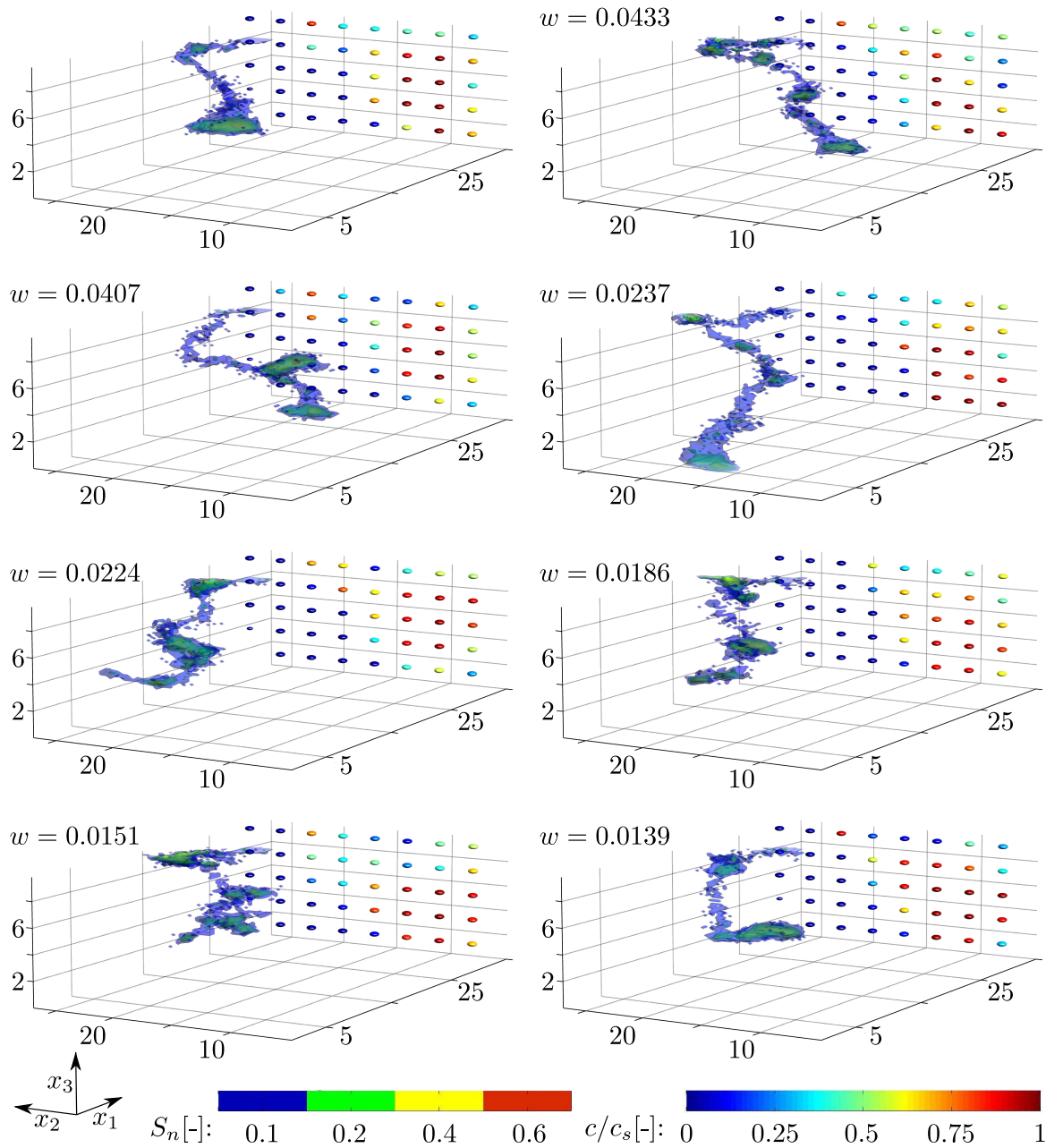


Figure 5.8: Comparison of several conditional CSAs with the CSA of the synthetic reality: The synthetic reality is plotted at the top row on the left side and conditional CSAs are plotted from right to left and top to bottom with decreasing a posteriori weights

However, it is not the scope of this thesis to discuss the explanatory power of these CSA characteristics on mass discharge or depletion time. Instead, this section assesses the explanatory power of concentration observations on these CSA characteristics. In the following, the inference performance for these two CSA characteristics on a few scenarios is discussed. Although this is not sufficient to make general conclusions on the inference performance of CSA characteristic, these examples demonstrate that CSA characteristics can actually be inferred from concentration observations in a Bayesian fashion.

Fig. 5.9 and Fig. 5.10, illustrate the inference of the CSA characteristics on three different scenarios, for  $A_{x_2, x_3}$  and GTP, respectively. First, in Figs. (5.9)-(5.10)a (left side), the applied  $c_o$  results from the same synthetic reality that has also been used for the inversion in the two previous sections. In Figs. (5.9)-(5.10) b and c, the inference result is shown for arbitrary CSA characteristics drawn from their prior distributions (blue line). The a posteriori distributions after the first two conditioning steps (on  $\mathbf{K}_o$  and  $\mathbf{K}_{n,o}$ ), serve here as prior for the inference with respect to  $c_o$ , which is shown in Figs. (5.9)-(5.10). The CSA characteristic of the CSA that caused the concentration observations  $c_o$  are always indicated by the green line. Nevertheless, the CSA and its characteristics are actually unknown.

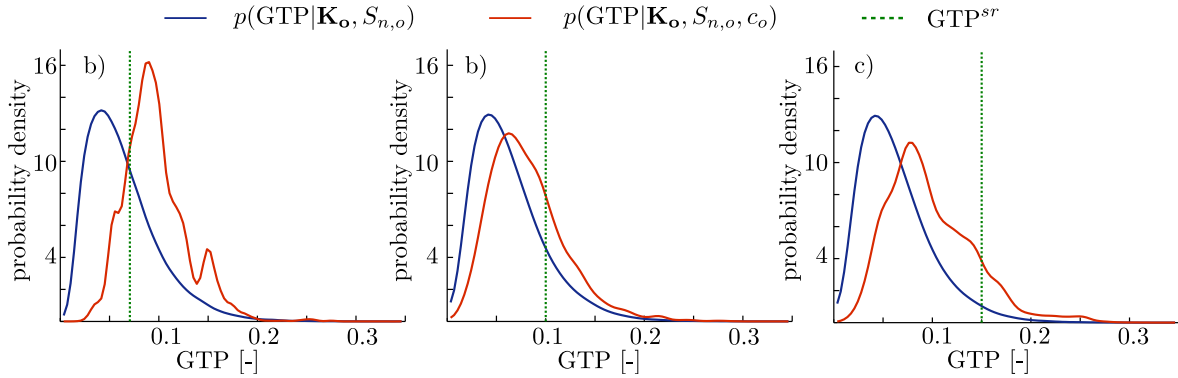


Figure 5.9: Inference of the ganglia to pool ratio (GTP) of the unknown CSA: the intermediate a posteriori distribution (blue line), the updated a posteriori distribution (red line), and the value of the synthetic reality (green line) are shown for three different scenarios with three different synthetic realities (a, b, and c).

Each a posteriori distribution (red line) in Figs. (5.9)-(5.10) (a, b, and c) is pulled towards the value of the CSA characteristic (green line) that has led to the concentration distribution at the downstream control plane (see Fig. 5.4). While in Fig. 5.9a the uncertainty is clearly reduced in the a posteriori distribution (red line) compared to the prior (blue line), the uncertainty of the a posteriori distribution in Figs. 5.9b and c is increased. The latter is due the low a priori probability of the selected synthetic reality. This a well known effect of a strong a priori distribution (here: physically constrained), a small data set (here: 45 concentration observations), and a large measurement error (here:  $\sigma_\varepsilon^2 = 0.04$ ) [e.g. *Gelman et al.*, 2000].

In Fig. 5.10a the uncertainty on  $A_{x_2, x_3}$  is clearly reduced. Low values of  $A_{x_2, x_3}$  are almost precluded, while higher discharge areas than the one used to generate  $c_o$  are more likely. The latter could be again an effect of the permeability field of the chosen synthetic reality. It

causes streamlines that transport the released solutes far out in the transversal direction. If the synthetic reality represents an outlier in the a priori distribution by an extremely large value of  $A_{x_2,x_3}$ , as shown in Fig. 5.10c, the a posteriori distribution shifts towards this outlier but still shows higher probabilities for lower (and thus less extreme) values for  $A_{x_2,x_3}$ .

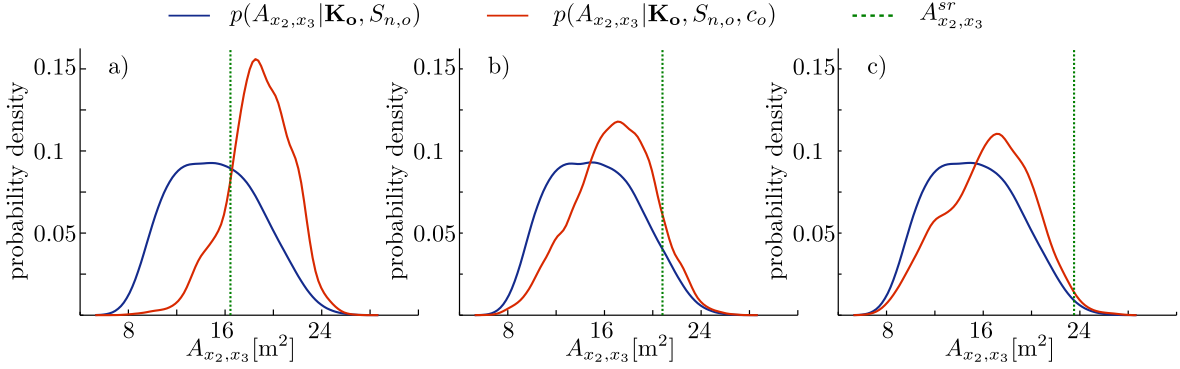


Figure 5.10: Inference of the area spanned by the unknown CSA in  $x_2$  and  $x_3$  direction: the intermediate a posteriori distribution (blue line), the updated a posteriori distribution (red line), and the value of synthetic reality (green line) are shown for three different scenarios (a, b, and c)

### 5.5.3 Conditional probability distribution of mass discharge predictions

In Chap. 3 I assessed the uncertainty of mass discharge and source depletion time predictions and the consistency of the proposed model framework in five different scenarios. Unfortunately, for the most consistent scenario (S1) the mass discharge *pdf* covered several orders of magnitudes. Obviously, this huge uncertainty is not a desirable basis for decision making purposes. Therefore, I will use the reverse-inverse method developed in this section to increase the information on mass discharge on the basis of the observed measurements ( $\mathbf{K}_o$ ,  $\mathbf{S}_{n,o}$ , and  $c_o$ ).

Fig. 5.11 shows the a priori logarithmic mass discharge *pdf* of Scenario 1 from Chap. 3 (black line) and the corresponding a posteriori *pdf* after conditioning on  $\mathbf{K}_o$ ,  $\mathbf{S}_{n,o}$  (blue line) and the updated a posteriori *pdf* after conditioning also on  $c_o$  (red line). The mass discharge uncertainty is significantly reduced after the overall conditioning on  $\mathbf{K}_o$ ,  $\mathbf{S}_{n,o}$ , and  $c_o$  (see red line in Fig. 5.11). This is not the case at the intermediate state of conditioning after the first two conditioning steps with respect to  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$  (blue line). In combination,  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$ , namely the observed DNAPL presence in DC 7 and the observed capillary barrier directly below this DNAPL presence in DC 7 and DC 8 (see Fig. 5.4), yield a high tendency of DNAPL pooling within that region (see Fig. 5.6). Thus, many random CSA realizations which are conditional on  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$  possess a large pool in that region containing a large fraction of the total DNAPL mass. Yet, CSAs with such large pools within that region are rejected when they are filtered against the downstream concentration observations  $c_o$  in the last conditioning step. This leads first to an increasing probability of low mass discharges

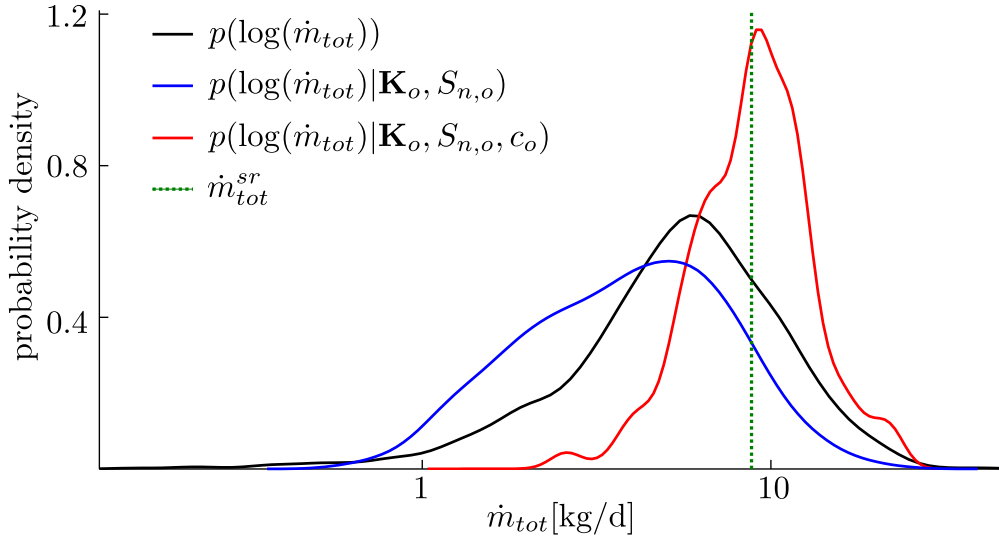


Figure 5.11: Comparison of the a priori *pdf*  $p(\log(\dot{m}_{tot}))$  (S1 in Chap. 3, the intermediate a posteriori *pdf*  $p(\log(\dot{m}_{tot})|\mathbf{K}_o, \mathbf{S}_{n,o})$ , and the updated a posteriori *pdf*  $p(\log(\dot{m}_{tot})|\mathbf{K}_o, \mathbf{S}_{n,o}, c_o)$  of log mass discharge predictions. For illustration, the mass discharge of the synthetic reality  $\dot{m}_{tot}^{sr}$  is indicated by the green line.

(blue line), and then to an dramatically diminished probability of such low mass discharges and an significantly increased probability of higher mass discharges in the last conditioning step (red line). This is because the CSAs who are identified as possible explanation of the observed concentration values all require large passages of ganglias in order to reach regions of low  $x_2$  values, where high concentration values were observed (see Fig. 5.4). Hence, the identified CSAs have substructures with low shadowing effects and thus larger dissolution rates. This is consistent with the observed *pdfs* for the CSA characteristics in Figs. (5.9)-(5.10) (left side) where low GTPs and  $A_{x_2, x_3}$  are observed to be more likely after the conditioning on  $\mathbf{K}_o$  and  $\mathbf{S}_{n,o}$ , than after the conditioning on  $c_o$ .

#### 5.5.4 Discussion on inversion performance

The small wiggles of the final a posteriori *pdfs* (red lines in Figs. (5.9)-(5.11)) indicate that the effective sample size after Bayesian filter is not yet sufficiently large enough for well-founded conclusions. The effective sample size (ESS) is defined as  $ESS = 1 / \sum_i w_i^2$ , with  $w_i$  being the  $i^{\text{th}}$  a posteriori weight. In the presented example, the 190 000 realization conditioned on  $K_o$ ,  $S_{n,o}$ , and  $c_o$  lead to an effective sample size of 123. Hence, in this specific scenario with the arbitrary chosen synthetic reality, the ensemble size that is available for the Bayes filter should be increased furthermore. Yet, all results in Secs. 5.5.1-5.5.3 indicate a clear trend of the a posteriori *pdfs* that is unlikely to change significantly when the Monte Carlo based inversion has converged. In general, to allow quantitative statements on the inversion performance of the presented revers-inverse method, various synthetic realities should be tested in a Monte Carlo simulation. However, this is not the scope of this thesis.



## 5.6 Summary and conclusions

This chapter derived and assessed a novel reverse-inverse methodology to efficiently and accurately infer CSA and aquifer parameters and their characteristics. The approach is to infer the interdependent unknown CSA and aquifer parameters jointly using all available data in a sequential order that ensures both computational efficiency and physical and statistical accuracy. The proposed method considers a priori knowledge on CSAs and aquifer parameters, measurement uncertainty, and the non-uniqueness of CSA and aquifer combinations that cause the observed concentration values in a probabilistic consistent manner. The key challenge is to maintain the strong non-linearities and the non-uniqueness of this inverse problem. To that end, a bundle of sophisticated inverse techniques are assembled and applied to the novel simulation framework which has been introduced in Chap. 3. The proposed method is based on Bayesian inference and the adjoint state method, which have been described in Sec. 5.2.

An efficient and accurate reverse particle tracking random walk method (RPTRW) is introduced in Sec. 5.4, in order to utilize the dramatic reduction of computational costs when solving for the adjoint states for inversion problems whose number of unknowns greatly increases the number of observed values. The advantages of a Lagrangian technique compared to the standard Eulerian approaches are at least as beneficial when solving for the adjoint states as for the forward transport problem (see Chap. 4).

The key conceptual step is utilizing the linearity of the transport equation and evaluating adjoint state sensitivities to obtain a swift transfer function from concentration observations to the likelihood of physically consistent combinations of CSA and aquifer parameters. Following the Bayesian approach to inversion, probabilities of CSA and aquifer parameters which are a priori constrained to be physically plausible are evaluated. To that end, physical knowledge on the DNAPL source formation process and its interdependency to aquifer properties is utilized to restrict the possible outcomes of contaminant distribution. This physically-based approach of joint CSA and  $\mathbf{K}$  inversion restricts the search space dramatically and thus improves and accelerates the inversion performance significantly. If physical and statistical consistency is maintained, observation of aquifer parameters improves the identification of CSA. Vice versa, observations of dissolved concentration may improve the characterization of the aquifer.

Using observed soil properties, pure-phase DNAPL presence, and dissolved-phase concentrations values, it was shown that the proposed method is able to reveal valuable knowledge on the unknown CSA. The obtained pattern of the conditional probability density field of DNAPL presence, reveals a clear trend of pure-phase DNAPL distribution towards the unknown CSA. There was a clear difference whether only observations of aquifer parameter and pure-phase DNAPL presence or also observed dissolved-phase concentration was used to infer CSAs. This suggests that concentration data possess strong explanatory power on CSAs. This observed benefit of concentration data for CSA inversion is only profitably utilized in a physically and statistically consistent framework.

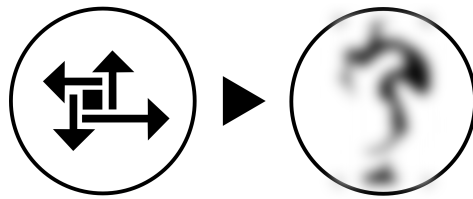
Further, I demonstrated how this method can be used to infer CSA characteristics, such as the ganglia to pool ratio (GTP) or the area of the CSA that is orthogonal to the mean flow

direction. Finally, the method was used to reduce the uncertainty of mass discharge predictions. The uncertainty of mass discharge predictions, which was analyzed in Chap. 3, could be reduced significantly with the proposed reverse-inverse method again mainly due to the downstream concentration observations. This improved knowledge on DNAPL presence, CSA characteristics, and mass discharge could be used to optimize further investigations of the contaminated site or directly to decide whether or not remediation is required and to design a remediation strategy.

### Key conclusions

- The proposed reverse-inverse methodology allows for joint inversion of complex CSAs and aquifer parameters in a physically and statistically consistent framework.
- Due its reverse-inverse concept and its Bayesian framework, this method efficiently and accurately accounts for all relevant interdependencies of a consistent model framework, treats parameter, structural, and observation uncertainty, and reflects the non-uniqueness of the underlying inverse problem.
- The restriction to physically plausible CSAs and aquifer parameters and their joint inversion reduces the solution space and improves the inversion performance significantly.
- The method enables efficient assimilation of various data types observed in the field. The inversion performance with and without the use of concentration values differed dramatically, which suggests that concentration data is highly informative for the CSA inversion.

## 6 Plume deformation and mixing of the dissolved contaminant



The previous chapter inferred information on the CSA using various field observations. This revealed that dissolved-phase concentration values are highly informative in terms of CSA identification and in terms of uncertainty reduction of mass discharge predictions. Hence, this type of data is highly valuable when contaminated sites are investigated. However, mixing and plume deformation of dissolved contaminants annihilates the information content of these concentration observations. This is the motivation to study the driving mechanisms of the mixing process in the current chapter.

Dilution and mixing is enhanced by hydromechanical dispersion stemming from the medium's heterogeneity. The mixing is driven by the generation of concentration gradients, e.g., due to the shear and stretching action of spatially fluctuating flow fields, and their dissipation by micro-scale diffusion. In general, dilution and reaction rates are underestimated when the relevant fine-scale heterogeneities are averaged out in the model-scale, and thus the interaction of heterogeneities, local-scale flow and transport processes are not adequately accounted for [Dentz *et al.*, 2011]. Therefore, profound knowledge on the local-scale mechanisms of mixing affecting the effective dilution and reaction rates on larger scales is required for large-scale numerical and analytical flow and transport models.

This topic has attracted attention across a wide field of disciplines, including research in geophysical flows in the atmosphere and the ocean [e.g., Weiss and Provenzale, 2008] and in porous media flows [e.g., Kapoor and Kitanidis, 1998; Cirpka *et al.*, 2011]. The enhanced mixing is of particular interest for soil remediation strategies, since it can be beneficial for bio-degradation and in-situ chemical oxidation. The same processes accelerate the dissipation of peak concentration values and annihilate the information content of concentration values for CSA identification purposes.

Concentration observations are scalar information of the contaminant plume emanating from the unknown CSA. Hence, there is a drastic difference whether the observed concentration values stem from locations where the shape of the plume is strongly governed by the CSA geometry or the plume is already highly deformed and blurred, due to mixing

and dilution. Thus, the performance of CSA identification and characterization is strongly dependent on the design of the field campaign that yields the data which is used for the inversion. Concentration values possess high information content on small-scale structures of the CSA, if they are observed close to the source zone, and on the large-scale CSA geometry, if they are obtained a little bit further downstream. In contrast, for reliable mass discharge estimates concentration observations with large distance to the source zone are more promising, since the chance of missing a peak concentration values diminishes with larger distances.

Unfortunately, it is hardly possible to identify the locations that reveal high information content in the concentration values beforehand. Optimizing the field campaign's design in a brute-force Monte Carlo framework drastically increases the computational effort, since not only the location of measurement is unknown but also the future measurement values themselves [e.g. *Leube et al.*, 2012]. Alternatively, profound knowledge on the mechanisms that relate to the deformation and mixing process may help to identify measurement locations which are promising to reveal concentration values with high information content on the CSA or on the emanating mass discharge. Therefore, the current chapter studies the mechanisms that relate to the deformation and mixing process in the current chapter and prepares a parametric basis of metrics that quantify the mixing process and mechanism that drive these mixing events. To this end, I apply again the Lagrangian concept on the transport equation using the numerical routines for particle tracking random walk (PTRW) which have been developed and applied in the previous chapters Chaps. 3-5. Building up on this work in future studies, these findings can be used to infer the impact and interplay of various mixing mechanisms on the mixing process. This chapter is based on and partly published in *de Barros et al.* [2012].

## 6.1 Goal, approach and contributions

### 6.1.1 Goal and approach

**Goal** The goal is to examine which and how kinematical and topological features of heterogeneous flow systems affect the mixing of a passive scalar. The focus is on investigating the role of shear and strain deformation, and vorticity for solute mixing in order to identify appropriate hydrodynamic kinematical measures that best indicate the mixing potential of a system. These findings can then be used to identify the potential of mixing and thus to identify promising observation locations (in terms of distance from the CSA) appropriate for the inversion task.

**Approach** This chapter studies the transport behavior of a passive scalar in a spatially variable incompressible flow through porous media. Irregular flow patterns cause spreading of contaminant plumes, leading to erratic distributions of the concentration and hence increasing the contact surface with the surrounding fluid [*Kitanidis*, 1994]. This mechanism leads to enhanced mixing as a result of the interaction between spatial variability of fluid

flow with diffusion and local-scale dispersion [Cirpka *et al.*, 2011]. To address these points, individual aspects of fluid deformation and how they correlate with the dilution index [Kitanidis, 1994] are studied. This chapter will show how specific kinematical measures of steady two- and three-dimensional flow fields can be linked directly to the increase of mixing and dilution of a passive scalar. Therefore, the local mechanisms of mixing in two-dimensional flow for transient solute transport are studied first, then the findings of local mechanisms are transferred to the global mixing dynamics by defining an effective global flow metric, and finally these findings are transferred to three-dimensional flow fields in a stochastic analysis of these mixing mechanisms.

### 6.1.2 Relation to the state-of-the-art

The impact of flow variability on scalar spreading and mixing has been a topic of intensive research in the porous media community [e.g., Kapoor and Kitanidis, 1998; Weeks and Sposito, 1998; Fiori and Dagan, 2000; Dentz *et al.*, 2011; Jha *et al.*, 2011]. The erratic features of flow lines cause the fluid parcel to shear and strain, while at the same time it can cause the rotation of fluid parcels [Kapoor, 1997]. The influence of pure velocity shear on scalar mixing in porous media has also been investigated recently by Bolster *et al.* [2011]. Therefore, the deformation and rotation of fluid parcels represent one of the key drivers for enhanced mixing, and its characterization can help in improving our overall understanding of mixing in porous formations such as aquifers. This is important in applications such as site remediation and natural attenuation of a contaminated system, quantifying concentration uncertainty [Dentz and Tartakovsky, 2010] and reactive transport [Sanchez-Vila *et al.*, 2007].

**Contributions** Although the relation to mixing properties of individual flow-kinematic measures, such as shear strain [Bolster *et al.*, 2011] and vorticity [Kapoor, 1997], has been identified, the relative importance of strain and rotation properties in terms of mixing enhancement is an open research question. Therefore, this chapter analyzes the correlation of mutual interplaying kinematic measures and topological characteristics of the flow system to enhanced mixing.

## 6.2 Basics

**Flow kinematics** To classify the heterogeneous flow field in terms of mixing, elemental components of velocity gradients and their relation to mixing are studied. The velocity gradient is the Jacobian:  $\varepsilon = \nabla \mathbf{v}$ .  $\varepsilon$  can be decomposed into a symmetric part  $\mathbf{S}$  and an asymmetric part  $\mathbf{\Omega}$ :  $\varepsilon = \mathbf{S} + \mathbf{\Omega}$ . While the symmetric part quantifies the rate of strain and hence is called the rate-of-strain tensor:

$$\mathbf{S}_{ij}(\mathbf{x}) = \frac{1}{2} \left( \frac{\partial v_i(\mathbf{x})}{\partial x_j} + \frac{\partial v_j(\mathbf{x})}{\partial x_i} \right), \quad (6.1)$$

the asymmetric part quantifies the rate of rotation and hence is called the rate-of-rotation tensor:

$$\Omega_{ij}(\mathbf{x}) = \frac{1}{2} \left( \frac{\partial v_i(\mathbf{x})}{\partial x_j} - \frac{\partial v_j(\mathbf{x})}{\partial x_i} \right). \quad (6.2)$$

The symmetric part contains stretching- and shear-deformation rates and relates to the total viscous dissipation rate, denoted by  $\theta_s$ :  $\theta_s(\mathbf{x}) \propto \mathbf{S}_{ij}(\mathbf{x})\mathbf{S}_{ji}(\mathbf{x})$ . The three components of the normal strain rate, denoted  $\alpha$  in the following, and of shear strain rate, denoted  $\eta$  in the following, are given by the diagonal and off-diagonal entries of  $\mathbf{S}$ , respectively. The asymmetric part contains the three components of vorticity, denoted  $\omega$  in the following, and relates to the total rotation, denoted by  $\eta$ :  $\eta(\mathbf{x}) \propto \Omega_{ij}(\mathbf{x})\Omega_{ji}(\mathbf{x})$ . This chapter focuses on the relative contribution of stretching deformation, shear deformation and vorticity in each principal direction to an increase of dilution state. This yields 9 mechanisms that may contribute to the mixing process of a passive scalar in three-dimensional flow systems. Following *Okubo* [1970] in two-dimensional studies of divergence free flow fields, this can be simplified to three kinematical measures:

1. stretching deformation  $\alpha = 2\varepsilon_{11}$ ,
2. shear deformation  $\eta = \varepsilon_{12} + \varepsilon_{21}$ , and
3. vorticity  $\omega = \varepsilon_{12} - \varepsilon_{21}$

**Dilution index** Throughout this chapter, mixing and dilution is quantified in terms of the dilution index  $E(t)$  as proposed by *Kitanidis* [1994],

$$E(t) = \exp[S(t)], \quad S(t) = - \int p(\mathbf{x}, t) \log [p(\mathbf{x}, t)] \, d\mathbf{x}, \quad (6.3)$$

where  $p(\mathbf{x}, t)$  is the normalized concentration  $p(\mathbf{x}, t) = c(\mathbf{x}, t) / \int c(\mathbf{x}, t) d\mathbf{x}$ , and  $S(t)$  describes the system entropy. The dilution index is a measure for the effective volume occupied by a dissolved substance and it therefore also represents a measure for the mixing state of a system. The current chapter will study how the mixing mechanisms discussed above relate to the dilution index  $E$  and to related metrics, which are introduced in Secs. 6.3-6.4.

### 6.3 Scalar mixing in two-dimensional spatially heterogeneous flow fields<sup>1</sup>

This section investigates the mechanisms that lead to enhanced scalar mixing in steady two-dimensional heterogeneous flow fields. A local-scale analysis of scalar transport relates the flow topology in terms of the *Okubo-Weiss* parameter [*Okubo*, 1970] to scalar mixing, which is quantified in terms of the entropy of the scalar distribution. Based on these findings, the

<sup>1</sup>This study has been elaborated in cooperation with Felipe P. J. de Barros (Sonny Astani Department of Civil and Environmental Engineering, University of Southern California), and Marco Dentz (Spanish National Research Council, IDAEA-CSIC, Barcelona, Spain), and is published in *de Barros et al.* [2012].

global mixing dynamics are studied with numerical simulations of flow and transport in a heterogeneous porous medium in a Lagrangian framework. In order to assess the impact of the local flow topology on a distributed contaminated plume, an effective Okubo-Weiss function as the average of all positive local Okubo-Weiss values that are occupied by the plume will be defined in the current section. The task of this effective Okubo-Weiss function is to reflect the mixing dynamics which are quantified in terms of the evolution of the system entropy. Hence, it serves as an effective measure of the mixing potential of a heterogeneous, two-dimensional flow system.

Transport of a scalar  $c(\mathbf{x}, t)$  in a spatial random flow field  $\mathbf{v}(\mathbf{x})$  is governed by the advection-dispersion equation:

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} + \mathbf{v}(\mathbf{x}) \cdot \nabla c(\mathbf{x}, t) - D \nabla^2 c(\mathbf{x}, t) = 0. \quad (6.4)$$

For simplicity, a constant and isotropic local dispersion is considered and quantified by the dispersion coefficient  $D$ . Porosity is constant and set to one. The steady state velocity field has zero divergence,  $\nabla \cdot \mathbf{v}(\mathbf{x}) = 0$ .

In the following, the spatial position is non-dimensionalized by the characteristic velocity length scale  $\ell$  and time by the advection time scale  $\tau_v = \ell/\bar{v}$ , in which  $\bar{v}$  is a typical, characteristic flow velocity (e.g., average velocity):

$$\mathbf{x} = \ell \hat{\mathbf{x}}, \quad t = \tau_v \hat{t}. \quad (6.5)$$

Thus, (6.4) reads in dimensionless form as:

$$\frac{\partial \hat{c}(\hat{\mathbf{x}}, \hat{t})}{\partial \hat{t}} + \hat{\mathbf{v}}(\hat{\mathbf{x}}) \cdot \nabla \hat{c}(\hat{\mathbf{x}}, \hat{t}) - \hat{D} \nabla^2 \hat{c}(\hat{\mathbf{x}}, \hat{t}) = 0, \quad (6.6)$$

where,

$$c(\mathbf{x}, t) = \ell^{-d} \hat{c}(\ell \hat{\mathbf{x}}, \tau_v \hat{t}), \quad \mathbf{v}(\mathbf{x}) = \bar{v} \hat{\mathbf{v}}(\ell \hat{\mathbf{x}}), \quad \hat{D} = \frac{D}{\bar{v} \ell}, \quad (6.7)$$

and  $d$  denotes the dimensionality of space. For simplicity of notation, the hats are omitted in the following.

### 6.3.1 Local mechanisms

In order to study the local mechanisms of creation and destruction of scalar gradients, the evolution of a solute pulse that originates from a point-like injection at  $\mathbf{x} = 0$  at time  $t = 0$  is considered. To this end, the transport problem is transformed into the coordinate system moving with  $\mathbf{v}(\mathbf{x})$  [Tennekes and Lumley, 1972]:

$$\mathbf{x} = \mathbf{x}(t) + \mathbf{x}', \quad \mathbf{x}(t) = \int_0^t \mathbf{v}[\mathbf{x}(t')] dt'. \quad (6.8)$$

The concentration in the moving coordinate system is denoted  $g(\mathbf{x}', t) \equiv c[\mathbf{x}' + \mathbf{x}(t), t]$  and satisfies:

$$\frac{\partial g(\mathbf{x}', t)}{\partial t} + \{\mathbf{v}[\mathbf{x}(t) + \mathbf{x}'] - \mathbf{v}[\mathbf{x}(t)]\} \cdot \nabla' g(\mathbf{x}', t) - D \nabla'^2 g(\mathbf{x}', t) = 0. \quad (6.9)$$

Hence,  $g(\mathbf{x}', t)$  studies the second spatial moments of the initially point-like injection (i.e., extensions around the plumes centroid), regardless of the plumes positions that are caused by advective transport. For simplicity of notation, the primes are dropped in the following. As long as the extension of  $g(\mathbf{x}, t)$  due to diffusion is not larger than a typical velocity length scale  $\ell$ , that is, for times that are shorter than the diffusion time scale  $\tau_D = \ell^2/D$ , Eq. (6.9) can be approximated as:

$$\frac{\partial g(\mathbf{x}, t)}{\partial t} + [\boldsymbol{\varepsilon}(t)\mathbf{x}] \cdot \nabla g(\mathbf{x}, t) - D\nabla^2 g(\mathbf{x}, t) = 0. \quad (6.10)$$

Where  $\boldsymbol{\varepsilon}(t)$  is the deformation tensor:

$$\varepsilon_{ij}(t) = \frac{\partial v_i[\mathbf{x}(t)]}{\partial x_j}. \quad (6.11)$$

This section considers solute transport in a two-dimensional flow system (i.e.,  $d = 2$ ). Thus,  $\nabla \cdot \mathbf{v}(\mathbf{x}) = 0$  gives for the normal stress and strain components  $\varepsilon_{22}(t) = -\varepsilon_{11}(t)$ . Please note that the deformation tensor changes with travel distance.

Following *Okubo* [1970], the stretching deformation  $\alpha$ , vorticity  $\omega$ , and shear deformation  $\eta$  in two-dimensional flows are defined as discussed in Sec. 6.2:

$$\alpha = 2\varepsilon_{11}, \quad \omega = \varepsilon_{21} - \varepsilon_{12}, \quad \eta = \varepsilon_{21} + \varepsilon_{12}. \quad (6.12)$$

Thus, the deformation tensor  $\boldsymbol{\varepsilon}$  can be rewritten as:

$$\boldsymbol{\varepsilon} = \frac{1}{2} \begin{pmatrix} \alpha & \eta - \omega \\ \eta + \omega & -\alpha \end{pmatrix}. \quad (6.13)$$

For the initial condition represented by a Dirac function  $\delta(\mathbf{x})$ :  $g(\mathbf{x}, t = 0) = \delta(\mathbf{x})$  and an infinite domain,  $g(\mathbf{x}, t)$  from (6.10) is Gaussian:

$$g(\mathbf{x}, t) = \frac{\exp\left\{-\mathbf{x} \cdot [2\boldsymbol{\kappa}(t)]^{-1} \mathbf{x}\right\}}{2\pi \det[\boldsymbol{\kappa}(t)]}, \quad (6.14)$$

which is characterized by zero mean and the spatial variance matrix  $\boldsymbol{\kappa}(t)$  due to the moving coordinate system that is defined for  $g(\mathbf{x}, t)$  in Eq. (6.9). Hence, following *Tennekes and Lumley* [1972] the spatial variance matrix  $\boldsymbol{\kappa}(t)$  satisfies:

$$\frac{d\boldsymbol{\kappa}(t)}{dt} = \boldsymbol{\varepsilon}(t) \cdot \boldsymbol{\kappa}^T(t) + \boldsymbol{\kappa}(t) \cdot \boldsymbol{\varepsilon}^T(t) + 2D\mathbf{I} \quad (6.15)$$

where  $\mathbf{I}$  denotes the identity matrix. The dilution index for the Gaussian distribution (6.14) is given by:

$$E(t) = 2\pi e\sqrt{\det[\boldsymbol{\kappa}(t)]} \quad (6.16)$$

Within a distance shorter than a typical velocity length scale  $\ell$ , or equivalently, for times smaller than the advection time scale  $\tau_u$ , the deformation properties are assumed to be constant,  $\boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}$ . Then solving for Eq. (6.15) and inserting Eq. (6.13) leads to the following explicit solutions for the second spatial moments  $\kappa_{ij}(t)$ :

$$\kappa_{11}(t) = \frac{2D}{\theta^3} \left\{ (\alpha^2 + \eta^2 - \eta\omega) \sinh(\theta t) + \theta\alpha \cosh(\theta t) - \theta[\alpha - (\eta - \omega)\omega t] \right\} \quad (6.17)$$



$$\kappa_{22}(t) = \frac{2D}{\theta^3} \{ (\alpha^2 + \eta^2 + \eta\omega) \sinh(\theta t) - \theta\alpha \cosh(\theta t) + \theta[\alpha - (\eta + \omega)\omega t] \} \quad (6.18)$$

$$\kappa_{12}(t) = \frac{2D}{\theta^3} [\theta\eta \cosh(\theta t) + \alpha\omega \sinh(\theta t) - \theta(\eta + \alpha\omega t)]. \quad (6.19)$$

Here,  $\theta$  is defined as the square-root of the Okubo-Weiss parameter [Okubo, 1970; Weiss and Provenzale, 2008]  $\Theta = -4 \det(\varepsilon) \equiv \theta^2$ . It is given by:

$$\theta = \sqrt{\theta_s^2 - \omega^2}, \quad \theta_s = \sqrt{\alpha^2 + \eta^2}. \quad (6.20)$$

where  $\omega^2$  is enstrophy, which quantifies the intensity of vorticity, and  $\theta_s$  quantifies the strength of stretching and shear deformation. The Eigenvalues of the variance matrix  $\kappa(t)$  can be obtained from Eqs. (6.17)-(6.19) which leads to:

$$\lambda_{1/2}(t) = \frac{2D}{\theta} \left\{ \frac{\theta_s^2}{\theta^2} \sinh(\theta t) - \frac{\omega^2}{\theta^2} \theta t \pm \frac{\theta_s}{\theta} \sqrt{[\cosh(\theta t) - 1]^2 + \frac{\omega^2}{\theta^2} [\sinh(\theta t) - \theta t]^2} \right\}. \quad (6.21)$$

They describe the extensions of the concentration distribution in the principal directions. This eigenvalues are completely controlled by the two parameters  $\omega$  and  $\theta_s$ , as can be seen in Eq. (6.21).

The character of the variance matrix depends on the sign of  $\Theta$ . For negative Okubo-Weiss parameter ( $\Theta < 0$ ),  $\theta$  is imaginary (due to  $\theta = \sqrt{\Theta}$ ) and the hyperbolic functions in Eqs. (6.17)-(6.19) transform to trigonometric functions, which are bounded between  $\pm 1$ . In this case, the leading behavior of the  $\kappa_{ij}(t)$  is linear with time. For positive Okubo-Weiss parameter ( $\Theta > 0$ ), the hyperbolic functions indicate an exponential increase of the  $\kappa_{ij}(t)$  with time, and with  $\theta$ , e.g.:

$$\cosh(\theta t) = 1/2 [\exp(\theta t) + \exp(-\theta t)] \quad (6.22)$$

Thus, for  $\Theta > 0$ , the volume occupied by the solute will increase significantly faster than for  $\theta < 0$ . Physically, when  $\Theta > 0$ , shear and normal strain dominates, whereas for  $\Theta < 0$ , vorticity dominates. In the following, the impact of local stress-strain as well as shear and vorticity on the local mixing properties is studied.

In the absence of vorticity,  $\omega = 0$  and thus  $\theta = \theta_s$ , the Eigenvalues read as:

$$\begin{aligned} \lambda_1(t) &= \frac{2D}{\theta_s} [\exp(\theta_s t) - 1] \\ \lambda_2(t) &= \frac{2D}{\theta_s} [1 - \exp(-\theta_s t)]. \end{aligned} \quad (6.23)$$

This means that there is stretching in one principal direction and compression in the other. In this situation, there is an equilibrium between compression and diffusive extension [Batchelor, 1959], and the asymptotic extension in the direction of compression defines the corresponding Batchelor scale,  $\ell_B = \lim_{t \rightarrow \infty} \sqrt{\lambda_2(t)} = \sqrt{2D/\theta_s}$ . In the presence of vorticity, the system behaves differently. In this case, an equilibrium between compression and diffusion cannot be established due to the rotation of the plume: Rotation prevents a persistent compression of the plume in a given direction.

From (6.21) and (6.16), the compact expression for  $E(t)$  is obtained:

$$E(t) = \frac{4\pi De}{\theta} \sqrt{\frac{2\theta_s^2}{\theta^2} \cosh(\theta t) - \frac{2\theta_s^2}{\theta^2} - (\omega t)^2}. \quad (6.24)$$

For  $\Theta > 0$ , and  $\theta t \gg 1$ , the dilution index increases exponentially fast with time as:

$$E(t) \sim \frac{4\pi De\theta_s}{\theta^2} \exp\left(\frac{\theta t}{2}\right), \quad (6.25)$$

while for  $\Theta < 0$ , the asymptotic behavior for  $\omega t \gg 1$  is linear with time:

$$E(t) \sim \frac{4\pi Dewt}{\sqrt{\omega^2 - \theta_s^2}}. \quad (6.26)$$

Note that  $E(t)$  is always larger than for a homogeneous medium, for which  $E_0(t) = 4\pi eDt$  [Kitanidis, 1994]. From (6.24), the rate of increase of the logarithm of the dilution index is obtained as:

$$\frac{d \log[E(t)]}{dt} = \theta \left[ \frac{\sinh(\theta t) - \frac{\omega\theta}{\theta_s^2} \omega t}{2 \cosh(\theta t) - 2 - \frac{\theta_s^2}{\theta^2} (\omega t)^2} \right]. \quad (6.27)$$

For  $\Theta < 0$  and  $\sqrt{-\Theta}t \gg 1$ , the rate of increase of  $\log[E(t)]$  behaves linear with time:

$$\frac{d \log[E(t)]}{dt} \sim \frac{1}{t}, \quad (6.28)$$

which is the same as the one obtained in a constant flow field [Kitanidis, 1994]. For  $\Theta > 0$  and  $\theta t \gg 1$ , the rate of increase of  $\log[E(t)]$  is constant:

$$\frac{d \log[E(t)]}{dt} \sim \frac{\theta}{2}. \quad (6.29)$$

This result indicates that the local mixing and dilution dynamics are controlled by the value of  $\theta$ . The derivations presented above are valid for times smaller than the advective time scale. In the following, the evolution of the dilution index for transport in a heterogeneous porous medium as a function of the local values of  $\theta$  is studied.

### 6.3.2 Global mixing

This section analyzes the evolution of  $E(t)$  of a solute plume that moves through a heterogeneous porous medium, and specifically, its relation to the local value of the Okubo-Weiss parameter  $\Theta(\mathbf{x})$ . Flow is given by the Darcy equation  $\mathbf{v}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x})$ , where  $K(\mathbf{x})$  is the hydraulic conductivity and  $h(\mathbf{x})$  the hydraulic head.  $K(\mathbf{x})$  is modeled as a lognormally distributed random field such that  $Y(\mathbf{x}) = \log K(\mathbf{x})$  is a Gaussian random field. Second-order stationary  $K(\mathbf{x})$  fields with a Gaussian covariance function are generated as described in Sec. 2.5. The geometric mean conductivity is set to  $K_g = 10^{-4} \text{ ms}^{-1}$ , the variance is  $\sigma_Y^2 = 4$ , and the correlation length (isotropic) is  $\ell = 10 \text{ m}$ .

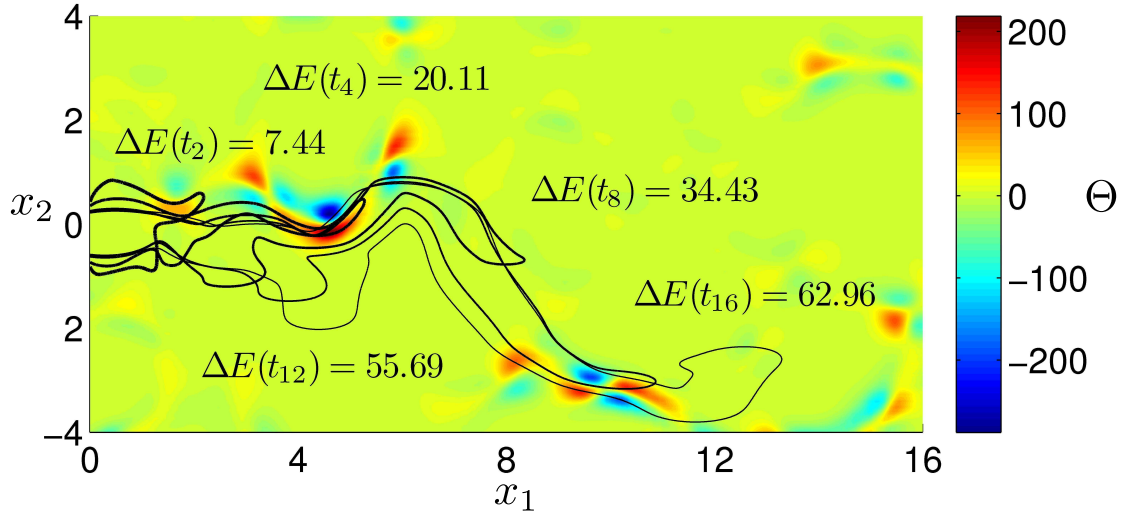


Figure 6.1: Map of the local values of the Okubo-Weiss parameter  $\Theta(\mathbf{x})$  for aquifer realization 1, and the  $p = p_0(t)$  isolines (6.30) at different times, together with the difference of the dilution indices at successive times.

The domain spans an area of  $20\ell \times 12\ell$ . The horizontal boundaries are no-flow boundaries. The vertical boundaries are defined as Dirichlet boundaries that enforce an average head gradient  $G = 10^{-1}$  from left to right. Thus, the mean flow velocity is  $\bar{v} = K_g G = 10^{-5} \text{ ms}^{-1}$ , and the advection time scale is  $\tau_v = 10^6 \text{ s}$ . The flow problem is solved with the standard Galerkin finite element code described in *Nowak et al.* [2008]. An area of  $\ell^2$  is discretized by  $10^4$  regular quadratic elements. As indicated in (6.5), all lengths are measured in terms of the correlation length  $\ell$  and times in terms of the advection time  $\tau_v$ .

The transport equation (6.4) is solved by random walk particle tracking based on the equivalent Langevin equation  $d\mathbf{x}(t)/dt = \mathbf{v}[\mathbf{x}(t)] + \sqrt{2D}\boldsymbol{\xi}(t)$ , in which  $\boldsymbol{\xi}(t)$  is a Gaussian white noise characterized by zero mean and unit variance (this is analogous to Chap. 4 and Sec. 5.4). The dimensionless diffusion coefficient is set to  $D = 10^{-5}$ . The concentration field is expressed in this framework as  $c(\mathbf{x}, t) = \langle \delta[\mathbf{x} - \mathbf{x}(t)] \rangle$ , where the angular brackets stand for the white noise average. Initially, 812040 particles are uniformly placed within a rectangle of dimensions  $l_1 \times l_2$  with  $l_1 = 1$  and  $l_2 = 2$ , centered in  $(x_1, x_2) = (1/2, 0)$ . Thus, the initial concentration is  $c_0 = 1/2$ . The left domain boundary is at  $x_1 = -2$ , the right at  $x_1 = 18$ , the horizontal boundaries are at  $x_2 = \pm 6$ . The boundaries do not affect transport for the simulation times chosen. To quantify the mixing state at specific time steps, we reconstruct the  $c(\mathbf{x}, t)$  from the particle locations with a kernel density estimator as described in Sec. 4.2. In the following, the global mixing properties are studied in 4 different realizations of the random medium defined above.

Figure 6.1 illustrates a map of  $\Theta(\mathbf{x})$  for realization 1 of the random flow field  $\mathbf{v}(\mathbf{x})$ , the isolines of  $p(\mathbf{x}, t) = p_0(t)$ , at five different times and the difference of the dilution index of the plumes at successive times  $\Delta E(t_i) = E(t_i) - E(t_{i-1})$ . Here, and in the following realizations, a constant time step  $\Delta t = 2/5$  so that  $t_i = i\Delta t$  is chosen. The concentration isoline

$p_0(t)$  defines a concentration threshold such that:

$$\int p(\mathbf{x}, t) H[p(\mathbf{x}, t) - p_0(t)] d\mathbf{x} = 0.99. \quad (6.30)$$

The plume bends in regions of negative  $\Theta(\mathbf{x})$ , for which vorticity  $\omega(\mathbf{x})$  dominates over shear. One can observe significant increases of the dilution index when the plume comprises areas characterized by a large positive  $\Theta(\mathbf{x})$  at times  $t_{12}$ , and  $t_{16}$ .

The relation between the global mixing properties and the values of the Okubo-Weiss parameter within the plume volume is illustrated in a more quantitative manner in Figure 6.2. According to (6.29), a significant local increase in dilution can be expected for positive values of the Okubo-Weiss parameter. In order to study the impact of the local flow topology on the plume scale, the effective Okubo-Weiss function  $\Theta^e(t) > 0$  is defined as the average of  $\Theta(\mathbf{x}) > 0$  over the area occupied by the plume

$$\Theta^e(t) = \frac{\int \Theta(\mathbf{x}) H[\Theta(\mathbf{x})] H[p(\mathbf{x}, t) - p_0(t)] d\mathbf{x}}{\int H[\Theta(\mathbf{x})] H[p(\mathbf{x}, t) - p_0(t)] d\mathbf{x}}, \quad (6.31)$$

where  $H(x)$  is the Heaviside step function. The concentration threshold  $p_0(t)$  is defined by (6.30). Figure 6.2 shows  $\Theta^e(t)$  for four different realizations of  $K(\mathbf{x})$ . The effective  $\Theta^e(t)$  is a measure for the mixing potential due to the 'hot spots' (areas with  $\Theta(\mathbf{x}) > 0$ ) within the area covered by the solute.

Figure 6.2b shows the time evolution of  $\log[E(t)]$  for four different realizations of  $K(\mathbf{x})$ . At small times, the dilution indices for the different realizations are approximately equal and given by the value for a homogeneous medium. For increasing time, their evolutions are different and clearly related to the evolution of  $\Theta^e(t)$ . For realization 1,  $\Theta^e(t)$  shows an increase at the third and fourth time step ( $t = 1.2$  &  $1.6$ ) to relatively high values, which is directly reflected in a steep increase of  $\ln[E(t)]$ , as expected from the local scale analysis. As  $\Theta^e(t)$  decreases to smaller values with increasing time, the slope of  $\log[E(t)]$  likewise becomes slightly smaller.

In realization 2, small values of  $\Theta^e(t)$  between the third and sixth time steps (from  $t = 1.2$  to  $t = 2.4$ ) are reflected by a milder slope of  $\log[E(t)]$  than for realization 1. As  $\Theta^e(t)$  increases, the slope of  $\log[E(t)]$  also increases. At late times, the values of  $\Theta^e(t)$  become smaller and the  $\log[E(t)]$  curve accordingly becomes flatter. A clear correlation between the behavior of  $\Theta^e(t)$  and  $\log[E(t)]$  can also be observed in realization 3. Small values of  $\Theta^e(t)$  at the first seven time steps (from  $t = 0.4$  to  $t = 2.8$ ) are accompanied by a moderate rate of increase of  $\log[E(t)]$ . Consistently large values of  $\Theta^e(t)$  are obtained at larger time steps ( $t > 3$ ), which corresponds to consistently larger rate of increase in  $\log[E(t)]$ . For realization 4, values of  $\Theta^e(t)$  of around 0 are observed within the first eight time steps (from  $t = 0.4$  to  $t = 3.2$ ). The growth rate is similar to the one for a homogeneous medium as indicated by the local scale analysis, see (6.28). Then between time step nine and fifteen ( $t > 3.5$ ), when  $\Theta^e(t)$  increases moderately to larger values, the rate of increase of  $\log[E(t)]$  also increases to a larger value.

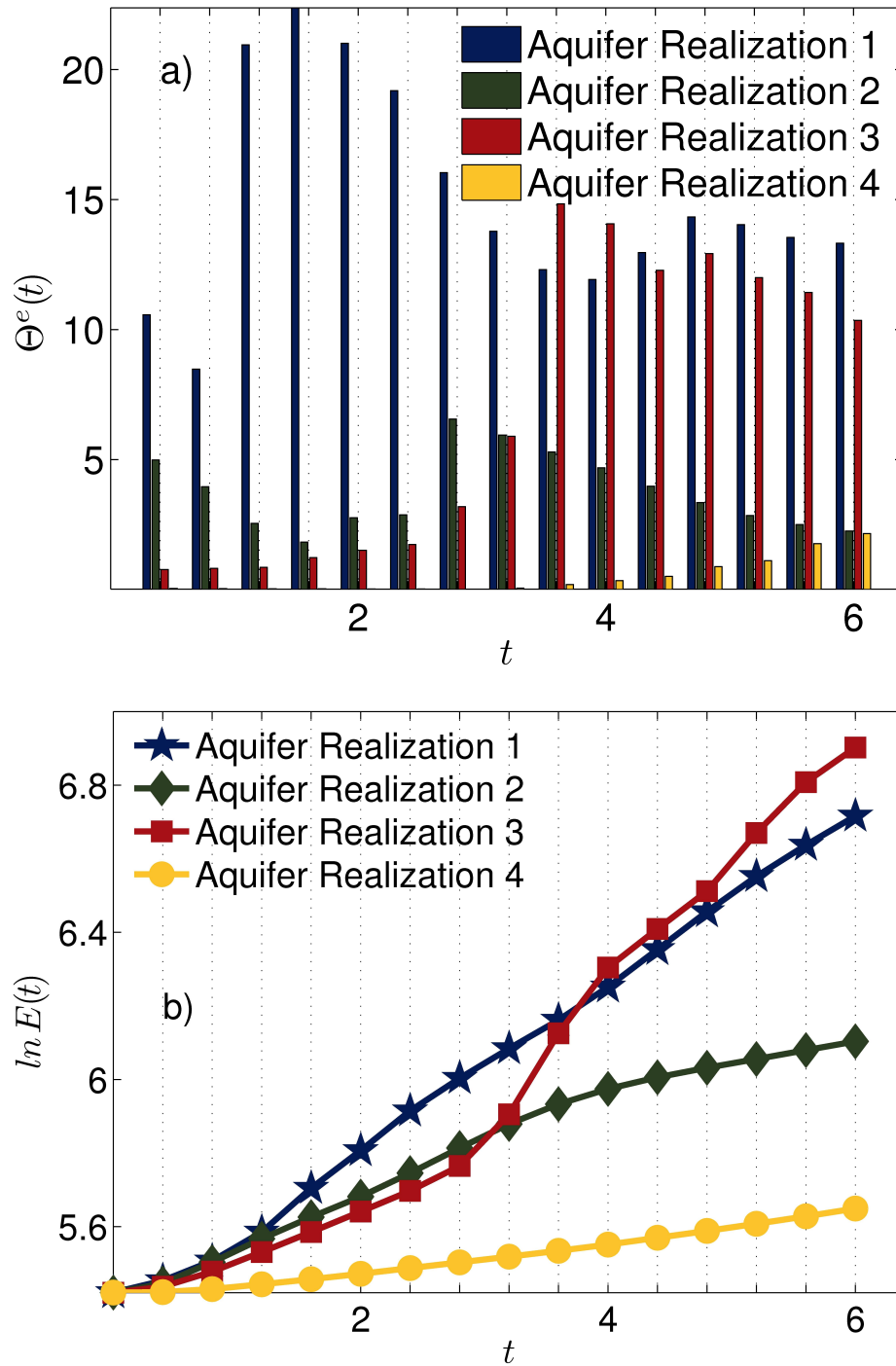


Figure 6.2: (a) Effective Okubo-Weiss parameter  $\Theta^e(t)$ , (6.31), versus time; (b) System entropy  $S(t) = \log[E(t)]$  versus time.

Overall, these examples demonstrate that the evolution of the effective Okubo-Weiss function  $\Theta^e(t)$  reflects very well the mixing dynamics as quantified in terms of the evolution of the system entropy  $S(t) = \log[E(t)]$ , and thus serves as an effective measure of the mixing potential of a two-dimensional, heterogeneous flow system.

## 6.4 Mechanisms of enhanced mixing in three-dimensional heterogeneous flow fields

In this section I will outline an approach to transfer the findings of the two-dimensional study (Sec. 6.3) to three-dimensional flow systems. Therefore, mixing and dilution of a dissolved-phase plume in three-dimensional heterogeneous porous media within stationary dissolved-phase transport is studied. The analysis of two-dimensional flow systems showed how specific hydrodynamic features of a Darcian flow, namely normal strain, shear strain, and vorticity can be linked directly to the increase of mixing and dilution of a passive scalar. The interplay of these mechanisms, however, may differ in three-dimensional flow systems. In three-dimensional domains, flow trajectories may twist and warp. This might re-shuffle the initial spatial structure and thus generate even more contact surface with the surrounding fluid. In general, kinematical mechanisms such as shear and normal strain deformations gain an additional dimension to work with.

Numerical simulations will illustrate how these mechanisms are also among the key drivers for enhanced mixing in three-dimensional porous media. The question is raised, under what circumstances and to which extent can the additional degrees of freedom in the hydrodynamic mechanisms further enhance mixing and dilution in spatially heterogeneous problems. In this numerical study, metrics of the mixing behavior for a scalar plume emitted from a continuous source are quantified. Within a stochastic analysis, the impact of several hydrodynamic features on the pre-defined mixing metrics is indicated.

### 6.4.1 Numerical experiment

The previous study showed that a passive scalar, being transported through a heterogeneous porous medium, experiences various types of local mixing events. Looking only at a small passage of the heterogeneous medium, there are only a few mixing events that can be directly linked to the flow topology of this small passage. To identify the features of flow topology that lead to enhanced mixing, individual mixing events are studied in a numerical experiment. To this end, an experimental design for solute transport within a cuboid aquifer is conceptualized and simulated. In this numerical experiment, the hypothetical aquifer consists of three perfectly connected chambers ( $S_1$ ,  $S_2$ ,  $S_3$ ) that possess different material properties and transport parameters. The second chamber ( $S_2$ ) is designed such that natural conditions are anticipated and enhanced mixing is expected, while the first ( $S_1$ ) and third ( $S_3$ ) chambers only represent an initial smoothing and a post-relaxation phase, respectively. This set-up analyzes the effect of individual mixing events on the direct and long-term dilution process. The initially smoothed signal (in  $S_1$ ) is deformed in  $S_2$  and finally relaxed in  $S_3$ .

Within the relaxation part ( $S_3$ ), the intensity, persistence, and total enhanced dilution due to the induced mixing in  $S_2$  is observed.

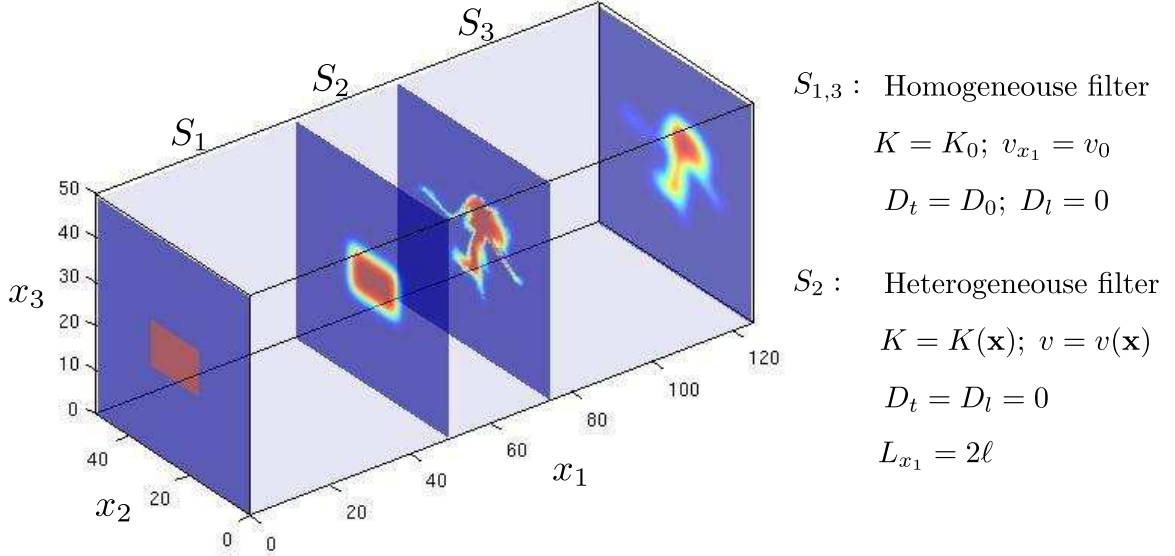


Figure 6.3: Experimental setup of the numerical simulations

The homogeneous filter in  $S_1$  and  $S_3$  with a constant hydraulic conductivity  $K_0$  leads to a uniform velocity  $v_0$ . Hence, the stationary transport problem can be solved with Greens functions. The hydraulic conductivity  $K(\mathbf{x})$  in  $S_2$  is a random second-order stationary field with  $\sigma_Y^2$  and  $\ell$  as variance and integral scale, respectively, of the Gaussian covariance function. The length of  $S_2$  in mean flow direction is  $L_x = 2\ell$ . Thus, the deformation of the plume in  $S_2$  is similar to a local deformation event of irregular flow in an aquifer. The flow in  $S_2$  is solved with a standard Galerkin FEM method as described in *Nowak et al.* [2008]. With a Lagrangian mapping from the inlet to the outlet of  $S_2$ , I observe the deformation that the plume experiences within  $S_2$ . In  $S_3$ , a different dilution process for each realization of  $K(\mathbf{x})$  is observed and compared to the reference case with homogeneous  $K_0$  also in  $S_2$ . Finally, I relate the mixing enhancement to the kinematic flow mechanism observed in  $S_2$ .

#### 6.4.2 Metrics of scalar mixing enhancement

In order to quantify the strength of various local mixing events on the basis of the dilution index, three metrics are introduced in the following and illustrated in Fig. 6.4

**Initial enhancement** The initial enhancement assesses the direct dilution response of the induced perturbation within  $S_2$ . Hence, the initial dilution rate observed at the inlet of  $S_3$  is compared to the dilution rate at the same location but without the deformation and mixing

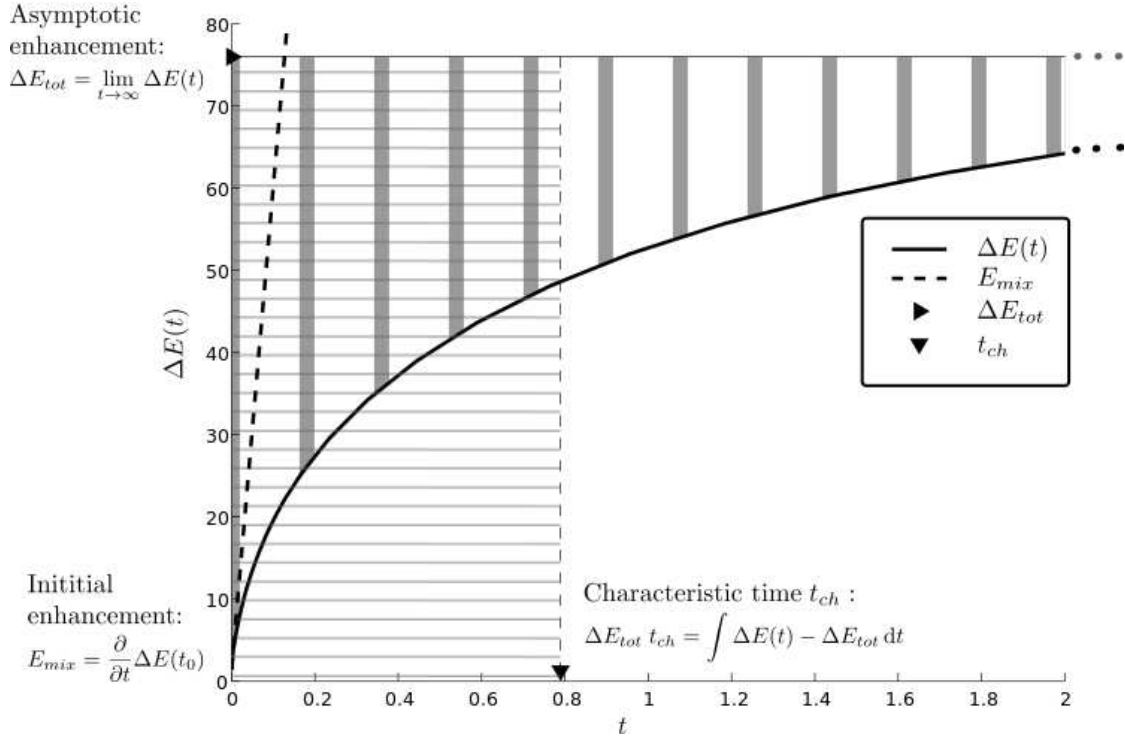


Figure 6.4: Illustration of the three proposed metrics to quantify mixing enhancement

in  $S_2$ . This initial enhancement  $E_{mix}$  quantifies the intensity of mixing events observed in  $S_2$ :

$$E_{mix} = \frac{\partial E(t_0)}{\partial t}. \quad (6.32)$$

**Asymptotic enhancement** The asymptotic enhancement assesses the temporally integrated increase of the dilution state. It therefore measures the total impact on the dilution process of the mixing initiated in  $S_2$ :

$$E_{tot} = \lim_{t \rightarrow \infty} \Delta E(t). \quad (6.33)$$

**Typical dissipation time** The dissipation time is a characteristic time of the duration of the dilution enhancement, which is linked to the persistence of individual mixing events. This dissipation time  $t_{ch}$  is measured under the condition that no additional plume deformation occurs which is the case in chamber  $S_3$ :

$$t_{ch} = \int_{t_0}^{t=\infty} 1 - \frac{\Delta E(t)}{E_{tot}} dt. \quad (6.34)$$



### 6.4.3 Stochastic analysis

It is expected that the strain mechanisms are again among the main drivers for enhanced mixing, as was observed in the two-dimensional study. However, the role of the rotation mechanism may change significantly whether it is analyzed in two- or in three-dimensional flow systems. This is because streamlines are allowed to twist in three-dimensional flow fields, which may intrude pure-aqueous phase within regions of high concentration values and vice versa.

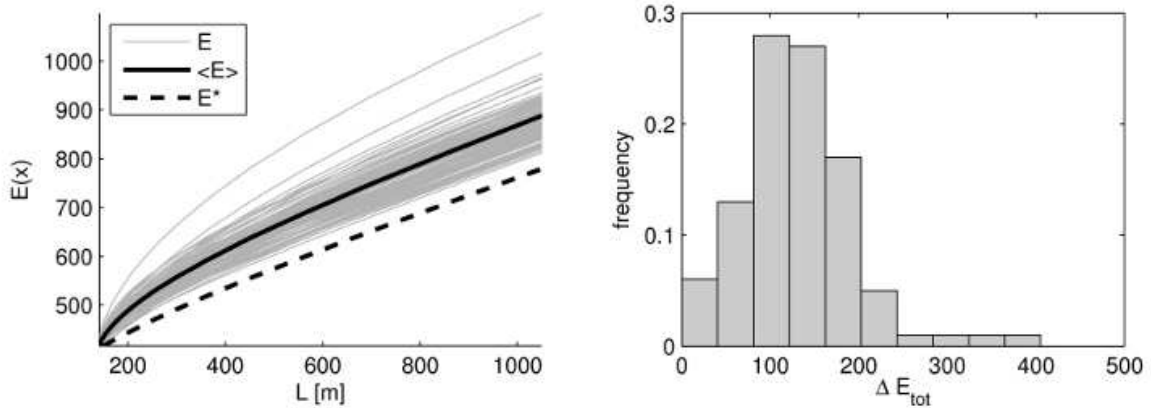


Figure 6.5: Mixing enhancement for various realizations of  $S_2$  resulting from a Monte Carlo simulation, left: ensemble of dilution curves over time in  $S_3$  (gray lines), average dilution over time (black and solid line) both due to heterogeneity in  $S_2$ , and the dilution with homogeneous  $S_2$  (dashed line); right: frequency of asymptotic mixing enhancement

The plume deformation within  $S_2$  leads to a significant dilution boost if the plume observed a mixing event within  $S_2$ . Whether the plume observes a mixing event (i.e. a deformation that leads to enhanced dilution) or not was obtained by the effective Okubo-Weiss function  $\Theta^e(t)$  in the two-dimensional study of Sec. 6.3. A contaminant plume that migrates through heterogeneous aquifers observes such mixing events in repeated batches, as observed in Sec. 6.3. The heterogeneous filter in  $S_2$  is designed such that it possesses only a few mixing events (in the best case only one or none), such that we can measure the impact of individual mixing events. Hence, the intensity of mixing events varies heavily from realization to realization. The initially smoothed plume signal entering  $S_2$  can be seen as a lower limit of dilution potential, when  $S_2$  is homogeneous. In other words, each realization of a heterogeneous  $S_2$  should lead to higher dilution rates in  $S_3$  compared to the homogeneous  $S_2$ .

In Fig. 6.5, the uncertainty of enhanced mixing results is illustrated for an example with constant heterogeneity within the ensemble variance  $\sigma_Y^2 = 4$  and correlation lengths  $\ell = 2L_2$  as shown in Fig. 6.3. It can be seen that a minimum enhanced mixing is guaranteed by the gap between gray lines and the black dashed line. Further, the asymmetry of the histogram (on the right hand side of Fig. 6.5) and the low density of gray lines at higher dilution indexes

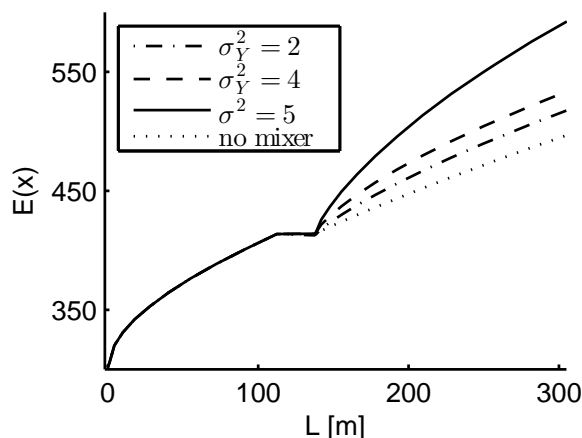


Figure 6.6: Average dilution curves due to heterogeneity in  $S_2$  for different variances of the geostatistical model

in the left graph of Fig. 6.5, suggest that extreme mixing events are rare. Fig. 6.6 shows the average dilution curve of three Monte-Carlo simulations with different variances of the geostatistical model. The mixing enhancement is proportional, on average, with the aquifer heterogeneity. Here, the variance  $\sigma_Y^2$  is increased stepwise:  $\sigma_Y^2 = 2, 4, 5$ .

In Fig. 6.7 the asymptotic mixing enhancement  $E_{tot}$  is scattered against mechanisms of the irregular flow field that may govern the mixing enhancement. These 9 metrics are volumetric averages of specific quantities of the local rate-of-deformation tensor: normal strain, shear strain, and vorticity. The three components (due to  $d = 3$ ) of each mechanism are analyzed individually in order to study their relative importance and mutual interplay with respect to increased growth rate of the dilution index. It is evident that none of these metrics are able to explain the overall mixing behavior in the sense of a first-order sensitivity. Nevertheless, each one is correlated with the asymptotic mixing by an individual pattern. Each graph in Fig. 6.7 spans a band of a linear function whose factor is unknown. Since this counts for more or less all 9 graphs, the factor of each mechanism could be dependent on all the other 8 metrics.

Fig. 6.7 indicates that each mechanism (stretching, shearing, and vorticity) effects the mixing process. However, the individual influence on the mixing events for each mechanism is not yet quantified. Hence, the specific kinematic- and topological-flow features that drive the mixing enhancement in three dimensional flow systems is still unknown. These first results could still serve future studies to quantify the impact of these mixing mechanisms on the expected plume deformation and enhanced mixing and dilution for various three-dimensional irregular flow systems. The individual (first-order sensitivities) and interactive (higher-order sensitivities) contribution of each mechanism to mixing could be identified by a sensitivity analysis, e.g., via a nonlinear Bayesian filtering scheme [e.g. *Leube et al.*, 2012]. Therefore, Sobol sensitivity indices [*Sobol'*, 1990; *Homma and Saltelli*, 1996] could be evaluated to compare the relative impact of these 9 mixing mechanism and of their combinations on

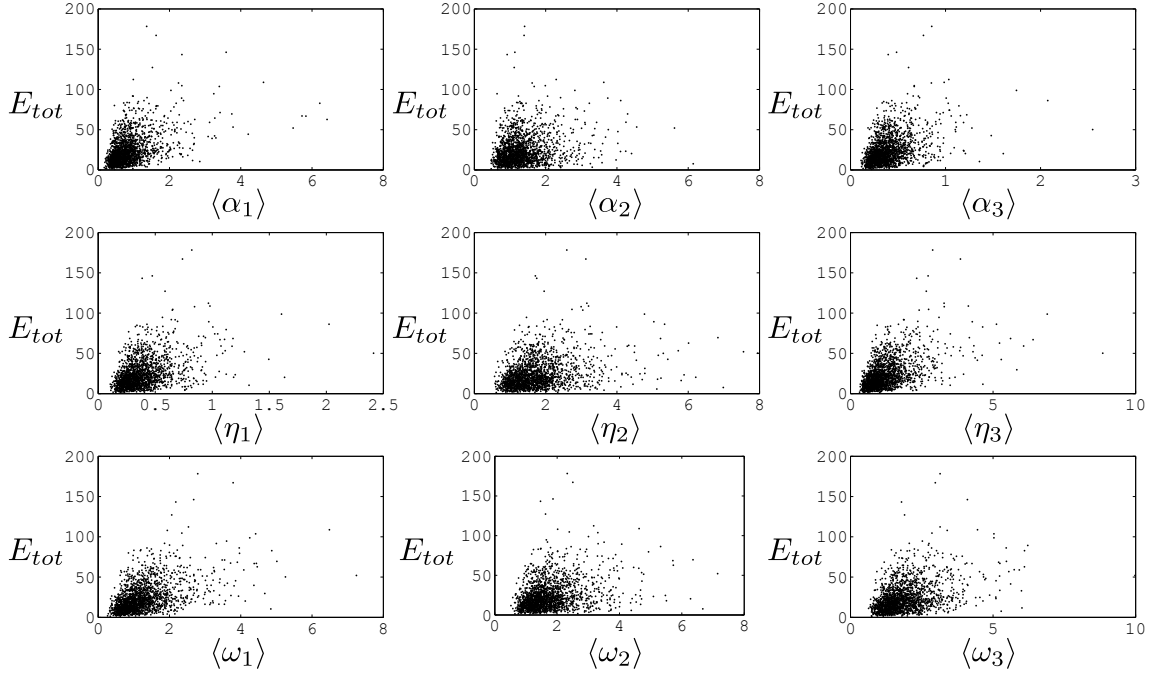


Figure 6.7: Asymptotic enhanced mixing versus various mixing mechanisms

the mixing metrics which have been introduced in Sec. 6.4.2.

#### 6.4.4 Summary and conclusions

In summary, a local analysis to understand the flow mechanisms that lead to increased mixing in heterogeneous media was performed. The mixing state of the system is described in terms of the dilution index  $E(t)$  and the mixing dynamics in terms of the rate of increase of its logarithm. The local mixing dynamics in two-dimensional flow systems are controlled by the Okubo-Weiss parameter  $\Theta(\mathbf{x})$ , which describes the relative importance between local strain and rotation properties of the flow field and thus characterizes the flow topology. As mixing is an inherently local-scale process [Le Borgne *et al.*, 2011], this analysis gives valuable insight into the global mixing properties. The results illustrate how  $\Theta(\mathbf{x})$  dictates the growth rate of the dilution index at the local and global scale. The defined effective Okubo-Weiss function  $\Theta^e(t)$  quantifies the flow topology within the area occupied by the plume. Numerical simulations clearly demonstrate that the global mixing properties are directly related to  $\Theta^e(t)$ .

I outlined an approach to transfer these findings to three-dimensional flow systems. To that end, I described a stochastic analysis in a numerical experiment, that relates individual local mixing events to the kinematical measures of flow. First results of this approach clearly indicate, that stretching, shear strain, and vorticity correlate with the enhanced dilution.

In order to quantify the role of each local flow mechanism, a multivariate regression analysis is required. Having quantified the relative importance of these mixing mechanisms, these findings can be used to design promising field campaigns for data collection and profound remediation strategies. The design of such field campaigns may differ significantly according to their objectives, i.e. field campaigns whose data is expected to reduce the uncertainty of mass discharge predictions are not necessarily promising to identify characteristics of the unknown CSA geometry. If, in future studies, an optimal design is evaluated with a brute-force Monte Carlo algorithm [e.g., *Leube et al.*, 2012], this can be compared to results of the current study, in order to analyze the utility and predictive power of various kinematical and topological measures of the flow system to the corresponding mixing and dilution enhancement and towards promising designs of field campaigns.

Further, these insights may have a significant impact for the quantification of mixing and reaction processes in porous media, because they may link the local and global mixing properties of the solute body to the topology of the flow field and thus to the medium's heterogeneity. Hence, profound knowledge on the interplay of heterogeneity, local-scale flow, and transport processes is of particular interest for flow and transport models simulating the fate of contaminants.

### Key conclusions

- In two-dimensional flows, both local and global mixing properties directly relate to the Okubo-Weiss parameter, which describes the relative importance between local strain and rotation properties of the flow system and thus characterizes the topology of two-dimensional flow systems.
- For three-dimensional flow system, the kinematical measures (stretching and shear strain, and vorticity) again correlate with the enhanced dissipation of concentration gradients.
- Due to this enhanced mixing and dissipation of concentration gradients, the quality of CSA inversion (applied in Chap. 5) strongly depends on the locations of concentration observations. Profound knowledge of these mechanisms could be used to conceptualize field campaigns whose data is highly informative for the inversion of CSA characteristics and mass discharge estimates.

## 7 Summary, conclusions and outlook

Groundwater contamination with dense-non-aqueous-phase liquids (DNAPL) pose an everlasting risk to ecosystems. Their high frequency and the difficulties of their remediation demand rational decisions based on a profound risk assessment. This is still an ongoing challenge because describing the behavior of DNAPLs in the subsurface is far from trivial. To that end, reasonable and well-designed predictive simulation tools and an adequate treatment of the relevant uncertainties is required. The overall goal of this thesis was to develop and study a physically and statistically consistent model framework that provides a sound basis for rational decisions arising in the assessment of contaminated sites.

To obtain such a sound basis, statistically consistent knowledge is desired on the contaminant source architecture (CSA), mass discharge, source depletion time, and on the process of enhanced mixing of dissolved contaminants in heterogeneous aquifers. However, the literature did not provide a model framework that reasonably simulates and infers physically-based random CSAs, predicts probability density functions of mass discharge and source depletion times, and aspires an overall physical and statistical consistency. Furthermore, the driving mechanisms of enhanced mixing in heterogeneous aquifers and their relation to the dissipation of concentration gradients was an open research question. This led to three main challenges which have been tackled within my thesis:

- (I) What is an adequate level of model complexity? And how to obtain a physically and statistically consistent model framework for reliable mass discharge and source depletion time predictions that are adapted to the presence of uncertainty?
- (II) How to identify CSAs and how to reduce the uncertainty of CSA characteristics, mass discharge, and source depletion times with available field data? A key question for this purpose is: how to assimilate concentration observations in an efficient inverse model?
- (III) What are the driving mechanisms of enhanced mixing and plume deformation and how to cope with the related loss of information content of concentration observations in a heterogeneous subsurface environment?

To cope with these challenges, I postulated three theses and demonstrated their significance, validity, and novelty throughout the thesis:

- (I) The model framework must at least account for the heterogeneity of aquifers, the irregularity of flow fields, realistic and thus complex-shaped CSAs, the three-dimensionality of natural systems, adequate physical interlinkages of the key parameters at the adequate spatial and temporal scales, and it must at least treat the uncertainty of aquifer parameters and of the CSA.

- (II) Joint identification of CSAs and aquifer parameters based on concentration observations can be achieved via non-linear and non-unique Bayesian inversion. An accurate and efficient inverse method for this task can be obtained by applying the method of adjoint states and utilizing the linearity of the transport equation.
- (III) The enhanced mixing of dissolved DNAPL and the solute plume deformation in heterogeneous aquifers significantly influences the inference quality of CSAs from downstream concentration observations. Knowledge on the driving processes of enhanced mixing allows to choose adequate measurement designs.

## 7.1 Summary of working steps

In order to elaborate on the postulated theses, I developed a model framework that avoids as many as reasonably possible simplifications such as: homogenization, upscaling, simplified contaminant source zones (e.g., block-shaped), chemical equilibrium assumptions on excessively large scales, and one- or two-dimensional simplified conceptualizations. Further, the model framework accounts for structural, parameter, and measurement uncertainties. To this end, I defined and conducted the following five steps:

1. The well-established pore-scale concept of invasion percolation algorithms was adapted to the Darcy scale and modified to account for the uncertainty of CSAs, which is due to unknown aquifer properties at small scales (below Darcy scale) and instabilities of the drainage process when DNAPL is displacing the aqueous phase (Sec. 3.2).
2. A novel particle tracking random walk method (PTRW) that can simulate Dirichlet and third-type boundary conditions was developed and benchmarked in Sec. 3.3 and Chap. 4. This method discretizes the boundary conditions by smooth volumes (so called emitters), and evaluates dissolution rates emitter wise, such that the sum of all emitter-wise dissolution rates equals the total mass discharge emanating from the CSA. Hence, one model outcome of the new PTRW method is a matrix system containing the interdependencies of all emitters forming the CSA.
3. Utilizing this matrix system, a quasi-steady-state source depletion concept was developed in Sec. 3.4. Depleted emitters are erased in that matrix system and the new dissolution rates and mass discharge follow immediately. Only when the CSA depletion affects the groundwater flow through relative permeabilities does this matrix system need to be updated by the PTRW method.
4. A novel method for the joint inversion of CSAs and aquifer parameters in a Bayesian framework was developed in Chap. 5. This method solves for adjoint sensitivities of source terms with respect to concentration observations via reverse particle tracking random walk (RPTRW). With this, Bayesian inference of CSA and aquifer parameters was performed.
5. Finally, the thesis explored mixing mechanisms that control the CSA inference quality. Due to aquifer heterogeneity and irregular groundwater flow, the initial structure of the plume emanating from a CSA is deformed and blurred.

## 7.2 Summary of conclusions

### **A novel physically and statistically consistent model framework was developed and assessed in Chap. 3**

- For reliable mass flux and source depletion time probability estimates one needs at least to account for adequate aquifer heterogeneity and randomness which is associated with non-uniform groundwater flow fields, complexity and uncertainty of CSAs, and for adequate physical interlinkages for the governing parameters.
- CSA geometry plays a key role for characterizing mass discharge and source depletion behavior. Therefore, decisions on the basis of simplified models for DNAPL contaminated sites should be taken with caution.
- Relative permeability exerts a secondary influence on these processes compared with the CSA. However, its effect on depletion times is non-neglectable. Thus, the only acceptable simplification (out of all tested simplifications in this thesis) seems to be a quasi-steady-state approach that approximates reasonably well the time evolution of the CSA and of the related relative permeabilities

### **A novel method for implementing Dirichlet and third-type boundary conditions in PTRW simulations was developed and benchmarked in Chap. 4**

- Dirichlet and third-type boundary conditions in PTRW simulations is challenging, but possible using the developed PTRW method.
- The requirements for a consistent model framework as discussed in Chap. 3 are fully met. According to these requirements, the method can account for complex-shaped and adequately discretized boundary conditions (e.g., by smooth-emitter volumes), is suitable for the quasi-steady-state depletion concept, and avoids artificial numerical dissipation of concentration values at high Péclet numbers.
- The new PTRW method outperforms other well-established methods like, the multiple analytical superposition technique (MASST) and a Petrov-Galerkin FEM.

### **A novel CSA identification approach was developed in Chap. 5**

- The proposed reverse-inverse methodology allows for joint inversion of complex CSAs and aquifer parameters in a physically and statistically consistent framework.
- Due its reverse-inverse concept and its Bayesian framework, this method efficiently and accurately accounts for all relevant interdependencies of a consistent model framework, treats parameter, structural, and observation uncertainty, and reflects the non-uniqueness of the underlying inverse problem.
- The restriction to physically plausible CSAs and aquifer parameters and their joint inversion reduces the solution space and improves the inversion performance significantly.

- The method enables efficient assimilation of various data types observed in the field. The inversion performance with and without the use of concentration values differed dramatically, which suggests that concentration data is highly informative for the CSA inversion.

### Plume deformation and mixing of dissolved contaminants was studied in Chap. 6

- In two-dimensional flows, both local and global mixing properties directly relate to the Okubo-Weiss parameter, which describes the relative importance between local strain and rotation properties of the flow system and thus characterizes the topology of two-dimensional flow systems.
- For three-dimensional flow system, the kinematical measures (stretching and shear strain, and vorticity) again correlate with the enhanced dissipation of concentration gradients.
- Due to this enhanced mixing and dissipation of concentration gradients, the quality of CSA inversion (applied in Chap. 5) strongly depends on the locations of concentration observations. Profound knowledge of these mechanisms could be used to conceptualize field campaigns whose data is highly informative for the inversion of CSA characteristics and mass discharge estimates.

## 7.3 Overall conclusions

This work conveys five key messages:

**1. Physical consistency:** Assessing the fate and transport of DNAPL contaminants in the subsurface, all crucial physical or statistical dependencies between heterogeneous aquifer parameters, CSA formation, dissolution and depletion, and advective-dispersive transport within the source zone and towards the concentration observations need to be taken into account at their appropriate spatial and temporal scale, observing the three-dimensionality of the physical space.

**2. Complexity of CSAs and statistical consistency:** The complex CSA and the heterogeneous permeability are two equally important sources of uncertainty. These propagate uncertainty and spatial variability onto all subsequent model states and model outcomes. A realistic description of the CSA is required. Otherwise, meaningless mass flux estimates with excessive error bounds or substantial bias may result. Due to uncertainties and heterogeneities of aquifer parameters at small scales (below the Darcy scale), CSAs can only be described statistically.

**3. Increase in prediction confidence due to physical and statistical consistency:** The physically-based stochastic CSA model reduces the overall prediction uncertainty, because it is a valuable source of information restricting the range of possible outcomes. This provides a strong prior for the Bayesian inference of CSA and aquifer parameters.



**4. Assimilating observed data increases confidence for decision purposes:** The remaining prediction uncertainty can be tackled by conditioning on observed data with the developed efficient reverse-inverse methodology for joint CSA and  $\mathbf{K}$  inference. This yields convenient and realistic confidence intervals of predicted contaminated site behavior and characteristics.

**5. Plume deformation and mixing of the dissolved contaminant:** Mixing dynamics can be related to local strain and rotation properties of the flow system. In two-dimensional flows the effective Okubo Weiss function directly relates to the dissipation rate of concentration gradients.

## 7.4 Outlook

The thesis work highlighted need for further research in various areas. Out of these, I selected six preferential research tasks that directly build upon my thesis.

**Analyzing the performance of the proposed CSA identification method:** The reverse-inverse CSA identification approach should be tested for different measurement designs, and in various applications. Chap. 5 inferred several features (probability pattern of DNAPL presence, emanating mass discharge from the CSA, and CSA characteristics) for one arbitrary chosen synthetic reality. To quantify the inversion performance a Monte Carlo simulation of these inversions with various synthetic realities should be performed and analyzed.

**Identification of aquifer properties under uncertain physically-based CSAs:** Although the proposed reverse-inverse method identifies CSAs and aquifer parameters jointly, Chap. 5 only demonstrated and discussed the identification of CSA properties. Hence, the influence of complex and uncertain CSAs on the inference of aquifer properties, and vice versa, the influence of various spatial structures of the aquifer on the inference of CSAs, could be studied. To this end, more complex spatial features of the permeability fields than those provided by the multivariate Gaussian distribution should be applied. For example, theoretical copulas could be constructed to serve as spatial random functions, as recently studied by *Guthke* [2013].

**Applying the new inverse method to contaminated sites or laboratory experiments:** Testing the performance of the proposed reverse-inverse method on actual field (or laboratory) data is one of the most interesting tasks for future studies that directly builds upon this thesis. The statistical inference approach should be applied on data sets that are either obtained from field campaigns, laboratory experiments, or quite realistic multi-phase multi-component simulations. This is superior to using synthetic realities out of the same model that provides the propositions for the inverse method. Inferring the properties of an unknown CSA that emerged naturally is also an opportunity to test the random CSA generator proposed in Sec. 3.2. The latter is obviously only possible if the data set is obtained from a realistic scenario, where DNAPL was spilled above an heterogeneous aquifer and was able to percolate naturally. Examples of such data sets are provided in [*Brewster et al.*,

1995; Zhang *et al.*, 2007]; the data set by Rivett *et al.* [2001] is not obtained under natural spill conditions but could be also used to study the performance of the reverse-inverse method.

Further, the proposed reverse-inverse CSA identification is based on a probabilistic framework. The results and the benefit of statistical inference (multiple conditional realizations that enable uncertainty quantification) versus best estimates against the necessary additional run-time, or against necessary model simplifications to reduce run-time, could be compared.

**CSA characteristics:** Another interesting scope for future studies is to find CSA characteristics with high explanatory power on mass discharge and source depletion time, and that support decisions such as which remediation strategies could be most promising. In Chap. 5 the ganglia to pool ratio and the CSA's discharge area with respect to the mean groundwater flow are inferred. However, the mass discharge *pdfs* for various model simplifications observed in Chap. 3 suggest that the explanatory power of such simple CSA characteristics is questionable. Hence, the proposed study may reveal the need for new CSA characteristics. For example, CSAs could be characterized through their fractal geometry and new CSA characteristics could also consider the groundwater flow topology.

**Mixing of the dissolved contaminant in three-dimensional flow systems:** Finally, the thesis work explored mixing mechanisms that control the CSA inference quality due to enhanced dilution which blurs the initial structure and information content of the dissolved-phase plume. Continuing the study of three-dimensional mixing enhancement may reveal crucial insights valuable for the performance of CSA identification techniques. To transfer the findings of the two-dimensional study in Sec. 6.3 to three-dimensional flow systems, I outlined an approach that studies local mixing mechanisms via a stochastic analysis of a three-dimensional numerical experiment. First results indicate that again stretching and shear deformation, and vorticity may directly relate to an increase in the dissipation of concentration gradients of dissolved contaminants. To study the individual and the joint contribution of various kinematical and topological flow metrics to mixing, a multivariate sensitivity analysis should be performed in a future study. Knowing the driving mechanisms of mixing may help to infer measurement designs for proper CSA identification and thus supplement optimal design strategies.

**Optimal design of field campaigns under uncertain physically-based CSAs:** Another crucial research challenge is to evaluate which observation data contains the most valuable information to infer the relevant properties of a contaminated site, e.g., mass discharge, CSA geometry, or aquifer parameters. Considering physically-based CSAs and their uncertainty in optimal design studies is a challenge because of the non-linearities and non-uniqueness of the underlying inverse problem. Yet the efficiency of the reverse-inverse approach in Chap. 5 could be utilized within a linearization-free optimal design framework [e.g., Leube *et al.*, 2012]. This may improve the inference performance significantly via reducing the costs of data collection campaigns, and/or via high uncertainty reduction of the desired parameters. Furthermore, applying physically-based and uncertain CSAs in optimal design studies may reveal more system-specific information and improve the reliability of proposed measurement designs, and therefore increase the benefits of data collection campaigns.

## Bibliography

- Ababou, R., A. C. Bagtzoglou, and A. Mallet (2010), Anti-diffusion and source identification with the RAWscheme: A particle-based censored random walk, *Environ. Fluid Mechan.*, 10(1), 41–76, doi:10.1007/s10652-009-9153-4.
- Abriola, L. M. (1989), Modeling multiphase migration of organic chemicals in ground-water systems - A review and assessment, *Environ. Health Persp.*, 83, 117–143, doi:10.2307/3430652.
- Alcouffe, R. E., A. Brandt, J. E. Dendy, Jr., and J. W. Painter (1981), The multi-grid method for the diffusion equation with strongly discontinuous coefficients, *J. Scien. Stat. Comp. (SIAM)*, 2(4), 430–454, doi:10.1137/0902035.
- Anderson, M. (1988), The Dissolution and Transport of Dense Non-Aqueous Phase Liquids in Saturated Porous Media, Ph.D. thesis, University of Oregon, Eugene, OR.
- Anderson, M., R. Johnson, and J. Pankow (1992a), Dissolution of dense chlorinated solvents into ground water: 1. Dissolution from well-defined residual source, *Ground Water*, 30(7), 250–256, doi:10.1111/j.1745-6584.1992.tb01797.x.
- Anderson, M. R., R. L. Johnson, and J. F. Pankow (1992b), Dissolution of dense chlorinated solvents into groundwater. 3. Modelling contaminant plumes from fingers and pools of solvent, *Environ. Sci. Technol.*, 26(5), 901–908, doi:10.1021/es00029a005.
- Aral, M. M., J. Guan, and M. L. Maslia (2001), Identification of contaminant source location and release history in aquifers, *J. Hydrol. Eng.*, 6(3), 225–234, doi:10.1061/(ASCE)1084-0699(2001)6:3(225).
- Bagtzoglou, A. C., D. E. Dougherty, and A. F. B. Tompson (1992), Application of particle methods to reliable identification of groundwater pollution sources, *Water Resour. Mang.*, 6(1), 15–23, doi:10.1007/BF00872184.
- Bárdossy, A., and J. Li (2008), Geostatistical interpolation using copulas, *Water Resour. Res.*, 44, W07412, doi:10.1029/2007WR006115.
- Basu, N. B., P. S. C. Rao, I. C. Poyer, M. D. Annable, and K. Hatfield (2006), Flux-based assessment at a manufacturing site contaminated with trichloroethylene, *J. Contam. Hydrol.*, 86(1), 105–127, doi:10.1016/j.jconhyd.2006.02.011.
- Basu, N. B., P. S. C. Rao, R. W. Falta, M. D. Annable, J. W. Jawitz, and K. Hatfield (2008), Temporal evolution of DNAPL source and contaminant flux distribution: Impacts of source mass depletion, *J. Contam. Hydrol.*, 95(3-4), 93–109, doi:10.1016/j.jconhyd.2007.08.001.

- Batchelor, B., J. Valdes, and V. Araganth (1998), Stochastic risk assessment of sites contaminated by hazardous wastes, *J. Environ Eng.*, 124(4), 380–388, doi:10.1061/(ASCE)0733-9372(1998)124:4(380).
- Batchelor, G. (1959), Small-scale variation of convected quantities like temperature in turbulent fluid part 1. general discussion and the case of small conductivity, *J. Fluid Mech.*, 5, 113–133, doi:10.1017/S002211205900009X.
- Bedinger, M. S. (1961), Relation between median grain size and permeability in the Arkansas River Valley, *Tech. rep.*, U.S. Geol. Surv. Professional Paper 424-C, Arkansas.
- Benekos, I. D., O. A. Cirpka, and P. K. Kitanidis (2006), Experimental determination of transverse dispersivity in a helix and a cochlea, *Water Resour. Res.*, 42, W07406, doi:10.1029/2005WR004712.
- Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, 2 ed., Springer, New York.
- Binning, P., and M. Celia (2002), A forward particle tracking EulerianLagrangian Localized Adjoint Method for solution of the contaminant transport equation in three dimensions, *Advan. Water Resour.*, 25(2), 147–157, doi:10.1016/S0309-1708(01)00051-3.
- Blunt, M. J., M. D. Jackson, M. Piri, and P. H. Valvatne (2002), Detailed physics, predictive capabilities and macroscopic consequences for pore-network models of multiphase flow, *Adv. Water Res.*, 25(8-12), 1069–1089, doi:10.1016/S0309-1708(02)00049-0.
- Bolster, D., M. Barahona, M. Dentz, D. Fernandez-Garcia, X. Sanchez-Vila, P. Trinchero, C. Valhondo, and D. M. Tartakovsky (2009), Probabilistic risk analysis of groundwater remediation strategies, *Water Resour. Res.*, 45, W06413, doi:10.1029/2008WR007551.
- Bolster, D., M. Dentz, and T. Le Borgne (2011), Hypermixing in linear shear flow, *Water Resour. Res.*, 47, W09602, doi:10.1029/2011WR010737.
- Brewster, M. L., A. P. Annan, J. P. Greenhouse, B. H. Keuper, G. R. Olhoeft, J. D. Redman, and K. A. Sander (1995), Observed migration of a controlled DNAPL release by geophysical methods, *Ground Water*, 33(6), 977–987, doi:10.1111/j.1745-6584.1995.tb00043.x.
- Broholm, K., S. Feenstra, and J. A. Cherry (2005), Solvent release into a sandy aquifer. 2. Estimation of DNAPL mass based on a multiple-component dissolution model., *Environ. Sci. Technol.*, 39(1), 317–324, doi:10.1021/es0306462.
- Brooks, A. N., and A. T. Corey (1964), Hydraulic Properties of Porous Media, in *Hydrol. Papers*, Fort Collins, Colorado State University.
- Brooks, A. N., and T. J. R. Hughes (1982), Streamline upwinding/Petrov-Galerkin formulations for convective dominated flows with particular emphasis on the incompressible Navier-Stokes equation, *Comp. Meth. App. Mech. Eng.*, 32(1), 199–259.
- Byers, V. S., A. S. Levin, D. M. Ozonoff, and R. W. Baldwin (1988), Association between clinical symptoms and lymphocyte abnormalities in a population with chronic domestic exposure to industrial solvent-contaminated domestic water supply and a high incidence of leukaemia, *Cancer Immunology Immunotherapy*, 27, 77–81.

- Celia, M. A., T. F. Russell, I. Herrera, and R. E. Ewing (1990), An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation, *Adv. Water Res.*, 13(4), 187–206, doi:10.1016/0309-1708(90)90041-2.
- Chambers, J. E., M. H. Loke, R. D. Ogilvy, and P. I. Meldrum (2004), Noninvasive monitoring of DNAPL migration through a saturated porous medium using electrical impedance tomography., *J. Contam. Hydrol.*, 68(1-2), 1–22, doi:10.1016/S0169-7722(03)00142-6.
- Chaudhuri, A., and M. Sekhar (2006), Stochastic modeling of solute transport in 3-D heterogeneous porous media with random source condition, *Stoch. Environ. Res. Risk Assess.*, 21(2), 159–173, doi:10.1007/s00477-006-0053-6.
- Chaudhuri, A., and M. Sekhar (2008), Modelling of solute transport in a mild heterogeneous porous medium using stochastic finite element method: Effects of random source conditions, *Int. J. Numer. Meth. Fluids*, 56(5), 557–586, doi:10.1002/ftd.1541.
- Christ, J. A., L. D. Lemke, and L. M. Abriola (2005), Comparison of two-dimensional and three-dimensional simulations of dense nonaqueous phase liquids (DNAPLs): Migration and entrapment in a nonuniform permeability field, *Water Resour. Res.*, 41, W01007, doi:10.1029/2004WR003239.
- Christ, J. a., C. A. Ramsburg, K. D. Pennell, and L. M. Abriola (2006), Estimating mass discharge from dense nonaqueous phase liquid source zones using upscaled mass transfer coefficients: An evaluation using multiphase numerical simulations, *Water Resour. Res.*, 42, W11420, doi:10.1029/2006WR004886.
- Christ, J. A., L. D. Lemke, and L. M. Abriola (2009), The influence of dimensionality on simulations of mass recovery from nonuniform dense non-aqueous phase liquid (DNAPL) source zones, *Adv. Water Res.*, 32(3), 401–412, doi:10.1016/j.advwatres.2008.12.002.
- Christ, J. A., C. A. Ramsburg, K. D. Pennell, and L. M. Abriola (2010), Predicting DNAPL mass discharge from pool-dominated source zones., *J. Contam. Hydrol.*, 114(1-4), 18–34, doi:10.1016/j.jconhyd.2010.02.005.
- Christakos, G. (1992), *Random Field Models in Earth Sciences*, 2 ed., Dover Publications, New York.
- Cirpka, O. A., and P. K. Kitanidis (2000), Characterization of mixing and dilution in heterogeneous aquifers by means of local temporal moments, *Water Resour. Res.*, 36(5), 1221–1236, doi:10.1029/1999WR900354.
- Cirpka, O. A., A. Olsson, Q. Ju, M. A. Rahman, and P. Grathwohl (2006), Determination of transverse dispersion coefficients from reactive plume lengths, *Ground Water*, 44(2), 212–221, doi:10.1111/j.1745-6584.2005.00124.x.
- Cirpka, O. a., F. P. J. de Barros, G. Chiogna, M. Rolle, and W. Nowak (2011), Stochastic flux-related analysis of transverse mixing in two-dimensional heterogeneous porous media, *Water Resour. Res.*, 47, W06515, doi:10.1029/2010WR010279.

- Class, H. (2001), Theorie und numerische Modellierung nichtisothermischer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien, Ph.D. thesis, Universität Stuttgart, Stuttgart.
- Clement, T. P., Y. Sun, B. S. Hooker, and J. N. Petersen (1998), Modeling multispecies reactive transport in ground water., *Ground Water Monit. Remed.*, 18(2), 79–92, doi:10.1111/j.1745-6592.1998.tb00618.x.
- Cohen, R. M., A. H. Vincent, J. W. Mercer, C. R. Faust, and C. P. Spalding (1994), *Methods for Monitoring Pump-and-Treat Performance*, 102 pp., EPA/600/R-94/123, R. S. Kerr Environ. Res. Lab., Ada, Oklahoma.
- Dagan, G., and S. P. Neuman (1997), *Subsurface Flow and Transport: A Stochastic Approach*, 1 ed., Cambridge University Press, New York.
- Daus, A. D., E. O. Frind, and E. A. Sudicky (1985), Comparative error analysis in finite element formulations of the advectiondispersion equation, *Advan. Water Resour.*, 8(2), 86–95, doi:10.1016/0309-1708(85)9005-3.
- de Barros, F. P. J., and W. Nowak (2010), On the link between contaminant source release conditions and plume prediction uncertainty, *J. Contam. Hydrol.*, 116(1-4), 24–34, doi:10.1016/j.jconhyd.2010.05.004.
- de Barros, F. P. J., M. Dentz, J. Koch, and W. Nowak (2012), Flow topology and scalar mixing in spatially heterogeneous flow fields, *Geophys. Res. Lett.*, 39, L08404, doi:10.1029/2012GL051302.
- de Barros, F. P. J., D. Fernández-García, D. Bolster, and X. Sanchez-Vila (2013), A risk-based probabilistic framework to estimate the endpoint of remediation: Concentration rebound by rate-limited mass transfer, *Water Resour. Res.*, 49(4), 1929–1942, doi:10.1002/wrcr.20171.
- Dekker, T. J., and L. M. Abriola (2000), The influence of field-scale heterogeneity on the infiltration and entrapment of dense nonaqueous phase liquids in saturated formations, *J. Contam. Hydrol.*, 42(2-4), 187–218, doi:10.1016/S0169-7722(99)00092-3.
- Dentz, M., and D. M. Tartakovsky (2010), Probability density functions for passive scalars dispersed in random velocity fields, *Geophys. Res. Lett.*, 37, L24406, doi:10.1029/2010GL045748.
- Dentz, M., T. Le Borgne, A. Englert, and B. Bijeljic (2011), Mixing, spreading and reaction in heterogeneous media: a brief review, *J. Contam. Hydrol.*, 120-121, 1–17, doi:10.1016/j.jconhyd.2010.05.002.
- DiCarlo, D. a. (2013), Stability of gravity-driven multiphase flow in porous media: 40 Years of advancements, *Water Resour. Res.*, 49(8), 4531–4544, doi:10.1002/wrcr.20359.
- Dietrich, C. R., and G. N. Newsam (1993), A fast and exact method for multidimensional Gaussian stochastic simulations, *Water Resour. Res.*, 29(8), 2861–2870, doi:10.1029/93WR01070.

- DiFilippo, E. L., and M. L. Brusseau (2011), Assessment of a simple function to evaluate the relationship between mass flux reduction and mass removal for organic-liquid contaminated source zones, *J. Contam. Hydrol.*, doi:10.1016/j.jconhyd.2010.12.011.
- Doherty, J., L. Brebber, and P. Whyte (1994), PEST: Model-independent parameter estimation., *Tech. rep.*, Watermark Numerical Computing.
- Douglas, J. J., and T. F. Russel (1982), Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, *J. Numer. Anal. (SIAM)*, 19(5), 871–885, doi:10.1137/0719063.
- Eden, M. (1961), A two-dimensional growth process, in *4th Berkley Symp. Math. Stat. Prob.*, edited by J. Neyman, pp. 223–239, University of California Press, Berkeley.
- Einarson, M. D., and D. M. Mackay (2001), Predicting impacts of groundwater contamination, *Environ. Sci. Technol.*, 35(3), 66–73, doi:10.1021/es0122647.
- Environmental Protection Agency (EPA) (2005), Guidelines for Carcinogen Risk Assessment, *Tech. Rep. March*, Risk Assessment Forum U.S., Rep. EPA/630/P-03/001F, Washington, DC.
- Enzenhöfer, R., T. Bunk, and W. Nowak (2014), Nine steps to risk-informed wellhead protection and management: A case study., *Ground water*, doi:10.1111/gwat.12161.
- Enzenhöfer, R., W. Nowak, and R. Helmig (2012), Probabilistic exposure risk assessment with advectivedispersive well vulnerability criteria, *Adv. Water Res.*, 36, 121–132, doi: 10.1016/j.advwatres.2011.04.018.
- Erning, K., D. Schäfer, A. Dahmke, A. Luciano, P. Viotti, and P. M. P. (2009), Simulation of DNAPL infiltration into groundwater with differing flow velocities using TMVOC combined with PETRASIM, in *TOUGH Symp. 2009.*, pp. 1–7, LBNL, Berkeley, CA.
- Ewing, R. P., and B. Berkowitz (1998), A generalized growth model for simulating initial migration of dense non-aqueous phase liquids, *Water Resour. Res.*, 34(4), 611–622, doi: 10.1029/97WR03754.
- Ewing, R. P., and B. Berkowitz (2001), Stochastic pore-scale growth models of DNAPL migration in porous media, *Adv. Water Res.*, 24(3-4), 309–323, doi:10.1016/S0309-1708(00)00059-2.
- Fagerlund, F., A. Niemi, and T. Illangasekare (2006), Modelling NAPL source zone formation in stochastically heterogeneous layered media - a comparison with experimental results, in *TOUGH Symp. 2006*, pp. 1–21, LBNL, Berkeley, CA.
- Falta, R. W. (2008), Methodology for comparing source and plume remediation alternatives., *Ground water*, 46(2), 272–85, doi:10.1111/j.1745-6584.2007.00416.x.
- Feenstra, S. (1996), Aqueous Concentration Ratios to Estimate Mass of Multi-Component NAPL Residual in Porous Media, Ph.D. thesis, University of Waterloo, Ontario, Canada.

- Feenstra, S., and J. A. Cherry (1996), Diagnosis and assessment of DNAPL sites., in *Dense Chlorinated Solvents and other DNAPLs in Groundwater*, edited by J. F. Pankow and J. A. Cherry, pp. 395–473, Waterloo Press, Portland.
- Feenstra, S., and N. Guiger (1996), Dissolution of dense non-aqueous phase liquids (DNAPLs) in the subsurface., in *Dense Chlorinated Solvents and other DNAPLs in Groundwater*, edited by J. F. Pankow and J. A. Cherry, 1 ed., chap. 7, pp. 203–232, Waterloo Press, Portland.
- Fernàndez-Garcia, D., and X. Sanchez-Vila (2010), Optimal reconstruction of concentrations, gradients and reaction rates from particle distributions, *J. Contam. Hydrol.*, 120-121, 99–114, doi:10.1016/j.jconhyd.2010.05.001.
- Ferziger, J. H., and M. Perić (1996), *Computational methods for fluid dynamics*, 3 ed., Springer, Berlin.
- Fetter, C. W. (1999), *Contaminant Hydrogeology*, 2 ed., Prentice Hall, Upper Saddle River, NJ.
- Feyen, L., J. J. Gómez-Hernández, P. J. Ribeiro, K. J. Beven, and F. De Smedt (2003), A Bayesian approach to stochastic capture zone delineation incorporating tracer arrival times, conductivity measurements, and hydraulic head observations, *Water Resour. Res.*, 39(5), 1126, doi:10.1029/2002WR001544.
- Fiori, A., and G. Dagan (2000), Concentration fluctuations in aquifer transport: A rigorous first-order solution and applications, *J. Contam. Hydrol.*, 45(1-2), 139–163, doi: 10.1016/S0169-7722(00)00123-6.
- Fishburn, P. C. (1970), *Utility Theory for Decision Making*, 1 ed., Wiley, RESEARCH ANALYSIS CORP MCLEAN VA, New York.
- Flemisch, B. (2013), Tackling Coupled Problems in Porous Media: Development of Numerical Models and an Open Source Simulator, Ph.D. thesis, University of Stuttgart, Stuttgart.
- Franssen, H. J. H., J. Gómez-Hernández, and A. Sahuquillo (2003), Coupled inverse modelling of groundwater flow and mass transport and the worth of concentration data, *J. Hydrol.*, 281(4), 281–295, doi:10.1016/S0022-1694(03)00191-4.
- Freeze, A. R., and D. B. McWhorter (1997), A framework for assessing risk reduction due to DNAPL mass removal from low permeability soils., *Ground Water*, 35(1), 111–123, doi: 10.1111/j.1745-6584.1997.tb00066.x.
- Freeze, R. A. (1975), A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, 11(5), 725–741, doi: 10.1029/WR011i005p00725.
- Frind, E. O., J. W. Molson, and D. L. Rudolph (2006), Well vulnerability: a quantitative approach for source water protection., *Ground Water*, 44(5), 732–42, doi:10.1111/j.1745-6584.2006.00230.x.



- Fritz, J., I. Neuweiler, and W. Nowak (2009), Application of FFT-based algorithms for large-scale universal kriging problems, *Math. Geoscienc.*, 41(5), 509–533, doi:10.1007/s11004-009-9220-x.
- Fure, A. D., J. W. Jawitz, and M. D. Annable (2006), DNAPL source depletion: linking architecture and flux response., *J. Contam. Hydrol.*, 85(3-4), 118–40, doi: 10.1016/j.jconhyd.2006.01.002.
- Gamerman, D., and H. F. Lopes (2006), *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2 ed., Chapman & Hall/CRC, Boca Raton, FL USA.
- Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2000), *Bayesian Data Analysis*, 5 ed., Chapman & Hall/CRC, Boca Raton.
- geostatistics for evaluation of mass discharge uncertainty at contaminated sites Troldborg, M. o. B., W. Nowak, I. V. Lange, M. C. Santos, P. J. Binning, and P. L. Bjerg (2012), Application of Bayesian geostatistics for evaluation of mass discharge uncertainty at contaminated sites, *Water Resour. Res.*, 48, W09535, doi:10.1029/2011WR011785.
- Gerhard, J. I., and B. H. Kueper (2003), Influence of constitutive model parameters on the predicted migration of DNAPL in heterogeneous porous media, *Water Resour. Res.*, 39, W1279, doi:10.1029/2002WR001570.
- Glass, R., E. Webb, and S. Conrad (1995), An up-scaled, buoyant invasion percolation model for use in delineating subsurface DNAPL location, in *ASME/AICL National heat Transfer Conference*, vol. 87, pp. 281–285, Portland, OR.
- Glass, R. J., S. H. Conrad, and W. Peplinski (2000), Gravity-destabilized nonwetting phase invasion in macroheterogeneous porous media: Experimental observations of invasion dynamics and scale analysis, *Water Resour. Res.*, 36(11), 3121–3137, doi: 10.1029/2000WR900152.
- Glass, R. J., S. H. Conrad, and L. Yarrington (2001), Gravity-destabilized nonwetting phase invasion in macroheterogeneous porous media : Near-pore-scale macro modified invasion percolation simulation of experiments, *Water Resour. Res.*, 37(5), 1197–1207, doi: 10.1029/2000WR900294.
- Gomez-Hernandez, J. J., A. Sahuquillo, and J. E. Capilla (1997), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory., *J. Contam. Hydrol.*, 203, 162–174.
- Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993), Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings-F*, 140(2), 107–113.
- Gorelick, S. M., B. Evans, and I. Remson (1983), Identifying sources of groundwater pollution: An optimization approach, *Water Resour. Res.*, 19(3), 779–790, doi: 10.1029/WR019i003p00779.

- Grant, G. P., and J. I. Gerhard (2007), Simulating the dissolution of a complex dense non-aqueous phase liquid source zone: 2. Experimental validation of an interfacial area-based mass transfer model, *Water Resour. Res.*, 43, W12409, doi:10.1029/2007WR006039.
- Grant, G. P., J. I. Gerhard, and B. H. Kueper (2007), Multidimensional validation of a numerical model for simulating a DNAPL release in heterogeneous porous media., *J. Contam. Hydrol.*, 92(1-2), 109–28, doi:10.1016/j.jconhyd.2007.01.003.
- Guthke, P. (2013), Non-multi Gaussian Spatial Structures: Process Driven Natural Genesis, Manifestation, Modeling Approaches, and Influences on Dependent Processes, Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.
- Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), MODFLOW 2000, *Tech. rep.*, U.S. GEOLOGICAL SURVEY.
- Härdle, W., M. Müller, S. Sperlich, and A. Werwatz (2004), *Nonparametric and Semiparametric Models*, 1st ed., Springer, Berlin, Germany.
- Haslauer, C. P., P. Guthke, A. Bárdossy, and E. a. Sudicky (2012), Effects of non-Gaussian copula-based hydraulic conductivity fields on macrodispersion, *Water Resour. Res.*, 48, W07507, doi:10.1029/2011WR011425.
- Hassan, A. E., and M. M. Mohamed (2003), On using particle tracking methods to simulate transport in single-continuum and dual continua porous media, *J. Hydrol.*, 275(3-4), 242–260, doi:10.1016/S0022-1694(03)00046-5.
- Healy, R. W., U. S. G. Survey, and T. F. Russell (1993), A finite-volume Eulerian-Lagrangian localized adjoint method for solution of the advection-dispersion equation, *Water Resour. Res.*, 29(7), 2399–2413, doi:10.1029/93WR00403.
- Heberton, C. I., T. F. Russell, L. F. Konikow, and G. Z. Hornberger (2000), A three-dimensional finite-volume Eulerian-Lagrangian localized adjoint method (ELLAM) for solute-transport modelling, *Tech. rep.*, US Geology Survey Water-Ressources Investigation Report, Reston, Virginia.
- Held, R. J., and T. H. Illangasekare (1995), Fingering of dense nonaqueous phase liquids in porous media: 1. Experimental investigation, *Water Resour. Res.*, 31(5), 1213–1222, doi:10.1029/95WR00428.
- Helmig, R. (1997), *Multiphase Flow and Transport Processes in the Subsurface*, 343 pp., Springer, Berlin, doi:10.1007/978-3-642-60763-9.
- Hendricks Franssen, H. J., and W. Kinzelbach (2008), Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, *Water Resour. Res.*, 44(9), n/a–n/a, doi:10.1029/2007WR006505.
- Herrera, I., R. E. Ewing, M. A. Celia, and T. F. Russell (1993), Eulerian-Lagrangian Localized Adjoint Method: the theoretical framework, *Num. Methods Partial Diff. Eq.*, 9(4), 431–457, doi:10.1002/num.1690090407.

- Ho, Y. C., and R. C. K. Lee (1964), A Bayesian approach to problems in stochastic estimation and control, *Autom. Control, IEEE Transact.*, 9(4), 333 – 339, doi:10.1109/TAC.1964.1105763.
- Homma, T., and A. Saltelli (1996), Importance measures in global sensitivity analysis of non-linear models, *Reliab. Eng. Syst. Safe.*, 52(1), 1–17.
- Hosseini, A. H., C. V. Deutsch, C. a. Mendoza, and K. W. Biggar (2011), Inverse modeling for characterization of uncertainty in transport parameters under uncertainty of source geometry in heterogeneous aquifers, *J. Hydrol.*, 405(3-4), 402–416, doi:10.1016/j.jhydrol.2011.05.039.
- Huang, C., B. X. Hu, X. Li, and M. Ye (2009), Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source, *Stoch. Environ. Res. Risk Assess.*, 23(8), 1155–1167, doi:10.1007/s00477-008-0289-4.
- Illangasekare, T. H., J. L. Ramsey, K. H. Jensen, and M. B. Butts (1995), Experimental study of movement and distribution of dense organic contaminants in heterogeneous aquifers, *J. Contam. Hydrol.*, 20(1), 1–25, doi:10.1016/0169-7722(95)00045-W.
- Imhoff, P. T., P. R. Jaffe, and F. Pinder (1994), An experimental study of complete dissolution of a nonaqueous phase liquid in saturated porous media, *Water Resour. Res.*, 30(2), 307–320, doi:10.1029/93WR02675.
- Jawitz, J. W., A. D. Fure, G. G. Demmy, S. Berglund, and P. S. C. Rao (2005), Groundwater contaminant flux reduction resulting from nonaqueous phase liquid mass reduction, *Water Resour. Res.*, 41, W10408, doi:10.1029/2004Wr003825.
- Jha, B., L. Cueto-Felgueroso, and R. Juanes (2011), Fluid mixing from viscous fingering, *Phys. Rev. Lett.*, 106, doi:10.1103/PhysRevLett.106.194502.
- Johnson, R. L., and F. Pankow (1992), Dissolution of dense chlorinated solvents into groundwater. 2. Source functions of pools of solvents, *Environ. Sci. Technol.*, 26(5), 896–901, doi:10.1021/es00029a004.
- Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, Academic Press, New York.
- Kapoor, V. (1997), Vorticity in three-dimensionally random porous media, *Transp. Por. Media*, 26(1), 109–119, doi:10.1023/A:1006593712793.
- Kapoor, V., and P. Kitanidis (1998), Concentration fluctuations and dilution in aquifers, *Water Resour. Res.*, 22(1), 91–119, doi:10.1007/BF00974313.
- Kass, R. E., and L. Wasserman (1994), Formal rules for selecting prior distributions: a review and annotated bibliography, *Journal of the American Statistical Association*, 91(435).
- Kavanaugh, M., P. S. C. Rao, L. Abriola, J. Cherry, G. Destouni, R. Falta, D. Major, J. Mercer, C. Newell, T. Sale, S. Shoemaker, R. Siegrist, G. Teutsch, and K. Udell (2003), The DNAPL remediation challenge: Is there a case for source depletion?, *Tech. rep.*, (No. EPA/600/R-03/143), Washington, DC.

- Kavanaugh, M. C., R. Deeb, J. Nyman, S. D. Lloyd Bo, and M. A. Widdowson (2011), Improved field evaluation of NAPL dissolution and source longevity, *Tech. Rep. October*, ARCADIS/Malcolm Pirnie, Emeryville, CA.
- Kennedy, M. C., and A. O'Hagan (2001), Bayesian calibration of computer models, *J. R. Statist. Soc. B*, 63(3), 425–464, doi:10.1111/1467-9868.00294.
- Kinzelbach, W. (1988), The random walk method in pollutant transport simulation, in *Groundwater flow and quality modelling*, edited by E. Custodio, A. Gurgui, and J. L. Ferreria, pp. 227–246, D. Reidel, Dordrecht.
- Kitanidis, P. K. (1986), Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resour. Res.*, 22(4), 499–507, doi:10.1029/WR022i004p00499.
- Kitanidis, P. K. (1994), The concept of the Dilution Index, *Water Resour. Res.*, 30(7), 2011–2026, doi:10.1029/94WR00762.
- Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour. Res.*, 31(10), 2411–2419, doi:10.1029/95WR01945.
- Kitanidis, P. K. (1996), On the geostatistical approach to the inverse problem, *Adv. Water Res.*, 19(6), 333–342, doi:10.1016/0309-1708(96)00005-X.
- Kitanidis, P. K. (1997), *Introduction to Geostatistics: Applications to Hydrogeology*, 1st ed., Cambridge University Press Cambridge.
- Koch, J., and W. Nowak (2014a), A method for implementing Dirichlet and third-type boundary conditions in PTRW simulations, *Water Resour. Res.*, 50, doi:10.1002/2013WR013796.
- Koch, J., and W. Nowak (2014b), Simulation framework for reliable mass discharge and source depletion time predictions of DNAPL contaminants in heterogeneous aquifers under uncertainty, *Water Resour. Res.*, submitted.
- Kokkinaki, A., D. M. O'Carroll, C. J. Werth, and B. E. Sleep (2013), Coupled simulation of DNAPL infiltration and dissolution in three-dimensional heterogeneous domains: Process model validation, *Water Resour. Res.*, 49(10), 7023–7036, doi:10.1002/wrcr.20503.
- Konikow, L. F., D. J. Goode, and G. Z. Hornberger (1996), A three-dimensional method-of-characteristics solute-transport model (MOC3D), *Tech. rep.*, U.S. Geological Survey (USGS), Reston, Virginia.
- LaBolle, E. M., G. E. Fogg, and A. F. B. Tompson (1996), Random-walk simulation of transport in heterogeneous porous media: Local mass-conservation problem and implementation methods, *Water Resour. Res.*, 32(3), 583–593, doi:10.1029/95WR010137.
- Le Borgne, T., M. Dentz, P. Davy, D. Bolster, J. Carrera, J. R. de Dreuzy, and O. Bour (2011), Persistence of incomplete mixing: A key to anomalous transport, *Phys. Rev. Lett.*, 84, doi:10.1103/PhysRevE.84.015301.

- Lemke, L. D., and L. M. Abriola (2006), Modeling dense nonaqueous phase liquid mass removal in nonuniform formations: Linking source-zone architecture and system response, *Geosphere*, 2(2), 74–82, doi:10.1130/GES00025.1.
- Lemke, L. D., L. M. Abriola, and P. Goovaerts (2004), Dense nonaqueous phase liquid (DNAPL) source zone characterization: Influence of hydraulic property correlation on predictions of DNAPL infiltration and entrapment, *Water resources research*, 40(1), W01511.
- Lenormand, R. (1990), Liquids in porous media, *J. Physics Condens. Matter*, 2, 79–88, doi:10.1088/0953-8984/2/S/008.
- Leube, P. C., A. Geiges, and W. Nowak (2012), Bayesian assessment of the expected data impact on prediction confidence in optimal sampling design, *Water Resour. Res.*, 48, W02501, doi:10.1029/2010WR010137.
- Levich, V. G. (1962), *Physicochemical hydrodynamics*, Prentice-Hall, Englewood Cliffs.
- Li, K. B., P. Goovaerts, and L. M. Abriola (2007), A geostatistical approach for quantification of contaminant mass discharge uncertainty using multilevel sampler measurements, *Water Resour. Res.*, 43, W06436, doi:10.1029/2006WR005427.
- Li, X., W. Wu, and O. C. Zienkiewicz (2000), Implicit characteristic Galerkin method for convection-diffusion equations, *Int. J. Num. Meth. Eng.*, 47, 1689–1708, doi:10.1002/(SICI)1097-0207(20000410)47:10<1689::AID-NME850>3.3.CO;2-N.
- Luce, R. D., and H. Raiffa (1957), *Games and Decisions: Introduction and Critical Survey*, 1st ed., John Wiley & Sons, New York.
- Luciano, A., P. Viotti, and M. P. Papini (2010), Laboratory investigation of DNAPL migration in porous media, *J. Hazard. Mater.*, 176(1), 1006–1017, doi:10.1016/j.jhazmat.2009.11.141.
- Mackay, D. M., P. V. Roberts, and J. A. Cherry (1985), Transport of organic contaminants in groundwater. Distribution and fate of chemicals in sand and gravel aquifers, *Environ. Sci. Technol.*, 19(5), 384–92, doi:10.1021/es00135a001.
- Marinovich, M. (2006), Source Zone and Model Parameter Influence on NAPL Plume Development in a Heterogeneous Aquifer, Ph.D. thesis, University of Western Australia, Perth, Australia Matheron.
- Matheron, G. (1963), Principles of geostatistics, *Economic Geology*, 58(8), 1246–1266, doi:10.2113/gsecongeo.58.8.1246.
- Matern, B. (1986), *Spatial Variation*, 2nd ed., Springer, Berlin, doi:10.3109/13880208809053896.
- Maxwell, R. M., S. D. Pelmulder, A. F. B. Tompson, and W. E. Kastenberg (1998), On the development of a new methodology for groundwater-Driven health risk assessment, *Water Resour. Res.*, 34(4), 833–847, doi:10.1029/97WR03605.

- McWhorter, D. B., and T. C. Sale (2003), Reply to comment by P. S. C. Rao and J. W. Jawitz on: Steady state mass transfer from single-component dense nonaqueous phase liquids in uniform flow fields; by T. C. Sale and D. B. McWhorter, *Water Resour. Res.*, 39, 1069, doi:10.1029/2002WR001423.
- Mercer, J. W., and R. M. Cohen (1990), A review of immiscible fluids in the subsurface: properties, models, characterization and remediation, *J. Contam. Hydrol.*, 6(2), 107–163, doi:10.1016/0169-7722(90)90043-G.
- Michalak, A. M., and P. K. Kitanidis (2004), Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling, *Water Resour. Res.*, 40, W08302, doi:10.1029/2004WR003214.
- Miller, C. T., M. M. Poirier-McNeill, and A. S. Mayer (1990), Dissolution of trapped nonaqueous phase liquids: Mass transfer characteristics, *Water Resour. Res.*, 26(11), 2783–2796, doi:10.1029/WR026i011p02783.
- Neupauer, R. M., and R. Lin (2006), Identifying sources of a conservative groundwater contaminant using backward probabilities conditioned on measured concentrations, *Water Resour. Res.*, 42, W03424, doi:10.1029/2005WR004115.
- Neupauer, R. M., and J. L. Wilson (2001), Adjoint-derived location and travel time probabilities for a multidimensional groundwater system, *Water Resour. Res.*, 37(6), 1657–1668, doi:10.1029/2000WR900388.
- Neupauer, R. M., and J. L. Wilson (2002), Backward probabilistic model of groundwater contamination in non-uniform and transient flow, *Advan. Water Resour.*, 25(7), 733–746, doi:10.1016/S0309-1708(02)00073-8.
- Newell, C. J., S. K. Farhat, D. T. Adamson, and B. B. Looney (2011), Contaminant plume classification system based on mass discharge., *Ground water*, 49(6), 914–9, doi:10.1111/j.1745-6584.2010.00793.x.
- Nowak, W. (2005), Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface, Ph.D. thesis, University of Stuttgart, Stuttgart, Germany.
- Nowak, W. (2009), Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator, *Water Resour. Res.*, 45, W04431, doi:10.1029/2008WR007328.
- Nowak, W., and A. Litvinenko (2013), Kriging and spatial design accelerated by orders of Magnitude: combining low-rank covariance approximations with FFT-techniques, *Math. Geoscienc.*, 45(4), 411–435, doi:10.1007/s11004-013-9453-6.
- Nowak, W., R. L. Schwede, O. A. Cirpka, and I. Neuweiler (2008), Probability density functions of hydraulic head and velocity in three-dimensional heterogeneous porous media, *Water Resour. Res.*, 44, W08452, doi:10.1029/2007WR006383.
- Nowak, W., F. P. J. de Barros, and Y. Rubin (2010), Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain, *Water Resour. Res.*, 46, W03535, doi:10.1029/2009WR008312.

- Okubo, A. (1970), Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences, in *Deep Sea Res. Oceanographic Abstrac.*, vol. 17, pp. 445–454, Elsevier, doi:10.1016/0011-7471(70)90059-8.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994), Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, 263(5147), 641–646, doi: 10.1126/science.263.5147.641.
- Page, J. W. E., K. Soga, and T. Illangasekare (2007), The significance of heterogeneity on mass flux from DNAPL source zones: An experimental investigation, *J. Contam. Hydrol.*, 94(3-4), 215–234, doi:10.1016/j.jconhyd.2007.06.004.
- Pankow, J. F., and J. A. Cherry (1996), *Dense Chlorinated Solvents and other DNAPLs in Groundwater*, 513 pp., Waterloo Press, Portland.
- Parker, J. C., and E. Park (2004), Modeling field-scale dense nonaqueous phase liquid dissolution kinetics in heterogeneous aquifers, *Water Resour. Res.*, 40, W05109, doi: 10.1029/2003WR002807.
- Parker, J. C., and M. T. van Genuchten (1984), Flux Averaged and Volume Averaged Concentrations in Continuum Approaches to Solute Transport, *Water Resour. Res.*, 20(7), 866–872, doi:10.1029/WR020i007p00866.
- Pironneau, O. (1982), On the transport-diffusion algorithm and its application to the Navier-Stokes equations, *Num. Math.*, 332, 309–332, doi:10.1007/BF01396435.
- Press, J. S. (2012), *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*, 2 ed., Dover Publications, New York.
- Rao, P. S. C., and J. W. Jawitz (2003), Comment on: Steady state mass transfer from single-component dense nonaqueous phase liquids in uniform flow fields; by T. C. Sale and D. B. McWhorter, *Water Resour. Res.*, 39(3), 1068, doi:10.1029/2001WR000599.
- Reeves, P. C., and M. a. Celia (1996), A functional relationship between capillary pressure, saturation, and interfacial area as revealed by a pore-scale network model, *Water Resour. Res.*, 32(8), 2345–2358, doi:10.1029/96WR01105.
- Rivera, J. (2013), Investigating the Influence of Relative Permeability on DNAPL Dissolution and Depletion at Contaminated Sites, Master's thesis, University of Stuttgart, Stuttgart, Germany.
- Rivett, M. O., S. Feenstra, and J. A. Cherry (2001), A controlled field experiment on groundwater contamination by a multicomponent DNAPL: creation of the emplaced-source and overview of dissolved plume development, *J. Contam. Hydrol.*, 49(1), 111–149, doi: 10.1016/S0169-7722(00)00191-1.
- Robinson, B., Z. V. Dash, and G. Srinivasan (2010), A particle tracking transport method for the simulation of resident and flux averaged concentration of solute plumes in groundwater models, *Comp. Geosci.*, 14(4), 779–792, doi:10.1007/s10596-010-9190-6.

- Rubin, Y. (2003), *Applied Stochastic Hydrology*, 1st ed., Oxford University Press, New York, USA.
- Rubin, Y., and G. Dagan (1987), Stochastic identification of transmissivity and effective recharge in steady groundwater flow 1. Theory, *Water Resour. Res.*, 23(7), 1185–1192, doi:10.1029/WR023i007p01185.
- Russell, T. F., and M. a. Celia (2002), An overview of research on EulerianLagrangian localized adjoint methods (ELLAM), *Adv. Water Res.*, 25(8-12), 1215–1231, doi:10.1016/S0309-1708(02)00104-5.
- Russell, T. F., C. I. Heberton, L. F. Konikow, and G. Z. Hornberger (2003), A finite-volume ELLAM for three-dimensional solute-transport modeling, *Ground Water*, 41(2), 258–272, doi:10.1111/j.1745-6584.2003.tb02589.x.
- Saba, T., and T. H. Illangasekare (2000), Effect of groundwater flow dimensionality on mass transfer from entrapped nonaqueous phase liquid contaminants, *Water Resour. Res.*, 36(4), 971–979, doi:10.1029/1999WR900322.
- Saenton, S., and T. H. Illangasekare (2004), Determination of DNAPL entrapment architecture using experimentally validated numerical codes and inverse modeling, *Comp. Meth. Water Res.*, 1(55), 776–778, doi:10.1016/S0167-5648(04)80098-4.
- Saenton, S., and T. H. Illangasekare (2007), Upscaling of mass transfer rate coefficient for the numerical simulation of dense nonaqueous phase liquid dissolution in heterogeneous aquifers, *Water Resour. Res.*, 43, W02428, doi:10.1029/2005WR004274.
- Saenton, S., and T. H. Illangasekare (2013), Effects of incomplete remediation of NAPL-contaminated aquifers: experimental and numerical modeling investigations, *Appl. Water Sci.*, 3(2), 401–414, doi:10.1007/s13201-013-0090-5.
- Saenton, S., T. H. Illangasekare, K. Soga, and T. a. Saba (2002), Effects of source zone heterogeneity on surfactant-enhanced NAPL dissolution and resulting remediation end-points, *J. Contam. Hydrol.*, 59(1-2), 27–44, doi:10.1016/S0169-7722(02)00074-8.
- Salamon, P., D. Fernández-García, and J. J. Gómez-Hernández (2006), A review and numerical assessment of the random walk particle tracking method, *J. Contam. Hydrol.*, 87(3-4), 277–305, doi:10.1016/j.jconhyd.2006.05.005.
- Sale, T. C., and D. B. McWhorter (2001), Steady state mass transfer from single-component dense nonaqueous phase liquids in uniform flow fields, *Water Resour. Res.*, 37(2), 393–404, doi:10.1029/2000WR900236.
- Sanchez-Vila, X., M. Dentz, and L. Donado (2007), Transport-controlled reaction rates under local non-equilibrium conditions, *Geoph. Res. Lett.*, 34, L10404, doi:10.1029/2007GL029410.
- Scheidegger, A. E. (1974), *The Physics of Flow Through Porous Media*, 8, 3rd ed., University of Toronto Press, Toronto.

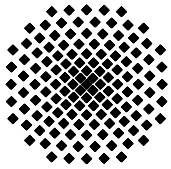


- Schwede, R. L., and O. a. Cirpka (2009), Use of steady-state concentration measurements in geostatistical inversion, *Adv. Water Res.*, 32(4), 607–619, doi: 10.1016/j.advwatres.2009.01.010.
- Schwede, R. L., and O. A. Cirpka (2010), Stochastic evaluation of mass discharge from pointlike concentration measurements, *J. Contam. Hydrol.*, 111(1-4), 36–47, doi: 10.1016/j.jconhyd.2009.10.011.
- Schwille, F., and J. F. Pankow (1988), *Dense Chlorinated Solvents in Porous and Fractured media - Model experiments*, 2nd ed., Lewis Publishers, Chelsea, MI.
- Sciortino, A., T. C. Harmon, and W. W.-G. Yeh (2000), Inverse modeling for locating dense nonaqueous pools in groundwater under steady flow conditions, *Water Resour. Res.*, 36(7), 1723–1735, doi:10.1029/2000WR900047.
- Scott, D. W. (1992), *Multivariate Density Estimation: Theory, Praticce, and Visualization*, 1st ed., 305 pp., John Wiley & Sons, New York, USA.
- Seyedabbasi, M. A., C. J. Newell, D. T. Adamson, and T. C. Sale (2012), Relative contribution of DNAPL dissolution and matrix diffusion to the long-term persistence of chlorinated solvent source zones, *J. Contam. Hydrol.*, 134-135, 69–81, doi:10.1016/j.jconhyd.2012.03.010.
- Shepherd, R. G. (1989), Correlations of permeability and grain size, *Ground Water*, 27(5), 633–638, doi:10.1111/j.1745-6584.1989.tb00476.x.
- Sherwood, T. K., R. L. Pigford, and C. R. Wilke (1975), *Mass Transfer*, 1 ed., McGraw-Hill, New York, USA.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, 1st ed., Chapman and Hall, London, UK.
- Skaggs, T. H., and Z. J. Kabala (1994), Recovering the release history of a groundwater contaminant, *Water Resour. Res.*, 30(1), 71–79, doi:10.1029/93WR02656.
- Skaggs, T. H., and Z. J. Kabala (1995), Recovering the history of a groundwater contaminant plume: Method of quasi-reversibility, *Water Resour. Res.*, 31(11), 2669–2673, doi: 10.1029/95WR02383.
- Smith, G. D. (1985), *Numerical Solution for Partial Differential Equations: Finite Difference Methods*, 1st ed., Oxford University Press, New York, USA.
- Snodgrass, M. F., and P. K. Kitanidis (1997), A geostatistical approach to contaminant source identification, *Water Resour. Res.*, 33(4), 537–546, doi:10.1029/96WR03753.
- Sobol', I. M. (1990), On sensitivity estimation for nonlinear mathematical models, *Math. Model.*, 2(1), 112–118.
- Soga, K., J. W. E. Page, and T. H. Illangasekare (2004), A review of NAPL source zone remediation efficiency and the mass flux approach., *J. Hazard. Mater.*, 110(1-3), 13–27, doi: 10.1016/j.jhazmat.2004.02.034.

- Sun, A. Y. (2007), A robust geostatistical approach to contaminant source identification, *Water Resour. Res.*, 43, W02418, doi:10.1029/2006WR005106.
- Sun, A. Y., S. L. Painter, and G. W. Wittmeyer (2006), A robust approach for iterative contaminant source location and release history recovery., *J. Contam. Hydrol.*, 88(3-4), 181–96, doi:10.1016/j.jconhyd.2006.06.006.
- Sun, N. Z. (1994), *Inverse Problems in Groundwater Modeling*, vol. 6, 1st ed., 337 pp., Kulwer Academic Publishers, Dordrecht, The Netherlands.
- Tartakovsky, D. M. (2007), Probabilistic risk analysis in subsurface hydrology, *Geophys. Res. Lett.*, 34, L05404, doi:10.1029/2007GL029245.
- Tennekes, H., and J. L. Lumley (1972), *A First Course in Turbulence*, 1st ed., MIT Press, Cambridge, Mass.
- Tobler, W. R. (1970), A computer movie simulating urban growth in the Detroit region, in *Economic Geography*, vol. 46, pp. 234–240, Clark University, Worcester, MA.
- Troldborg, M. (2010), Risk Assessment Models and Uncertainty Estimation of Groundwater Contamination from Point Sources.
- Troldborg, M., G. Lemming, P. J. Binning, N. Tuxen, and P. L. Bjerg (2008), Risk assessment and prioritisation of contaminated sites on the catchment scale, *J. Contam. Hydrol.*, 101(1-4), 14–28, doi:10.1016/j.jconhyd.2008.07.006.
- Troldborg, M., W. Nowak, N. Tuxen, P. L. Bjerg, R. Helmig, and P. J. Binning (2010), Uncertainty evaluation of mass discharge estimates from a contaminated site using a fully Bayesian framework, *Water Resour. Res.*, 46, doi:10.1029/2010WR009227.
- US Environmental Protection Agency (EPA) (2007), Trichloroethylene (TCE), TEACH Chemical Summary, *Tech. rep.*
- Van der Vorst, H. A. (1992), Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *J. Scien. Stat. Comp. (SIAM)*, 13(2), 631–644, doi:10.1137/0913035.
- Van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society America J.*, 44(5), 892–898.
- Wang, P. P., and C. Zheng (2005), Contaminant transport models under random sources, *Ground Water*, 43(3), 423–433, doi:10.1111/j.1745-6584.2005.0034.x.
- Weeks, S., and G. Sposito (1998), Mixing and stretching efficiency in steady and unsteady groundwater flows, *Water Resour. Res.*, 34(12), 3315–3322, doi:10.1029/98WR02535.
- Weiss, J., and A. Provenzale (2008), *Transport and Mixing in Geophysical Flow*, 1st ed., Springer Verlag, New York, USA.
- Weiss, N. A. (2006), *A Course in Probability*, 1 ed., Addison Wesley, Boston, USA.

- Wen, X. H., and J. J. Gómez-Hernández (1996), The constant displacement scheme for tracking particles in heterogeneous aquifers, *Ground Water*, 34(1), 135–142, doi:10.1111/j.1745-6584.1996.tb01873.x.
- White, F. (2003), *Fluid Mechanics*, 5th ed., 866 pp., McGraw-Hill, New York, USA.
- Wilkinson, D., and J. F. Willemsen (1983), Invasion percolation: a new form of percolation theory, *J. Phys. A: Math. Gen.*, 16, 3365–3376, doi:10.1088/0305-4470/16/14/028.
- Woodbury, A., E. Sudicky, T. J. Ulrych, and R. Ludwig (1998), Three-dimensional plume source reconstruction using minimum relative entropy inversion, *J. Contam. Hydrol.*, 32(1-2), 131–158, doi:10.1016/S0169-7722(97)00088-0.
- Yeh, H.-D., T.-H. Chang, and Y.-C. Lin (2007), Groundwater contaminant source identification by a hybrid heuristic approach, *Water Resour. Res.*, 43, W09420, doi:10.1029/2005WR004731.
- Zeng, L., L. Shi, D. Zhang, and L. Wu (2012), A sparse grid based Bayesian method for contaminant source identification, *Advan. Water Resour.*, 37, 1–9, doi:10.1016/j.advwatres.2011.09.011.
- Zhang, C., C. J. Werth, and A. G. Webb (2007), Characterization of NAPL Source Zone Architecture and Dissolution Kinetics in Heterogeneous Porous Media Using Magnetic Resonance Imaging, *Environ. Sci. Technol.*, 41(10), 3672–3678, doi:10.1021/es061675q.
- Zhang, C., H. Yoon, C. J. Werth, A. J. Valocchi, N. B. Basu, and J. W. Jawitz (2008), Evaluation of simplified mass transfer models to simulate the impacts of source zone architecture on nonaqueous phase liquid dissolution in heterogeneous porous media, *J. Contam. Hydrol.*, 102(1-2), 49–60, doi:10.1016/j.jconhyd.2008.05.007.
- Zienkiewicz, O. C., and R. L. Taylor (1977), *The Finite Element Method*, 3rd ed., McGraw-Hill, London, UK.
- Zinn, B., and C. F. Harvey (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields, *Water Resour. Res.*, 39(3), 1051, doi:10.1029/2001WR001146.





## Institut für Wasser- und Umweltsystemmodellierung Universität Stuttgart

Pfaffenwaldring 61  
70569 Stuttgart (Vaihingen)  
Telefon (0711) 685 - 64717/64749/64752/64679  
Telefax (0711) 685 - 67020 o. 64746 o. 64681  
E-Mail: [iws@iws.uni-stuttgart.de](mailto:iws@iws.uni-stuttgart.de)  
<http://www.iws.uni-stuttgart.de>

### Direktoren

Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy  
Prof. Dr.-Ing. Rainer Helmig  
Prof. Dr.-Ing. Silke Wieprecht

### Vorstand (Stand 19.08.2013)

Prof. Dr. rer. nat. Dr.-Ing. A. Bárdossy  
Prof. Dr.-Ing. R. Helmig  
Prof. Dr.-Ing. S. Wieprecht  
Prof. Dr. J.A. Sander Huisman  
Jürgen Braun, PhD  
apl. Prof. Dr.-Ing. H. Class  
Dr.-Ing. H.-P. Koschitzky  
Dr.-Ing. M. Noack  
Jun.-Prof. Dr.-Ing. W. Nowak, M.Sc.  
Dr. rer. nat. J. Seidel  
Dr.-Ing. K. Terheiden

### Emeriti

Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Jürgen Giesecke  
Prof. Dr.h.c. Dr.-Ing. E.h. Helmut Kobus, PhD

### Lehrstuhl für Wasserbau und Wassermengenwirtschaft

Leiter: Prof. Dr.-Ing. Silke Wieprecht  
Stellv.: Dr.-Ing. Kristina Terheiden  
**Versuchsanstalt für Wasserbau**  
Leiter: Dr.-Ing. Markus Noack

### Lehrstuhl für Hydromechanik und Hydrosystemmodellierung

Leiter: Prof. Dr.-Ing. Rainer Helmig  
Stellv.: apl. Prof. Dr.-Ing. Holger Class  
**Jungwissenschaftlergruppe: Stochastische  
Modellierung von Hydrosystemen**  
Leiter: Jun.-Prof. Dr.-Ing. Wolfgang Nowak, M.Sc.

### Lehrstuhl für Hydrologie und Geohydrologie

Leiter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy  
Stellv.: Dr. rer. nat. Jochen Seidel  
**Hydrogeophysik der Vadosen Zone**  
(mit Forschungszentrum Jülich)  
Leiter: Prof. Dr. J.A. Sander Huisman

### VEGAS, Versuchseinrichtung zur Grundwasser- und Altlastensanierung

Leitung: Jürgen Braun, PhD, AD  
Dr.-Ing. Hans-Peter Koschitzky, AD

## Verzeichnis der Mitteilungshefte

- 1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963
- 2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964
- 3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächen-tragwerke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Randwert- und Lastwertmatrizen*, 1965
- 4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

- 5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht, und*  
Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gesteigungskosten und jeweilige Vorteile, sowie Unny, T.E.: Schwingungsuntersuchungen am Kegelstrahlschieber, 1967*
- 6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschiffahrt an diesem Wert, 1967*
- 7 *Sonderheft anlässlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von* Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967
- 8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse, 1968*
- 9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund, 1968*
- 10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften, 1968*
- 11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem, 1969*
- 12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität, 1969*
- 13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschiffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle, 1969*
- 14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion, 1970*
- 15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper, 1970*
- 16 Mobasseri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standsicherheit, 1970*
- 17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasser-rückhaltebecken, 1970*

- 18 Gál, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, vergriffen
- 19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, vergriffen
- 20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971
- 21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gál, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972
- 22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972
- 23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972
- 24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, vergriffen
- 25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, vergriffen
- 26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972
- 27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972
- 28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschifffahrt durch Schwimmerhebwerke*, 1973
- 29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973
- 30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973
- 31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973
- 32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974

- 33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974
- 34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975
- 35 *Sonderheft anlässlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von:* Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975
- 36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotations-symmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975
- 37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976
- 38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976
- 39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6
- 40 Ruppert, Jürgen: *Über den Axialwirbelkammverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X
- 41 Hutarew, Andreas: *Beitrag zur Beeinflußbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, vergriffen
- 42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6
- 43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4
- 44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2
- 45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0



- 46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9
- 47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7
- 48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5
- 49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, vergriffen
- 50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwasserreinleitungen*, 1982, ISBN 3-921694-50-7, vergriffen
- 51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5
- 52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, vergriffen
- 53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, vergriffen
- 54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X
- 55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, vergriffen
- 56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6
- 57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, vergriffen
- 58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangbelasteter Baukörper*, 1985, ISBN 3-921694-58-2
- 59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, vergriffen
- 60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, vergriffen
- 61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2

- 62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebene Grundwassermodellen*, 1986, ISBN 3-921694-62-0
- 63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9
- 64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7
- 65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5
- 66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3
- 67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1
- 68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufererosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, vergriffen
- 69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, vergriffen
- 70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1
- 71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X
- 72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*, 1991, ISBN 3-921694-72-8
- 73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6
- 74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, vergriffen
- 75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2
- 76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0

- 77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, vergriffen
- 78 Rommel, Michael: *Verwendung von Kluftdaten zur realitätsnahen Generierung von Kluftnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7
- 79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen
- 80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen
- 81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern*, 1993, ISBN 3-921694-81-7
- 82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versucheinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5
- 83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen
- 84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1
- 85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karst-aquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, vergriffen
- 86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, vergriffen
- 87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation -*, 1996, ISBN 3-921694-87-6
- 88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4
- 89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, vergriffen
- 90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, vergriffen
- 91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, vergriffen

- 92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien*, 1997, ISBN 3-921694-92-2
- 93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, vergriffen
- 94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen* 1997, ISBN 3-921694-94-9
- 95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, vergriffen
- 96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5
- 97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, ISBN 3-921694-97-3
- 98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, vergriffen
- 99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden*, 2000, ISBN 3-933761-02-2
- 100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4
- 101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, vergriffen
- 102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0
- 103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9
- 104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7
- 105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien*, 2001, ISBN 3-933761-08-5
- 106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3
- 107 Josef, Reinhold.: *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7

- 108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5
- 109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erddämmen*, 2001, ISBN 3-933761-12-3
- 110 Lang, Stefan: *Parallele numerische Simulation instationärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1
- 111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X
- 112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten*, 2002, ISBN 3-933761-15-8, vergriffen
- 113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation*, 2002, ISBN 3-933761-16-6
- 114 Silberhorn-Hemminger, Annette: *Modellierung von Kluftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluftgenerierung*, 2002, ISBN 3-933761-17-4
- 115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen*, 2003, ISBN 3-933761-18-2
- 116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung*, 2003, ISBN 3-933761-19-0
- 117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water*, 2003, ISBN 3-933761-20-4
- 118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2
- 119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0
- 120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9
- 121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7
- 122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -*, 2003, ISBN 3-933761-25-5
- 123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance*, 2003, ISBN 3-933761-26-3

- 124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003*, 2003, ISBN 3-933761-27-1
- 125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip*, 2003, ISBN 3-933761-28-0
- 126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8
- 127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1
- 128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X
- 129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flusstauhaltungen*, 2004, ISBN 3-933761-32-8
- 130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Upscaling, Mass-Conservative Discretization and Efficient Multilevel Solution*, 2004, ISBN 3-933761-33-6
- 131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004*, Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4
- 132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2
- 133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers*, 2004, ISBN 3-933761-36-0
- 134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface*, 2005, ISBN 3-933761-37-9
- 135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7
- 136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers*, 2005, ISBN: 3-933761-39-5
- 137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments*, 2005, ISBN 3-933761-40-9

- 138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7
- 139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Sub-surface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5
- 140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media*, 2005, ISBN 3-933761-43-3
- 141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen*, 2005, ISBN 3-933761-44-1
- 142 Hundecha, Hirpa Yeshewatersfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X
- 143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser*, 2005, ISBN 3-933761-46-8
- 144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, 3-933761-47-6
- 145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden*, Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4
- 146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes*, 2006, ISBN 3-933761-49-2
- 147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebecken-systemen*, 2006, ISBN 3-933761-50-6
- 148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4
- 149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2
- 150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias: *VEGAS-Statuskolloquium 2006*, Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0
- 151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9
- 152 Fischer, Markus: *Beanspruchung eingeeerdeter Rohrleitungen infolge Austrocknung bindiger Böden*, 2006, ISBN 3-933761-55-7

- 153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5
- 154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models*, 2006, ISBN 3-933761-57-3
- 155 Bielinski, Andreas: *Numerical Simulation of CO<sub>2</sub> sequestration in geological formations*, 2007, ISBN 3-933761-58-1
- 156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssystems für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3
- 157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation*, 2007, ISBN 3-933761-61-1
- 158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines*, 2007, ISBN 3-933761-62-X
- 159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8
- 160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage*, 2007, ISBN 3-933761-64-6
- 161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4
- 162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2
- 163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0
- 164 Götzing, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9
- 165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias: *VEGAS – Kolloquium 2007*, Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7
- 166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0



- 167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9
- 168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7
- 169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling*, 2008, ISBN 3-933761-73-5
- 170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3
- 171 Papafotiou, Alexandros: *Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1
- 172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7
- 173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4
- 174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1
- 175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8
- 176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach*, 2008, ISBN 978-3-933761-80-4
- 177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1
- 178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen*, 2009, ISBN 978-3-933761-82-8
- 179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa*, 2009, ISBN 978-3-933761-83-5
- 180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding*, 2009, ISBN 978-3-933761-84-2
- 181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers*, 2009, ISBN 978-3-933761-85-9

- 182 Kopp, Andreas: *Evaluation of CO<sub>2</sub> Injection Processes in Geological Formations for Site Screening*, 2009, ISBN 978-3-933761-86-6
- 183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on CO<sub>2</sub> and water (two-phase) flow in porous media*, 2009, ISBN 978-3-933761-87-3
- 184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0
- 185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7
- 186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3
- 187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0
- 188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7
- 189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modeling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4
- 190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1
- 191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8
- 192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5
- 193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2
- 194 Bliedernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins*, 2011, ISBN 978-3-933761-98-9
- 195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -*, Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2010, ISBN 978-3-933761-99-6

- 196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia*, 2010, ISBN 978-3-942036-00-9
- 197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen*, 2010, ISBN 978-3-942036-01-6
- 198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins*, 2010, ISBN 978-3-942036-02-3
- 199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow*, 2011, ISBN 978-3-942036-03-0
- 200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation*, 2011, ISBN 978-3-942036-04-7
- 201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas*, 2011, ISBN 978-3-942036-05-4
- 202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen*, 2011, ISBN 978-3-942036-06-1
- 203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks*, 2011, ISBN 978-3-942036-07-8
- 204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology*, 2011, ISBN 978-3-942036-08-5
- 205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken*, Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2
- 206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour*, 2011, ISBN 978-3-942036-10-8
- 207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta*, 2011, ISBN 978-3-942036-11-5
- 208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam*, 2011, ISBN 978-3-942036-12-2
- 209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone*, 2012, ISBN 978-3-942036-13-9

- 210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking*, 2012, ISBN 978-3-942036-14-6
- 211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach*, 2012, ISBN 978-3-942036-15-3
- 212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media*, 2012, ISBN 978-3-942036-16-0
- 213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung*, 2012, ISBN 978-3-942036-17-7
- 214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish*, 2012, ISBN 978-3-942036-18-4
- 215 De Boer, Cjstmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1
- 216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7
- 217 Lebreuz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4
- 218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of CO<sub>2</sub> Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1
- 219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8
- 220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5
- 221 Walter, Lena: *Uncertainty studies and risk assessment for CO<sub>2</sub> storage in geological formations*, 2013, ISBN 978-3-942036-25-2
- 222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects*, 2013, ISBN 978-3-942036-26-9
- 223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes*, 2014, ISBN 978-3-942036-27-6
- 224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3
- 225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0

- 226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6
- 227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3
- 228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0
- 229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7
- 230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4
- 231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1
- 232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8
- 233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: [www.iws.uni-stuttgart.de](http://www.iws.uni-stuttgart.de) zur Verfügung.





